Terminology Structuring Through the Derivational Morphology

Natalia Grabar[1] and Thierry Hamon[2]

[1] Université Paris Descartes, Faculté de Médecine; Inserm, UMR_S 729; SPIM, 75006 Paris, France
natalia.grabar@spim.jussieu.fr
[2] LIPN – UMR CNRS 7030, Université Paris-Nord, Avenue J.B. Clément, 93430 Villetaneuse, France
thierry.hamon@lipn.univ-paris13.fr

Abstract. In this work, we address the deciphering of semantic relations between terms in order to build structured terminologies. We study particularly the contribution of morphological clues. Among linguistic operations proposed by the morphology, we analyze affixation and suppletion. We show interpretative schemata emerging from morphologically formed lexemes and corresponding terminological relations. Morphology appears to be a useful tool for the deciphering of semantic relations between terms.

1 Introduction

Terminological resources allow to encode knowledge of a given technical or scientific area. The content of such resources is known to depend on the application needs [1], furthermore collected terms can be organized in a more or less sophisticated way: a simple list of terms, term records in term banks, structured terminology, etc. In this work, we focus on the building of structured terminologies from textual corpora. We rely particularly on the deciphering of semantic relationships between terms based on their morphological structure. We address here mainly affixation clues and show how they can be used for terminology structuring. Using morphological operations for such a task is suggested by the fact that these operations are one of the basics for the formation of vocabulary.

We start with the presentation of types of semantic relationships between terms in structured terminologies (sec. 2), and approaches currently used for the deciphering of such relationships (sec. 3). We then present the material we need for the deciphering of semantic relations through morphological operations (sec. 4), and we describe and analyze clues provided by morphology as well as types of relationships which can be induced with these clues (sec. 5). We apply this method on terms from two areas: medicine and cogeneration[1] [2]. We finally conclude and draw some perspectives (sec. 6).

[1] Cogeneration is a technology used for the generation of electricity. It allows to combine the generation of electricity and of heat (hot water, steam, ...)

T. Salakoski et al. (Eds.): FinTAL 2006, LNAI 4139, pp. 652–663, 2006.

2 Semantic Relationships in Structured Terminologies

The structuring of terms can be obtained using different types of semantic relationships. According to the nature of related terms, we distinguish three types of relationships [3]: taxonomic, lexical and transversal.

Taxonomic relationships organize terms within a tree. Two kinds of taxonomic relations can be distinguished: hierarchical and partitive. *Hierarchical* (`is-a`, *subsumption*, *hyperonymy* or *hyponymy*) relationships link a generic term to its specific terms. These relationships are most present in structured terminological resources and traditionally forms their backbone. Examples above are from medical terminology SNOMED [4]:

pneumonie (*pneumonia*)	`is-a`	*bronchopneumonie* (*bronchopneumonia*)
pneumopathie inflammatoire (*pneumonitis*)	`is-a`	*bronchopneumonie* (*bronchopneumonia*)
bronchopneumonie (*bronchopneumonia*)	`is-a`	*maladie de l'appareil respiratoire* (*diseases of the respiratory system*)

Partitive (*meronymy*, *mereology*, `part-whole` or `part-of`) relationships are often used to describe artefacts and living organisms through the enumeration of their constituant parts. When they are assimilated to hierarchical relationships, they can ensure the hierarchical structuring of terms as well:

poumon (*lung*) `part-of` *appareil respiratoire* (*respiratory system*)

Lexical relationships are established between terms which are subsumed by the same hyperonym. These relationships correspond to two types: synonymy and antonymy. They are less frequently identified in terminologies than taxonomic relationships. *Equivalence* relationships (*synonymy*) link terms which refer to the same entity. For instance, the terms *pneumonie* (*pneumonia*) and *pneumopathie inflammatoire* (*pneumonitis*) are synonymous within the medical terminology SNOMED. Synonymy can also relate variants of a given term. *Opposite* (*adverse* or *antonymy*) relationship relies on co-hyponyms which are not synonyms [5]. This relationship exists for instance between cogeneration terms *électricité nucléaire* (*nuclear power*) and *électricité non nucléaire* (*not nuclear power*) [2].

