Conceptual Structuring through Term Variations

Béatrice Daille
IRIN
University of Nantes
France
daille@irin.univ-nantes.fr

Abstract

Term extraction systems are now an integral part of the compiling of specialized dictionaries and updating of term banks. In this paper, we present a term detection approach that discovers, structures, and infers conceptual relationships between terms for French. Conceptual relationships are deduced from specific types of term variations, morphological and syntagmatic, and are expressed through lexical functions. The linguistic precision of the conceptual structuring through morphological variations is of 95 %.

1 Introduction

Term extraction systems are now an integral part of the compiling of specialized dictionaries and updating of term banks. Several tools exist either for extracting or structuring terminology (see (Cabré et al., 2001) for a review of the current systems). Systems for identifying conceptual relationships are generally based on external evidence: (Cortès and Cabré, 2002) define a catalogue of linguistic markers to detect conceptual relationship such as similarity or inclusion relationships. Similarity is detected by the prototype linguistic expression: *to be similar to*. Other systems relies on internal evidence. (Morin and Jacquemin, 1999; Hamon and Nazarenko, 2001) structure complex terms (multi-word terms) with the help of lexical databases or general dictionaries. The relationships handled are limited to synonymy and hyperonymy. (Grabar and Zweigenbaum, 2000) identify morphological families of word forms applied on a medical thesaurus with a precision of 92 % such as (*arthrite* "arthritis", *arthrose* "arthrosis", *arthropathie* "arthropathy", ...). They also make deductions on their conceptual relationship with other lexical units and their contribution to the overall knowledge organization of a specialized field, namely medicine. The conceptual relationships identified are synonymy, reference and hyperonymy.

In this paper, we present a term detection approach that discovers, structures, and infers conceptual relationships between terms for French. Conceptual relationships are deduced from specific types of term variations, morphological and syntagmatic, and are expressed in terms of lexical functions (Wanner, 1996). Term variations have already proved their reliability in information retrieval: (Jacquemin, 2001) considers different types of terminological variants including syntactic variations to perform accurate text indexing. In the remaining sections, we present some conceptual systems, provide a linguistic typology of term variations and describe the two steps of our approach. In the final part, we give results and briefly discuss them.

2 Conceptual systems

Terms are generally classified using partitive and generic relationships to be presented in a thesaural structure. But other relationships exist, the so-called *complex relationships* (Sager, 1990, pages 34-35) which are domain and application dependent. Examples of such complex relationships are:

FALLOUT is *caused by* NUCLEAR EXPLOSION
COAL-MINE is *a place for* COAL-MINING

Studying term formation, (Kageura, 2002) introduces intra-term relationships dealing with complex terms and defining the role of the determinant with respect to the head noun. In *computer program*, *program* is the head noun and *computer* the determinant. As *program* is intended for *computer*, a "destination" intra-relationship occurs between *program* and *computer*. (L'Homme, 2002) used lexical functions to represent various types of relationships via an unique formalism for a computerized version of a dictionary. These relationships are of several types:

- paradigmatic such as the generic (**Gener**) or antonymy (**Anti**) relationships:
  Gener(*retail sale*) = *sale*,
  Anti(*retail sale*) = *wholesaling*;
- derivational such as nominalizations ($\mathbf{S}_0$, $\mathbf{V}_0$):
  $\mathbf{S}_0$(*to program*) = *programmer*,
  $\mathbf{V}_0$(*to program*) = *programming*;
- syntagmatic (**Bon**, $\mathbf{Real}_1$):
  Bon(*software*) = *performing*,
  $\mathbf{Real}_1$(*programme*) = *to run* [DET∼].

These conceptual relationships are assigned manually to terms or sets of terms. We propose to automatically assign conceptual relationships to complex terms through their variations.

3 Concept identification through Term variations

For complex terms identification, it is necessary to first define syntactic structures which are potentially lexicalisable. These complex sequences are so-called "base-terms".

