



**HAL**  
open science

## French compounds

Florence Villoing

► **To cite this version:**

Florence Villoing. French compounds. *Probus*, 2012, 24 (1), pp.29-60. 10.1515/probus-2012-0003 .  
hal-01361754

**HAL Id: hal-01361754**

**<https://hal.science/hal-01361754v1>**

Submitted on 7 Sep 2016

**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.

# French compounds

FLORENCE VILLOING

## *Abstract*

*This article focuses on compounding as a process of word formation within the theoretical framework of lexeme-based morphology. It provides a systematic analysis of the two types of compounding in French: native compounding, the main type, and neoclassical compounding, which is quite marginal. It presents the various rules: native compounds are prototypically constructed of two lexemes and form a third one; they are predominantly endocentric; the governing constituent and the compound head, if any, is on the left and controls the semantic relations between the two constituents, whether coordinated, attributive or subordinating. Neoclassical compounds are prototypically constructed of bound neoclassical elements and form adjectives; they are often exocentric; the governing constituent is on the right. Inflection in native compounds is complex. Several areas of the analysis remain unresolved, particularly regarding the boundaries between morphological/syntactic compounds.*

## **1. Introduction<sup>1</sup>**

Compounding is a productive morphological procedure in French that operates alongside derivation (suffixation, prefixation and conversion).

- 
1. This study has benefited from MorboComp, an international research project on compounds, devised and directed by Sergio Scalise, based at the University of Bologna, as well as from numerous discussions with Dany Amiot, Antonietta Bisetto, Françoise Kerleroux, and Fiammetta Namer, for whom I owe special thanks for reading and commenting on previous versions. I also extend all my thanks to Dana Cohen for her dynamic contribution in the elaboration of the English version and the reviewer of the first version of this paper for his useful remarks.

The term is not used consistently in the literature and designates various complex lexical units analyzed through a range of approaches. Overall, two broad approaches can be identified. The first considers any complex lexicalized unit composed of two or more terms as a compound word, regardless of the formation process (morphological or syntactic). The primary focus of this approach are the lexical properties of the compound word – their syntactic independence and integration into the lexicon. The second approach focuses on the compounding processes. This approach typically examines (i) the formation of syntactic and morphological compounds (Section 3.1), and within the morphological component, (ii) the boundary between compounding and derivation (Section 3.2).

This article belongs to the second approach. It focuses on compounding as a process of word formation within the theoretical framework of lexeme-based morphology, which considers morphology and syntax as two separate grammatical components. Only morphologically-formed compounds are considered in this study, excluding lexicalized complex syntactic sequences (for the criteria behind this distinction, see Section 3.1).

The purpose of this article is to present the morphological compounding rules of French, as well as their phonological, categorial and semantic constraints. The article provides a systematic analysis of two major types of compounding in French: native compounding, which is the main type, and neoclassical compounding, quite marginal. These types are distinguished through two central criteria – the type of input unit and the order of internal structure between the two constituents. A native compound (1) is prototypically formed of two lexemes of the current lexicon of French, without any linking element; the internal order of constituents is XY, where X is the governing element (underlined in the examples).<sup>2</sup>

- (1) *poisson-chat*    *wagon-fumeur*    *ouvre-boîte*  
 fish-cat            car-smoker            open-can  
 ‘catfish’            ‘smoking car’        ‘can-opener’

A neoclassical compound (2) is prototypically composed of two bases of Greek or Latin origin, that are not syntactically autonomous in French, connected by a linking element; the internal order of constituents is YX, where X is the governing element.

- (2) *ludo-thèque*    *homi-cide*    *cyno-céphale*  
 game-library    human-kill    dog-head

---

2. Abbreviations in the glosses: A = adj.; DF = definite; F = fem.; FUT = future; IMP = imperative; INF = infinitive; N = noun; M = masc.; P = preposition; PASS = passive; PL = plural; POSS = possessive; PRS = present; PST = past; SG = singular; V = verb; 1,2,3 = first, second, third person.

The article presents the properties of compounds, focusing on the units of input (Section 2) and output (Section 3), on the different construction patterns (Section 4.1) and on the semantic relations between constituents in each pattern (Section 4.2 and Section 5). The final section (Section 6) examines the relations between compounding and inflection.

## 2. Input: basic units

Guevara and Scalise's (2009) generalization that compounding is universally composed of two major lexical categories applies perfectly to French. Nevertheless, the literature on compounding traditionally questions the status of the units involved: word, lexeme, stem or root. Under the assumption that morphology and syntax are independent of each other, each of the two component uses its own distinct units as input, lexemes in morphology (abstract units of lexical structure) and words in syntax (elements with syntactically relevant properties). There is some consensus in considering lexemes as the prototypical building blocks of morphological compounding (Corbin 1992; Bauer 2001; Fradin 2003, 2009; Booij 2005, Scalise and Vogel 2010; Montermini 2010). Since lexemes are phonologically realized as one or more stems (Aronoff 1994; Stump 2001; Bonami and Boyé 2003), it is expected that the units that appear in compounds correspond to two stems of lexemes. Most French data conforms with this analysis. Nevertheless, many types of compounds display less canonical structures: components that are words (inflected units) rather than stems or even non-autonomous syntactic units, unconnected to any lexeme.

### 2.1. Lexemes

2.1.1. *Lexemes in native compounds.* Native compounds are formed of two lexemes, which belong, by definition, to the major categories (noun, verb, adjective), and are uninflected. The phonological form in which these lexemes appear within the compound corresponds to one of the lexeme stems. Thus, no constituent is marked by inflection; no modality, tense, person or aspect marking on the verb in VN compounds (3a), no number on the N (3b), and no gender or number on adjectives (3c), disregarding cases of agreement.

- (3) a. *ouvre<sub>V</sub>-boîte<sub>N</sub>*  
       'can-opener'  
     b. *homme<sub>N</sub>-grenouille<sub>N</sub>*  
       'frogman'

- c. *vert*<sub>A</sub>-*pomme*<sub>N</sub>  
‘apple-green’

Although widely accepted, this analysis is not self-evident, particularly since stems are typically homonyms of words in the inflectional paradigm of the lexeme. There is a long-standing debate on the nature of the verbal component of Verb-Noun compounds in French, for example, since its form is homonymous with words of the singular present tense or of the singular imperative. The first assumptions, inherited from the comparative grammarians of the 19<sup>th</sup> century, favored an inflected form of the verb, in the present indicative or in the imperative (see Villoing 1999, 2000). The development of the conceptual framework of lexeme-based morphology and the clear distinction drawn within this approach between lexemes, words and stems, allows us to clearly identify the verb in VN compounds as a stem of the lexeme (Corbin 1992; Villoing 2009; adopted by Bonami et al. 2009; Fradin 2009).

2.1.2. *Lexemes in neoclassical compounds.* The input units of neoclassical compounding are strictly morphological and have no syntactic realization (consequently leading to the variety of labels found in the literature (see Iacobini 2004)). These components lack autonomy due to two factors:

- (i) Components borrowed from a classical source, primarily Greek (4a), but also Latin (4b), and remain more or less homomorphic with their Greek or Latin etymon (‘*archéoconstituants*’ in Corbin and Paul 1999). These elements appear both in compounding and (sometimes) in derivations (4c).
- (4) a. [*micro*][*céphale*] [*méga*][*lithe*]  
small<sub>A</sub>-head<sub>N</sub> large<sub>A</sub>-stone<sub>N</sub>  
b. [*hom*]*i*[*cide*] [*aqu*]*i*[*feré*]  
human<sub>N</sub>-kill<sub>V</sub> water<sub>N</sub>-contain<sub>V</sub>  
c. [*aquat*]*ique* [*phob*]*ie* [*céphal*]*ée*  
water<sub>N</sub>-SUF fear<sub>N</sub>-SUF head<sub>N</sub>-SUF ‘headache’
- (ii) Some components correspond to truncated French lexemes (*fractoconstituants* in Corbin and Paul 1999; *combining forms* in English references).<sup>3</sup> Thus, truncated non-autonomous constituents function as

---

3. Fracto-constituents must be distinguished from suffixoids (‘secreted affixes’) and from constituents appearing in blends (see Section 3; cf. Corbin and Paul 1999; Fradin 2000) and also from lexemes derived through truncation, which, unlike fracto-constituents, are syntactically independent, like all French lexemes (*auto*<sub>N</sub>, truncation of *automobile* ‘car’; *micro*<sub>N</sub>, truncation of *microphone*; *photo*<sub>N</sub>, truncation of *photographie* ‘photography’; *télé*<sub>N</sub>, truncation of *television* ‘TV’).

representations of the lexeme in the compound. This is illustrated by final *-ol* for *alcool* (5a) and *-one* (5b) representing *hormone* in chemical terminology, and in the non-scientific domain, initial *euro-* for *Europe*, *péto-* for *pétrole*, *afro-* for *africain*, *socio-* for *social*, *anarcho-* for *anarchiste* (6), final *-tique* for *informatique*, and *-ciel* for *logiciel* (7).

- (5) a. *menthol, éthanol*  
b. *cortisone, aestrone*
- (6) *eurodéputé* ‘euro-MP’, *pétrochimie* ‘petrochemistry’, *afro-cubain*, *socio-culturel*, *anarcho-capitaliste*
- (7) *bureautique* ‘office automation’, *robotique* ‘automation’, *didacticiel* ‘educational software’, *ludiciel* ‘game software’

Here too, in the scholarly lexicon, truncated lexemes can serve as the basis for affixation: e.g., *phosphate*, *phosph-* for *phosphore*, *hydrure* ‘hydride’, *hydr-* for *hydrogène* ‘hydrogen’ (cf. Cottez 1988).

