

# Integrating Neoclassical Combining Forms into a Lexeme-Based Morphology

Dany Amiot

*Université d'Artois, Arras & Grammatica*

Georgette Dal

*UMR 8163 STL ("Savoirs, Textes, Langage"), CNRS  
& Universities of Lille 3 and Lille 1*

## 1. Introduction

The aim of this paper is to account for a part of the so-called combining forms (CFs) – such as *anthrop-*, *-logue*, *lud-* in ANTHROPOLOGUE ‘anthropologist’ and LUDOTHÈQUE ‘game library’ –, in present-day French in the context of a lexeme-based morphology. We will consider only one type of CFs, the so-called neoclassical CFs (Bauer 1983, Lüdeling & al. 2002) or classical CFs (Fradin 2000), which can be characterised by the following four properties:

- Their lexematicity in source languages: in Latin or Greek, they were usually lexemes with grammatical words associated (e.g. LUDUS, from Thomas d’Aquin: “*Ludus est necessarium ad conversationem humanae vitae*”).
- The absence of syntactic realisation in the target language: in French (English, German, and so on), neoclassical CFs can only appear as bound constituents of lexemes, without receiving associated grammatical words in the target language. E.g. in French: \**J’ai vu un anthrop(e) avec un chapeau* (\*I saw an anthrop with a hat); \**Les enfants aiment les lud(e)s* (\*Children love luds).
- The kind of vocabulary they serve to form: generally, the complexes in which they appear belong to the learned vocabulary of scientific or technical fields: medicine (LAPAROSCOPIE ‘laparoscopy’), biology (BACTÉRIOLOGIE ‘bacteriology’), physics (LITHOSPHERE ‘lithosphere’), technology (CHRONOGRAPHE ‘chronograph’), etc.
- The presence of a linking vowel (*o* or *i*) between the two constituents in the phonological context /... C<sub>f</sub>C<sub>i</sub>.../ where C<sub>f</sub> and C<sub>i</sub> are consonants in, respectively, final position of the first constituent and initial position of the second constituent: for French, it is generally claimed that *-o-* appears when at least one of the constituents has a Greek origin – *ludothèque*, *cassetothèque* –, and *i* when at least one has a Latin origin – *omnivore*, *herbicide*. But *-o-* is more common than *-i-* because it also appears when only one constituent is of Greek origin (such as in LUDOÉDUCATIF ‘edutainment’, CYTOCHIMIE ‘cytochemistry’) or with non-neoclassical CFs (as in AFRO-CUBAIN ‘Afro-Cuban’, ELECTROAIMANT ‘electromagnet’).

This set of properties seems to establish a well-defined class of homogeneous elements, but we show that they are not as homogeneous as they may appear. For example, not all the neoclassical CFs have the same positional constraints: some of them can appear in initial or final position (*anthrop* in ANTHROPOPHAGE ‘anthropophagite, cannibal’ vs. AFRICANANTHROPE ‘Fossil of pre-hominian discovered in Eastern Africa’); some only in initial position (*micro-* in MICROORGANISME ‘micro-organism’ or MICROAMPERE ‘microampere’; and others only in final position (*vore* in OMNIVORE ‘omnivore’ or PUBLIVORE ‘fond of publicity / publicity hound’).

Our examination of previous studies on neoclassical CFs (§ 2.) indicates that such analyses are anything but homogeneous, even though the authors have basically tried to prove the existence of a special category for these elements. Our perspective is slightly different from these approaches: categorization of CFs will be not our main aim; what we will do is to examine whether the basic units of lexeme-based morphology, lexemes and exponents of lexeme construction rules (affixes, non-segmental or supra-segmental phenomena) can account for neoclassical CFs, or whether a new type of unit is required.

The assumptions we defend here are (i) neoclassical CFs are not a homogeneous category, they are not all of the same type and cannot be analysed in the same way; (ii) the notions provided by lexeme-based morphology are sufficient to analyse neoclassical CFs. As a result, we pose the following questions: Do neoclassical CFs belong to the French (English, German, etc.) language or to Latin and/or Greek? Are they part of an international stock common to most Indo-european languages?

## 2. Previous Approaches

Previous analyses of neoclassical CFs vary according to the criteria taken into account: the bound nature of these elements, whether their position is fixed or not, their semantic nature (lexical or grammatical), their phonological properties, etc. Another important factor is whether these criteria are considered separately or in conjunction. Overall, these analyses have led to four main results: neoclassical CFs are (i) affixes, (ii) roots/stems<sup>1</sup>, (iii) roots/stems in some cases and affixes in others, (iv) neither affixes nor roots.

