

Neoclassical compounds as relational adjectives

Marine Lasserre, Fabio Montermini
CLLE-ERSS, CNRS & University of Toulouse
marine.lasserre@gmail.com, fabio.montermini@univ-tlse2.fr

1. Introduction

One way to interpret the meaning of a complex lexeme is to observe the parts of which it is constituted and the interactions between them. This compositional analysis is adopted, more or less explicitly, in the majority of studies on morphological derivation. The present study deals with mismatches that can occur between form and meaning, taking as example a special class of complex lexemes in French, so-called neoclassical compounds and we propose that some of them can be analyzed as pure relational adjectives. In particular, we focus our attention on French derivatives ending in *-logique*, which, we claim, should be considered as the output of a single constructional strategy, rather than as constructed by the sequential adjunction of the element *-logue* and of the suffix *-ique*. As our main objective here is to study the lexicon and in particular the mechanisms by which it is formed, we focus our attention on neologisms, i.e. on non ‘dictionarized’ words. Moreover, we defend a context-based analysis, since we propose to characterize the semantics of complex lexical items not on the basis of abstract labels given by the linguist, but on the basis of their contexts of use. Behind this way of dealing with the semantics of complex lexemes there is a basic assumption which is analogous to that of distributional semantics (see Lenci 2008 and Sahlgren 2008 for an overview), according to which it is possible to find a correlation between the distribution of two words and their semantic distance. We also strongly defend the idea that taking context into account is fundamental for studying the meaning of complex lexemes, and consequently of derivational constructions. The observation of single lexemes in isolation is not sufficient to give us a precise and realistic view of how derivational morphology works, in particular for those languages for which large textual resources are now available and easy to access.

This paper is organized as follows: in Section 2 we present some general characteristics of the units which are classified as neoclassical compounds in the literature, focusing in particular on the use as pure relational adjectives of some of them in everyday language; Section 3 presents the data on which this study is conducted, i.e. a Web-based corpus of NA pairs in French, where the adjective is a constructed one ending in *-logique*; in Section 4 this corpus is compared with two other corpora based, respectively, on the French Wikipedia and on ten years of the newspaper *le Monde*; finally, Section 5 contains some concluding remarks, and sketches an explanation for the tendency for adjectives containing a neoclassical element, such as those in *-logique*, to evolve as purely relational ones.

2. The problem: neoclassical compounds as relational adjectives

The complex lexemes we consider here belong to the class of so-called neoclassical compounds (for a more precise description, cf. Amiot & Dal 2007, 2008; Villoing 2012). Most neoclassical compounds have been originally coined in scientific and technical vocabularies, in many European languages, from the 17th century on. Today they are also found massively in everyday language (Lasserre 2013). They differ from native compounds in that they involve what we call a neoclassical element, which was a lexeme in Latin or Ancient Greek but has lost its autonomy. Neoclassical elements do not receive a unified treatment in the literature, in particular concerning their status: are they lexemes, affixes, or something in between? We are not going to solve this question in this paper, but we suspect, like Amiot &

Dal (2007), that all neoclassical elements do not belong to a homogeneous class. In particular, some of them which are really frequent and are used to form large series of lexemes (like *-crate*, *-logue* or *-cide*), can be considered as exponents of constructions: they are the formal expression of a semantic and syntactic operation, just like canonical affixes.

It can be easily observed that in many neoclassical compounds, all formally identified elements do not necessarily have the expected meaning.

(1)

- a. *mine carbonifère* (138)
mine coal + contain
'coal mine'
- b. *dictature pétrolifère* (559)
dictatorship petrol + contain
'petro-dictatorship'

(2)

- a. *personne bibliophile* (14)
person books + lover
'book lover'
- b. *connaissances cinéphiles* (825)
knowledge film + lover
'knowledge of film'

(3)

- a. *région viticole* (529,000)
region wine + culture
'wine-making region'
- b. *gastronomie ostréicole* (386)
gastronomy oyster + culture/farm
'oyster gastronomy'

In (1a)-(2a)-(3a) we show various contexts in which a neoclassical element (*-fère*, *-phile* and *-cole*) has the expected meaning (respectively 'production', 'lover' and 'culture/farm'), while in (1b)-(2b)-(3b) we show contexts in which the same neoclassical elements have unexpected meanings¹. In parentheses, we specify the number of occurrences of each collocation on Google. In these latter examples, only the distribution gives us a key to interpret the meaning of the lexemes in question. Moreover, if in (3a) the meaning of *viticole* is related to the meaning of the lexeme *viticulture* 'vine-growing', in (3b) the meaning of *ostréicole* is not related to *ostréculture* 'oyster-farming' but directly to *huître* 'oyster'. The neoclassical element *-cole* does not display the expected meaning 'culture, farm' but only marks the relation between the base lexeme *huître* and the head noun *gastronomie*. In consequence, we argue that these complex words are close to relational adjectives which involve suffixes whose only function is to relate a noun and an adjective.