The non-autonomous status of these constituents of neoclassical compounds reawakens the question regarding their linguistic status. Can these components be analyzed as lexemes? Since fracto-constituents are related to French lexemes phonologically, categorially and semantically, they are treated as one of the forms of French lexemes. Within a lexeme-based morphology, these are stems of lexemes, whether suppletive or not (Montermini 2010). Some archeo-constituents are treated in the same way. Many of them can be analyzed as allomorphic or suppletive bases of French lexemes. As such, they are stems of lexemes. But what status should be attributed to archeo-constituents that cannot be linked to any French lexeme (for example, because their semantic values do not match any French lexeme: *électr-* referring to electrical energy, *graph-* referring to multiple semantic values that cannot be combined under the lexeme *écrire* ‘write’). Can we analyze such archeo-constituents as lexemes despite their lack of syntactic independence? This approach can be adopted, following Corbin (1985, 1992), Fradin (2000, 2003), Iacobini (2004), Amiot and Dal (2007), Namer (2009), Montermini (2010), given that archeo-constituents do behave like lexemes: (i) they have a referential capacity to denote entities, properties and relations (Namer 2009); (ii) they belong to one of the major categories noun, verb or adjective. These would be particular lexemes, specified as [+bound] in their lexical representation (Corbin 1992; followed by Montermini 2010).

While neoclassical compounding prototypically operates on non-autonomous bases, it can also create compounds from French lexemes. Therefore, identification of neoclassical compounding relies on the second criterion, the internal order of components (the reverse to the order of native compounding).

- (8) *insecticide*<sub>A/N</sub>,<sup>4</sup> *macromolécule*<sub>N</sub>, *alcoolothérapie*<sub>N</sub> ‘alcohol<sub>n</sub>-therapy<sub>n</sub>’,  
*anarcho-capitaliste*<sub>A</sub>, *droito-faciste*<sub>A</sub> ‘right-wing fascist’

## 2.2. Words

However, some compounds reveal inflected forms of the lexeme. This is illustrated by VN compounds that include a plural N, marked phonologically (9a) and in writing (9b):

- (9) a. *pique-bœufs*<sub>N</sub>, *protège-yeux*<sub>N</sub> *coupe-œufs*<sub>N</sub>  
 pick<sub>v</sub>-bull<sub>N,PL</sub> protect<sub>v</sub>-eye<sub>N,PL</sub> cut<sub>v</sub>-eggs<sub>N,PL</sub>  
 ‘oxpecker’ ‘eye-protector’ ‘egg-cutter’  
 b. *sèche-mains*<sub>N</sub>, *ramasse-miettes*<sub>N</sub>, *presse-fruits*<sub>N</sub>  
 dry<sub>v</sub>-hand<sub>N,PL</sub> collect<sub>v</sub>-crumb<sub>N,PL</sub> press<sub>v</sub>-fruit<sub>N,PL</sub>  
 ‘hand-dryer’ ‘crumb-collector’ ‘juicer’

These words are not the result of syntactic marking, but of inherent inflection that is semantically motivated (cf. Booij 1996, 2005). In other words, the plural in (9) depends either on the lexicon, where the compound is listed in this form (cf. *protège-yeux* nonexistent with a singular noun \**protège-oeil*) or on the speaker’s choice to mark a semantic value associated with a plurality (signaling the objects denoted by N as plural).

## 2.3. Linking elements

In addition, neoclassical compounds are generally characterized by the presence of linking elements, or “a liaison vowel” /i/ or /o/, between the two components of the compound, which appears in certain phonological contexts (end of the first constituent and the initial of the second constituent); /i/ appears in Latinate components (10a) and /o/ in Greek components (10b) or if it corresponds to a fracto-constituent of a French lexeme (10c), /i/ or /o/ (the most common) when the components are French lexemes or proper nouns (10d). These linking elements are not necessary in other contexts, regardless of the nature of the constituents (11):

- (10) a. *insecticide*  
 b. *anglophone*  
 c. *afro-cubain*  
 d. *riziculture*, *alcoolothérapie*, *sarkophobe*  
 (11) [*méga*][*lithe*], [*quadru*][*pède*], [*ciné*][*phile*], [*télé*][*vore*]

4. Examples that lacks glosses correspond to the same lexeme in English.

### 3. Output: delimitation

#### 3.1. Compounding vs. lexicalized phrases

Within a modular approach to grammar, where morphology and syntax are two autonomous components, a number of criteria have been proposed to distinguish between morphological compounding and syntactic compounding (cf. Scalise 1984, 1992; Matthews 1991; Anderson 1992; Aronoff 1994; Corbin 1992, 1997; Bisetto and Scalise 1999; Fradin 2003, 2009; a.o.). Here, we adopt the distinguishing criterion proposed by Corbin (1992, 1997) and adopted by Fradin (2003, 2009): compounding is a lexical morphological operation if it obeys the lexical constraints that organize the morphological component. Thus, a sequence of words is a morphological compound if it is a lexeme (noun, verb or adjective) constructed from other lexemes, according to a non-syntactic mode of organization. It gains the status of a lexical unit. Therefore, the following are not considered compound (contra Zwanenburg 1992):

- (i) complex units composed of non-lexemes, such as complex prepositions (12a) and complex conjunctions (12b);
  - (ii) lexicalized phrases that behave like lexical units (13);
  - (iii) lexicalized syntactic constructions that behave like lexical units: NPs (14), PPs (15), VPs (16);
  - (iv) lexicalized syntactic constructions – NPs of the structure NA (17a) or AN (Old French) (17b) – that behave like lexical units;
  - (v) NPs of the structure NN (18).
- (12) a. *par-dessus*      *en dehors de*  
           ‘from above’    ‘apart from’
- b. *de sorte que*    *bien que*  
           ‘such that’     ‘albeit’
- (13) *rendez-vous*<sub>N</sub>            *qu’en-dira-t-on*<sub>N</sub>  
 go.<sub>2IMP</sub>-you                what-P-say.<sub>3FUT</sub>-pron  
 ‘appointment, date’    ‘gossip’
- (14) *fil de fer*<sub>N</sub>            *brosse à dents*<sub>N</sub>    *arc-en-ciel*<sub>N</sub>  
 wire.<sub>N</sub>-P-iron.<sub>N</sub>    brush.<sub>V</sub>-P-teeth.<sub>N</sub>    arc.<sub>N</sub>-P-sky.<sub>N</sub>  
 ‘iron wire’        ‘toothbrush’        ‘rainbow’
- (15) *sous verre*<sub>N</sub>            *sans-papier*<sub>N</sub>,  
 under.<sub>P</sub>-glass.<sub>N</sub>        without.<sub>P</sub>-paper.<sub>N</sub>  
 ‘coaster’                ‘illegal immigrant’
- (16) *boit-sans-soif*<sub>N</sub>            *va et vient*<sub>N</sub>  
 drink-without-thirst    go-and-come  
 ‘drunk’                    ‘back and forth’



- (17) a. *poids lourd*      *chambre froide*      *coffre-fort*  
 weight<sub>N</sub>-heavy<sub>A</sub>      room<sub>N</sub>-cold<sub>A</sub>      box<sub>N</sub>-strong<sub>A</sub>  
 ‘truck’      ‘cold-storage’      ‘safe’
- b. *moyen âge*      *rouge-gorge*      *chauve-souris*  
 middle<sub>A</sub>-age<sub>N</sub>      red<sub>A</sub>-throat<sub>N</sub>      bald<sub>A</sub>-mouse<sub>N</sub>  
 ‘middle ages’      ‘robin’      ‘bat’
- (18) *avocat ami*      *opération prix*      *case départ*  
 lawyer friend      operation price      case departure  
 ‘lawyer friend’      ‘transaction price’      ‘square one’

In contrast, the VN, NN, AA and AN compounds presented in Section 4 are morphologically-formed, as they meet the morphological properties of formation and exhibit syntactic anomalies: in VN compounds, the absence of a determiner between the verb and the noun, and a diverse range of semantic relations between the verb and noun (*ouvre-boîte* ‘can opener’), the absence of a coordinating conjunction between the constituents in coordinated NN (*horloger-bijoutier* ‘jeweler-watchmaker’) and AA (*aigre-doux* ‘sweet and sour’) compounds, hyponymic interpretation in all other NN compounds (*poisson-chat* ‘catfish’, *pause-café* ‘coffee-break’), the presence of an adjectival rather than a nominal head in AN compounds (*bleu-ciel* ‘sky-blue’).

Nonetheless, the boundary between the two domains is harder to establish in NN sequences, since these can result from both syntactic and morphological processes. Several criteria have been proposed to distinguish the two types (David 1993; Rainer and Varela 1992; Noailly 1990; a.o.). The most convincing are based on the interpretative patterns associated with NN sequences, that are less constrained in syntactic structures. Thus, Corbin (1992, to appear), Fradin (2003, 2009) and Van Goethem and Amiot (2009) show that, unlike NN compounds, the constituents of syntactic NN sequences (18) have no lexical meaning and are not interpreted as hyponymic or coordinated. However, this criterion is difficult to apply (as illustrated by the various analyses for *roman photo*: a syntactic structure in Fradin 2009, a morphological compound in Corbin to appear) and the diversity of attested types constantly leads to its reexamination.

### 3.2. Compounds output: lexemes?

Morphological compounds are prototypically defined as the combination of phonological, syntactic and semantic properties of two lexemes (Lx1 and Lx2), forming a third (Lx3). However, the output of morphological compounds sometimes present behaviors that are not typical of lexemes. For example, Gaeta and Ricca (2009) have shown that in many cases the objects formed by

compounding are not lexemes in the sense that they are not going to lexicalise (such as evaluative derivatives used with pragmatic function in French). This is the case with french compounds in which the interpretation is context dependent or related to a particular situation. For example, the VN compound *porte-chapeau* ‘wear-hat’ may refer to a person that wear an hat in a particular situation; the neoclassical NN compound *obamania* refers to the current enthusiasm for U.S. President; neither does neologisms lexicalised. Another example are compounds presenting some typical behaviors of syntactic objects (Lieber and Scalise 2007), revealing that the border between the different components of grammar are sometimes fuzzy. This is particularly striking with Romance VNs (Ricca 2010) and french compounds in (19) and (20):

- (19) *fou*re-*tout*<sub>N</sub>    *mang*e-*tout*<sub>N</sub>    *gu*é*rit*-*tout*<sub>N</sub>  
 fill-all            eat-all            heal-all  
 ‘hold-all’        ‘type of bean’    ‘heal-all’
- (20) *couche*-*tard*<sub>N</sub>    *lève*-*tôt*<sub>N</sub>        *couche*-*dehors*<sub>N</sub>    *pass*e-*partout*<sub>N</sub>  
 sleep-late        rise-early        layer-outside        pass-everything  
 ‘night owl’    ‘early riser’    ‘external layer’    ‘master-key’

These structures involve a verb and a second element (an indefinite pronoun in (19), an adverb in (20)) and could be readily classified as lexicalized phrases, given their compliance with syntactic structure<sup>5</sup> (Fradin 2003, 2009). However, they are best treated as a morphological compound, given their conformity with the semantics of VN compounds (Section 4.2.3.2; see Villoing 2009; Corbin to appear). Indeed, the complex sequences (19) and (20) involve the same semantic types as the VN compounds (nouns designating objects), and denote a referent through marking a salient property taken as habitual or functional. In addition, the semantic relations between the verb and the pronoun (19) conform to the predicate-patient relations typical of VN compounding, and the semantic relations between verb and adverb (20) are not so distinct, since a predicate-modifier relation can be said to hold, as the adverb indicates location (*couche*-*dehors*, like in the VN compound *traîne*-*buisson*), or temporality (*lève*-*tôt*, *couche*-*tard*, like in the VN compound *réveille*-*matin*). We can conclude with Corbin (to appear) that the absence of a syntactic anomaly in complex lexicalized sequences is not sufficient to rule out morphological compounding and prefer a syntactic analysis.