- (i) Neoclassical CFs are considered affixes when the only criterion taken into account is their boundness; Williams (1981) or Bauer (1979) propose this kind of analysis. For Bauer (1979) for example, *anglo-*, *bio-* or *electro-* are prefixes in, respectively, *Anglo-Indian*, *biochemistry* and *electrocardiogram* and *-crat*, and *-phile* are suffixes in *bureaucrat* and *audiophile*. So, words like *biocrat* or *electrophile* appear “to be made up of a prefix and a suffix, but ha[ve] no root” (*op. cit.*, 509).

---

<sup>1</sup> In the presentation of the previous approaches, we do not distinguish between the two terms, which are often used as equivalent by the authors we refer to.

Analyses of this sort have been criticized, for example by Scalise (1984: 75-76) and by L. Bauer himself<sup>2</sup>, who challenge the idea that a complex could be formed by joining a prefix to a suffix.

When scholars take into account other criteria, such as semantic and positional criteria, the results of the analysis are different, but not necessarily homogeneous:

- (ii) Booij (1992) for example considers that neoclassical CFs are not affixes but “non-native roots” following two basic patterns: “root + X” (where X is an existing non-native word, such as *fisica* in *astrofisica*) and “root + root” (e.g. *psychografie*). Plag (2003) adopts a similar analysis: for him, even when a CF has a fixed position, it is never an affix, but a bound root.
- (iii) S. Scalise or D. Corbin both make a clear distinction between affixes on the one hand and neoclassical CFs (stems for Scalise 1984, archeoconstituants for Corbin 2001) on the other hand, even for elements with a Greek (or Latin) origin that do not constitute syntactic units in French (English, etc.). In Corbin (2001), the distinction is based on a semantic criterion: for her, affixes have an instructional meaning (they are operators), while archeoconstituants have a descriptive meaning<sup>3</sup>. She uses this criterion to distinguish between *archi-*, *pré-* or *iso-*, affixes, and *phile*, *graph*, *aero*, *gastro*, archeoconstituants. The same criterion allows her to distinguish between two *micro-* in French: the prefix, when *micro-* is joined to a measurement noun (indicating that the measurement is divided by 106, such as in *microseconde*) and the archeoconstituant with an adjectival meaning in other cases (e.g. in *micro-organisme* which refers to a ‘very small organism’). The only difference between the prefix and the archeoconstituant is the kind of meaning they have: instructional or lexical. In a similar vein, Iacobini (2004) analyses Italian neoclassical CFs. He, too, considers them heterogeneous, but he distinguishes between three subcategories: neoclassical CFs with a lexical meaning, prefixes with a classical origin, and a third type, which is not labelled, like *-crate*, *-voro* or *-fero*, on the borderline between lexemes and suffixes. Like suffixes, this third type of CFs occur only in a final position and they are productive. Like lexemes, they have a lexical content. For example, in *carnivoro* ‘meat eater’, *calorifero* ‘heater’, *-voro* and *-fero* can be regarded as verbs with an argument corresponding to the left

---

<sup>2</sup> Bauer (1983: 214) notes that the “notion of a prefix and a suffix occurring with no root thus leads to a contradiction”.

<sup>3</sup> Corbin (2001: 44): “Cette façon de décrire le sens des affixes et leur intervention dans la construction du sens des unités construites permet de les différencier des autres unités infralexicales entrant dans la construction des mots que sont ce que j’appelle les *archéoconstituants*, c’est-à-dire les constituants empruntés au latin et au grec (ex. *brachy-* ‘court’, *anthropo-* ‘homme’, *-cide* ‘qui tue’), et les *fractoconstituants*, c’est-à-dire les représentants tronqués et lexicalisés comme tels d’unités française (ex. *euro-* = *Europe*) dans *eurocorps*, *eurodéputé*, *euromissile*, etc.): *archéoconstituants* et *fractoconstituants* ont un sens de nature descriptive et non instructionnelle”.

‘Describing in this way the meaning of the affixes and their semantic role in the construction of complex units allows them to be distinguished from the other infralexical units involved in word-building and that I call *archeoconstituents*, i.e. constituents borrowed from Latin or Greek (e.g. *brachy-* ‘short’, *anthropo-* ‘man’, *-cide* ‘that kills’) and *fractoconstituents*, i.e. constituents that have been shortened and lexicalised directly from French lexical units (e.g. *euro-* = *Europe*) in *eurocorps* ‘euro-body’, *eurodéputé* ‘euro-MP’, *euromissile* ‘*ibid*’, etc.): *archeoconstituents* and *fractoconstituents* have a descriptive and not an instructional meaning.’

constituent. The complexes in which they appear resemble the synthetic compounds of Germanic languages (*meat-eater* / *carnivoro*) and the V+N compounds of romance languages (*portacenere* ‘ashtray’ / *calorifero*). For Iacobini, this type of CFs form compounds, but they do not seem to be roots.

