A Reference Generative Lexicon for French

Namer, Fiammetta; Bouillon, Pierrette; Jacquey, Evelyne

A Reference Generative Lexicon for French

Fiammetta Namer
Université Nancy2 et ATILF
CLSH - 23 bd Albert 1er
BP3397
54015 Nancy Cedex, France
fiammetta.namer@univ-nancy2.fr

Pierrette Bouillon
ETI/TIM/ISSCO
Université de Genève
40, Bvd du Pont d'Arve,
1205 Genève, Suisse
pierrette.bouillon@issco.unige.ch

Evelyne Jacquey
ATILF-CNRS
44 avenue de la Libération
BP30687
54063 Nancy Cedex,
France
evelyne.jacquey@atilf.fr

Abstract

This paper describes an original approach which aims at building a reference semantic lexicon for French. Its main characteristic is that of relying on morphological properties. The method combines morphological analysis results from large-scale lexical resources (i.e. word lists from the 'Trésor de la Langue Française informatisé' (TLFi)) with already tested acquisition methodologies on lexical information. The representation format, within the Generative Lexicon framework, has been chosen for its expressiveness and conciseness features. This approach allows us to consider building a reference lexicon for French, which is fundamentally homogeneous and also has large coverage. A feasibility study of the described method provides a projection of expected results, from both quantitative and qualitative points of view.

Keywords

lexical acquisition, reference lexicon for French, generative lexicon model, word formation, corpora, semantics.

Introduction

In this article, we present a methodology to automatically build a French lexicon in the Generative Lexicon (GL) formalism. It relies on two different acquisition methods: linguistic rules based on morphological constraints (called Lexeme Formation Rules, henceforth LFRs, cf. (Aronoff, 1994)(Fradin, 2003)) and learning rules, from the machine readable version of the *Trésor de la Langue Française* dictionary (henceforth TLFi[1]) and/or corpora. The latter are used either for non-constructed words, when LFRs do not apply, or when they provide underspecified or ambiguous results.
With this combined approach, we want to ensure the coherence of the lexicon. LFRs define general semantic structures that apply to similar word classes and can therefore be systematically mapped onto corresponding GL structures. For example, all deverbal adjectives (such as WASHABLE or PERISHABLE) denote an expected property, whatever the type of verb, be it agentive (1a) or unaccusative (1b). This property can be represented in a GL structure where the predicate is encoded in the telic role.

(1) a. washable shirt
b. perishable good

In the following, we present this methodology in detail. We first situate it in a more general context, to show its specificity. Then we illustrate it through concrete examples that show the respective contributions of LFRs, and lexical and textual resources. Finally, we provide a quantitative evaluation of the expected results.

2 Motivation

The development of any lexical resource should face two questions (Bouillon, Busa, 2001):

1) Can the resource be reproduced using a set of consistent criteria? This question is important since a positive answer would

[1] TLFi : "Trésor de la Langue Française informatisée": electronic version of the 16 volumes French dictionary 'Trésor de la langue Française', freely available at URL : http://atilf.atilf.fr/tlf.htm

mean that the principles underlying the resource are independent of domain and people's intuition.

2) Does the framework provide a good basis for developing large-scale resources on systematic grounds?

Compared with other similar projects (e.g. Acquilex, cf. (Copestake *et al.*, 1993), *Simple* (Busa *et al.*, 2001) and *Clips*, cf. (Calzolari *et al.*, 2003)) our approach allows us to answer these questions in a original way. In order to acquire the GL structures, we reuse an existing morphological analyser, called Dérif (cf. (Namer, 2002; 2003; 2005b)). The GL structures are then derived from the set of morphological and semantic features extracted by this morphological analyser, with two positive consequences. Firstly, the GL representations are defined by the morphological process involved that motivates them. Secondly, all the words that belong to the same morphological type will receive the same GL representation, which ensures the global coherence of the resource.