As Roché (2006) – among others – argues, the suffix *-el* in *présidentiel*, built on *président*, marks a formal operation, a categorial operation but not a semantic operation. In (1b)-(2b)-(3b), the neoclassical elements *-fère*, *-phile* and *-cole* are not different in that sense from the suffix *-el*. Rainer (2013) considers that the use of the Latin element *-fère* for marking relational adjectives is due to the fact that the original Latin meaning is lost in the competence

¹ We consider that *carbon-*, *bibli-*, *viti-* and *ostréi-* are suppletive stems of the base lexemes, respectively, CARBON 'coal', LIVRE 'book', VIGNE 'vineyard' and HUÎTRE 'oyster'

of the majority of speakers who reinterpret ambiguous uses, an explanation we can extend to other cases of neoclassical compounds.

In this paper, we focus on the semantic characterization of the final sequence *-logique*, and we argue that, like *-cole* or *-fère*, it can be used to form relational adjectives. This element is part of the triad *-logue* ‘-logist’ / *-logie* ‘-logy’ / *-logique*. It has two distinct origins in Ancient Greek: the noun *logos* which means ‘discourse’ and the verb *legein* which means ‘collect, gather’. These two origins are not distinguished anymore and all words in *-logie* are perceived as constructed with a unique element.

If we look at the entries of the *TLFi* dictionary, summarized in Table 1, we observe that the element *-logie* has six main meanings, the element *-logue* two, and the element *logique* is just seen as forming adjectives corresponding to nouns in *-logie*. In fact, these meanings are not clear-cut, and we find several ambiguities: *étymologie* may correspond to meanings 1 and 6, *terminologie* to meanings 1 and 4 and so on. Furthermore, the availability of each meaning is not comparable: for instance, meaning 6 is very frequent whereas meaning 5 is no more found in neologisms.

Table 1: Correspondences between *-logue*, *-logie* and *-logique* in the *TLFi*.

	1	2	3	4	5	6
<i>-logie</i>	‘discourse’ <i>étymologie</i>	‘expression’ <i>dactylogie</i> ‘discourse with hands’	‘works in <i>n</i> parts’ <i>trilogie</i>	‘collection’ <i>terminologie</i>	‘behaviour’ <i>misologie</i> ‘refusal of speaking’	‘study’ <i>cardiologie</i>
<i>-logue</i>	‘practioner’ <i>philologie</i>					‘specialist’ <i>cardiologue</i>
<i>-logique</i>	adjectives corresponding to nouns in <i>-logie</i>					

From a formal point of view, the nouns in *-logie* could be considered as being formed on the nouns in *-logue* by a suffix having the form *-ie* (phonologically /i/) which triggers a modification of the preceding consonant. The ending *-ie* is particularly frequent in neoclassical composition: *-phile* / *-philie*, *-phobe* / *-phobie*, *-scope* / *-scopie*, *-trophe* / *-trophie*, *-mane* / *-manie*, etc. However, if this is a genuine case of derivation, its direction is reversed compared to the other means used for constructing nouns of specialists in French. Specialist nouns formed with the suffixes *-iste* and *-ien* take nouns of disciplines as bases, but the nouns of disciplines in *-logie* would have the specialist nouns in *-logue* as bases:

(4)

- a. *violon* ‘violin’ > *violoniste* ‘violinist’
- b. *musique* ‘music’ > *musicien* ‘musician’
- c. *cardiologie* ‘cardiology’ < *cardiologue* ‘cardiologist’

A cross-linguistic observation shows that, in English, the specialist noun in *-logist* seems built on the discipline noun in *-logy*, just as specialist nouns in *-ist* and *-ien*. These considerations make us think that the lexemes in *-logue* and those in *-logie* are not linked by an oriented derivational process. Consequently, we argue here that the relation between *-logie* nouns and *-logique* adjectives is not a derivation either and that they constitute two poles (out of at least three) of a more complex lexical network which at least also contains words in *-logue*.

In order to support this hypothesis, let us analyze the lexeme *météorologique*, ‘meteorological’. In a purely concatenative approach, we would analyze this complex word as

the relational adjective corresponding to the noun *météorologie* ‘meteorology’, identifying three elements inside it: *météoro-* ‘weather’, *-logie-* ‘study of’ and *-ique* ‘related to’.

However, the different interpretations this lexeme can have in the expressions in (5) make us think that the sequence *-logique* forms not only adjectives related to a *-logie* noun, but also adjectives directly related to the base.

(5)

- a. *recherches météorologiques* ‘meteorological research’
=> *météorologique*: ‘related to the meteorology as the study of the weather’
- b. *prévisions météorologiques* ‘meteorological predictions, weather forecast’
=> *météorologique*: ‘related to the weather as an object of study’
- c. *tempête météorologique* ‘meteorological storm, weather storm’
=> *météorologique*: ‘related to the weather’

If we now turn more specifically to the relations between the three elements *-logie*, *-logue* and *-logique*, we found different scenarios. First, a *X-logique* adjective is canonically related to a discipline noun in *X-logie*, as we see in (6) where *cardiologique* is undoubtedly constructed on *cardiologie*.