- (iv) Warren (1990) proposes a different analysis. While both roots and CFs have a lexical content, CFs are not roots because they do not correspond to a free form. Nor can they be considered affixes, because they belong to open classes. Moreover, they are different from suffixes because they have a lexical meaning, and from prefixes because (a) they don’t need “have productive force” (p. 123), (b) they do not have the same phonetic shape, and (c) they correspond to a model-word (e.g. *phyto-* / Gk: *phyton*). So, neoclassical CFs are irreducible to the categories of the other systems of word formation.

When we look at conceptualisations of the relationship between neoclassical compounding and “ordinary” compounding, we see that these are no more homogeneous than the analyses of neoclassical CFs. Two attitudes can be distinguished:

- the differences between the two types of formation are emphasized; cf. e.g. Warren (1990) or Plag (2003). Plag, for example, stresses the fact that neoclassical compounds have formal properties (combinatory and phonological properties, the presence of the linking vowel between the two elements) “that distinguish them from the other types of compounds” (*op. cit.*: 159).
- their similarities are emphasized, in two different ways:
  - o Some scholars like Booij (1992) or Scalise (1984) stress the fact that neoclassical compounds share some fundamental properties with other compounds. For example, Booij (1992: 56) claims that the non-native compound “conforms to the general Dutch pattern of compounding in that the second constituent is the head”.
  - o Another way to bring out similarities in the two types of compounding is to show that neoclassical CFs share properties with other elements, not only with affixes or lexemes, but also with other phenomena such as blending, clipping or secretion. The works of Bauer (1998), Iacobini (2004) or Lüdeling & al. (2002) follow this approach. For Bauer (1998: 419-420) for example, “neoclassical compounding is a name for a relatively but not completely arbitrary subdivision of word-creation space and should be read as being a prototype rather than a clear-cut category”. For the author, word creation is conceived as a three-dimensional space (whose three parameters are: belonging to the patrimonial lexicon (native vs. foreign), formation type (simplex / derivative / compound); degree of shortening) and *neoclassical compound* is used to label a part of this space.

Lüdeling & al. (2002: 253) have very different theoretical presuppositions (in their opinion there is no difference between stems and affixes) but their results are very

similar to those of Bauer (at least for German): “no clear-cut principled difference can be found” between neoclassical and native word formation, because neither phonological properties nor differences in the combinability or in the productivity of these elements allow them to be distinguished from native elements. According to these authors, the relationship between neoclassical compounding and native compounding has to be conceived as a continuum.

### 3. The Tools of Lexematic Morphology

In this section, we will examine whether lexematic morphology is equipped to account for neoclassical CF’s. First, we will define successively the notions of lexeme and of affix in the framework of lexematic morphology.

#### 3.1. Lexemes

Since Matthew (1974), the lexeme is generally considered to be an abstract lexical unit that possesses the following properties:

- (i) it belongs to an open list and is a member of a major lexical category, *i.e.* it is a noun, a verb or an adjective. Sometimes, the category of adverb is added, especially complex adverbs based on adjectives; e.g., in French, the class of adverbs suffixed by *-ment* (cf. Fradin 2003: 73<sup>4</sup>).
- (ii) semantically, a lexeme has a constant and entirely specified meaning (cf. Fradin & Kerleroux forthcoming).
- (iii) it has a phonological representation.

While properties (ii) and (iii) do not raise problems, property (i) does: How can a syntactic category be attributed to an element that does not correspond to a grammatical word? It is possible to answer this question when a CF can be used as a base for suffixation, such as *hydr* or *phob* in the adjectives *hydrique* ‘hydric’, *phobique* ‘phobic’: since in French the suffix *-ique* is used to form adjectives on nominal bases (e.g. *colère* / *colérique* ‘anger/quick tempered’, *scène* / *scénique* ‘stage<sub>N/A</sub>’), *hydr* and *phob* can be analysed as nouns. But not all neoclassical CFs serve as a base for derivation, cf. e.g. *micro-* or *-cide*. We will return to this question, § 4.2.

We should also point out, and it will be useful for some of our analysis, that a lexeme can have one or several roots / stems (hence *radicals*), some of them not being visible to the syntax: in French, for example, a verb such as *démontrer* ‘demonstrate’ has two radicals: *démontr-* and *démonstr-*; the peculiarity of the latter being that it never has syntactic realisations, and only appears in morphologically complex lexemes, such as *démonstration* ‘demonstration’ or *démonstratif* ‘demonstrative’.

---

<sup>4</sup> If the suffixation in *-ment* is derivational (see Dal 2007).

### 3.2. *Affix*

Within this theoretical framework, an affix is not the same sort of element as a lexeme. Morphological objects are not the result of the concatenation of morphemes but the result of the application of a rule to lexemes. Affixes are therefore the exponents of rules – such as reduplication, apophony, and so on –, which can be characterised as realising phonetically and graphically a semantic function. Thus, “affix” is a simpler way of saying ‘exponent of lexeme construction rules’ (LCR), and LCRs can be regarded as generalizations between two sets of lexemes, one more complex than the other.