5. Assuming that the reflexive pronoun was historically lost in the verbs *se coucher* and *se lever* during the lexicalisation process, although it would have been expected to remain (<sup>o</sup>*un se couche-tard*, <sup>o</sup>*un se lève-tôt*), contrary to morphologically derived nominals constructed on the basis of a pronominal verb that systematically lose its pronoun (*s'autodétruire* > *autodestruction*; *se désister* > *désistement*; *s'empreser* > *empressement*; *s'évanouir* > *évanouissement*).

3.3. *Compounding vs. derivation*

The boundary between compounding and derivation is a recurring issue in defining processes of morphological construction. The traditional criterion is the autonomy of the constituents: compounds are composed of two free forms and derivations are composed of a free form and a bound form. This criterion raises many well-known problems (e.g., Scalise 1984, 1992; Corbin 1985).

The first problem concerns bound forms – whether archeo-constituents or fracto-constituents – in neoclassical compounds. According to the boundedness criterion, these constituents must be analyzed as affixes (Williams 1981; Bauer 1979). However, neoclassical elements can combine with each other (Section 4.1.2.2); if they are analyzed as affixes, such sequences create prefix + suffix constructions without a lexical base – forms that defy the most fundamental definition of derivation. In response to this difficulty, many studies have begun to recognize that neoclassical elements include some elements that must be analyzed as prefixes and others which must be considered lexical. Various criteria have been proposed to distinguish affixes from lexical archeo-constituents (Amiot and Dal 2007; Corbin 1985; Fradin 2000, 2003; Iacobini 1999, 2004; Namer 2009; Warren 1990). Here, we focus on the most central criteria: (i) structural properties (an archeo-constituent can occupy both positions of a compounding pattern and can serve as a basis for affixation; both are impossible for an affix); (ii) semantic properties (an archeo-constituent has a denotational semantic content of a descriptive nature while an affix is an exponent of lexeme construction rule that has a functional, instructional meaning). But these criteria remain difficult to use and the difference between affix and lexical archeo-constituent is very fuzzy. For example, according to these criteria, Corbin (1992, 2001) distinguish two statuses for *micro-* and *nano-* in French: the prefixes *micro-* and *nano-* attach to a measurement noun with an instructional meaning indicating that the measurement is divided by  $10^6$  for *micro-* (21a) and  $10^9$  for *nano-* (21b), while the archeo-constituents *micro*<sup>-6</sup> and *nano*- have an adjectival sense ‘small’ (22a and b). However, this criterion of adjectival meaning ‘small’ does not imply that the constituent is an adjective: small dimension may be also be express by suffixes in French (cf. *-ette* in *maisonnette* ‘small house’).

- (21) a. *micro-seconde, micro-ampère*  
 b. *nanofarad, nanoseconde*
- (22) a. *micro-organisme, microcéphale*  
 b. *nanocephale, nanocormie*

---

6. Amiot and Dal (2007) go further and show that *micro-* in (21a) also functions like a prefix.

The boundedness criterion is also problematic with respect to free forms in the process of grammaticalisation. Some constructed lexemes, that are readily accepted as compounds at a first glance, exhibit recurring components on the right or left, the status of which is problematic; these may correspond to constituents of compounds (lexemes) to affixes. According to the semantic criterion (no conceptual meaning), Corbin (1992, 2001) has shown that the French lexemes *petit*, *grand*, *beau* and *arrière* have been grammaticalised in certain contexts where they function as prefixes creating family terms: *grand* (23) reflects a relation of ancestry, *petit* (24) a relation of descent, *beau* (25) indicates a relation through marriage and the adverb *arrière* (26) allows the construction of terms designating individuals in further generations beyond grandparents.

- (23) *grand-parents*    *grand-mère*    *grand-oncle*  
 ‘grandparents’    ‘grandmother’    ‘great-uncle’
- (24) *petits-enfants*    *petite-fille*    *petit-fils*  
 ‘grandchildren’    ‘granddaughter’    ‘grandson’
- (25) *beau-père*    *belle-famille*    *beau-frère*  
 ‘stepfather’    ‘step-family’    ‘step-brother’
- (26) *arrière-grands-parents*    *arrière-petite-fille*  
 ‘great grandparents’    ‘great-granddaughter’

On the other hand, Van Goethem and Amiot (2009) and Amiot and Van Goethem (2010) show, based on typical affix properties, that the form *bébé* in (27) can be considered a prefixoid and is undergoing grammaticalisation, while *fleuve* in (28) remains a noun.

- (27) *bébé-phoque*    *bébé-voiture*    *bébé-maison*  
 ‘baby-seal’    ‘baby-car’    ‘baby-house’
- (28) *roman-fleuve*,    *discours fleuve*  
 ‘very long novel’    ‘lengthy discourse’

Lesselingue and Villoing (2002) and Villoing (2002) also show that complex lexemes like *aide*-N can be divided into two sets: VN compounds (29) and prefixes with the prefixoid *aide*- (30).

- (29) *aide-mémoire*    *aide-ouïe*    *aide-nourrice*  
 ‘memorandum’    ‘hearing aide’    ‘nurse aide’
- (30) *aide-infirmière*    *aide-jardinier*    *aide-maçon*  
 ‘nursing assistant’    ‘gardner’s assistant’    ‘mason’s assistant’

3.4. *Compounding vs. blending*

Neoclassical compounding can form complex lexical units with fracto-constituents by exploiting the semantic relations between constituents of the same type (such as hyponymy), a property shared with blends (31). This fact poses a problem in delineating the boundary between the two processes.

(31) *informatique, franglais, photocopillage*

Following various analyses on French blending phenomena (Grésillon 1984; Fradin 1997, 2000; Corbin to appear), we can establish a number of differences, listed below.

- Blends are necessarily composed of two truncated constituents; this is a defining property for such words, and rarely appears in neoclassical compounds.
- The truncation of constituents in blends is constrained by the phonological segment shared by both constituents (32), which is not the case for scholarly compounds.

(32) *informatique* = informma(tion) (auto)matique; *FRANGLAIS* = fran (çais) + (an)glais; *photocopillage* = photocoppie + pillage

- The interpretation of neoclassical compounds is related to the position of the governing constituent, as it is in native compounding. This is not the case with blendings.
- Blends are limited in their categorial combinations to NN and AA sequences, while neoclassical compounds show a wider range.
- Blends are not formed through regular productive patterns (unlike neoclassical compounds); they are constructed one at a time and appear primarily in written texts; they are the result of creative language games, hence their traditional assignment to extra-grammatical morphology.

4. **Types of compounds**

Every compounding rule involves a categorical pattern (Section 4.1) and a semantic relationship between the components (coordination, subordination, attributivity, Section 4.2). In some cases, the reference and syntactic properties of the compound may be deduced from one of the constituents (endocentric compounds) and in others they cannot (exocentric compounds) (Section 5.). Each of these elements is discussed below, systematically distinguishing native compounding from neoclassical compounding.

## 4.1. Compounding patterns

4.1.1. *Native compounding.* As noted in Section 2.1, French native compounds take the major lexical categories (noun, verb, adjective<sup>7</sup>) as input and produce nouns (output of the most productive patterns) as well as adjectives. Table 1 below presents all productive patterns of French native compounding.

Table 1. *Productive patterns of native compounds*

		Input 2		
		N	A	V
Input 1	N	<i>horloger-bijoutier<sub>N</sub></i> <i>poisson-chat<sub>N</sub></i> <i>timbre-poste<sub>N</sub></i>	–	–
	A	<i>gris ardoise<sub>A</sub></i> <i>rouge-brique<sub>A</sub></i> <i>vert Véronèse<sub>A</sub></i>	<i>aigre-doux<sub>A</sub></i> <i>gris-bleu<sub>A</sub></i> <i>sourd-muet<sub>A</sub></i>	–
	V	<i>lèche-vitrine<sub>N</sub></i> <i>ouvre-boîte<sub>N</sub></i> <i>casse-pieds<sub>A</sub></i>	–	–

The most productive patterns in contemporary French are the  $VN_{A/N}$  and  $NN_N$ , and to a lesser degree,  $AA_A$  and  $AN_A$  (Corbin 1997, 2004; Fradin 2003, 2009). The endocentric<sup>8</sup>  $NN_N$ ,  $AA_A$ , and  $AN_A$  patterns produce only words that match the category of the head component (underlined henceforth): the  $\underline{NN}_N$  pattern produces nouns, the  $\underline{AA}_A$  and  $\underline{AN}_A$  patterns produce adjectives. Only the exocentric VN pattern can produce both nouns (its primary output; e.g. *ouvre-boîte* ‘can-opener’) and adjectives (some of which can be used as nouns too; e.g., *casse-pieds* ‘a bore’). Compounding forms lexemes, and native compounds can themselves serve as the input for new compounding processes (*porte<sub>V</sub>-cure-dents<sub>N</sub>* ‘toothpick-holder’).