3.1 Base terms and their linguistic Variations

For French, the syntactic structures or patterns of base-terms are:

Noun1 Adj *emballage biodégradable (biodegradable package)*

Noun1 (Prep (Det)) Noun2 *ions calcium (calcium ion) protéine de poissons (fish protein)*, *chimioprophylaxie au rifampine (rifampicin chemoprophylaxis)*

Noun1 à Vinf *viandes à griller (grill meat)*

These base structures are not frozen structures and accept variations. Terminological variation in texts is now a well-known phenomenon estimated from 15 % to 35 %, depending on the domain reflecting by the texts and the different kinds of variants handled. For acquisition, it is essential to identify extensively all the concepts represented by terms in textual data. Thus, only term variants which can preserve the base-term semantics and thus refer to the same concept are taken into account in a first step. Two sequences such as *histamine présente dans le vin (histamine which is present in wine)* et *histamine du vin (histamine of the wine)* refer to the same term *histamine du vin (wine histamine)*; but, the sequences *produit à surgeler (product to be frozen)* and *produit surgelé (frozen product)* refer to two different terms linked by an aspectual relationship.

We present now a linguistic typology of base-term variations for French:

Graphical case differences and presence of a optional hyphen inside the Noun1 Noun2 structure.

Inflexional orthographic variants gathering together inflexional variants that are predictable such as *conservations de produit (product preservations)* or unpredictable such as *conservation de produits (product preservation)*.

Shallow syntactic The shallow syntactic variations modify the function words of the base-terms. There are three kinds of internal syntactic variations:

is-1 variations of the preposition: *chromatographie en colonne (column chromatography)* ↔ *chromatographie sur colonne (chromatography on column)*;

is-2 optional character of the preposition and of the article: *fixation azote (nitrogen fixation)* ← *fixation d'azote (fixation of nitrogen)* → *fixation de l'azote (fixation of the nitrogen)*;

is-3 predicative variants: the predicative role of the adjective: *pectine méthylée (methylate pectin)* → *ces pectines sont méthylées (these pectins are methylated)*.

Syntactic The shallow syntactic variations modify the internal structure of the base-terms:

S-1 Internal modification variants: insertion inside the base-term structure of

- a modifier such as the adjective inside the Noun1 Prep Noun2 structure: *lait de brebis (goat's milk)*, *lait cru de brebis (milk straight from the goat)*;
- a nominal specifier inside the Noun Adj. These specifiers belongs to a closed list of nouns such as *type, origine, couleur (colour)*: *protéine végétale* "vegetable protein" $\rightarrow$ *protéine d'origine végétale* "protein of vegetable origin".

S-2 Coordinational variants: head or expansion coordination of base term structures and enumeration:
analyse de particules "particule analysis" $\rightarrow$ *analyse et le tri de particules* "particle sort and analysis"
alimentation humaine "human feeding" $\rightarrow$ *alimentation animale et humaine* "human and animal feeding".

Morphosyntactic The Morphosyntactic variations modify the internal structure of the base-terms and its components are liable to morphological modification (including derivation).

M-1 Morphology : the preposition inside a candidate term of Noun1 Prep Noun2 structure is equivalent to a prefix applying on Noun2: *pourrissement après récolte (rot after harvest)* $\leftrightarrow$ *pourrissement post-récolte (post-harvesting rot)*;

M-2 Derivational morphology: a derivational variation that keeps the synonymy of the base term implies a relational adjective: *acidité du sang (acidity of the blood)* $\leftrightarrow$ *acidité sanguine (blood acidity)*. This morphosyntactic variation could be associated with a syntactic variation: the sequence: *alimentation destinée à l'homme et à l'animal* "food destined to man and to animal" is a variation of the base-term: *alimentation animale* "animal food".

Two other types of variation could have been included in this typology: paradigmatic and anaphorical variations. The first one relies on the substitution principle of distributional linguistics (Harris, 1968). One or two words of the base-term could be substituted by one of their synonyms without modifying the syntactic structure (Hamon and Nazarenko, 2001). The second one gathers elliptical anaphora and acronyms.

3.2 Variations reflecting conceptual relationships

All these variations are those which could preserve synonymy with the base term. They can, of course, include semantic discrepancies and can refer either to two base terms or to a base term and a conceptually linked term. Thus, two different prepositions lead to two base terms: *transmission par satellite (satellite transmission)* $\neq$ *transmission entre satellites (transmission between satellites)* and internal modification (see variation S-1a) refers to a overcomposed term: *huile essentielle de sapin (fir essence)* is a hyponym of *huile essentielle (essence)* and not a variation of *huile de sapin (fir oil)*.