- (6) *La Fédération Française de Cardiologie apporte son soutien à la recherche cardiologique*
‘The French Cardiology Federation supports cardiological research’
[<http://www.fedecardio.org/qui-sommes-nous/financer-la-recherche/bourses-fedecardio>]

However, *X-logique* adjectives may also be related to a specialist noun in *X-logue* (7):

- (7) *Je sais uniquement que j'étais myope à moins 4.75 à l'oeil gauche avec de l'astigmatie (dont je ne connais pas le degré exact étant donné que je ne comprends pas le jargon ophtalmologique et que je ne sais pas lire une ordonnance d'ophtalmologue*
‘... I don't understand the ophthalmological jargon and I cannot read a prescription made by an ophthalmologist’
[<http://www.ophtalmologie.fr/operation-myopie/hopital-public-resultat-decevant-quelle-solution-t3840.html>]

This relation is more marginal and ambiguous, since *jargon ophtalmologique* can be interpreted as the jargon used in the discipline of ophthalmology and the jargon used by ophthalmologists. Finally, and most interestingly, a *X-logique* adjective can be directly related to the base.

- (8) *On est tellement habitués aux tempêtes métaphoriques (financières, médiatiques) qu'on est dépourvus face à une tempête météorologique.*
‘We are so used to metaphorical storms (financial, mediatic) that we are unable to face a meteorological storm.’
[http://www.liberation.fr/chroniques/2009/01/31/-_306669]

In (8), we cannot link the *X-logique* adjective to a *X-logie* (or *X-logue*) noun. The sequence *-logique* is therefore the exponent of a construction giving relational adjectives, just like the canonical relational suffixes *-ique* or *-el-log-* has not, or has lost, its semantic values ‘study’, ‘discourse’ or ‘collection’.

However, as nothing is so clear-cut in the observation of the lexicon, we also observe many cases in which the semantic interpretation is more ambiguous.

- (9) *L'examen ophtalmologique de routine comprend un certain nombre d'évaluations pour réaliser un bilan complet de votre vue.*
 'The routine eye examination includes a number of tests in order to obtain a global evaluation of your sight.'
 [http://www.docvadis.fr/croixrousseophtalmo/page/les_maladies_de_l_oeil/les_examens_complementaires_1.html]

In (9) the collocation *examen ophtalmologique* may be analyzed as 'examination within the discipline of ophthalmology', 'made by an ophthalmologist' or just as 'examination of the eyes'.

Furthermore, some fully lexicalized lexemes, like *biologique* 'biological/organic', progressively changed into qualifying adjectives.

- (10)
- a. *recherche biologique* 'research on biology'
 - b. *phénomène biologique* 'biological phenomenon'
 - c. *agriculture biologique* 'organic farming'

In (10a)-(10b) *biologique* is a relational adjective, related to biology in (10a) and directly constructed on the base *bio-* 'life' (10b). However, in (10c), the adjective cannot be specifically linked to any other lexeme and only qualifies a particular type of farming.

3. The relational use of neological relational adjectives

In order to analyse the availability of *-logique* and its semantics, we collected a corpus of adjectives containing this final sequence which are not recorded in the main French dictionaries (*TLFi* and *Le Grand Robert*). The corpus was made up by systematically searching on the Web the corresponding *-logique* adjective for a list of *-logue / -logie* lexemes collected from the French Google ngrams². For each of these adjectives (329 overall) we searched a sample of contexts on the Web, in order to dress a list of head nouns for the NA pairs in which they appear (henceforth the "corpus of neologisms"). We collected up to 35 different contexts per adjective (roughly corresponding to the first two pages of results provided by a search engine). Overall, our corpus of neologisms contains 2,279 NA pairs with 927 different head nouns. In (11), we give a sample of the head nouns encountered most frequently with *-logique* adjectives in the corpus of neologisms:

(11) <i>étude</i>	'study'	87
<i>recherche</i>	'research'	62
<i>analyse</i>	'analysis'	42
<i>approche</i>	'approach'	38
<i>science</i>	'science'	30
<i>connaissance</i>	'knowledge'	25
<i>donnée</i>	'data'	22
<i>littérature</i>	'literature'	20
<i>aventure</i>	'adventure'	17
<i>théorie</i>	'theory'	16
<i>travail</i>	'work'	16
<i>méthode</i>	'method'	15
<i>problème</i>	'problem'	15

² <http://storage.googleapis.com/books/ngrams/books/datasetsv2.html>.

<i>aspect</i>	‘aspect’	14
<i>examen</i>	‘exam’	14
<i>expérience</i>	‘experience’	13
<i>perspective</i>	‘perspective’	13
<i>savoir</i>	‘knowledge’	13
<i>tradition</i>	‘tradition’	13
<i>réflexion</i>	‘reflection’	12

Before going further into the analysis, we should spend some words about the use of the Web as a linguistic resource. As it is evident from the data above, even if we limit the collection of data to the pages indexed in the first four pages of results provided by a search engine (two in the singular and two in the plural form), we obtain a number of contexts which is much larger than what we can obtain from corpora which are among the largest available for French. The choice of limiting the analysis to the first four pages (maximum 40 contexts / *-logique* adjective) was due to the need of manually verifying each context (e.g. concerning the linguistic skills of the writer), a task which can only be accomplished on a limited number of data. However, we consider that this limitation does not affect the results, since (i) for the great majority of *-logique* adjectives in our corpus, search engines do not provide more than two pages of results (95/329 adjectives have more than 10 contexts, and 8/329 have more than 20); (ii) even if it is well known that the results provided by search engines are organized according to their own criteria, these criteria are certainly not susceptible of creating a linguistic bias for the research we wanted to carry out.