Category combinations are restricted, and not all are possible. Combinations that do not appear in the table are unacceptable for one of two reasons:

- (i) The resulting compounds correspond to lexicalized syntactic constructions (as noted in the distinction between syntax and morphology drawn in the introduction).

7. The use of adverbs as input for compounding in French remains an open question (see Section 3.2).

8. For discussion of endocentricity and exocentricity, cf. Section 5.

- (ii) The units stem from patterns that are no longer productive (illustrated in Table 2).

Table 2. *Non-productive patterns of native compounds*

		Input 2		
		N	A	V
Input 1	N	–	–	<i>[main][tenir]<sub>V</sub></i> <i>[sau][poudrer]<sub>V</sub></i>
	A	–	<i>nouveau-né<sub>A/N</sub></i>	<i>[clair][semer]<sub>V</sub></i>
	V	–	<i>gagne-petit<sub>A/N</sub></i> <i>pête-sec<sub>A/N</sub></i>	<i>cache-cache<sub>N</sub></i> <i>pousse-pousse<sub>N</sub></i>
	ADV	<i>[mal][chance]<sub>N</sub></i> <i>[mal][heur]<sub>N</sub></i>	<i>[bien][heureux]<sub>N</sub></i>	<i>[mal][traiter]<sub>V</sub></i> <i>[mal][mener]<sub>V</sub></i>

4.1.2. *Neoclassical compounding.* Neoclassical compounding flourished in Europe of the 18th and 19th centuries, with the creation of modern scientific nomenclature, in medicine, chemistry, botany and other specialized jargons, such as the language of law, or of economy (Fradin, 2003; Iacobini, 2004; Dardano, 2003; and see Namer, 2009 for an overview). Scholars of the period noted that the formation of scholarly terminology was based on Greek and Latin compounding, in which these scholars were proficient (cf. Benveniste, 1974; Cottez, 1988). Scholarly terminology of this type was initially created in French, based on the standard listed in Diderot and D’Alembert’s encyclopedia, then borrowed into other languages (primarily Romance, Germanic and Slav languages; cf. Darmsteter 1877), and adapted to their respective phonological systems (cf. It. *biometrica*; Fr. *biométrie*; Eng. *biometrics*; Gr. *Biometrie*; Sp. *biometria*; Iacobini 2004: 69). Scientific terminology entered the general lexicon through the distribution of scientific knowledge in education and the media, thanks to the interest in the scientific disciplines. Although some neoclassical compounds are borrowed from ancient languages (cf. Greek and Latin: *philosophe*, *quadrupede*, *homicide*), the modes of neoclassical compounding are now part of the morphological system of French (but remains marginal compare with native compounds) and enable the creation of a few lexemes that go beyond the scientific nomenclature (cf. *cinéophile*, *télévore*, *insecticide*, *anglophone*) (Fradin 2003). The rules of neoclassical compounding are clearly distinguished from native compounding and from compounding in the languages of their origin (cf. the association of constituents from different origins, alterations of the rules of Greek or Latin compounding; e.g., Benveniste (1974: 163–170) on the construction of *microbe*, Darmsteter (1894:

253) on the construction of *oxygène* and *hydrogène*, Rainer (2008) on the construction of *neo-*).

Establishment of the patterns of neoclassical compounding requires the clarification of the category of archeo-constituents. The syntactic dependence of elements in this category prevents the use of the standard criteria of syntactic distribution and inflectional marking. However, the majority of studies generally distinguish archeo-constituents as a category of their own. Several criteria have been proposed (cf. Namer and Villoing 2005 for a summary), the combination of which enables us to analyze much of the relevant data (although the border between neoclassical lexeme and affix remains difficult in establishing).

The table below presents the possible patterns of French neoclassical compounding used in standard language.<sup>9</sup>

Table 3. Regular patterns of neoclassical compounding in French

		Input 2				
		N		A	V	
Input 1	N	<i>cynocephale</i> <sub>A</sub> <i>androgynie</i> <sub>A</sub> <i>théiforme</i> <sub>A</sub>	<i>ludothèque</i> <sub>N</sub> <i>velodrome</i> <sub>N</sub> <i>alcootest</i> <sub>N</sub>	<i>photosensible</i> <sub>A</sub> <i>seropositif</i> <sub>A/N</sub>	<i>xenophobe</i> <sub>A/N</sub> <sup>10</sup> <i>anglophile</i> <sub>A/N</sub> <i>televore</i> <sub>A</sub>	<i>homicide</i> <sub>A/N</sub> <i>océanographe</i> <sub>A/N</sub>
	A	<i>macrocéphale</i> <sub>A</sub> <i>rectangle</i> <sub>A</sub> <i>longiligne</i> <sub>A</sub>	<i>mégalithe</i> <sub>N</sub> <i>microorganismes</i> <sub>N</sub>	<i>afro-cubain</i> <sub>A</sub> <i>anarcho-capitaliste</i> <sub>A</sub> <i>médico-social</i> <sub>A</sub>		
	V	–	–	–	–	–

Neoclassical compounding can combine archeo-constituents (of the same origin: *macro*<sub>greek</sub>*céphale*<sub>greek</sub>, *homi*<sub>latin</sub>*cide*<sub>latin</sub>, or another: *ludo*<sub>latin</sub>*thèque*<sub>greek</sub>), archeo-constituents and fracto-constituents (*lud*<sub>latin</sub>*iciel*, *euro-phile*<sub>greek</sub>), archeo- or fracto-constituents and French lexemes (*photo*<sub>greek</sub>*sensible*<sub>fr</sub>, *vélo*<sub>fr</sub>*drome*<sub>greek</sub>, *anarcho-capitaliste*<sub>fr</sub>, *robo*<sub>fr</sub>*tique*) and (rarely) french lexemes (*alcool*<sub>fr</sub>*o-thérapie*<sub>fr</sub>, *droit*<sub>fr</sub>*o-fasciste*<sub>fr</sub>). The only option that is not attested seems to be the combination of two fracto-constituents (cf. Corbin, to appear).

9. Scientific terminology can employ very specific modes of formation that differ widely from the patterns of neoclassical compounding found in standard language (see Namer 2009 for a review and additional references). Such constructions are related to specific terminological needs of particular scientific disciplines (Corbin and Paul 1999; Iacobini 2004), and are not discussed here.

10. The notation <sub>ADJ/N</sub> or <sub>N/ADJ</sub> indicates that the resultant compounds function as adjectives or nouns depending on context; the sequence indicates the more frequent use.



All major categories are possible as input (noun, verb, adjective). Almost all combinations are possible with nouns and adjectives, but verbal compounds are only possible if the verb is the second component (serving as governor)<sup>11</sup>. As for output categories, neoclassical compounding systematically produces adjectives (regardless of the input categories), as well nouns (sometimes both, depending on context), with the exception of the AA<sub>A</sub> pattern, which produces only adjectives. No pattern produces compound verbs<sup>12</sup>.

#### 4.2. *Classification: subordination, attributivity, coordination*

Semantic relations between compound constituents are varied. Conventionally, the semantic relations of components are marked by letters: X marks the semantically (and syntactically, in endocentric compounds, cf. Section 5) governing element and Y marks its governed counterpart.<sup>13</sup> Hierarchically structured compounds exhibit two possible sequences: the native French sequence is always XY, the governor preceding the governed constituent. Neoclassical compounding exhibits the reverse sequence, YX. Compounds with semantically equal constituents are represented as XX'.

The semantico-grammatical relations between constituents have been the topic of many studies and various classifications have been proposed. Bisetto and Scalise (2005) present a critical discussion and propose a classification based on two related criteria, the semantico-grammatical criterion and the presence of the head. They distinguish three main types of constituent relations: coordination, attribution and subordination. We follow this classification in examining the constituent relations of native and neoclassical compounds.

4.2.1. *Coordinated compounds*.<sup>14</sup> The denotation of a compound is the sum of the denotations of both constituents, each having its own referent. There is no dependence relation between the constituents, and their semantic significance is analogous. The sequence of components is not related to their semantics, but based on prosodic or pragmatic constraints.

(33) XX'<sub>X</sub> compounds:  $X1 \cup X2 = X3$

11. VN compounds are all borrowed from Greek, sometimes via Latin [*mis*][o][*gyne*], [*mis*][*anthrope*], [*phil*][o][*sophe*], [*phil*][*anthrope*]. Some are copies of these borrowed terms, such as [*phil*][o][*russe*], [*phag*][o][*cyte*] 'cell eating'.

12. Verbs such as [*hydro*][*masser*], [*photo*][*protéger*], [*aéro*][*glisser*], [*thermo*][*réagir*] are analysed by Namer (to appear) as back-formation, based on a neoclassical NN pattern ([*hydro*][*massage*], [*photo*][*protection*], [*aéro*][*glissement*] 'hovercrafting', [*thermo*][*réaction*]).

13. In other terminology, X is labeled the determiner (*le déterminant*) and Y is the determined (*le déterminé*).

14. Also known as "additives" or classified as "*dvandva*" according to the traditional terminology of Sanskrit grammar (used by Benveniste).

4.2.1.1. Coordinated compounds of the type NN. Real coordinated  $NN_N$  compounds are rare in French (Fradin 2009). These are typically compounds that denote countries (34a) and fields of study (34b).

- (34) a. *Bosnie-Herzégovine*<sub>N</sub>  
 b. *histoire-géographie*<sub>N</sub> *latin-grec*<sub>N</sub>

These true coordinated compounds contrast with coordinated compounds as in (35), which denote not the sum of two referents, but a single referent combining the properties of both components. In this case, the denotation of the compound corresponds to the intersection of the denotations of both constituents ( $X1 \cap X2 = X3$ ). Thus, a *horloger-bijoutier* is a *horloger* who is also a *bijoutier*, while the field of *histoire-géographie* (34b) includes both history and geography. Nouns of this type typically refer to social activities (35a), machines, objects or locations (35b). It is remarkable that both nouns generally have the same semantic properties (human, instrument, location) and are built by the same morphological rule, if any.