Transversal relationships relate terms which are located in different branches of hierarchical tree. When these relationships are under-specified they are addressed as `see-also` relationships. They can also be used to finely depict the domain knowledge. In postcoordinated terminologies these relationships ensure defining complex terms on the basis of their known semantic primitives and composition rules [6,7]. For instance, the diagnosys *pneumonie* (*pneumonia*) can be defined as a morphological affection *inflammation* (*inflammation*) which is located in a body part *poumon* (*lung*):

pneumonie (*pneumonia*) → `is-a` → *inflammation* (*inflammation*)
pneumonie (*pneumonia*) → `located-in` → *poumon* (*lung*)

3 Approaches for Mining Relationships Between Terms

Two types of approaches for the deciphering of semantic relationships between terms are usually distinguished: external approaches based on the analysis of contexts in which terms occur, and internal approaches based on the analysis of the internal structure of terms. Both are related to given types of relationships and are necessarily overseen by human expertise. Automatic tools based on these approaches are often designed for the detection of taxonomic relationships, whereas synonymy pays less attention. As for transversal relationships, they are neglected but can take advantage of the acquisition of taxonomic and synonymous relationships, being sometimes their side effect.

External approaches rely on text corpora. On one hand, they aim at the detection of expressions and phrases sensitive to contain given type of relationships between two terms, i.e. markers and lexico-syntactic patterns [8,9,10,11,12]. On the other hand, they aim at the detection of common contexts or cooccurrences of terms and then group them into homogeneous classes [13,14,15]. They can provide with various semantic relationships.

Among internal approaches for terms structuring, we distinguish lexical inclusion analysis for the deciphering of taxonomic relationships [16,17,18], transformation rules for detection of synonyms and morpho-syntactic variants of terms [19,20], and analysis of morphological structure of words. As we are particularly interested in this last work, we present it in more detail.

Lexical functions [21], can thus be used for the encoding of semantic relationships between terms [22] or even for their automated detection [23,24]:

$\mathbf{Real}_1(\textit{logiciel}) = \textit{excuter, faire tourner}$
$\quad(\mathbf{Real}_1(\textit{program}) = \textit{to run})$
$\mathbf{S}_0(\textit{programmer}) = \textit{programmation}$
$\quad(\mathbf{S}_0(\textit{to program}) = \textit{programming})$

In these examples, lexical functions indicate the nature of semantic relationships between lexemes ($\mathbf{Real}_1$ for `realizes`, $\mathbf{S}_0$ for `subject`) and can lead to semantic and terminological relationships between them. Furthermore, the work described in [25] presents denominal adjectives as clues for the detection of semantic relations between terms. The aim of our work is to propose a more complete investigation of morphological clues for terminology structuring. We rely on the analysis of these clues which are suggested by derivational morphology, and specifically by affixation and suppletion.

4 Material

The main condition needed for the deciphering of semantic relationships between terms through the morphology is the understanding of morphological operations as well as their impact on the semantics of formed lexemes. This can be performed when studying linguistic morphological description of languages. Indeed, derivational morphology proposes a set of linguistic operations which allow the

creation of "new" lexemes (nouns, adjectives, verbs, etc.). We adopt here the approach of morphological analysis of lexemes as proposed in [26], which is particularly suitable for natural language processing, as it allows to explain the construction of the meaning of lexemes. In this approach, the meaning is conveyed by morphological operations and operators (affixes, bases, compounds), and the meaning of complex lexemes is constructed at the same time as their form. Derivational morphology proposes the following main types of operations:

- *Affixation* (*derivation*), which combines bases (*artery*, *stenosis*) and affixes (suffixes, such as *-al*, *-ic*, prefixes or infixes):
  {*artère*, *artér*iel} ({*artery*, *arter*ial})
  {*sténose*, *sténot*ique} ({*stenosis*, *stenot*ic})
- *Conversion* describes formations where morphologically related lexemes have the same form but different syntactic categories:
  {*muqueuse*/A, *muqueuse*/N} ({*mucous*/A, *mucous membrane*/N})
  {*wound*/N, *wound*/V}
- *Compounding* combines at least two bases and forms compound lexemes:
  artery and *scopy* gives *arterio*scopy