We coded all the head nouns in the corpus according to their compatibility with one of the meanings commonly assumed for *-logie* and listed in Table 1. The three meanings attested are ‘study of’, which is found in about two thirds of the head nouns in our corpus, ‘discourse on X’ and ‘collection of X’ which, on the other side, are quite marginal. Moreover, about a third of the head nouns in the corpus are not compatible semantically with any of the meanings of Table 1, or, if they are, they are too generic and undefined to be unambiguously linked with one of them (cf. *plan* or *niveau* in such expressions as *sur le plan X* ‘on the X plan’ or *au niveau X* ‘at a X level’). The exact figures are given in (12):

(12)

a.	‘study of X’	657
	<i>étude</i> ‘study’	
	<i>recherche</i> ‘research’	
b.	others	183
	<i>aventure</i> ‘adventure’	
	<i>moment</i> ‘moment’	
c.	generic	63
	<i>plan</i> ‘plan’	
	<i>niveau</i> ‘level’	
d.	‘discourse on X’	16
	<i>discussion</i> ‘discussion’	
	<i>débat</i> ‘debate’	
e.	‘collection of X’	9
	<i>inventaire</i> ‘inventory’	
	<i>patrimoine</i> ‘patrimony’	

Although, as we have already observed, the distinction between the three main meanings of *-logue / -logie* is not always neat, we may assume – without going further into this matter – that there is a hierarchy in the semantic instruction of this morphological construction (13).

(13) ‘study of X’ > ‘discourse on X’ > ‘collection of X’

A further observation of the head nouns in the NA pairs related to the ‘study of X’ meaning (whose interest will become clear below) allows producing a more fine-grained semantic classification of them (semantic classes are given in the order of importance):

(14)

- a. object / concept
donnée ‘data’
théorie ‘theory’
- b. activity
étude ‘study’
recherche ‘research’
- c. property
connaissance ‘knowledge’
savoir ‘knowledge’
- d. output / result
découverte ‘discovery’
ouvrage ‘work’
- e. place / institution
centre ‘centre’
institut ‘institute’
- f. human
expert ‘expert’
confrère ‘colleague’

Unlike the words in *-logue / -logie*, however, those in *-logique* display, at least in our corpus of neologisms, a large proportion of examples whose semantics is not compatible with any of the ‘canonical’ meanings of these neoclassical compounds. It is precisely on this class of items that we want to focus our attention, and we do the claim that they correspond to a class of purely relational adjectives.

As far as the semantic relation between a *logique* adjective and its base is concerned, those observed in our corpus of neologisms correspond to those existing for dictionaryed items:

(15)

- a. ‘related to the science of X’ (*Xlogie + -ique*)
... *chacun saura rendre grâce à ces infatigables mineurs du savoir discologique.*
(‘knowledge in records’)
[<http://www.pop-hits.net/article-les-promesses-de-l-aube-episode-2012-starring-jerry-mengo-96203143.html>]
- b. ‘related to the specialists of the science of X’ (*Xlogue + -ique*)
Un album-concept est un terme discologique qui traduit la volonté de la part d'un artiste ou d'un groupe de créer une œuvre filée tout au long d'un disque.
(‘term of music industry’)
[http://fr.wikipedia.org/wiki/Vocabulaire_du_rock]

c. ‘related to X’ (X + *-ique*)

Le tout autour et avec les excellentissimes choix discologiques de Sebastien Tison, dont la playlist figure dans le même blog.

(‘choices of records’)

[<http://patwhite.com/node/4314>]

We consider that we can take this as a proof of the fact that in French we have three distinct but parallel constructions whose exponents are, respectively, *-logue*, *-logie* and *-logique*, and that each of them is available with new bases. The last one, in particular, may construct derivatives whose meaning is distant from the canonical ones observed with lexicalized *-logue* / *-logie* lexemes, and which, in general, is hard to characterize, due to the variety of meanings of the bases this construction applies to, and of the NA pairs its outputs enter in. This already seems an argument for classifying these adjectives among the relational ones. Moreover, it is quite frequent that in this case the noun chosen as a base designates an object for which there is no socially recognized science or discipline, and consequently no institutionalized name for it and / or its specialists, if not in an ironic or fictional use. In this case too, we may sketch a semantic characterization of the most frequent head nouns.

(16)

a. ACTIVITY / PERIOD OF TIME

aventure bistrologique (< bistro) ‘adventure in a bistro’

soirée vinologique (< vin) ‘wine party (evening)’

b. PROPERTY

vitesse escargotologique (< escargot) ‘snail speed’

rhétorique injurologique (< injure) ‘rhetoric of insults’

c. CONCRETE OBJECT

bougie aromacologique (< arôme) ‘aromatic candle’

kit bobologique (< bobo) ‘first-aid kit’

d. HUMAN

coach tarologique (< tarot) ‘tarot coach’

Interestingly, the most frequent semantic classes for these lexemes correspond quite well to those observed for the meaning ‘study of’, although we cannot see any trace of this meaning in the examples above. However, although observing their proportion is interesting, as it allows making a correlation between the two sets of lexemes, one may also claim that the meanings in (16) correspond to quite general semantic classes, thus suggesting that a precise semantic characterization does not give any special cue in understanding the behavior of these adjectives. We take it as a further argument in favor of the hypothesis that they are relational adjectives.