In return, we hope for practical and theoretical outcomes. On the theoretical side, we first complement Derif with a deeper semantic representation. Remember that the morphological contribution to the construction of lexical meaning is only partial: a LFR only provides fundamental elements of the semantic interpretation that has to be specified by the context (cf. (Aronoff, 1980)). The qualia structure will reorganise them using a theoretical vocabulary rich enough to understand how words are composed and with explanatory force. Secondly, we confirm the GL theoretical basics, by answering the following question: does a GL provide a good framework to describe all aspects of lexical meaning for different types of words and on a large scale? Ideally all the properties that can be predicted by LFRs should be represented formally by the qualia structure. From this structure, different information can then be extracted to serve different practical goals. In particular, (Bouillon *et al.*, 2000; Claveau *et al.*, 2001; Claveau *et al.*, 2003) showed how qualia noun-verb relations can be used to improve information retrieval.

An open question is of course whether this resource will be concretely used by people. We think that by linking this resource to an existing morphological analyser specially conceived for unknown words we made it particularly suitable to answer the fundamental question of language creativity. In the following we present this method in detail.

3 Methodology for lexicon acquisition

Acquiring lexical entries makes the two approaches described below work together. The former is based on the use of morphological knowledge (section 3.1), the latter on corpus-based learning methods (section 3.2). These two methods interact to complement each other and to validate, invalidate or define each method results (section 3.3).

3.1 Acquiring lexical entries through Lexeme Formation Rules (LFR)

This approach makes use of results which are first provided by the morphological parser DeriF (Namer, 2002; 2003; 2005b), then reformatted to be modelled on GL notations, according to (Jacquey, Namer, à paraître; Namer, Jacquey, 2003; soumission) extensions.

3.1.1 DeriF results

DeriF simulates Lexeme Formation Rules (LFRs) and thus provides a lexeme L with its morphological base B[2]. L-to-B pairing is completed with annotations that reflect constraints LFRs prototypically impose on L and/or B. For instance, Figure 1 illustrates how the LFR for *dé-* prefixed deadjectival verbs encodes the analysed L lexeme: $\text{DESSOULER}_{\text{VERB}}$ (*sober up*) and its B base: SOUL_{ADJ} (*drunk*). Apart form the L-to-B relation itself (lines 1, 2), the rule provides the related lexemes with several features: B has to be a **qualifying** adjective and describes a **temporary** property (line 3). The verb (line 4) is **dynamic**, either **causative transitive**, thus with a **cause** subjet ("le café salé dessoule Max": *salty coffee sobers Max up*) or **intransitive resultative** -also known as anticausative-, thus with a **theme** subjet ("Max dessoule" : *Max sobers up*).

[2] DériF's analysis of L includes a verification step in which L is checked against an appropriate exception list, containing non-compositional lexemes, that are no longer analysable as morphologically constructed.

1	DESSOULER/VERB (*sober up*) ==> SOUL,ADJ (*drunk*)/dé:prefix
2	"(suppress – deprive from) SOUL character"
3	SOUL/ADJ:(predicative,_,temporary)
4	DESSOULER/VERB:(dynamic,trans., [cause,theme],causative)|| (dynamic,intrans.,[theme], resultative)

Fig 1 : DeriF analysis of $\text{DESSOULER}_{\text{VERB}}$

3.1.2 Translating results into GL format

According to what DeriF produces two GL formatted entries – that is, both L's and B's, if L and B are morphologically related by a LFR - are in the best case generated.