- (35) a. *horloger-bijoutier*<sub>N</sub> *auteur-compositeur*<sub>N</sub>  
 ‘watchmaker-jeweler’ ‘author-composer’  
*député-maire*<sub>N</sub> *ingénieur chimiste*<sub>N</sub>  
 ‘representative-mayor’ ‘engineer-chemist’  
 b. *moissonneuse-batteuse*<sub>N</sub> *lecteur-enregistreur*<sub>N</sub> *bar-tabac*<sub>N</sub>  
 ‘combine harvester’ ‘player-recorder’ ‘bar tobacconist’

This pattern is rarely used for neoclassical compounds. Like native compounds, coordinated neoclassical  $NN$  compounds (36) denote a single referent that combines the properties associated with each component. Thus, an *androgynie* denotes a man or a woman who possesses the sexual properties of the opposite sex.

- (36) *androgynie*<sub>A</sub>, *rhino-pharynx*<sub>N</sub>

4.2.1.2. Coordinated compounds of the type AA. The two constituents of native coordinated compounds refer to contrasting properties or states (37), and the resultant adjectives refer to the sum of these properties/states.

- (37) a. *aigre-doux*<sub>A</sub>  
 ‘sour-sweet’  
 b. *sourd-muet*<sub>A</sub>  
 ‘deaf-mute’  
 c. *gris-bleu*<sub>A</sub>  
 ‘gray-blue’

The semantic relation is the same with neoclassical compounds (38). For example, a *franco-allemand* treaty (38a) is equally French and German, and an *anarcho-capitaliste* philosophy (38b) is a philosophy that equally values anarchic (*anarcho-*) and capitalist ideas.

- (38) a. *franco-allemand*<sub>A</sub>  
           ‘Germano-French’  
       b. *anarcho-capitaliste*<sub>A</sub>  
       c. *socio-culturel*<sub>A</sub>, *médico-social*<sub>A</sub>, *politico-commercial*<sub>A</sub>

4.2.2. *Attributive compounds.* Attributive compounds, like subordinating compounds, differ from coordinated compounds in the unequal relations between the two components: the X component has a governing role with respect to the Y component. In attributive compounds, the denotation of the compound is a subset of the denotation of the governor X (thus, endocentric). As for the semantic relationship between X and Y, these compounds are defined by the attributive nature of the relations: X determines the reference of the compound and Y specifies or modifies the reference by assigning it a property.

4.2.2.1. Attributive compounds of the type NN. In native attributive N1N2 compounds, N2, the governed constituent, provides a criterion for distinguishing a sub-category within the class to which N1 refers (for example, *poisson-chat* in (39) names a sub-species of fish. The N2 is interpreted non-literally, since it does not attribute all its semantic properties, only some salient or stereotypical properties of its denoted category. Thus, a *poisson chat* is a fish that displays a property stereotypically associated with cats – whiskers. Therefore, the compound presents a metaphorical reading.

- (39) *poisson-chat*<sub>N</sub>    *chou-fleur*<sub>N</sub>    *homme-grenouille*<sub>N</sub>  
       fish-cat            cabbage-flower    man-frog  
       ‘catfish’            ‘cauliflower’    ‘frogman’

Only pragmatic knowledge enables us to infer which semantic property of N2 is relevant in the compound. Any conceivable pragmatic relation between the two constituents can be used to establish this meaning (Corbin 1992, 1997, to appear).

Neoclassical attributive N1N2 compounds exhibit the same semantic properties, but the governing order is reversed. Thus, the nouns in (40a) refer to a type of N2 to which certain characteristics of N1 are attributed: a *pithecanthrope* is a human showing properties of an ape. The more common adjectival compounds in (40b) combine their salient properties in a similar way.

- (40) a. *pithécanthrope*<sub>N</sub>  
ape-human  
b. *antropomorphe*<sub>A</sub> *palmipède*<sub>A</sub> *cynocéphale*<sub>A</sub>  
human-form palm-feet dog-head

4.2.2.2. Attributive compounds of the type AN. Adjectival compounds of the AN type are also attributive compounds. This structure is unique to native compounds and includes only color terms. The adjectival constituent is the governor, referring to the color to which the compound denotation belongs. The noun is the governed element, denoting a particular shade of the color denoted by the governing adjective. Thus, the compound does not select all semantic properties associated with the noun, but only an iconic property. In *bleu-ciel*, *ciel* ‘sky’ represents a specific shade of light blue, just as slate in *gris ardoise* refers to a certain shade of gray; *vert Véronèse* is a shade of green found in the paintings of the artist Veronese.

- (41) *bleu ciel*<sub>A</sub>, *gris ardoise*<sub>A</sub>, *jaune citron*<sub>A</sub>, *vert Véronèse*<sub>A</sub>

4.2.2.3. Attributive compounds of the type AN. The AN structure is governed by the nominal constituent, while the adjectival constituent provides a property that distinguishes a subset of the denotation of that noun. This pattern is limited to neoclassical compounding and produces adjectival (42a) and nominal (42b) compounds. Thus, a *macrocéphale* is identified by the property of having a large (*macro-*) head (*-céphale*) and a *mégalthé* is a large (*mega-*) stone (*-lithé*). A similar pattern is found in native compounding, but is a syntactic rather than a morphological process (cf. Section 3.1).

- (42) a. *macrocéphale*<sub>A</sub>, *rectangle*<sub>A</sub>, *électro-aimant*<sub>A</sub> ‘electromagnet’,  
*orthodoxe*<sub>A</sub>  
b. *mégalthé*<sub>N</sub>, *macromolécule*<sub>N</sub>, *microorganisme*<sub>N</sub>, *tragi-comédie*<sub>N</sub>

4.2.3. *Subordinating compounds.* Bisetto and Scalise (2005) define the subordinating relation between two constituents as a relation of complementhood of various types. Thus, the governed noun may designate any type of specification with respect to the governing noun.

4.2.3.1. Subordinating compounds of the type NN. In native compounds of the N1N2 type (43), N2 may specify some telic relation with respect to N1 (43a): *café* refers to the purpose of the *pause* in *pause café*, *poste* refers to the recipient of *timbre* in *timbre poste*, and *fumeur* refers to the users of *wagon* in *wagon fumeur*. But various other relations are possible (43b). Many other relations are possible, e.g., a *bébé-éprouvette* is a baby conceived in a test-tube.

- (43) a. *pause-café*<sub>N</sub>    *timbre-poste*<sub>N</sub>    *wagon-fumeur*<sub>N</sub>  
           ‘coffee break’    ‘postage stamp’    ‘smoking car’  
 b. *bébé-épreuve*<sub>N</sub>  
           ‘test-tube-baby’

Neoclassical compounds of the *N1N2* type (44) display similar semantic relations. *N2* provides specification with respect to *N1* that may be telic in nature (44a): a *vélodrome* is a path (*-drome*) for bicycles (*velo-*), a *gynécologue* is a specialist (*-logue*) in the illnesses of women (*gynéco-*). Other, less specific, subordination relations are also possible (44b): *hydrothérapie* is a therapy employing water (*hydro-*).

- (44) a. *velodrome*<sub>N</sub>    *alcootest*<sub>N</sub>    *gynécologue*<sub>N</sub>  
           ‘bicycle-course’    ‘breathalyzer’  
 b. *hydrothérapie*<sub>N</sub>, *eurodéputé*<sub>N</sub> ‘euro-MP’, *radioguidage*<sub>N</sub>

4.2.3.2. Subordinating compounds of the types *VN/NV*. Compounds including a governing verb and a governed noun are very productive in native compounding (*VN* compounds). Regardless of the category of the compound (nominal or adjectival), or its denotation (prototypically, an instrument [*ouvre-boîte*<sub>N</sub>, *essuie-main*<sub>N</sub>], more rarely an agent [*trouble-fête*<sub>N</sub>], but also an event [*lèche-vitrine*<sub>N</sub> ‘window shopping’ lit. lick-window], a location [*coupe-gorge*<sub>N</sub>] or a patient [*gobe-mouton*<sub>N</sub> ‘things that sheep swallow’ lit. swallow-sheep]), the verb is generally a transitive verb that expresses an activity and the *N* is an argument of the verb. The most common argument relation is predicate-patient (45):

- (45) *lèche-vitrine*<sub>N</sub>    *trouble-fête*<sub>A/N</sub>    *ouvre-boîte*<sub>N</sub>    *coupe-gorge*<sub>N</sub>  
           lick-window    disturb-party    open-tin    cut throat  
           ‘window-shopping’    ‘killjoy’    ‘can-opener’    ‘cut-throat alley’

Some *VN* compounds exhibit other, rarer, argument relations between the *N* and the *V* (contradicting syntactic analyses posed in terms of direct complementhood of the verb): instrument/method (*cuit-vapeur*<sub>N</sub> ‘steam cooker’ lit. ‘cook steam’), a temporal period (*réveille-matin*<sub>N</sub> ‘alarm clock’ lit. ‘wake-morning’), location (*traîne-buisson*<sub>N</sub> ‘bird who shuffles along under shrubs’ lit. ‘shuffle shrub’), or agent (*croque-monsieur*<sub>N</sub> ‘ham and cheese toast’ lit. ‘crunch sir’). Other non-prototypical *VN* compounds include intransitive verbs (*traîne-buisson*<sub>N</sub>, *trotte-bébé*<sub>N</sub> ‘child walker’) or non-agentive verbs (*cache-coeur*) (for more details, see Villoing 2009).

If we assume to analyze the second component as a verb (and not an adjective such as Corbin to appear), neoclassical compounds in (46) are *NV* com-

pounds. They are typically adjectival.<sup>15</sup> Nouns of this structure present similar denotations to those found with native compounds: event (*infanticide*<sub>N</sub>), agent (*biographe*<sub>N</sub> ‘biographer’), instrument (*insecticide*<sub>N</sub>). The major difference between native and neoclassical compounds is that neoclassical compounding allows stative verbs which are impossible in native VN compounds. Consequently, the argument relation between verb and noun is not systematically a predicate-patient relation, as is the prototypical case with NV compounds with dynamic verbs (cf. Namer and Villoing 2005).