Suppletion, which is not a morphological operation, allows relating and substituting bases which have the same meaning but are provided by different languages, often Greek and Latin. Suppletion can appear both in affixation and compounding operations. For instance, *estomac* (*stomach*) (latin word) can be substituted with *gastr-* (greek word) and *foie* (*liver*) with *hepat-* (greek word):

{*estomac*, *gastr*ique} ({*stomach*, *gastr*ic})
{*foie*, *hépat*ique} ({*liver*, *hepat*ic})

Affixation is a basic morphological operation widely used in many languages. But both compounding and suppletion, being related to the use of greek and latin words, are often reserved for some specialized languages, like medecine, agronomics or biology. We assume that all morphological operations can potentially lead to the deciphering of semantic relations between terms and are then useful for terminology structuring.

Morphological parsing and the analysis of lexemes can be reached with a list of morphological operators [27], morphological lexicon [28,29,30,31] or automated tools [32]. In this work, we mainly use a morphological lexicon built in a previous work [33] and some affixes which convey suitable semantic relations for terminology structuring. Both are related to suppletion and affixation operations. We use a medical dictionary [34] as reference knowledge for the validation of the morphologically built meaning of medical lexemes and terms.

5 Morphologically-Induced Terminology Structuring

5.1 Suppletion

Suppletion provides with relations of equivalence and pseudo-synonymy between two bases from different languages: *estomac* (*stomach*) and *gastr-*, *foie* (*liver*)

and *hepat-* in previous examples. By extension, affixed and compound lexemes in which suppletive bases are the only difference can be considered as equivalent or (quasi-)synonymous [35]. The semantic proximity between lexemes coined with suppletive bases is reinforced if involved morphological operations are equivalent. The following first example corresponds to the coining of verbs from two noun bases with the same meaning *pierre* (*stone*), the second example corresponds to the coining of adjectives from two bases which mean *estomac* (*stomach*):

1. *pért-* and *lith-* $\longrightarrow$ *pétr*ifier and *lith*ifier
2. *estomac* and *gastr-* $\longrightarrow$ *stomac*al (*stomach*al) and *gastr*ique (*gastr*ic)

Both examples allow the construction of lexemes with very close meanings.

5.2 Affixation

Affixation refers to the creation of lexemes by adding affixes to bases. The meaning of affixed lexemes is the result of influence of the meaning conveyed by affix on its base. We present here prefixation and suffixation operations and their possible semantic involvement in the terminology structuring.

Prefixation

Construction of opposite meaning. The opposite meaning can be constructed with negation prefixes *dé-*, *ir-*, *anti-*, *non-*, *in-*, or with privative prefixes *a-*, *dys-* [36]. The usefulness of some of such prefixes for terminology structuring has already been noticed in [23]. In the present work, we used a set of 52 word pairs linked together with such morphological operations (*i.e.*, {*accessible*, *inaccessible*}, {*fonction*, *dysfonction*}), and this allowed us relating 40 pairs of medical terms, among which 30 pairs comprise complex terms. All the induced pairs have been validated as correct:

- {*activateur du plasminogène*, *inactivateur du plasminogène*} ({*plasminogen activator*, *plasminogen inactivator*})
- {*kératose*, *dyskératose*} ({*keratosis*, *dyskeratosis*})
- {*continence fécale*, *incontinence fécale*} ({*fecal continence*, *fecal incontinence*})
- {*cycle mensuel normal*, *cycle mensuel anormal*} ({*normal menstrual cycle*, *abnormal menstrual cycle*})

Construction of meaning for spatial localization. Transversal relations for relative spatial localization can be detected through prefixes like *sur-*, *sous-*, *contre-*, *péri-* [36]. We used 99 word pairs linked with 12 such prefixes, and related 40 term pairs, among which 21 pairs with complex terms. All the induced pairs have been considered as correct:

- {*abcès rénal*, *abcès périrénal*} ({*renal abscess*, *perirenal abscess*})
- {*hyperplasie kystique*, *hyperplasie intrakystique*} ({*cystic hyperplasia*, *intracystic hyperplasia*})
- {*cervicite chronique*, *endocervicite chronique*} ({*chronic cervicitis*, *chronic endocervicitis*})
- {*région auriculaire*, *région sous-auriculaire*} ({*auricular region*, *infraauricular region*})

Other prefixes can lead to other meanings possibly useful for terminology structuring: temporal localization (*pré-*, *post-*), comparison (*super-*), etc.