To sum up, what we propose is that (i) among the semantic instructions of an adjective in *-logique*, there is one (corresponding to the examples *tempête météorologique* in (5c) and *choix discologiques* in (15c)) in which the meaning of the adjective is directly connected with the meaning of the base, with the *-log-* sequence providing no specific semantic contribution; (ii) the availability of this semantic value determines the possibility of having adjectives in *-logique* directly constructed on a noun with a purely relational meaning.

4. A distributional analysis of *-logique* adjectives: a comparison of three corpora

The classes and the figures listed above can give us a rough idea of the semantic value of *-ique* adjectives, in particular when they are new, non-lexicalized lexemes. However, due to the polysemy of several base nouns and head nouns and to the various ambiguities we pointed out, a precise semantic characterization of them is hard to obtain.

For this study, in particular, we compared our corpus of neologisms with the two resources *les voisins de Wikipédia* (henceforth vdW)³ and *les voisins du Monde* (vdM)⁴. These two resources are based, respectively, on the French Wikipedia (collected in 2008) and on 10 years of the newspaper *le Monde* (1991-2000). They are tagged for syntactic context, and are presented via an interface which allows searching the syntactic relation a word enters into.

We systematically searched the head nouns for all *-logique* adjectives in the two corpora. The results (including the Web-based corpus of neologisms) are presented in Table 2, and the ten most frequent head nouns in the vdW and vdM corpora are given in Table 3:

Table 2: Number of *-logique* adjectives, head Ns and NA pairs in the three corpora considered.

Corpus	Nr. Words	A in <i>-logique</i>	Head N	NA pairs
vdWikipédia	262 M	55	705	1,505
vdMonde	200 M	57	305	934
Web		329	927	2,279

Table 3: Ten most important head nouns in the vdW and vdM corpora.

vdW			vdM		
<i>point</i>	‘point’	19	<i>étude</i>	‘study’	15
<i>étude</i>	‘study’	17	<i>recherche</i>	‘research’	15
<i>recherche</i>	‘research’	17	<i>analyse</i>	‘analysis’	13
<i>plan</i>	‘plan’	15	<i>ordre</i>	‘order’	13
<i>aspect</i>	‘aspect’	13	<i>approche</i>	‘approach’	12
<i>système</i>	‘system’	13	<i>donnée</i>	‘data’	12
<i>chirurgie</i>	‘surgery’	13	<i>point</i>	‘point’	12
<i>donnée</i>	‘data’	12	<i>raison</i>	‘reason’	11
<i>problème</i>	‘problem’	12	<i>plan</i>	‘plan’	10
<i>niveau</i>	‘level’	11	...		

If our main hypothesis is correct, i.e. that *-logique* adjectives tend to acquire a purely relational function, we should observe a difference between the traditional corpora (which represent, respectively, a technical and a more formal register) and the data from the Web, which represent a less formal register, possibly more suitable to provide hints on how speakers use derivational morphology in a creative way. We can expect, in fact, that the traditional corpora contain a greater proportion of head nouns which are compatible with the core meaning(s) of the adjectives in *-logique*, whereas in the corpus of neologisms we can observe a large variety of meanings among head nouns. In other words, we expect to observe

³ <http://redac.univ-tlse2.fr/applications/vdw.html>

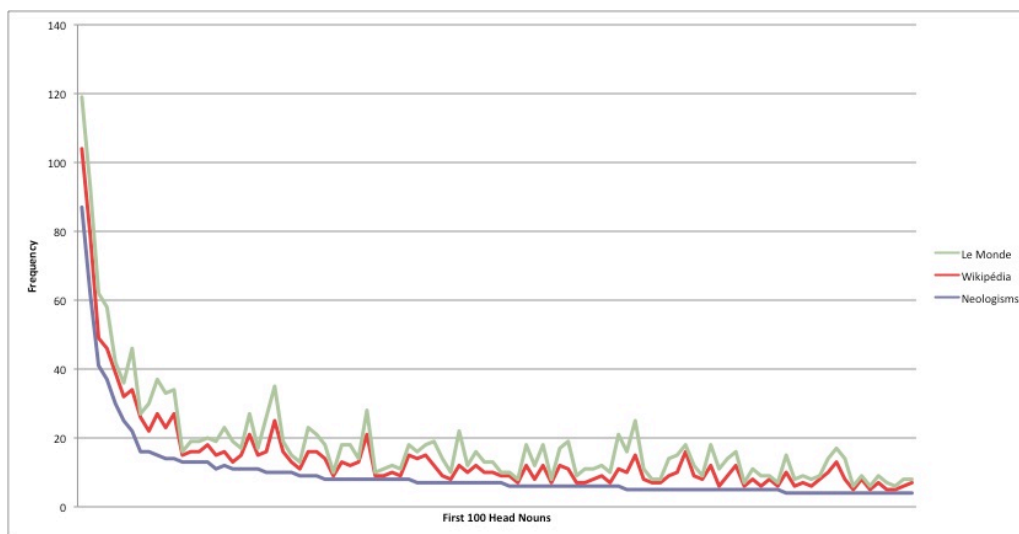
⁴ <http://redac.univ-tlse2.fr/applications/vdlm.html>

less distance between the vdW and vdM corpora than between the latter and the Web-based corpus of neologisms.