- (46) *carbonifère*<sub>A</sub>    *anglophobe*<sub>A</sub>    *mélomane*<sub>A</sub>    *noctiluque*<sub>A</sub>  
 ‘carboniferous’    ‘anglophobic’    ‘music lover’    ‘noctiluent’  
*insecticide*<sub>A</sub>    *biographe*<sub>N</sub>  
 ‘biographer’

4.2.3.3. Subordinating compounds of the type NA. Only neoclassical compounding can produce a few adjectives of the NA<sub>A</sub> structure, with a governing adjective and a governed noun that modifies the adjective. Thus, a *photosensible* or *photorésistant* organ is sensitive or resistant to changes in light (*photo-*), a *séropositive* person has a positive reaction to serological tests, a *halotolérant* organism can adapt to very high concentrations of salt.

- (47) *photosensible*<sub>A</sub>,    *photorésistant*<sub>A</sub>,    *halotolérant*<sub>A</sub>,    *séropositif*<sub>A</sub>,  
*aérosensible*<sub>A</sub>

### 4.3. Problem cases

4.3.1. *Morphological compounds, syntactic phrases or affix?* While most compounds can be classified as morphological compounds or syntactic structures based on the criteria noted above (Section 3.1), there are problem cases remaining.

The problem of classifying NN sequences arises again, particularly in sequences such as (48–50) that can be analyzed as lexicalized syntactic constructions (Section 3.1), as NN compounds (Section 4.2), or as derivations which include a constituent undergoing grammaticalisation in the process of becoming an affix (Section 3.3). Amiot and Van Goethem (2010) review the analysis of the *N-clé* set (48), and demonstrate the diverse treatments possible: syntactic construction in Noailly (1990), morphological NN compounding in Fradin (2009) and Van Goethem and Amiot (2009), affixoid derivation in Booij

15. Most compound adjectives have a nominal equivalent derived through conversion (*CARNIVORE*<sub>A</sub> > *CARNIVORE*<sub>N</sub>).

(2008). This case clearly demonstrates the difficulty in determining the status of the recurring constituent in these series.

(48) *notion clé*    *mot clé*    *personnage clé*  
 ‘key notion’   ‘key word’   ‘key figure’

(49) *soldes monstres*    *foule-monstre*    *succès monstre*  
 ‘monster sales’    ‘crazy crowd’    ‘amazing success’

(50) *voyage éclair*    *visite éclair*    *guerre éclair*  
 ‘flash voyage’    ‘flash visit’    ‘blitzkrieg’

4.3.2. *Categorization of native compound constituents.* The native compounds in (51) present an ambiguity, since they could be the result of either VN or NN compounding.

(51) *aide-chimiste*<sub>N</sub>    *appui-tête*<sub>N</sub>    *garde-barrière*<sub>N</sub>  
 ‘assistant chemist’   ‘headrest’   ‘level-crossing keeper’  
*soutien gorge*<sub>N</sub>  
 ‘bra’

The possible confusion between these compounding patterns stems from the ambiguity of the first component, which could be classified as a verbal or a nominal lexeme (*aide*, *appui*, *garde* and *soutien*). This homonymy stems from the conversion process that derives deverbal nouns selected by the same verbal stem as in VN compounding (the stem of the present indicative). As a result, the derived deverbal noun exhibits the same phonological properties as the theme of the corresponding verbal lexemes.

(52) *aider*<sub>V</sub>    >    *aide*<sub>N</sub>    *appuyer*<sub>V</sub>    >    *appui*<sub>N</sub>  
 ‘to help’    ‘an help’    ‘to press’    ‘a support’  
*garder*<sub>V</sub>    >    *garde*<sub>N</sub>  
 ‘to look after’    ‘a guard’

In addition to this homonymy, NN and VN compounds can create nouns that belong to the same semantic type as derived deverbal nouns (animate human). The ambiguity of the first component can be resolved through evaluation of the semantic compatibility between the compound and the compounding rules. Thus, *garde-fou* is a VN compound, since it denotes an instrument, while the derived deverbal noun *garde* denotes a human or an event; the compound noun *soutien-gorge* (support-breast) ‘bra’ must be analyzed as a VN compound, since the semantic relations between *soutien* and *gorge* is a process-argument relation, which is impossible in a NN compound (Villoing 2001; Corbin to appear).

4.3.3. *Categorization of constituents of neoclassical compounds.* The categorization of neoclassical compound constituents is a recurring question. The problem with neoclassical compounds whose last constituent includes a suffix is determining the semantic scope of the suffix (over the entire compound or over the final constituent only), which leads to a problem with the identification of the categories involved. Various views have been proposed in the literature, particularly regarding cases where the suffix derives denominal adjectives, as in (53).

(53) *cérébrospinal*<sub>A</sub>, *urogénital*<sub>A</sub>, *néonatal*<sub>A</sub>

Thus, neoclassical compounds with a final *-al* suffix are analyzed in various ways, depending on author and even for the same author. Thus, *cérébrospinal*<sub>A</sub> ‘relating to the spine and brain’ is analyzed as an NN compound with an *-al* suffix by Namer (2009), creating the structure [[*cérébro-spin*]-*al*]]. In contrast, *urogénital*<sub>A</sub> ‘relating to the urinary and genital systems’ is analyzed by Corbin (to appear) as a neoclassical AA compound, and the *-al* suffix takes scope only over the second constituent, leading to the structure [*uro*<sub>A</sub>-*génital*<sub>A</sub>], while *néonatal*<sub>A</sub> is treated by the same author as a suffixation on the compound *nouveau-né*<sub>N</sub>.

So, what about the compounds in (54)? Are they based on the NN pattern, followed by *-al*/*-el* or *-ique* suffixation, or are they NA compounds?

(54) *radio-électrique*<sub>A</sub>, *spatio-temporel*<sub>A</sub>, *hétérosexuel*<sub>A</sub>, *homosexuel*<sub>A</sub>

## 5. The compound head: endocentricity and exocentricity

The meaning of a compound depends on the meaning of its components and on the semantic relations between them (cf. Section 4.2), but also on the relations between the governing component (the compound head) and the compound itself. In French, as in many languages, there are two possibilities: (i) the compound includes a head which transmits its semantic and syntactic properties to the compound (endocentricity); (ii) the compound has no head constituent (exocentricity).

### 5.1. Endocentric compounds

The head constituent of an endocentric compound systematically functions as the semantic governing constituent (cf. 4.2). Thus, 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). Consequently, the



interpretation of a compound corresponds to a hyponymic relation between the compound and the first constituent for native compounds, and with the second constituent in neoclassical compounds (the denotation of the compound is a subset of the denotation of the compound head). The morphosyntactic properties of a compound (category, and for a noun also gender), are also inherited from the governing head constituent.

Endocentric patterns of native compounding include subordinating and attributive NN<sub>N</sub> compounds (55a), AN compounds (55b) and coordinated NN<sub>N</sub> (56a) and AA compounds (56b). Coordinated compounds are characterized by their two heads, which belong to the same category, and in nouns, also the same gender specification.

- (55) a. *poisson-chat*<sub>N</sub>, *pause-café*<sub>N</sub>  
 b. *BLEU CIEL*<sub>A</sub>, *JAUNE CITRON*<sub>A</sub>
- (56) a. *auteur-compositeur*<sub>N</sub>, *histoire-geographie*<sub>N</sub>  
 b. *aigre-doux*<sub>A</sub>, *franco-allemand*<sub>A</sub>

Endocentric patterns of neoclassical compounding include subordinating (*hippodrome*<sub>N</sub>), attributive (*pithécanthrope*<sub>N</sub>) and coordinate NN compounds (*rhino-pharynx*<sub>N</sub>), the NA pattern (*photosensible*<sub>A</sub>), the attributive AN pattern (*mégalthé*<sub>N</sub>) and the coordinated AA<sub>A</sub> pattern (*androgyné*<sub>A</sub>, *franco-allemand*<sub>A</sub>).

## 5.2. Exocentric compounds

The defining property of exocentric compounds is the absence of a head: the compound denotes an entity that does not correspond to a subset of the denotation of either constituent, and its morphosyntactic properties are not inherited from either constituent. Even if the constituents are semantically unequal, the semantic governor does not function as the compound head.

The native VN pattern (57) is systematically exocentric, regardless of the syntactic category of the compound (noun or adjective) or its denotation (see further discussion in Section 4.2.3.2): a *ouvre-boîte* does not denote a type of can nor a process of opening; *lèche-vitrine*, although denoting a type of ‘licking’ process, is not a verb.

- (57) *lèche-vitrine*<sub>N</sub>, *ouvre-boîte*<sub>N</sub>, *trouble-fête*<sub>A/N</sub>

There are, however, neoclassical patterns that produce adjectives without any adjectival governing constituent: the NV pattern (*carbonifère*<sub>A/N</sub>), the AN pattern (*macro-céphale*<sub>A</sub>), and the NN pattern (*plamipède*<sub>A</sub>), when these form adjectives.

## 6. Compounding and inflection

Inflection in French compounds is problematic and unstable (analyses vary depending on the author), and includes many exceptions. The source of difficulty lies in the fact that gender and number inflection are rarely marked phonologically in French, while spelling rules are often based on potentially mistaken etymological analysis, and many cases are arbitrarily decided by grammarians, based on semantic ‘feelings’ (Catach 1981; Grevisse 1986).

Only compounds whose constituents are merged are uniformly analyzed. In this case, the inflection of the compound is finally-marked, as it is in simple lexemes. This includes some rare native compounds (58) and most neoclassical compounds, whatever the patterns or the nature of the constituents (59). The scientific compounds whose constituents are linked by a hyphen follow the final-marking principle (60a), aside from some neoclassical NN and AA compounds with autonomous French bases, in which both constituents are marked (60b).