Suffixation

*Construction of agent and action nouns.*When constructing agent and action nouns, morphological rules apply suffixes to verbal bases V. Agent nouns are mainly formed with the suffix *-eur* (*-or*, *-er*), and receive the general semantic instruction *The agent that V*. Action nouns are formed with suffixes *-age*, *-ade*, *-erie*, *-ment*, *-tion* or *-ure* (*-age*, *-ing*, *-ment*, *-tion*), and receive the general semantic instruction *Action of V* or *Result of V*. These rules allow the detection of transversal relations among terms, called respectively `actor-of`, `action-of` and`result-of` [23]. The suffix *-eur* (*-or*, *-er*) is ambiguous and can match with substrings which are not affixes, like in the following examples:

vap<u>eur</u>s de métal (*metal fumes*)
tum<u>eur</u> à cellules géantes (*giant cell tum<u>or</u>*)

Despite this the suffix *-eur* allows the deciphering of semantic relations between terms with a good precision. Here are few examples from medicine (ME) and electricity (EL) areas:

ME - *activat<u>eur</u> du plasmogène* (*plasminogen activat<u>or</u>*)
- *buv<u>eur</u> modéré de boisson alcoolisée* (*alcoholic beverage moderate drink<u>er</u>*)
- *gros fum<u>eur</u> de cigarettes (plus de 20 cigarettes par jour)* (*heavy smok<u>er</u> (over 20 per day)*)
- *marqu<u>eur</u> lymphocytaire* (*lymphocyte mark<u>er</u>*)
- *dialys<u>eur</u> péritonéal* (*peritoneal dialyz<u>er</u>*)

EL - *product<u>eur</u> d'électricité* (*electricty produc<u>er</u>*)
- *capt<u>eur</u> solaire* (*solar panel*)
- *consommat<u>eur</u> éligible* (*eligible consum<u>er</u>*, *eligible custom<u>er</u>*)
- *disjonct<u>eur</u> de couplage* (*circuit break<u>er</u>*)
- *cogénérat<u>eur</u>* (*cogenerat<u>or</u>*)
- *construct<u>eur</u> de turbine* (*turbine manufactur<u>er</u>*)
- *compress<u>eur</u> gaz* (*gas compress<u>or</u>*)

As for suffixes that label process or its result they allow the detection of terms which introduce `action-of` and `result-of` relations:

ME - *bless<u>ure</u> par balle* (*gunshot wound*)
- *brûl<u>ure</u> avec carbonisation* (*burn injury with charring*)
- *bloc<u>age</u> congénital* (*congenital obstruct<u>ion</u>*, *congenital block<u>ing</u>*)
- *tampon<u>ade</u>* (*compress<u>ion</u>*, *compressed structure*, *tampon<u>ade</u>*)
- *tatou<u>age</u>* (*tattoo*)
- *ablat<u>ion</u>* (*excis<u>ion</u>*, *ablat<u>ion</u>*, *abscess<u>ion</u>*, *extirpat<u>ion</u>*)
- *absorpt<u>ion</u> intestinale anormale* (*abnormal intestinal absorpt<u>ion</u>*)

EL - *alimentation électrique* (*electricity supply*, *electricity supplying*)
- *production éolienne* (*production of aeolian energy*)
- *raccordement au réseau* (*connection to the mains*)
- *aspiration* (*sucking*, *suction*)
- *refoulement* (*delivery*)
- *protection électrique* (*electric protection*, *electric depositing*)

Construction of nouns with partitive meaning. The morphological rule for the construction of nouns with collective meaning operates suffixes: *-ade*, *-age*, *-ail(le)* and *-ure* [37], which sometimes correspond to *-ing* in English. Constructed nouns mean that they contain one or more occurrences of base noun. It possibly introduces the `part-of` relations [37,23]. These suffixes are ambiguous and require human validation of induced relations.