A first observation concerns what we can call the ‘density’ of contexts for *-logique* adjectives. If we consider, for instance, the vdW and the Web corpus, the first has an average of 12,81 head nouns per each *-logique* adjective, whereas the second has an average of 2,81. The bigger dispersion for *-logique* adjectives in the Web corpus can be explained on the basis of the fact that they function more like pure adjectives, for which the choice of head nouns is larger, as it is less constrained by semantic factors.

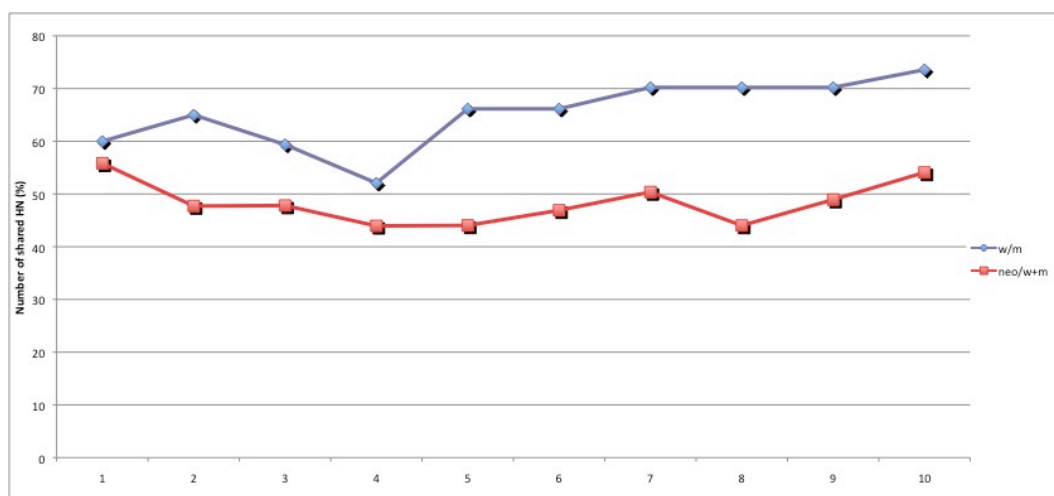
In order to obtain a more homogeneous set of data, we limited our analysis to the head nouns that appear in all three corpora (304 overall). Figure 1 shows the frequency in each corpus for the first 100 head nouns. Data were ordered according to their frequency (descending) in the Web corpus. As we can see, the two curves corresponding to the vdW and to the vdM corpora are much closer than the one referring to the Web corpus.

Figure 1: Frequency of the first 100 head nouns of *-logique* adjectives in the three corpora.



In a second time, we calculated the number of head nouns that *-logique* adjectives in the three corpora have in common, among those ranked from 1 to 100 by frequency, by portions of 10 (Figure 2).

Figure 2: Head nouns shared by the three corpora, among the 100 most frequent ones, in portions of 10.



As it can be seen, the vdW and vdM corpora systematically share more head nouns than with the corpus of neologisms, up to 73,57% among the first 100, versus 54,08% (in average, the precise figures are 63,88% between the corpus of neologisms and vdW and 44,28 between the corpus of neologisms and vdM). This suggests that in non-specialized, everyday language the use of *-logique* adjectives is less constrained, an hypothesis which is compatible with the idea of a progressive loss of the original semantic instruction linked with *-log-*.

5. Conclusion: integrating neoclassical compounds into a larger model of lexical creativity

Both the semantic observation of the corpus and the distributional analysis we provided go in the direction of our main hypothesis, i.e. that the sequence *-logique* is now the exponent of a construction in French which is autonomous, although it enters into a lexical network with at least two other constructions, namely those forming *-logue* and *-logie* nouns. Semantically, this construction is close to other constructions forming relational adjectives. That means that its main function is to transform a noun into an adjective, without any supplementary semantic load. We claimed that, for *-logique* adjectives, this is made possible by the fact that, even in the institutionalized portion of the lexicon, adjectives in *-logique* (i) can be often used to simply refer to the base noun and not to a discipline, (ii) display several systematic ambiguities among their various meanings. One question remains open though: why *-logique*, which originally is made up with a neoclassical compounding element and a relational suffix, has acquired this function? More in general why neoclassical compounding elements behave like this? In other words, why French has added new constructions, with new exponents, to those already existing for the construction of relational adjectives (e.g. *-ique*, *-ien*, *-al*, *-ier*, etc.). This seems to go against the idea of an economic organization of the lexicon and of the derivational morphology of languages, in which the ideal situation is that each meaning is constructed by one element only, and redundancy is avoided.