## 4. Neoclassical CFs in Lexeme-based Morphology

Taking the two basic notions of lexeme and exponent of rule / affix, we investigate whether it is possible to account for the different sorts of neoclassical CFs; to do so, we analyse four different CFs: *lud*, *anthrop*, *micro* and *logue*, which will serve to illustrate the different kinds of analyses it is possible to propose within the framework of lexematic morphology.

### 4.1. “*Lud*”: Radical *B* of a Lexeme with Multiple Radicals

*Lud-* has at least two special characteristics: it always appears in initial position and means ‘jeu’ (‘game/play’) in the complex lexemes in which it appears: **LUDIQUE** ‘relating to a game, ludic, playful’, **LUDOTHÈQUE** ‘game library’, **LUDICIEL** ‘game software’<sup>5</sup>.

Following Fradin (2003), who defines the lexeme as a multistratal entity including five types of information (graphemic (G), phonological (F), syntactical (SX), morphological (M) and semantic (S)), each independent of the others, we shall consider that *lud* is one of the radicals of the lexeme **JEU**. This analysis is supported by the fact that the forms *jeu* and *lud* appear in complementary distribution in complex lexemes: *jeu* appears in final position (**ANTIJEU** ‘a game which is the antithesis of what is normally considered a game’, **INTERJEU** ‘interplay’, **CONTRE-JEU** ‘play-back’), and *lud* in initial position (**LUDIQUE**, **LUDOTHÈQUE**, **LUDICIEL**).

Our hypothesis is that the choice of *jeu* or *lud* has a phonological motive and is part of a strategy to find the correct form of the output in lexical construction. The radical *jeu*, with its phonological pattern CV, is not a good input, especially because most French adjectival suffixes begin by a vowel (*-ique* [ik] but also *-aire* [ɛR], *-al* [al], *-eux* [ø], *-el* [ɛl], etc.). To avoid forms such as \**jeuïque* or \**jeuthèque*, the suppletive form *lud* (whose phonological pattern is CVC) is preferred.

In this account, *jeu* and *lud* differ only in their graphemic and phonological forms, but, since they belong to the same lexeme, the question of the categorical identity of *lud* (is it a noun or not?) simply does not arise.

---

<sup>5</sup> We except here **PRÉLUDE** (from Latin **PRAELUDIM**), **INTERLUDE** and **POSTLUDE**, which form a set, and **PRÉLUDER** and **ELUDER**, inherited from Latin **PRAELUDARE** and **ELUDARE**.

Figure 1 proposes a representation of this analysis. In Figure 1, according to Fradin (2003), “#” indicates the citational form of the radical; the circle prefixing *lud* indicates that this form does not appear independently in syntax; ‘res.: init.’, that *lud* is reserved for initial position:

JEU	A	B
(G)	jeu <sup>#</sup>	°lud <sup>#</sup>
(F)	[ʒø]	[lyd]
(SX)	← c a t : n →	
(M)		res.:init.
(S)	← j e u ’ →	

Figure 1.

The same treatment is proposed for:

- (i) other neoclassical CFs such as *pyr*, interpreted as a B radical of the lexeme FEU ([fø]), cf. ALLUME-FEU ‘fire-lighter’, COUPE-FEU ‘firebreak’, COUVRE-FEU ‘curfew’, CONTRE-FEU ‘fireback’, PARE-FEU ‘firebreak / fireguard’ vs PYROGÈNE ‘pyrogenic’, PYROGRAVURE ‘poker work’, PYROMANE ‘pyromaniac’, PYROTECHNIE ‘pyrotechnics’, PYROLYSE ‘pyrolysis’.
- (ii) other types of CFs, termed “fractomorphèmes” by Tournier (1985), “fractoconstituants” by Corbin & Paul (2000), “fractoformants” by Fradin (2000), *i.e.* constituents such as *pétro-* in PETRODOLLAR ‘petrodollar / arabodollar’: *pétro* is a B radical of the lexeme PÉTROLE.

4.2. “Anthrop”: Radical B of a Lexeme with Multiple Radicals or Sole Radical of a Lexeme with no Associated Grammatical Word?