(58) *portefeuille*|*des portefeuilles*    *tournevis*|*des tournevis*  
       ‘wallet.sg/pl’                                ‘screwdriver.sg/pl’

(59) *des ludothèques, des rizicultures, des mégalithes, des insecticides*

(60) a. *des rhino-pharynx, des alliances franco-françaises*  
       b. *des sociaux-démocrates*

However, most compounds have two constituents graphically separated by a hyphen or a space. Inflection may appear on both constituents or only on one, depending on the type of compound, corresponding to inflectional marking on the whole compound or on a single constituent. In some cases, inflection is not marked. Sometimes, the same structure can be inflected or uninflected, depending on factors that are rarely made explicit. Despite the minimal regularity observed, two major criteria can be used to determine the plural form of native French compounds: (i) The category of constituents: in principle, only nouns and adjectives can be inflected; verbs remain uninflected; (ii) The semantic relations between constituents: in principle, constituents of coordinated or attributive compounds are both inflected, while only the first constituent (the head) of subordinating compounds is inflected (if it is a noun or an adjective, but not a verb) (Catach 1981; Grevisse 1986).

Agreement in NN compounds varies according to the semantic relations between the constituents. In coordinated (61) and attributive NN compounds (62), plurality is marked on both constituents, while in subordinating NN compounds, it is only marked on the first element (64).

- (61) *des boulangers-pâtisseries, des moissonneuses-batteuses, des bars-tabacs*
- (62) *des poissons-chats, des hommes-grenouilles, des oiseaux-mouches, des choux-fleurs*
- (63) *des timbres-poste, des pauses-café, des wagons-poste, des bébés-éprouvette*

In coordinated adjectival compounds, number and gender are marked on both adjectival constituents.

- (64) *des femmes sourdes-muettes des paroles aigres-douces*  
 ‘deaf<sub>FPL</sub>-mute<sub>FPL</sub> women’ ‘sour<sub>PL</sub>-sweet<sub>FPL</sub> words<sub>FPL</sub>’

The questions of agreement is more delicate when the adjectival compound is composed of a color adjective; these compounds are invariable in principle:

- (65) *des chemises rose-pâle (\*roses-pâles)*  
 ‘pale-rose<sub>(\*PL)</sub> shirts’  
*des cheveux brun-clair (\*bruns-clairs)*  
 ‘bright-brown<sub>(\*PL)</sub> hairs’

Such invariance is also evident in AN compounds with color denoting adjectives (66). Yet, there is some variation of use in these compounds, since inflection may appear on the adjective in some cases (67) (Catach 1981).

- (66) *des cuisines vert-olive / vertes-olives*  
 ‘olive-green<sub>(PL)</sub> kitchens’  
*des tissus jaune-citron /\*jaunes citron*  
 ‘lemon-yellow<sub>(\*PL)</sub> fabrics’
- (67) *des élytres bleu-verts de l’encre bleu-noire*  
 ‘blue-green<sub>(MPL)</sub> elytrons’ ‘blue-black<sub>(FSG)</sub> ink’  
*des robes bleues ciel*  
 ‘sky-blue<sub>(FPL)</sub> dresses’

Since VN compounds are exocentric, no inflection is marked on the verb. Plural marking of the compound is finally-marked on the N, conforming to the standard pattern for simple nouns and adjectives in French.

- (68) *des ouvre-boîtes des coupe-gorges*  
 DET<sub>PL</sub> can-opener<sub>PL</sub> DET<sub>PL</sub> cut throat<sub>PL</sub>  
 ‘cut-throat alley’

However, in some cases, the compound is invariable, generally when the N is not a patient of the verb (69) but also depending on the semantic value of the noun (70).

- (69) *des réveille-matin*      *des croque-monsieur*      *des cuit-vapeur*  
 DET.PL alarm clock.(#PL)    DET.PL cheese toast.(#PL)    DET.PL steam-cooker.(#PL)  
 ‘wake-morning’
- (70) *des hache-viande*              *des coupe-vent*  
 ‘DET.PL meat-grinder.(#PL)’    ‘DET.PL windbreaker.(#PL)’  
*des porte-drapeau*  
 ‘DET.PL standard bearer.(#PL)’

Sometimes, double orthography is attested, revealing the fragility of the semantic “feeling” of plurality.

- (71) *des tourne-dique(s)*      *des brise-glace(s)*      *des taille-crayon(s)*  
 DET.PL turn-table.(PL)    DET.PL ice-breaker.(PL)    DET.PL pencil-sharpener.(PL)

In contrast, some nouns in VN compounds are always marked for plural (72). This is inherent inflection that marks semantic plurality without having an impact on agreement (cf. Section 2.1).

- (72) *des porte-avions*              *des sèche-cheveux*      *des protège-yeux*  
 DET.PL aircraft carrier.PL    ‘DET.PL hair-dryer.PL    DET.PL eye-protector.PL’

Many cases remain unanswered, especially when the first constituent is a homonymous verb or derived deverbal noun. Grammarians have imposed the presence or absence of plural marking according to their own analyses: thus, *appuie* ‘rest’ in *appuie-tête* ‘headrest’ is sometimes analyzed as a verb, and sometimes as a noun, leading to varied plural marking depending on the work cited (*des appuie-tête* DET.PL rest-head.SG vs. *des appuies-tête* DET.PL rest-head.PL). This is the same for the first constituent of *garde-meuble* ‘furniture-storage guard’ and *garde-côte* ‘coast guard’.

## 7. Conclusion

Taking into account the criteria distinguishing syntactic from morphological constructions, the following table provides a summary of productive compounding in French,<sup>16</sup> including both neoclassical and native compounding, established above.

16. Underlining the governing constituent and/or head of the compound.

Table 4. Productive compounding in French

	Endocentric		Exocentric	
	Native	Neoclassical	Native	Neoclassical
Coordinated	<u>NN</u> <i>histoire-géographie</i> <sub>N</sub> <i>auteur-compositeur</i> <sub>N</sub>	<u>NN</u> <i>rhinopharynx</i> <sub>N</sub>	–	<u>NN</u> <i>androgyne</i> <sub>A/N</sub>
	<u>AA</u> <i>aigre-doux</i> <sub>A</sub> <i>gris-bleu</i> <sub>A</sub>	<u>AA</u> <i>franco-allemand</i> <sub>A</sub> <i>medico-social</i> <sub>A</sub> <i>anarcho-capitaliste</i> <sub>A</sub>	–	–
Attributive	<u>NN</u> <i>poisson-chat</i> <sub>N</sub> <i>chou-fleur</i> <sub>N</sub>	<u>NN</u> <i>pithécantrope</i> <sub>N</sub>		<u>NN</u> <i>anthropomorphe</i> <sub>A</sub> <i>palmipède</i> <sub>A</sub> <i>cynocéphale</i> <sub>A</sub>
	<u>AN</u> <i>bleu-ciel</i> <sub>A</sub> <i>jaune-citron</i> <sub>A</sub>	–		
	–	<u>AN</u> <i>mégalithe</i> <sub>N</sub> <i>micro-organisme</i> <sub>N</sub> <i>macromolécule</i> <sub>N</sub>		<u>AN</u> <i>quadrupède</i> <sub>A</sub> <i>macrocéphale</i> <sub>A</sub> <i>longiligne</i> <sub>A</sub>
Subordinating	<u>NN</u> <i>pause-café</i> <sub>N</sub> <i>timbre-poste</i> <sub>N</sub> <i>bébé-éprouvette</i> <sub>N</sub>	<u>NN</u> <i>hippodrome</i> <sub>N</sub> <i>hydrothérapie</i> <sub>N</sub> <i>eurodéputé</i> <sub>N</sub>	<u>VN</u> <i>ouvre-boîte</i> <sub>N</sub> <i>lèche-vitrine</i> <sub>N</sub> <i>casse-pied</i> <sub>A/N</sub>	<u>NV</u> <i>biographe</i> <sub>A/N</sub> <i>insecticide</i> <sub>A/N</sub> <i>noctambule</i> <sub>A/N</sub>
	–	<u>NA</u> <i>photosensible</i> <sub>A</sub> <i>éco-compatible</i> <sub>A</sub> <i>séropositif</i> <sub>A</sub>		

Let us summarize the main points. Native compounding shows a preference for nouns, while neoclassical compounding shows a preference for adjectives. Native compounding is predominantly endocentric, although the highly-productive VN pattern is exocentric. Given the capacity of neoclassical compounding to create adjectives from any base, compounds of this type are more often exocentric. The governing constituent (and the compound head in endocentric compounds) is systematically on the left in native compounding, and necessarily on the right in neoclassical compounds. The governor controls the semantic relations between the two constituents, whether coordinated, attributive or subordinating. Inflection in compounds is formed as it is with simple

nouns in neoclassical compounds whose constituents are not graphically separated, but remains complex in native compounds. Only endocentric compounds exhibit some regularity with respect to inflection.

Several areas of the analysis remain unresolved, particularly regarding compounds in which one member bears a suffix without any clear semantic criteria to determine the scope of this suffix. There are also cases for which the morphological/syntactic boundaries remain unclear: in NN sequences, between compounds and derived lexemes when one of the components corresponds to a suffixoid. Finally, since only few studies have focused on neoclassical compounding, and since given the wider range of semantic relations exhibited between constituents in this process relative to native compounding, the classification proposed here may require further modifications in future.

Université Paris 8

Florence.Villoing@univ-paris8.fr

## References

- Amiot, Dany & Georgette Dal. 2007. Integrating neoclassical combining forms into a lexeme-based morphology. In Geert Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli & Sergio Scalise (eds.), *On-line proceedings of the 5<sup>th</sup> Mediterranean Morphology Meeting (MMM5)*, 323–336. University of Bologna.
- Amiot, Dany & Kristel Van Goethem. 2010. Le statut de *-clé* et de *sleutel-* dans *mot-clé* / *sleutelwoord*: une analyse unifiée? In Franck Neveu, Valelia Muni Toke, Thomas Klingler, Jacques Durand, Lorenz Mondada & Sophie Prévost (eds.), *CMLF 2010 – 2<sup>ème</sup> Congrès Mondial de Linguistique Française*. Paris: Institut de Linguistique Française. <http://dx.doi.org/10.1051/cmlf/2010205>.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark. 1994. *Morphology by itself*. Cambridge: The MIT Press.
- Bauer, Laurie. 1979. Against word-based morphology. *Linguistic Inquiry* 10(3). 508–509.
- Bauer, Laurie. 2001. Compounding. In Martin Haspelmath, Ekkehard Konig, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: an international handbook, Vol. 1*, 695–707. Berlin & New York: Mouton de Gruyter.
- Benveniste, Émile. 1974. Formes nouvelles de la composition nominale. In Émile Benveniste, *Problèmes de linguistique générale II*, 163–176. Paris: Gallimard.
- Bisetto, Antonietta & Sergio Scalise. 1999. Compounding: morphology and/or syntax?. In Lunella Mereu (ed.), *The boundaries of morphology and syntax*, 31–48. Amsterdam & Philadelphia: John Benjamins.
- Bisetto, Antonietta & Sergio Scalise. 2005. The classification of compounds. *Lingue e Linguaggio* IV(2). 319–330.
- Bonami, Olivier & Gilles Boyé. 2003. Supplétion et classes flexionnelles dans la conjugaison du français. *Langages* 152. 102–126.
- Bonami, Olivier, Gilles Boyé & Françoise Kerleroux. 2009. L'allomorphie radicale et la relation flexion-construction. In Bernard Fradin, Françoise Kerleroux & Marc Plénat (eds.), *Aperçus de morphologie du français*, 103–125. Saint-Denis: Presses Universitaires de Vincennes.