A first answer is that the use of *-logique* instead of, say, simply *-ique* fulfills various pragmatic functions. In many of the examples we found the adjectives in question are used in an ironic or creative manner. Moreover, there seems to be a strong tendency, at least in French, to correlate the length of a word and the fact that it belongs to a learned, prestigious register of the language, in particular when it contains a sequence that appears in genuine neoclassical compounds. There are, however, two other facts that we want to point out. On the one side, several studies on the morphology of French have shown that in this language, more than in others, word formation is strongly submitted to size constraint, and that complex words have preferred prosodic shapes, e.g. concerning the number of syllables of the derivate and/or of the base (Plénat 2009, 2011). We did not treat our corpus in a systematic way in this respect, but it seems quite clear to us that this explanation may hold for such derivates as *arbrologique* (< *arbre* ‘tree’), *biérolgique* (< *bière* ‘beer’), *rhumologique* (< *rhume* ‘cold’), in which a monosyllabic base attaches to a disyllabic exponent, as disyllabic derivates are in general dispreferred in French. Moreover, it has been shown (for instance by Roché 2009, 2011b) that in order to account for the output form of complex word, it is not only necessary to refer to semantic and phonological constraints, but also to purely lexical ones. A better derivate is one that fits into a dense lexical network. Figures 3, 4 and 5 show the proportion of adjectives in *-ique* in the *lexique.org* database⁵ according to the uni-, bi- and trigrams preceding the suffix.

⁵ <http://www.lexique.org>

Figure 3: *-ique* adjectives in the lexique.org database according to the preceding segments (unigrams).

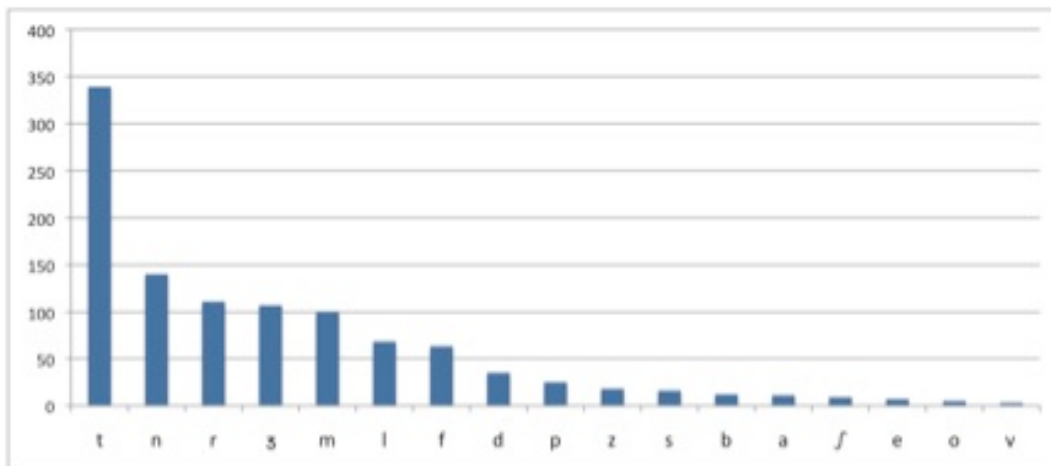


Figure 4: *-ique* adjectives in the lexique.org database according to the preceding segments (bigrams).

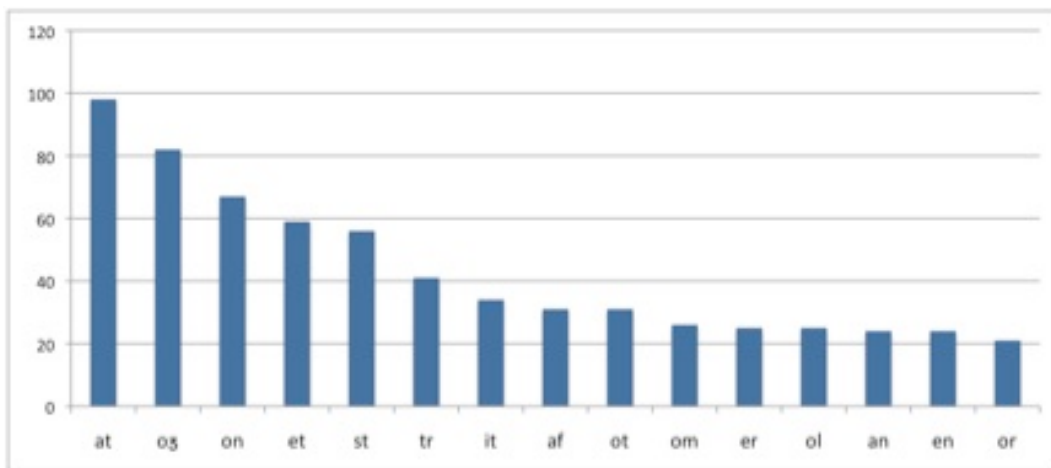
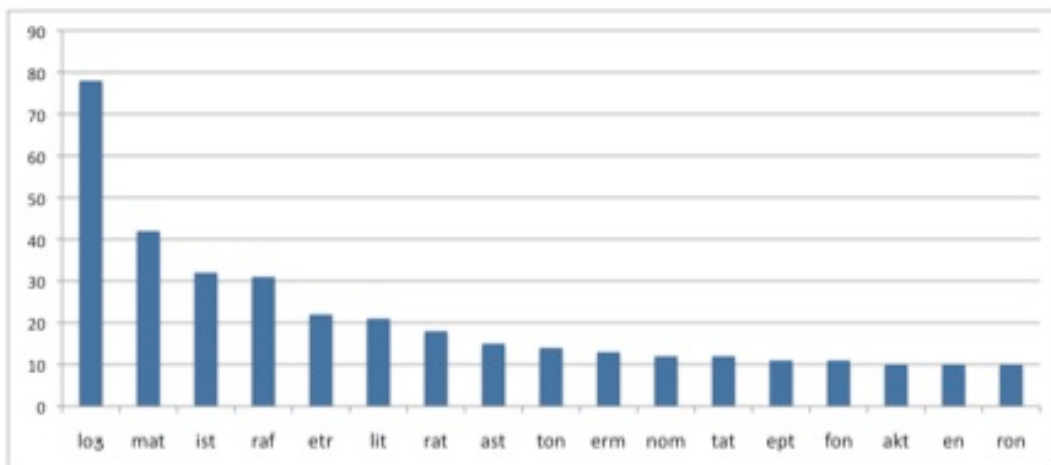