The case of *anthrop* is less straightforward. Though semantically it can be considered a suppletive radical for the lexeme HOMME, the argument based on a positional complementary distribution, used in the case of *lud* vs *jeu*, does not hold, since *anthrop* and *homme* can appear in initial position (ANTHROPOMÉTRIE ‘anthropometry’, HOMME-SANDWICH ‘sandwich man’) or in final position (MISANTHROPE ‘misanthrope’, SURHOMME ‘superman’). On the other hand, considering it as the only graphemic and phonological form of a lexeme ANTHROP leads to difficulties with the definition of the lexeme as the result of abstracting away inflectional marking (cf. the definitions of Fradin 2003: 102), since, by definition, a combining form never appears in syntax. Thus, this solution requires at least a revision of the notion of lexeme, such as that proposed by Booij (2002:141) or in Fradin and Kerleroux (forthcoming), who define the lexeme as “l’entité linguistique qui sert de base aux RCL”, (“the linguistic entity on which LCRs are based”) independently of any syntactic realization. From this point of view, *anthrop*, which can be suffixed by *-ique* (cf. ANTHROPIQUE ‘anthropic’) is a noun, since *-ique* coins relational adjectives on nominal bases (e.g. COLÉRIQUE<sub>A</sub> ‘quick-tempered’ < COLÈRE<sub>N</sub> ‘anger’; ALGÉBRIQUE<sub>A</sub> ‘algebraic’ < ALGÈBRE<sub>N</sub> ‘algebra’).

Yet considering *anthrop* as a B radical of a lexeme is, in our opinion, a better solution, since this allows us to deal with another suppletive form of HOMME, *homin-*. We believe that *anthrop-* and *homin-* can be considered, respectively, B and C radical of HOMME, the choice depending largely on the specialist language register and possibly even on the origin of the constituents: *homin-* tends to be used in biology (HOMINAL ‘*ibid.*’, HOMINICOLE ‘living in the human body’) or in zoology (HOMINIDÉS ‘hominidae’, HOMINIENS ‘hominidae’), in conjunction with constituents of Latin origin, whereas *-anthrop-* appears in a wider variety of fields – geography (ANTHROPOGÉOGRAPHIE ‘anthropogeography’), esotericism (ANTHROPOSOPHIE ‘anthroposophy’), anthropology (ANTHROPOGÉNÈSE ‘anthropogeny’), and so on – in conjunction with constituents of Greek origin. As for HOMME, it appears in native compounds.

The representation of such an analysis for *-anthrop-* and *homin-* appears in fig.2.

HOMME	A	B	C
(G)	homme <sup>#</sup>	°anthrop <sup>#</sup>	°homin <sup>#</sup>
(F)	[ɔm]	[ãntɾɔp]	[ɔmin]
(SX)		← c a t : n →	
(M)		res:gr.	res:lat.
			res:
			biology/zoology
(S)	←‘homme’ (as representative of mankind)→		

Figure 2.

However, we also propose (fig. 3) a representation of the other type of analysis, which consists in considering *-anthrop-* a lexeme without a grammatical word:

ANTHROP	(G)	°anthrop <sup>#</sup>
	(F)	[ãntɾɔp]
	(SX)	c a t : n
	(M)	
	(S)	anthrop’

Figure 3.

#### 4.3. “Micro”: The Exponent of an LCR

Regarding *micro-* as the exponent of an LCR, *i.e.* as a prefix, implies that this constituent has undergone a process of grammaticalization as defined by Olsen (2000: 901):

An originally free word that has entered into a compound can serve as the basis for an entire pattern of like compounds. Once such a pattern takes hold and becomes productive, the original constituent may begin to deviate from its free equivalent in form or meaning and develop into an affix-like element.



In fact, the evolution of *micro-* meets the criteria of grammaticalization (cf. e.g. Heine et al. 1991, Hopper and Traugott 1993, Lehmann 1995):

- (i) it originates from a Greek adjective, MIKROS, whose meaning was ‘small, short’
- (ii) it has lost its syntactic autonomy in French
- (iii) its meaning – or, more precisely, the meaning of the LCR it is associated with – is partly different from that of MIKROS in Greek.

*Micro-* always has a quantitative role:

- Most of the time, it applies to the entirety of what the noun located on its right denotes and indicates that its referent is smaller than the standard, such as in MICROFILM ‘microfilm’ or MICRO-ORGANISME ‘micro-organism’, which mean, respectively, ‘film / molecule of a very small size’. However, sometimes *micro-* does not apply to the entirety of the referent but only to one of its dimensions, such as in MICROCHIRURGIE ‘microsurgery’ or MICROÉCONOMIE ‘micro-economics’. MICROCHIRURGIE does not denote ‘minor surgery’ but the ‘branch of surgery concerned with very small living structures, done with very small instruments, often under a microscope’. As for MICROÉCONOMIE, it is the ‘branch of economics that studies limited economic phenomena’.
- When *micro-* is connected to a noun of measurement, cf. e.g. in MICROFARAD ‘*ibid*’ or MICROSECONDE ‘microsecond’, it means one millionth of the unit denoted by the noun, FARAD or SECONDE.