ME *ossature* (*skeletal system*, *skeleton*)
palmature congénitale (*congenital webbing*, *congenital membrane*)
vergeture (*linear atrophy*, *stretch marks*)
vomissure gastrique (*gastric vomitus*)
arcade sus-pyramidale du rein (*arcuate artery of kidney*)
cordage tendineux (*chordae tendineae*)

EL *sciure* (*sawdust*)
outillage (*tools*, *equipment*)
câblage électrique (*electric cabling*)

Construction of relational meaning: denominal adjectives. The rule, which coins denominal or relational adjectives, applies a set of suffixes to noun bases. Among these suffixes we have *-aire*, *-el*, *-al*, *-ique*, *-eux*, *-ien*, *-in*, *-ois* and *-é* (e.g. *-al*, *-ant*, *-ary*, *-ic*, *-ous*, *-ive* and *-'s*). Constructed adjectives receive the general semantic instruction *Relative to N* and allow then the indentification of semantic relations among base nouns and their derived adjectives:

{*aorte/Nom*, *aortique/Adj*} ({*aorta/Noun*, *aortic/Adj*})
{*germe/Nom*, *germinal/Adj*} ({*germ/Noun*, *germinal/Adj*})

But, when occurring in noun phrases, relational adjectives establish semantic relations among base nouns as well (*aorte* (*aorta*), *germe* (*germ*)) and their head nouns (*sténose* (*stenosis*), *cellule* (*cell*)) [38]:

sténose aortique (*aortic stenosis*)
cellule germinale (*germinal cell*)

These indications can be used for the deciphering of semantic relations between terms t_1 and t_2, especially when t_2 contains adjective formed with such suffixes on the basis of the head noun of the term t_1. [38] distinguishes two types of relations among relational adjective and its head noun. These relations are close to the `part-of` relation, namely belonging and possession:

- The belonging relation is constructed, in French, with suffixes *-é*, *-aire*, *-eux*, *-in* and *-ique*. The head noun corresponds to the whole entity, while the base noun of the adjective to its part:
  nerf denté (*dentate nerve*, *tooth*ed *nerve*): which means that the *Nerve is tooth-shaped* or *Nerve has teeth.*
- The possession relation is constructed, in French, with suffixes *-al*, *-aire*, *-el*, *-ien*, *-in*, *-ique* and *-ois*. The head noun corresponds to the part of an entity and the base noun of the adjective to whole entity:
  nerf dental (*dental nerve*): which means that the *Nerve is located in the tooth.*

In other words, in a noun phrase with relation of belonging base noun belongs to head noun, and in a noun phrase with possession base noun possesses head noun.

We applied these clues to medical and cogeneration terms and noticed that these semantic relations are particularly reliable when linking simple noun terms t_1 (*abdomen*) with noun phrase terms t_2 (*abdominal abscess*). When occuring in more complex terms, this operation has to be supported by strong syntactic analysis and then by manual validation. In table 1 we present few examples with medical terms. The first column contains simple term t_1 (base noun). The second column contains complex term t_2 (with the relational adjective). In the last column we indicate type of relations: p for possession and b for belonging.

Table 1. Examples of belonging/possession relations in medical terms

term t_1 (base noun)	term t_2 (with affixed adjective)	rel.
abdomen (*abdomen*)	⇒ *abcès abdominal* (*abdominal abscess*)	p
amygdale (*tonsil*)	⇒ *noyau amygdalien* (*amygdaloid nucleus*)	p
anévrisme (*aneurysm*)	⇒ *hématome anévrismal* (*aneurysmal hematoma*)	p
artère (*artery*)	⇒ *cône artériel* (*conus arteriosus*)	p
achromie (*achromasia*)	⇒ *mélanome achromique* (*amelanotic melanoma*)	b
actinomycose (*actinomycosis*)	⇒ *infection actinomycosique* (*actinomycotic infection*)	b
athérome (atheroma)	⇒ *embolie athéromateuse* (*atheromatous embolus*)	b

Furthermore, in specialized languages, such global semantic instructions (belonging and possession) can lead to more specific meaning. For instance, in examples of possession, when (1) the base noun of the adjective means a part of body or a tissue (*abdomen*, *amygdale*, etc.) and (2) when the head noun of the adjective means an illness or injury (*abcès*, *fibrosarcome*, etc.), the resulting relation corresponds typically to localisation:

abcès abdominal (*abdominal abscess*) `located-in` *abdomen* (*abdomen*),
noyau amygdalien (*amygdaloid nucleus*) `located-in` *amygdale* (*tonsil*)