Each entry specification detail level depends upon the kind of information DeriF provided. In the above example, the causal chain steps that constitute the complex event defined by $\text{DESSOULER}_{\text{VERB}}$ (Figure 2a) can be used to derive the qualia roles, namely: (1) **y** individual is SOUL (*drunk*) (presupposed initial state, encoded in AGENT|AGENT|AGENTIVE role) ; (2) either **x** agentive individual DESSOULE (*sobers up*) **y** (causation, in AGENT|AGENT role) or **y** DESSOULE (*sobers up*) (activity, in AGENT|FORM|FORM role) ; (3) **y** is no longer SOUL (*drunk*) (final state, FORM role). As far as the base adjective SOUL is concerned, the only feature inferable from *dé-* prefixation rule is that it refers to a property identified from an event point of view as a state (e) affecting an individual (y), (cf Figure 2b). For both DESSOULER and SOUL entries, Argument Structure and Event Structure values are instantiated from that of Qualia Structure.

(a) dessouler

FORM		not 1 (soul'(e1 :state,y :ind)
AGENT	FORM	FORM 2 dessouler-act'(e2 :act,y)
		AGENT 1
	AGENT	CAUSE(x :agent, 2)

(b) saoul

FORM	soul'(e :state,y :ind)

Fig 2 : Translating DESSOULER (a) and SOUL (b) meaning into QUALIA roles from DeriF parsing

3.2 Corpus learning based acquisition

DeriF only provides information when simple lexemes function as bases for complex morphological lexemes. In other cases, our methodology relies on a corpus-based learning technique. One particular corpus will be used, that of TLFi lexicographical data. A first experiment, realised on the part-of-speech and XML-tagged TLFi version, shows for instance that these dictionary definitions are regular enough for an easy detection of nouns referring to prototypically functional entities, which therefore instantiate the TELIC role within GL model. When looking at definitions such as "useful/allows to, intended/used for", one detects circa 14% nouns with a TELIC function (4279 of the 30544 TLFi nominal lexemes), which is a not insignificant result. Moreover, searching these verbal expressions usually brings back nouns such as "device, organ, instrument", which can help define FORMAL types (eg. "broom = housework $\underline{\text{ustensil}}_{\text{FORM}}$ used for $\underline{\text{sweeping}}_{\text{TELIC}}$"). These nouns (e.g. *utensil*) allow in turn finding new verbal expressions, and thus to detect new telic predicates, and so on... This preliminary validation confirms that the TLFi definitions corpus can be used in the framework of this French semantic lexicon automatic construction.

3.3 Crossing the two approaches

The two above presented acquisition methods can be combined in order to specify some pieces of information that LFR let underspecified. Underspecification has two possible origins. Either a rule does not propose sufficient clues to precisely instantiate lexical meaning (cf section 3.3.1); or it forms inherently ambiguous constructed lexemes (cf section 3.3.2). In both cases, a corpus-based approach contribution allows for the provision of missing information.

3.3.1 Morphology and under-specification: dé- prefixation

Let us consider the LFR for denominal *dé-* prefixed verbs. It selects two possible verb meanings, and thus two possible verb classes. The first verb class base noun describes the initial localisation of what is referred to by the verb theme. DE**TERRER**$_{VERB}$ (*dig*$_{VERB}$ *up*) = "remove something from **TERRE**$_{NOUN}$ (*earth*$_{NOUN}$)" is an instance of this class. On the contrary, the second verb class base noun DENOTES the entity which undergoes a change of localisation. This is observed, for instance, with DÉSOSSER$_{VERB}$ (*bone*$_{VERB}$) = "remove OS$_{NOUN}$ (*bones*$_{NOUN}$) from something". Clearly, DériF automatic analysis is unable to determine the (locatum/located) role of *dé-* prefixed verb's base nouns, since this role can be detected only by means of extralinguistic characteristics. So, *dé-* prefixed verb analysis is systematically accompanied by a disjunction of two definitions, each corresponding to a verb class. For instance, DETERRER$_{VERB}$ is given the following disjunctive gloss (line 2):

```
1 DÉTERRER/VERB (dig up) ==> TERRE,NOUN (earth)/dé:prefix
2 "Remove smth from TERRE || Remove TERRE from smth"
3 déterrer/VERBE: (dynamic,transitive ,[cause,theme], causative)
```