Since these dual interpretations are possible, Corbin (1992) proposes a dual analysis of *micro-*: *micro-* is an adjectival archeoconstituant (more or less a bound root) in the first interpretation (cases of MICROFILM or MICROCHIRURGIE) and it is a prefix in the second interpretation (case of MICROSECONDE). Yet Corbin (1992; cf. also 2001) considers that, at a more abstract level, it is fundamentally the same element, whose behaviour and semantic role vary according to the type of noun it is connected to.

Although this analysis is interesting, we believe that it is possible to account for the semantic behaviour of *micro-* in a simpler way, insofar as *micro-* basically behaves in French like *-et*, that is an exponent of an LCR. This is very clear if we compare complexes such as: CLOCHETTE (‘small bell’) and MICRO-ORGANISME (‘very small organism’); RÉFORMETTE (‘reform that concerns only limited problems’) and MICROÉCONOMIE (‘branch of economics that studies limited economic phenomena’). The meanings of the complexes formed by *-et* and *micro-* are very similar. Why then consider *-et* as an affix and not *micro-*? *Micro-* holds, in fact, as an affix:

- in all interpretations it has a quantitative meaning;
- variations in interpretation are due to the meaning of the base: whether it is a noun of measurement or not; whether it is a noun that refers to an entity with a spatial extension or not, etc. It does not differ in this way from a suffix like *-et*.

Possible arguments against such an analysis of *micro-* appear to be based on several factors: our etymological knowledge, the presence of the vowel *-o* at the end of *micro-* and the fact that the lexemes in *micro-* often belong to a learned lexicon, although this is not always the case, cf. *microdécision*, *micro-ordinateur*, etc. Yet if we want to account for the competence of a run-of-the-mill speaker, it seems better to consider *micro-* as a prefix in modern French.

For French, other constituents can be analysed in the same way: *macro-*, *mini-*, *mega-*, *maxi-*, etc. It is also possible to extend the analysis to foreign constituents such as (Eng.) *free*, (Germ.) *frei*, (Dutch) *vrij*, in, respectively TAX-FREE, FEHRLERFREI ‘lit. without error, perfect’, AUTOVRIJ lit. ‘without car, ‘no car traffic’: these constituents, which were originally adjectives, have also been grammaticalized and hold as exponents of an LCR that has formed adjectives whose meaning is characterised by the absence of what denotes the lexeme-base.

#### 4.4. *-logue in the names of specialists: a suffix?*

The behaviour of *-log-* ([lɔg]) is complex in modern French. It holds in two distinct ways:

- (i) It appears in lexemes such as LOGOPATHIE ‘impairment of speech, of language faculty’, LOGOMACHIE ‘quarrel on words’, DIALOGUE ‘dialogue’ or MISOLOGUE ‘somebody who hates argumentation, an enemy of the scientific method’. *-Log-* in this case appears in either initial (LOGOPATHIE, LOGOMACHIE) or final position (DIALOGUE, MISOLOGUE) and means ‘speaking, discourse’, *i.e.* it has kept the meaning the deverbal noun *logós* had in old Greek. In modern French, this *-log-* is still used to form learned complex lexemes, but only in initial position: LOGOPATHIE but also LOGOPHASIE ‘logophasia’, LOGOPHILE ‘logophile’, LOGOMORPHISME ‘logomorphism’, LOGOSPHERE ‘logosphere’, etc. All the lexemes in which it appears in final position (DIALOGUE, MISOLOGUE) are borrowed from Greek.
- (ii) It also appears, only in final position, in nouns for specialists, such as PSYCHOLOGUE ‘psychologist’, SISMOLOGUE ‘seismologist’, DERMATOLOGUE ‘dermatologist’ EGYPTOLOGUE ‘egyptologist’, DÉCLINOLOGUE ‘specialist in decline’<sup>6</sup>, etc. This *-logue* does not mean ‘speaking, discourse’, appears in the structure of a great number of lexemes and is currently being used to coin many neologisms like DÉCLINOLOGUE, FUTUROLOGUE ‘futurologist’ or BOBOLOGUE<sup>7</sup>.

We propose to consider these two *log* as different: [lɔg]<sub>1</sub>(= *-log-* < *logós*) holds more or less like *-anthrop-*, *i.e.* is a B radical of a lexeme, and we do not propose to

<sup>6</sup> This word is a neologism often used in the French media today in an ironic sense to refer to the “specialists / experts” who consider France as a land in decline and who are always negative in their judgments about it.

<sup>7</sup> *Bobo* in *bobologue* is a kind of acronym for *BO*urgeois ‘middle-class’ *BO*hême ‘bohemian / unconventional’ and denotes a social group. A *bobologue* is a “specialist” on the people in this social group.

study it in this work. We will focus on [lɔg]<sub>2</sub> (= *-logue* < *lógos*), which does not work in the same way as [lɔg]<sub>1</sub> but whose status has not yet been established.