But possession couples can also induce other types of semantic relations:

cône artériel (*conus arteriosus*) `conducts-to` *artère* (*artery*)
hématome anévrismal (*aneurysmal hematoma*) `produced-by` *anévrisme* (*aneurysm*)

When terms comprise more that one relational adjective, which builds both possession and belonging relations, interpretation schemas get more complex:

- *ganglion lymphatique abdominal* (*abdominal lymph node*):
ganglion (*node*) `contains` *lymphe* (*lymph*) `located-in` *abdomen*
- *vésicule cutanée acantholytique* (*acantholytic blister*):
(*vésicule* `located-in` *peau*) (*blister*) `caused-by` *acantholyse* (*acantholysis*)

Such analysis of medical terms can be qualified through the semantic axes of their primitives, for instance those from SNOMED, as it has been done with compound terms [39]. Term *ganglion lymphatique abdominal* (*abdominal lymph node*) can thus be described as the combination of axes M (morphology or illness) and T (topology or body parts): *ganglion* (*node*) from axis M, *lymphe* (*lymph*) from axis T, and *abdomen abdomen* from axis T.

Below, we give a few similar examples from the cogeneration area (possession `p` and belonging `b` relations):

term t_1 (base noun)	term t_2 (with affixed adjective)	rel.
atmosphère (*atmosphere*)	⇒ *polluant atmosphérique* (*atmospheric contaminants*)	`p`
industrie (*industry*)	⇒ *déchet industriel* (*industrial waste*)	`p`
troposphère (*troposphere*)	⇒ *ozone troposphérique* (*tropospheric ozone*)	`p`
gaz (*gas*)	⇒ *acide gazeux* (*acid gas*)	`b`
métallurgie (*metallurgy*)	⇒ *industrie métallurgique* (*metallurgical industry*)	`b`
carbone (*carbon*)	⇒ *gaz carbonique* (*carbonic gas*)	`b`
soufre (*sulfur*)	⇒ *rejet soufré* (*sulfur rejection*)	`b`

These examples can be analyzed in the same way and lead to specific terminological relations. For instance, with possession, when the head noun means chemical substance (*polluant*, *ozone*), the resulting relation is localisation:

polluant atmosphérique (*atmospheric contaminants*) `located-in` *atmosphère*
ozone troposphérique (*tropospheric ozone*) `located-in` *troposphère*

In all studied examples, affixation can lead to the detection of different semantic relations between terms: meronymy, antonymy and many transversal relations. It is obvious that the precision and completeness of used morphological resources define the quality and completeness of generated semantic relations between terms.

6 Conclusion and Perspectives

In this work we studied the contribution of morphological clues for the deciphering of semantic relations between terms in order to build structured terminologies. We particularly addressed clues given by affixation and suppletion. Morphology then appears to be useful for the deciphering of a large variety of semantic relations (synonymy, antonymie, taxonomy or transversal relations). We assume that the morphology allows complimenting results obtained with other methods for terminology structuring. This method has been applied to medical and cogeneration terms.

However such approach requires suitable morphological resources or a morphological analyser, and it especially requires the understanding of the meanings conveyed by morphological units and rules. This approach must be supported by linguistic research in the area of morphology. Furthermore, it is obvious that the precision and completeness of used morphological knowledge define the quality and completeness of generated semantic relations between terms.

We demonstrated that morphological operations, which allow the inducing of general semantic meanings of constructed lexemes, can indicate more specialized meanings specific to given scientific and technical areas. For instance, relational adjectives formed on the base of nouns receive the general meaning *Relative to N*, but can lead to possession and belonging relations, and further more to specific relations such as `located-in`, `caused-by`, etc. These final relations are suitable for the organization of knowledge from different areas. They allow particularly to describe complex terms through their atomic primitives.

In the medical area, numerous affixations are applied to nouns referring to body parts. When these affixed formations are parts of complex terms, they allow anchoring diseases, injuries, medical procedures, etc. in a given body part (*affection musculaire* (*disorder of muscle*), *anévrisme cardiaque* (*aneurysm of heart*), *angiome capillaire* (*capillary hemangioma*)). As the topography, or body part localizations, corresponds to widely used entities in medicine, it should be studied in a more detailed way.