Figure 5: *-ique* adjectives in the lexique.org database according to the preceding segments (trigrams).



As we see, the segments corresponding to *-log-* are among the most common for *-ique* adjectives, and the proportion becomes dramatic when the preceding trigram is taken into account, the sequence *-logique* representing 7,2% of all *-ique* adjectives (the second one, *-matique*, represents 3,8%). We can thus consider that, even from a purely phonological point

of view, the sequence *-log-* is activated in the speakers' competence, as being one of the most likely for preceding the suffix *-ique*.

More generally, the data and the analysis we presented militate against the possibility, for complex lexemes, to be segmented into a base and one or more affixes in a discrete way. Moreover, in line with what has already been proposed for other types of suffixation in French (e.g. by Roché 2011a on *-iste*), they suggest that the traditional view of complex words as exhaustive (in which all formal elements contribute equally to meaning) and complete (in which all pieces of meaning are conveyed by a specific formal element) is too simplistic. Thus, to conclude, the rules for word formation should be viewed mainly as means for inserting new complex words into existing lexical networks rather than one-way relations between a base and a derivate.

References

- Amiot D. & G. Dal (2007) Integrating neoclassical combining forms into a lexeme-based morphology. In G. Booij, L. Ducceschi, B. Fradin, E. Guevara, A. Ralli & S. Scalise (Eds.), *On-Line Proceedings of the Fifth Mediterranean Morphology Meeting*, 323-336.
- Amiot D. & G. Dal (2008) La composition néoclassique en français et l'ordre des constituants. In D. Amiot (ed.), *La composition dans une perspective typologique*. Arras: Artois Presses Université, 89-113.
- Lasserre M. (2013) Neoclassical compounds and language registers. In *Actes de la 4e édition des JéTou, 16 et 17 mai 2013*. Toulouse: Université de Toulouse 2, 98-108.
- Lenci A. (2008) Distributional semantics in linguistic and cognitive research. *Italian Journal of Linguistics* 20(1), 1-31.
- Plénat M. (2009) Les contraintes de taille. In B. Fradin, F. Kerleroux & M. Plénat (Eds.), *Aperçus de morphologie du français*. Saint-Denis: Presses Universitaires de Vincennes, 47-63.
- Plénat M. (2011) Enquête sur divers effets des contraintes dissimilatives en français. In M. Roché, G. Boyé, N. Hathout, S. Lignon & M. Plénat, *Des unités morphologiques au lexique*. Paris: Hermès-Lavoisier, 145-190.
- Rainer F. (2013) L'interprétation des composés latins du type *aurifer* en français. In G. Dal & D. Amiot (Eds.), *Repères en morphologie*, 182-214 [http://stl.recherche.univ-lille3.fr/textes_enligne/Reperes-Morphologie/Rainer-Reperes_morphologie_p182-214.pdf]. [Reprint of: Rainer F. (2002) L'intégration des composés latins du type *aurifer* en français. *Sillexicales* 3, 151-168.
- Roché M. (2006) Comment les adjectifs sont sémantiquement construits. *Cahiers de grammaire* 30, 373-387.
- Roché M. (2009) Pour une morphologie *lexicale*. In *Mémoires de la Société de Linguistique de Paris*. Nouvelle série 17, 65-87.
- Roché M. (2011a) Quel traitement unifié pour les suffixations en *-isme* et en *-iste*? In M. Roché, G. Boyé, N. Hathout, S. Lignon & M. Plénat, *Des unités morphologiques au lexique*. Paris: Hermès-Lavoisier, 69-143.
- Roché M. (2011b) Quelle morphologie? In M. Roché, G. Boyé, N. Hathout, S. Lignon & M. Plénat, *Des unités morphologiques au lexique*. Paris: Hermès-Lavoisier, 15-38.
- Sahlgren M. (2008) The distributional hypothesis. *Italian Journal of Linguistics* 20(1), 33-54.
- Villoing F. (2012) French compounds. *Probus* 24(1), 29-60.