In old Greek, *lógos*, which *-logue* comes from, was a bound form connected to the verb *légo* ‘speak, say’, and appeared in lexemes such as *ἀστρολόγος* (> fr. ASTROLOGUE ‘astrologist’), *μυθολόγος* (> fr. MYTHOLOGUE ‘mythologist’); these lexemes received an agentive interpretation (*ἀστρολόγος* ‘a man who speaks about stars’, *μυθολόγος* ‘somebody who composes myths / legends’). This type of noun was therefore considered to have an NV structure.

In modern French, such an analysis is no longer valid, at least when attempting to account for the competence of a run-of-the-mill speaker, frequently non-grecophone: today, nouns such as PSYCHOLOGUE, DÉCLINOLOGUE or FUTUROLOGUE are not perceived as having a verb in their structure, nor do they have an agentive interpretation. The behaviour of *-logue* seems to be very different from that of constituents like *-cide*, *-vore*, *-phile*, *-fère* etc. in, for example, insecticide ‘insecticide’, *publivore* ‘publicity hound’, *CINÉPHILE* ‘cinema lover, film buff’ or *FLORIFÈRE* ‘flowering’. In modern French, these constituents keep the verbal interpretation they had in old Greek / Latin and have a predicate-argument relationship with the constituent on their left. They correspond (cf. Iacobini (2004), quoted § 2) to the VN structure of native compounds (*florifère* / *porte-plume* ‘penholder’; *insecticide* / *tue-mouche* ‘fly swatter’). But, whereas the native compounds are in their great majority nouns<sup>8</sup>, neoclassical compounds are uniformly A compounds<sup>9</sup>.

If *-logue* is not a verb in nouns for specialists, what is its status: the radical of a lexeme or an exponent of an LCR?

The first assumption raises problems because in old Greek, *logos* was a bound form: thus, we have to hypothesize that there was a process of reanalysis in which, little by little, in the *Xlogue* lexemes, the verbal bound form was reinterpreted as a nominal form, whose meaning was first ‘speaker’, then ‘specialist’. From this perspective, the complex lexemes *Xlogue* have an N<sub>2</sub>N<sub>1</sub> structure, in which N<sub>1</sub> is the determined and N<sub>2</sub> the determiner; the complex denoting a hyponym of N<sub>1</sub>. Yet a problem still remains: is the noun *-logue* the stem of a lexeme that does not have a syntactic realisation or a B stem of SPÉCIALISTE?

The second assumption is easier to support: from a synchronic point of view, *-logue* is the exponent of an LCR that forms nouns for specialists, such as *-iste* in, for example *PIANISTE* ‘pianist’, *DENTISTE* ‘dentist’ or *CHIMISTE* ‘chemist’. From a certain point of view, it is possible to say that, in these lexemes, *-iste* also means ‘specialist’, which does not prevent us from considering it a suffix. Given that *-logue* and *-iste* hold in the same way, we consider *-logue* to be a suffix, like *-iste*, and in figure 4 we propose a representation of this analysis:

	SIMPLE	COMPLEX
(G)	...	...logue
(F)	[...]	[...lɔg]
(SX)	cat:n	cat::n
(S)	object	Specialist of the object denoted by the simple

*Figure 4.*

<sup>8</sup> On VN compounds in French, cf. Villoing (2002).

<sup>9</sup> For a comparison between VN and NV compounding in French, cf. Namer & Villoing (forthcoming).

## 6. Conclusion

The analysis proposed in this paper allows us to shed light on certain crucial points:

- Not all neoclassical CFs can be analysed in the same way; on this point, we are in agreement with Iacobini (2004).
- The tools of lexematic morphology are both sufficient and suitable for analysis of neoclassical CFs (at least those examined herein): these elements can be exponents of LCRs (*micro-* and *-logue*) or suppletive stems of a lexeme, used in constrained contexts (*lud-*, *-anthrop-*).
- Neoclassical CFs, although much is made of their heterolexicity, are easily integrated into the patrimonial lexicon.

However, our analysis also reveals (at least) two weaknesses: (i) in the current state of our knowledge, a distinction cannot be made between allomorphy and suppletion (in analysis of this type, everything is suppletion), and (ii) the problem of the difference in order between neoclassical compounding (determiner + determined) and native compounding (determined + determiner), in French as well as in other romance languages, remains to be solved, especially since we claim that neoclassical CFs are integrated into the patrimonial lexicon. We will deal with (ii) in Amiot and Dal (forthcoming)

There is nevertheless one factor that lexematic morphology cannot account for: the transitional stages that lead from a lexeme to an affix in diachronic evolution, because the lexemes and the exponents of rules (affixes) are conceived to be of a different nature. In the cases of *micro-* and *-logue*, the analysis did not raise any particular problems because these two elements are well grammaticalized in exponents of LCRs, but this is not always the case: e.g. CFs such as *-cide*, *-vore* or *-phage*, which share some characteristics with lexemes and others with affixes (cf. Iacobini 2004). The theory will certainly need to be refined if problems of this sort are to be solved.