One of the issues that were not addressed in this work is related to the polysemy and homonymy of affixes. For instance, the French language has two suffixes *-aire*: the first is applied to noun bases and used for the formation of relational adjectives {*cellule*, *cellulaire*}, the second is applied to verb bases and used for the formation of agent nouns {*contest(er)*, *contestataire*}. It is obvious that the homonymy, as well as the polysemy, of affixes raise an ambiguity in the resulting semantic relations between terms. In order to manage this ambiguity, linguistic features of studied operations and their affixes must be taken into account at different levels (syntactic, phonological, morphological and semantic) which should bring the first disambiguation of clues for terminology structuring.

In a further work, we plan to extend this study on compounding, which is widely used in sublanguages like medicine, biology, etc. The analysis of compound lexemes leads to a large set of semantic relations between terms, and is also useful for terminology structuring. In a further work, we also plan to apply morphological clues to terms from other areas.

References

1. Bourigault, D., Slodzian, M.: Pour une terminologie textuelle. In: Terminologie et Intelligence Artificielle (TIA), Nantes (1999)
2. Grabar, N., Jeannin, B.: Contribution de différents outils la construction d'une terminologie pour la recherche d'information. In Gréboval, C., ed.: Ingénierie des connaissances (IC), Rouen (2002) Poster.
3. Grabar, N., Hamon, T.: Les relations dans les terminologies structurées : de la théorie la pratique. Revue d'Intelligence Artificielle (RIA) **18**(1) (2004)

4. Côté, R.A.: Répertoire d'anatomopathologie de la SNOMED internationale, v3.4. Université de Sherbrooke, Sherbrooke, Québec. (1996)
5. Amsili, P.: L'antonymie en terminologie : quelques remarques. In: Terminologie et Intelligence Artificielle (TIA), Strasbourg (2003) 31–40
6. Spackman, K., Campbell, K.: Compositional concept representation using SNOMED: Towards further convergence of clinical terminologies. In: Journal of American Medical Informatics Association (JAMIA). (1998) 740–744
7. Zweigenbaum, P.: Encoder l'information médicale : des terminologies aux systèmes de représentation des connaissances. Information de Santé Innovation Stratégie (ISIS) **2**(3) (1999) 27–47
8. Hearst, M.A.: Automatic acquisition of hyponyms from large text corpora. In: Proceedings of the Fourteenth International Conference on Computational Linguistics, Nantes, France (1992) Disponible `http://sern.ucalgary.ca/KSI/KAW/KAW99/papers/Gery1/index.html`. Visité le 26/08/99.
9. Séguéla, P., Aussenac-Gilles, N.: Extraction de relations sémantiques entre termes et enrichissement de modèles du domaine. In: Actes d'Ingénierie des Connaissances (IC), Palaiseau, France (1999) 79–88
10. Morin, E.: Acquisition de patrons lexico-syntaxiques caractéristiques d'une relation sémantique. Traitement Automatique des Langues (TAL) **40**(1) (1999) 143–166
11. Garcia, D.: Structuration du lexique de la causalité et réalisation d'un système d'aide au repérage de l'action dans les textes. In: Terminologie et Intelligence Artificielle (TIA), Toulouse (1997) 7–26
12. Pearson, J.: Terms in Context. Volume 1 of Studies in Corpus Linguistics. John Benjamins, Amsterdam/Philadelphia (1998)
13. Grefenstette, G.: Explorations in automatic thesaurus discovery. Kluwer Academic Publishers (1994)
14. Nazarenko, A., Zweigenbaum, P., Habert, B., Bouaud, J.: Corpus-based extension of a terminological semantic lexicon. In: Recent Advances in Computational Terminology. John Benjamins (2001) 327–351
15. Assadi, H.: Construction d'ontologies partir de textes techniques – Application aux systèmes documentaires. Thèse de doctorat en informatique, Université de Paris 6, Paris, France (1998)
16. Bourigault, D.: Analyse syntaxique locale pour le repérage de termes complexes dans un texte. Traitement Automatique des Langues (TAL) (1993) 105–117
17. Bodenreider, O., Burgun, A., Rindflesch, T.C.: Lexically-suggested hyponymic relations among medical terms and their representation in the UMLS. In URI INIST CNRS, ed.: Terminologie et Intelligence artificielle (TIA), Nancy (2001) 11–21
18. Grabar, N., Zweigenbaum, P.: Lexically-based terminology structuring: Some inherent limits. In Chien, L.F., Daille, B., Kageura, K., Nakagawa, H., eds.: Proceedings of Second International Workshop on Computational Terminology (COMPUTERM 2002), Taipei, Taiwan, ACLCLP (2002) 36–42
19. Jacquemin, C.: A symbolic and surgical acquisition of terms through variation. In Wermter, S., Riloff, E., Scheler, G., eds.: Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing, Springer (1996) 425–438
20. Hamon, T., Nazarenko, A., Gros, C.: A step towards the detection of semantic variants of terms in technical documents. In: International Conference on Computational Linguistics (COLING-ACL'98), Université de Montréal, Montréal, Quebec, Canada (1998) 498–504