We propose to identify the conceptual relationships betwen base terms through syntactic or morphological clues. We use standard lexical functions to express the conceptual relationships. When there does not exist a lexical function to label a conceptual relationship, we introduce a new lexical function (*i.e* a non standard one). Standard lexical function are written in lower-case, non-standard in upper-case.

Syntactic The internal modification of the base structures mainly implies two types of semantic relationships:

- **Hyperonymy**: if it is a relational adjective that modifies the base term of N1 Adj or N1 Prep (Det) N2 structure, an hyperonymic relationship occurs between the base term and the modified one. The lexical function that captures hyperonymic relationships is the function **Spec** introduced by (Grimes, 1990):
Spec (*contraction isométrique* "isometric contraction") = *contraction musculaire isométrique* "isometric muscular contraction"
Spec (*agent bactérien* "bacterial agent") = *agent infectieux bactérien* "bacterial infectious agent"
- **Antonymy**: if it is an adverb of negation that

modifies the base term of N1 Adj structure, an antonymic relationship occurs between the base term and the modified one. This relationship of opposition is described with the function **Anti**:
Anti(*levure floculante* "flocculating yeast")= *levure non floculante* "non-flocculating yeast"

Morphosyntactic Semantic distinctions appear with base terms that are morphologically related to other base terms. Two base-terms w_1 w_2 and $w_1^{'}$ $w_2^{'}$ are considered as morphologically-related if one of the three following constraints are satisfied:

i. w_1 and $w_1^{'}$ are head nouns and are identical. w_2 and $w_2^{'}$ are expansions and are semantically related by the use of an affix;

ii. w_1 and $w_1^{'}$ are head nouns and are semantically related by the use of an affix. w_2 and $w_2^{'}$ are expansions and are identical;

iii. w_1 and $w_1^{'}$ are head nouns, w_2 and $w_2^{'}$ are expansions, either w_1 and $w_2^{'}$ are identical and w_2 and $w_1^{'}$ are semantically related by the use of a suffix such as *preserved food/food preservation*;

Some affixes that have been studied for French by (Corbin, 1987) provide clues to characterize the semantic link occurring between two morphologically-related candidate terms.

- **Antonymy**: the prefixes *ir, dé, non(-)* applying either on the head or expansion element on a base term whatever is its structure characterize an antonymic relationship. Examples are:
  Anti (*solubilisation micellaire* "micellar solubilization") = *insolubilisation micellaire* "micellar insolubilisation"
  Anti (*phénol polymérisé* "polymerized phenol") = *phénol non-polymérisé* "unpolymerized phenol"

- **Set of**: the suffixes *age, ade* applying on the head noun of base term attest of a "set of" relationship expressed with the function **Mult**:
  Mult (*plume de canard* "duck feather") = *plumage de canards* "duck feather" The two base-terms share the same pattern.

- **Result**: A "result" relationship is expressed with the function $\mathbf{N}_{Res}$ applying on nouns. This relation is induced either by:
  - the suffixes *age, ade, erie* applying on the head noun of base terms:
    $\mathbf{N}_{Res}$(*plumage de canards* "duck feather") = *plume de canard* "duck feather"
    $\mathbf{N}_{Res}$(*filetage du saumon* "salmon filleting") = *filet de saumon* "salmon fillet"; The two base-terms share the same pattern.
  - or by the suffixes *age, ade, erie, ment, tion, ure* associated with an inversion. We distinguish two cases:
    - ∗ if this morphological link involves a N Adj structure, the function $\mathbf{N}_{Res}$ applies:
      $\mathbf{N}_{Res}$(*conservation des aliments* "food preservation") = *aliment conservé* "preserved food";
    - ∗ if the morphological link involves a N *à* Vinf structure, we face a non-standard function where the term of N *à* Vinf structure expresses the state before the process. Thus, we introduce the new function N_{Init}:
      $\mathbf{N}_{Init}$(*conservation des aliments* "food preservation") = *aliment à conserver* "food to preserve";

- **Actor**: the suffixe *eur* applying on the head noun of a base term builds its actant expressed with the function $\mathbf{S}_1$: $\mathbf{S}_1$(*transport routier* "road transport" = *transporteur routier* "road haulier". The two base-terms share the same pattern.