## References

- Bauer, Laurie (1979). "Against Word-Based Morphology." *Linguistic Inquiry* 10/3, 508-509.
- Bauer, Laurie (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie (1998). "Is there a class of neoclassical compounds, and if so, is it productive?" *Linguistics* 36/3, 403-422.
- Booij, Geert (1992). "Compounding in Dutch." *Rivista di Linguistica* 4/1, 37-59.
- Booij, Geert (2002). *The Morphology of Dutch*. Oxford: Oxford University Press.
- Corbin, Danielle (1992). "Hypothèses sur les frontières de la composition nominale." *Cahiers de grammaire* 17, 26-55.
- Corbin, Danielle (2001). "Préfixes et suffixes : du sens aux catégories." *Journal of French Language Studies* 11/1, 41-69.
- Corbin, Danielle (2004). "French (Indo-European: Romance)." In G. Booij, Ch. Lehmann, J. Mugdan and S. Skopeteas (eds), *Morphologie / Morphology. Ein internationales Handbuch zur Flexion und Wortbildung / An International Handbook on Inflection and Word-Formation*. vol. 2, Berlin / New York: Walter de Gruyter, [Article 121], 1285-1299.
- Corbin, Danielle and Paul, Jérôme (2000). "Aperçus sur la créativité morphologique dans la terminologie de la chimie." *La banque des mots* 60, 51-68.
- Dal, Georgette (2007). "Les adverbes en -ment du français : flexion ou dérivation ?" In N. Hathout and F. Montermini (eds), *Morphologie à Toulouse. Actes du colloque international de morphologie 4<sup>e</sup> Décembrettes*. München: Lincom Europa, 121-147.
- Fradin, Bernard (2000). "Combining Forms, Blends and Related Phenomena." In U. Doleschal and A. Thornton (eds), *Extragrammatical and Marginal Morphology*. München: Lincom Europa, 11-59.
- Fradin, Bernard (2003). *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.
- Fradin, Bernard and Kerleroux, Françoise (forthcoming). "L'identité lexémique." In B. Fradin, F. Kerleroux and M. Plénat (éds.), *Aperçus de morphologie du français*. Paris: Presses Universitaires de Vincennes.
- Heine, Bernd, Claudi, Ulrike and Hünnemeyer, Friederike (1991). *Grammaticalization*. Chicago: Chicago University Press.
- Hopper, Paul and Traugott, Elizabeth C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Iacobini, Claudio (2004). "Composizione con elementi neoclassici." In M. Grossmann and F. Rainer (eds), *La formazione delle parole in italiano*. Tübingen: Max Niemeyer Verlag, 69-95.
- Kerleroux, Françoise (2006). "Les théories morphologiques à la fin du XX<sup>e</sup> siècle." In S. Auroux, K. Koerner, H.-J. Niederehe and K. Versteegh (eds), *Histoire des sciences du langage*. Berlin / New-York: de Gruyter, [article 242], 2313-2324.
- Lehmann, Christian (1995). *Thoughts on Grammaticalization*. München / Newcastle: Lincom Europa.
- Lüdeling, Anke, Schmid, Tanja and Kiokpasoglou, Sawwas (2002). "Neoclassical Word Formation in German." *Yearbook of Morphology 2001*, 253-283.
- Matthews, Peter H. (1974). *Morphology. An Introduction to the Theory of Word-Structure*. Cambridge: Cambridge University Press.
- Namer, Fiammetta and Villoing, Florence (forthcoming). "Have *Cutthroats* to Do with *Tracheotomes*? Distinctive Properties of VN vs NV Compounds in French." *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5) Fréjus 15-18 September 2005*. University of Bologna.
- Olsen, Susan (2000). "Composition." In G. Booij and al. (eds), *Morphologie / Morphology*. Berlin: Mouton de Gruyter, [article 87], 897-916.
- Plag, Ingo (2003). *Word-Formation in English*. Cambridge: Cambridge University Press.

- Scalise, Sergio (1984). *Generative Morphology*. Dordrecht (Holland) / Cinnaminson (U.S.A.): Foris Publications.
- Tournier, Jean (1985). *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris / Genève: Champion / Slatkine.
- Villoing, Florence (2002). *Les mots composés [VN]<sub>N/A</sub> du français: réflexions épistémologiques et propositions d'analyse*. Thèse de doctorat. Université de Paris X-Nanterre.
- Warren, Beatrice (1990). "The Importance of Combining Forms." In W. Dressler *et al.* (eds), *Contemporary Morphology*. Berlin / New York: Mouton de Gruyter, 111-132.
- Williams, Edwin (1981). "On the Notions 'Lexically Related' and 'Head of a Word'." *Linguistic Inquiry* 12, 245-274.