21. Melčuk, I.: Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-s'emantiques. Presses Universitaires de Montréal, Montréal (1984-1999)
22. L'Homme, M.C.: Fonctions lexicales pour repr'esenter les relations s'emantiques entre termes. Traitement automatique de la langue (TAL) **43**(1) (2002) 19–41
23. Daille, B.: Conceptual structuring through term variations. In: Proceedings of the ACL Workshop on Multiword Expressions : Analysis, Acquisition and Treatment. (2003) 9–16
24. Claveau, V., L'Homme, M.C.: Discovering specific semantic relationships between nouns and verbs in a specialized French corpora. In: Computerm. (2004)
25. L'Homme, M.C.: Adjectifs dérivés sémantiques (ADS) dans la structuration des terminologies. In: Journées d'étude Terminologie, Ontologie et représentation des connaissances, Lyon (2004)
26. Corbin, D.: Morphologie dérivationnelle et structuration du lexique. Volume 1. Presse universitaire de Lille, Lille (1987)
27. Creutz, M.: Unsupervised segmentation of words using prior distributions of morph length and frequency. In: Proc of ACL-03, Sapporo, Japan (2003) 280–287
28. Burnage, G.: CELEX - A Guide for Users. Centre for Lexical Information, University of Nijmegen (1990)
29. Dal, G., Namer, F., Hathout, N.: Construire un lexique dérivationnel : théorie et réalisations. In Amsili, P., ed.: Traitement Automatique des Langues Naturelles (TALN), Cargèse (1999) 115–124
30. Hathout, N., Namer, F., Dal, G.: An experimental constructional database: the MorTAL project. In Boucher, P., ed.: Morphology book. Cascadilla Press, Cambridge, MA (2001)
31. NLM: UMLS Knowledge Sources Manual. National Library of Medicine, Bethesda, Maryland. (2005) www.nlm.nih.gov/research/umls/.
32. Namer, F.: Acquisition automatique de sens partir d'opérations morphologiques en français : étude de cas. In: Traitement Automatique de la Langue Naturelle (TALN), Nancy (2002) 235–244
33. Grabar, N., Zweigenbaum, P.: A general method for sifting linguistic knowledge from structured terminologies. JAMIASUP (2000) 310–314
34. Manuila, L., Manuila, A., Lewalle, P., Nicoulin, M.: Dictionnaire médical. Masson, Paris (2001) 9^e édition.
35. Namer, F., Zweigenbaum, P.: Acquiring meaning for French medical terminology: contribution of morphosemantics. In: Annual Symposium of the American Medical Informatics Association (AMIA), San-Francisco (2004)
36. Amiot, D.: Préfixes ou prépositions ? Les cas de *sur-*, *sans-*, *contre-* et les autres. Lexique **16** (2001)
37. Aliquot-Suengas, S.: Référence collective/Sens collectif. La notion du collectif travers les noms suffixés du lexique français. Thèse de doctorat en linguistique, Université de Lille III, Lille, France (1996)
38. Mélis-Puchulu, A.: Les adjectifs dénominaux : les adjectifs de 'relation'. Lexique **10** (1991) 33–60
39. Pacak, M.G., Norton, L.M., Dunham, G.S.: Morphosemantic analysis of -itis forms in medical language. Methods in Medical Informatics (MIM) **19**(2) (1980) 99–105