Fig 3 : DériF analysis of DETERRER$_{VERB}$

This analysis illustrates several facts. On the one hand, no constraint on the verb base can be generalized: nouns selected by *dé-* LFR are either concrete (TERRE (*earth*) : DÉTERRER), or abstract (COURAGE : DÉCOURAGER (*discourage*)). On the other hand, the rule imposes transitivity on the resulting verb and describes a change-of-localisation accomplishment. To sum up, there are 3 constraints involved by denominal *dé-* prefixed verb formation rules : (1) verbs are transitive, causative, implying a cause **x** and a theme **y**; (2) verbal process denotes the act of the causer on the theme ("**x** causes something to happen to **y**"); (3) the verbal processes ends up in a final state affecting either **y** (such as with DÉTERRER) or the entity referred to by the base noun (such as with DÉSOSSER). In fact, with "**x** deterrer **y**" the relocated entity is **y**, which is removed from its initial ground[3], i.e. TERRE : *Max*$_x$ *déterre le coffre*$_y$ (Max$_x$ digs up the chest$_y$). Conversely, with "**x** désosser **y**", **y** denotes the initial ground from where the figure[3] entity referred to by OS is relocated : *Max*$_x$ *désosse le poulet*$_y$ (Max$_x$ bones the chicken$_y$). So, given some dénominal *dé-* préfixed verb, the LFRs at play here are unable to determine which of verb theme or base plays the figure role (and which, consequently, plays the ground role); therefore the rule cannot unambiguously define this verb with respect to its base. This is the reason why two definitions are systematically provided, such as with DÉTERRER, as indicated by Figure 3, line 2. Facing such ambiguities, which linguistic rules cannot solve, two cases may occur: either only one meaning is actually attested (this is the prototypical case), or both readings exist, but with different frequencies. Web corpus and TLFi definitions-based analysis may help either to remove ambiguities or to weigh each interpretation probability. For instance, with 'x déterrer y', the reading "x remove earth from y" is highly unlikely: it lacks from both TLFi definitions and Google first 100 result pages.

3.3.2 Morphology and exceptions : verb-to-nouns -oir suffixation rule

Whereas the previous example showed how dictionary definitions together with text content specify lexical information that morphology is only able to sketch, let us look now at the opposite (and more frequent) case : that of a corpus-based analysis helping the detection of exceptions to a specific LFR, namely that producing deverbal *–oir* suffixed nouns. Except for some nouns playing other thematic roles of the verb base predicate, for instance its patient (TIROIR (*drawer*) : "that what one *draws* (TIRER)"), *-oir* LFR mainly forms deverbal nouns referring to locations or instruments[4]. The difference between the two concepts is sometimes fine, as soon as the object referred to by the deverbal noun

[3] Figure and ground terms are those of (Talmy, 1983). Other linguists from cognitivism frameworks resp. use trajector and landmark (Langacker, 1987) or cible and site (Vandeloise, 1986).

[4] Among thematic roles definitions, that of (Fillmore, 1968) says that an instrument is "the inanimated force or object used by an Agent and causally involved in the action or state identified by the verb" and the location is "the place of an event".

possesses the required size to fulfil both roles. This is what is observed with most deverbal –*oir* nouns: ABREUVOIR (*watering place*), BALANÇOIRE (*swing*), ÉGOUTTOIR (*draining board*)[5] all denote entities which are used both as instruments, favouring the verbal process progress, and as process localizations. Other –*oir* suffixed deverbal nouns are polyreferential: they may refer to two distinct entities, such as HACHOIR[6] which denotes either a cutting instrument (*cleaver*) or the place where cutting takes place (*chopping board*). Finally, some nouns clearly identify exclusively either places: DORTOIR (*dormitory*), FUMOIR (*smoking room*) or instruments: RASOIR (*razor*). Figure 4 shows that DériF analysis of these nouns reflects the polysemy that characterizes most of them. On the other hand, almost nothing can be predicted for base verbs. They can be transitive: ABREUVER / ABREUVOIR, ergative: DESSOULER / DESSOULOIR ('*sobering up room*'), unergative: TROTTER (*trot*) / TROTTOIR(*sidewalk*), unaccusative: MOURIR (*die*) / MOUROIR ('*dying place*'). All in all, the only property common to all these verbs is that of being dynamic.

```
abreuvoir/NOUN ==> abreuver,VERB/oir:suffix

"Instrument of ABREUVER || Place of ABREUVER"

abreuver/VERB:          (dynamic,-,-,-)
```