- Booij, Geert. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1995*, 1–16. Dordrecht: Kluwer.
- Booij, Geert. 2005. *The grammar of words. An introduction to linguistic morphology*. Oxford: Oxford University Press.
- Booij, Geert. 2008. Composition et morphologie des constructions. In Dany Amiot (ed.), *La composition dans une perspective typologique*, 49–73. Arras: Artois Presses Université.
- Catach, Nina. 1981. *Orthographe et lexicographie: les noms composés*. Paris: Nathan.
- Corbin, Danielle. 1985. Les bases non-autonomes en français ou comment intégrer l'exception dans le modèle lexical. *Langue française* 66. 54–76.
- Corbin, Danielle. 1992. Hypothèses sur les frontières de la composition nominale. *Cahiers de grammaire* 17. 25–55.
- Corbin, Danielle. 1997. Locutions, composés, unités polylexématiques: lexicalisation et mode de construction. In Michel Martins-Baltar (ed.), *La locution, entre langue et usages*, 55–102. Fontenay-St-Cloud: ENS Editions.
- Corbin, Danielle. 2001. Préfixes et suffixes: du sens aux catégories. *Journal of French Language Studies* 11. 41–69.
- Corbin, Danielle. 2004. Français (Indo-européen: Roman). In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphology. An international handbook on inflection and word-formation, Vol. 2*, 1285–1300. Berlin & New York, Walter de Gruyter.
- Corbin, Danielle. To appear. *Le lexique construit*. Paris: Armand Colin.
- Corbin, Danielle & Jérôme Paul. 1999. Aperçus sur la créativité morphologique dans la terminologie de la chimie. *La banque des mots* 60. 51–68.
- Cottez, Henri. 1988. *Dictionnaire des structures du vocabulaire savant. Eléments et modèles de formation*. 4<sup>th</sup> edn. Paris: Dictionnaires Le Robert.
- Darmesteter, Arsène. 1877. *De la création des mots nouveaux dans la langue française et des lois qui la régissent*. Paris: Wieweg.
- Darmesteter, Arsène. 1894. *Traité de la formation des mots composés dans la langue française*. 1<sup>st</sup> edn. 1875. Paris: Honoré Champion.
- Dardano, Maurizio. 2003. Formazione delle parole nelle terminologie tecnico-scientifiche: introduzione. In Maria Grossmann & Franz Rainer (eds.), *La formazione delle parole in italiano*, 573–579. Tübingen: Niemeyer.
- David, Sophie. 1993. *Les unités nominales polylexicales, éléments de description et reconnaissance automatique*. Paris: Université de Paris 7 dissertation.
- Fradin, Bernard. 1997. Les mots-valises: une forme productive d'existants impossibles? In Danielle Corbin, Bernard Fradin, Benoît Habert, Françoise Kerleroux & Marc Plénat (eds.), *Mots possibles et mots existants, Vol. 1, Silexicales*, 101–110. Villeneuve d'Ascq: URA 382 – CNRS and Université de Lille 3.
- Fradin, Bernard. 2000. Combining forms, blends and related phenomena. In Ursula Dolsescha, Anna Maria Thornton (eds), *Extragrammatical and marginal morphology*, 11–59. Munich: Lincom Europa.
- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.
- Fradin, Bernard. 2009. Compounding in French. In Rochelle Lieber and Pavol Štekauer (eds.), *Oxford handbook of compounding*, 417–435. Oxford: Oxford University Press.
- Gaeta, Livio & Davide Ricca. 2009. *Composita solvantur*: Compounds as lexical units or morphological objects? *Italian Journal of Linguistics/Rivista di Linguistica* 21(1). 35–70.
- Grésillon, Almuth. 1984. *La règle et le monstre: le mot-valise*. Tübingen: Niemeyer.
- Grevisse, Maurice. 1986. *Le bon usage*. Paris: Duculot.

- Guevara, Emiliano & Sergio Scalise. 2009. Searching for universals in compounding. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 101–128. Berlin: Springer.
- Iacobini, Claudio. 1999. Distinguishing derivational prefixes from initial combining forms. In Geert Booij, Angela Ralli & Sergio Scalise (eds.), *Proceedings of 1<sup>st</sup> Mediterranean Morphology Meeting (MMM1)*, 132–140. University of Patras.
- Iacobini, Claudio. 2004. Composizione con elementi neoclassici. In Maria Grossmann & Franz Rainer (eds.), *La formazione delle parole in italiano*, 97–163. Niemeyer: Tübingen.
- Lieber, Rochelle & Sergio Scalise. 2007. The Lexical Integrity Hypothesis in a new theoretical universe. In Geert Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli & Sergio Scalise (eds.), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting*, 1–24. Bologna: University of Bologna. <http://mmm.lingue.unibo.it/proc-mmm5.php>.
- Lesselingue, Chrystèle & Florence Villoing. 2002. *Aide-jardinier, aide-sage-femme*, mots composés ou préfixés? In Peggy Afuta, Adil El Ghali & François Toussenet (eds.), *Actes de 7<sup>èmes</sup> rencontres de l'atelier des doctorants en linguistique (ADL'2002)*. Ecole doctorale de sciences du langage. Université Paris VII.
- Matthews, Peter H. 1991. *Morphology. An introduction to the theory of word-structure*. 2<sup>nd</sup> edn. Cambridge: Cambridge University Press.
- Montermini, Fabio. 2010. Units in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 77–92. Amsterdam & Philadelphia: John Benjamins.
- Namer, Fiammetta. 2009. *Morphologie, lexique et traitement automatique des langues: l'analyseur DériF*. London: Hermes Sciences Publishing.
- Namer, Fiammetta. To appear. Nominalisation et composition en français: d'où viennent les verbes composés? In Rafael Marin & Florence Villoing (eds.), *Lexique 20*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Namer, Fiammetta & Villoing Florence. 2005. Have cutthroats anything to do with tracheotomes? Distinctive properties of VN vs. NV compounds in French. In Geert Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli & Sergio Scalise (eds.), *On-line proceedings of the 5<sup>th</sup> Mediterranean Morphology Meeting (MMM5)*, 105–124. University of Bologna.
- Noailly, Michelle. 1990. *Le substantif épithète*. Paris: Presses Universitaires de France.
- Rainer, Franz. 2008. Neo- and neo-Latin. *Word Structure* 1(1). 53–64.
- Rainer, Franz & Soledad Varela. 1992. Compounding in Spanish. *Italian Journal of Linguistics/Rivista di Linguistica* 4(1). 117–142.
- Ricca, Davide. 2010. Corpus data and theoretical implications. With special reference to Italian VN compounds. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 237–254. Amsterdam & Philadelphia: John Benjamins.
- Scalise, Sergio. 1984. *Generative morphology*. Padova: Clesp.
- Scalise, Sergio. 1992. Compounding in Italian. *Italian Journal of Linguistics/Rivista di Linguistica* 4(1). 175–199.
- Scalise, Sergio & Irene Vogel (eds.). 2010. *Cross-disciplinary issues in compounding*. Amsterdam & Philadelphia: John Benjamins.
- Stump, Gregory T. 2001. *Inflectional morphology. A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Van Goethem, Kristel & Dany Amiot. 2009. Affixization processes in French and Dutch. Paper presented at the 7<sup>th</sup> Mediterranean Morphology Meeting, Nicosia (Cyprus), 10–13 September 2009.
- Villoing, Florence. 1999. La querelle, au XIX<sup>ème</sup> siècle, autour des mots composés du type [VN]<sub>N</sub>: quels enjeux pour la grammaire? *LINX* 40. 159–185.



- Villoing, Florence. 2000. Les principes de la morphologie dans la grammaire historique et comparée du XIX<sup>ème</sup> siècle: le témoignage des noms composés français du type [VN]<sub>N</sub>. *Moderne Sprachen* 44. 107–135.
- Villoing, Florence. 2001. Les *couche-dehors* et les *lève-tôt* sont-ils des construits morphologiques? Le rôle de la structure argumentale des lexèmes verbes dans la répartition morphologie/syntaxe. In Emmanuel Aïm, Kim Gerdes & Hi-Yon Yoo (eds.), *Actes de 6<sup>èmes</sup> rencontres de l'atelier des doctorants en linguistique (ADL'2001)*, 161–168. Paris: Université Paris VII.
- Villoing, Florence. 2002. *Les mots composés [VN]<sub>N/A</sub> du français: réflexions épistémologiques et propositions d'analyse*. Paris: Université Paris X-Nanterre dissertation.
- Villoing, Florence. 2009. Les mots composés VN. In Bernard Fradin, Françoise Kerleroux & Marc Plénat (eds.), *Aperçus de morphologie du français*, 75–197. Saint-Denis: Presses Universitaires de Vincennes.
- Warren, Beatrice. 1990. The importance of combining forms. In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds.), *Contemporary morphology*, 111–132. Berlin & New York: Mouton de Gruyter.
- Williams, Edwin. 1981. On the notions “lexically related” and “head of a word”. *Linguistic Inquiry* 12(2). 245–274.
- Zwanenbourg, Wiecher. 1992. Compounding in French. *Italian Journal of Linguistics/Rivista di Linguistica* 4(1). 221–240.