Fig 4 : DériF analysis of $\text{ABREUVOIR}_{\text{NOUN}}$

Representing abreuvoir properties within GL, according to the *–oir* LFR predictions requires the following facts to be expressed: (1) a noun designates an entity which possesses an **intrument•localisation** dotted type, which indicates its prototypical intrinsic polysemy; (2) it denotes one of its base verb participants; (3) this base verb instantiates the noun TELIC function, which means that the verb names the derived noun expected possible function. Any *–oir* suffixed noun is encoded according to these indications, which are meant to reveal the assumed default ambiguity which affects this noun type. This hypothesis has then to be checked against corpora content: does the noun co-occur with "avec" ('with') ? Is it found within a locative complement? Depending on the answers to these questions, either polysemy is confirmed, or the noun encoding has to be refined (to single **instrument** or **localisation** type).

4 Evaluation

We can already evaluate the expected results both from a qualitative and a quantitative point of view, in terms of the number of GL entries that can be generated from Dérif and type of corresponding GL representation. For now, at least 35.5% of the lexemes of the TLFi are analysable by Dérif (i.e 35,263 out of 99,445) in the form of 45.478 base-to-constructed lexeme relations[7]. This process requires around 85 LFRs. The rate of 35.5% can be explained by two factors : (1) not all LFRs have been implemented so far (verb to noun suffixation rules, for instance, are still missing), and (2) a large amount of TLFi lexical entries are not morphologically constructed. Fig. 5 shows the most frequently applied ones (the methodology developed to enhance DériF with new LFRs is explained in (Namer, 2003)).

Type of morphological derivation rule	Constructed Lexeme category : Rule (Base Category)
Suffixation	A :el, ique, if, al,eux, aire, ien, iste, (N), **able**(V), N : **eur**(V), ie, **ité** (A), V :**ifier**, **iser**(N,A), re (V)
Prefixation	V :**en**, a (A,N), **é, dé** (A,N,V), pré (V) A :**in**, hyper,sub,non (A), sur,anti,sub,sous, mono, poly, auto (N)
Conversion	N ->V, **A->V,** V->N, A -> N

Fig 5 : Rules Forming more than 80 TLFi lexemes

Among these rules, those in bold are already associated with constraints on the input and output, as illustrated in Fig. 1 (for adjs-to verbs *dé-* prefixation), 3 (for nouns-to-verbs *dé-* prefixation), 4 (for deverbal nouns in *-oir*) and

[5] base verbs being respectively ABREUVER (*water*), BALANCER (*swing*) and ÉGOUTTER (*drain*)

[6] Its verb base HACHER means either *mince* or *chop*

[7] Among them, 39.028 are derivation rules, and 6.450 are neo-classical compounding rules (see below).

6 below (for deverbal *–able* suffixed adjectives). They could be directly mapped to the corresponding GL entry that in turn can be refined by corpus learning.

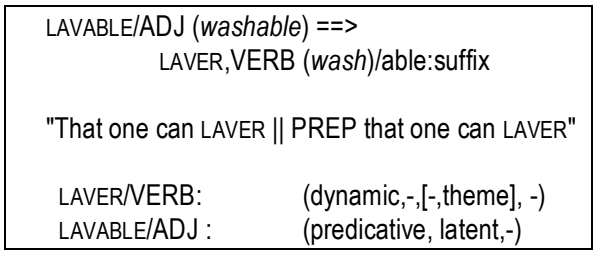

Fig 6 : DeriF Analysis of $\text{LAVABLE}_{\text{Adj}}$

The last rule in Fig. 6 is another example that illustrates the need for corpora. The default constraint in the rule indicates that the noun modified by the adjective is the theme of the base verb from which the adjective is derived. However (Hathout *et al.*, 2003) has showed that depending on the adjective (and thus the base verb) all arguments can play this role, cf. « un poisson/saison/étang pêchable » (*a fishable fish/season/lake*). The only way to complete the entry is therefore to find examples in corpora like: "pêcher un poisson" (*fish a fish*), "pêcher **dans** un étang" (*fish **in** a lake*), "pêcher **pendant** une saison" (*fish **during** a season*).

One of the advantages of Dérif is that it can be applied to any lexicon that can be used as input for this methodology and be used to increase the number of derived GL entries. For example, if we apply Dérif to a biomedical lexicon[8] results are quite different than those shown in Fig 5. Here, at least 59% of words (17,297 out of 29,273 lexical entries) are morphologically constructed. The lexeme formation rules involved mainly produce the so-called neo-classical compound lexemes. These lexemes are formed by compounding rules, that differ from standard compounding in several respects : rule semantics, compounds structure, but also the components involved (often greek or latin bases) and the textual domains (mainly specialized and technical ones). For instance, in the above-mentioned medical specialized lexicon, 13,237 out of the 21,757 morphologically constructed lexemes are neo-classical compounds. A recently performed quantitative evaluation of DériF against a Gold Standard ((Namer, Baud, 2007)) has shown a score of 77.3 % of correct analyses by DériF. Another experiment ((Namer, Baud, 2007),(Deléger *et al.*, 2007)) has proved that for specialized lexicons DériF LFRs can easily be translated for other languages; in particular, rules for so-called neo-classical compounds have been successfully transposed in English ((Deléger, Namer *et al.*, 2007))**.**

[8] This lexicon has been collected form several sources (MeSH, SNOMED among others), in the framework of two R&D French projects : UMLF and VumeF. For more details about the lexicon sources, see (Namer, 2005a)

5 Conclusion

In this paper, we presented a methodology for deriving a GL. It is original because it relies mainly on the morphological properties of the lexicon. On the practical side, this approach allows us to ensure the coherence of the produced lexicon. On the theoretical side, it shows how morphology can collaborate with other sources of knowledge (dictionaries, corpora) to derive deep representations of meaning. We don't know of any other attempts to take advantage of all of these different properties together. The expected results are (1) a GL lexicon, and (2) a set of tools to dynamically extract new entries; among these tools, we expect the further development of Dérif, the design of models to transform the output of Dérif into GL entries and that of extraction rules from corpora/TLFi.

References

Aronoff, Mark (1980). Contextuals. *Language* 56 4, 744-758.

Aronoff, Mark (1994). *Morphology by Itself*. Cambridge: MIT Press.

Bouillon, Pierrette, Busa, Federica (2001). Qualia and the Structuring of Verb Meaning. *The Language of Word Meaning*. Bouillon, P., Busa, F. Cambridge: Cambridge University Press**:** 177-197.

Bouillon, Pierrette, Fabre, Cécile, Sébillot, Pascale, Jacqmin, Laurence (2000). Apprentissage de ressources lexicales pour l'extension de requêtes. *TAL* 41 2, 367-393.

Busa, Federica, Calzolari, Nicoletta, Lenci, Alessandro (2001). Generative Lexicon and the Simple model, developing Semantic

Resources for NLP. *the Language of Word Meaning*. Bouillon, P., Busa, F. Cambridge: CUP**:** 333-349.

Calzolari, Nicoletta, Bertagna, Francesca, Lenci, Alessandro, Monachini, Monica (2003). New perspectives for lexical web resource in the Semantic Web Scenario. *Generative Approaches to the Lexicon*.

Claveau, Vincent, Sébillot, Pascale, Bouillon, Pierrette, Fabre, Cécile (2001). Acquérir des éléments du lexique génératif : quels résultats et à quels coûts ? *TAL* 42 3, 729-753.

Claveau, Vincent, Sébillot, Pascale, Fabre, Cécile, Bouillon, Pierrette (2003). Learning Semantic Lexicons from a Part-of-Speech and Semantically Tagged Corpus Using Inductive Logic Programming. *Journal of Machine Learning Research* 4, 493-525.

Copestake, Ann, Antonio, Sanfilippo, Briscoe, Ted, De Paiva, Valeria (1993). The ACQUILEX LKB : an introduction. *Inheritance, defaults and the lexicon*. Briscoe, T., Copestake, A., De Paiva, V. Cambridge: CUP**:** 148-163.

Deléger, Louise, Namer, Fiammetta, Zweigenbaum, Pierre (2007). Defining Medical Words: Transposing Morphosemantic Analysis from French to English. *MEDINFO*.

Fillmore, Charles (1968). The case for case. *Universals in Linguistic Theory*. Bach, E., Harms, R. New-York: Holt, Rinehart, and Winston**:** 1-88.

Fradin, Bernard (2003). *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.

Hathout, Nabil, Plénat, Marc, Tanguy, Ludovic (2003). Enquête sur les dérivés en *-able*. *Cahiers de Grammaire*. Hathout, N., Roché, M., Serna, N. Toulouse: ERSS. 28**:** 49-91.

Jacquey, Evelyne, Namer, Fiammetta (à paraître). Morphosémantique et modélisation : le cas des verbes dénominaux préfixés par *é-*. *Actes du Colloque "Le sens en linguistique" (25-27 mai 2003)*.

Langacker, Ronald (1987). *Foundations of Cognitive Grammar*. Standford: Standford University Press.

Namer, Fiammetta (2002). Acquisition automatique de sens à partir d'opérations morphologiques en français : études de cas. *Actes de Traitement Automatique du Langage Naturel (TALN) 2002*, 235-244.

Namer, Fiammetta (2003). Automatiser l'analyse morpho-sémantique non affixale : le système DériF *Cahiers de Grammaire*. Hathout, N., Roché, M., Serna, N. Toulouse: ERSS. 28**:** 31-48.

Namer, Fiammetta (2005a). Guessing the meaning of neoclassical compound within LG: the case of pathology nouns. *Generative Approaches to the Lexicon 2005, (may 19-21)*, 175-184.

Namer, Fiammetta (2005b). *La Morphologie Constructionnelle du Français et les Propriétés Sémantiques du Lexique - Mémoire présenté dans le cadre de l'habilitation à diriger des recherches*. UFR Sciences du Langage. Nancy: Université de Nancy2.

Namer, Fiammetta, Baud, Robert (2007). Defining and relating biomedical terms : towards a cross-language morphosemantics-based system. *International Journal of Medical Informatics* 76, 226-233.

Namer, Fiammetta, Jacquey, Evelyne (2003). Lexical Semantics and derivational morphology: the case of the popular *é-* prefixation in French *GL 2003 : 2nd International Workshop on Generative Approaches to the Lexicon (May, 15-17 2003)*, 115-122.

Namer, Fiammetta, Jacquey, Evelyne (soumission). Word Formation Rules and the Generative Lexicon: Representing noun-to-verb versus verb-to-noun Conversion. *Generative Lexicon Book*. Bouillon, P., Isahara, H., Kanzaki, K., Pustejovsky, J. Dordrecht: Kluwer.

Talmy, Leonard (1983). How language structures space. *Spatial orientation, Theory, Research, and Applications*. Pick, H., Acredolo, L. New York: Plenum**:** 225-282.

Vandeloise, Claude (1986). *L'espace en français: sémantique des prépositions spatiales* Paris: Les éditions du Seuil.