Other semantic relationships involving two base-terms with the same pattern are induced by prefixes. For those, we have to introduce new functions as:

- "again" relationship with the prefixes *re, ré*: **AGAIN**(*estérification enzymatique* "enzymatic esterification") = *réestérification enzymatique* "enzymatic reesterification";

- "before" relationship with the prefixe *pré(-)*: **BEFORE**(*traitement enzymatique* "enzymatic treatment") = *prétraitement enzymatique* "enzymatic pretreatment".

4 Automatic discovery and structuring

4.1 Linguistic structuring

The term extractor program takes as input a tagged and lemmatized corpus. The programme implements shallow parsing and morphological conflating. First, it scans the corpus, counts and extracts strings whose syntax characterizes base-terms or one of their variants. This collecting step uses local grammars based on regular expressions (Abney, 1997). These grammars use the morphosyntactic information associated with the words of the corpus by the tagger. The different occurrences referring to a base term or one of its variants are grouped as a pair formed by lemmas of the candidate base term. Second, morphological analysis is performed to confluate synomymic derivational variants of base terms such as *acidité du sang (acidity of the blood)* $\leftrightarrow$ *acidité sanguine (blood acidity)*. Stripping-recoding morphological rules adopt the following rule schemata:

$$A \rightarrow N \quad : \quad -S + M$$

where:

S is the relational suffix to be deleted from the end of an adjective. The result of this deletion is the stem **R**;

M is the mutative segment to be concatenated to **R** in order to form a noun.

For example, the rule [*-é +e*] says that if there is an adjective which ends with *é*, we should strip this ending from it and append the string *e* to the stem. The algorithm below resumes the successive steps for identifying relational adjectives:

1. Examine each candidate of Noun Adj structure;

2. Apply a transformational rule in order to generate all the possible corresponding base nouns.

3. Search the set of candidate terms for a pair formed with Noun1 (identical between a Noun1 (Prep (Det)) Noun2 and a Noun1 Adj structures) and Noun2 generated from step 2.

4. If step 3 succeeds, group the two base structures under an unique candidate term.

In Step 2, morphological rules generate one or several nouns for a given adjective. We generate a noun for each relational suffix class. A class of suffixes includes the allomorphic variants. This overgeneration method used in information retrieval by (Jacquemin, 2001) gives low noise because the base noun must not only be an attested form in the corpus, but must also appear as an extension of a head noun.

At the end of the linguistic processing, the term extractor proposes as output:

1. a list of pilot terms ranked from the most representative of the corpus to the least thanks to the Loglikelihood coefficient introduced by (Dunning, 1993).

2. for each pilot term, a XML structure is provided which gathers all the base structures and the variations encountered.

An example of such data is given in figure in Table 1.

4.2 Conceptual structuring

The conceptual structuring takes as input the data provided by the first step. First, we present the methodology employed to exploit variables of base terms. We then demonstrate the labelling of conceptual links through morphological analysis.

4.2.1 Treatment of modification variants

In the previous step, a first list of relational adjectives has been established thanks to their paraphrasic property. (Daille, 2001) demonstrated that candidate terms of N Adj structure where Adj is relational hold a more important naming potential than for the synonym form in N1 Prep N2. The absence of paraphrases, the non-paraphrasability, or a complex paraphrasability or a large derivational distance between the adjective and the noun do not allow exhaustive identification. We extend this list by exploiting the coordination variations of N Adj base terms. Indeed, a relational adjective holds the property to coordinate only with other relational adjectives. To summarise:

1. From *industrie de l'alimentation* "food industry" and *industrie alimentaire* "food industry",

Sorted list of candidate terms		
Score	**Pilot term**	**Index**
785	*acide gras*	926
722	*mise au point*	344
629	*acide aminer*	394
559	*matière grasse*	2002
512	*résultat obtenir*	155
472	***chromatographie gazeuse***	**1374**
469	*bactérie lactique*	118
. . .	. . .	. . .
XML structure associated to **chromatographie gazeuse**		
<CAND **ident=1374** freq=103>		
<NPN freq=9>	<BASE>	<TERM>*chromatographie du gaz*</TERM>
		<TERM>*chromatographie gaz*</TERM>
		<TERM>*chromatographie de gaz*</TERM>
	</BASE>	
</NPN>		
<NPNA freq=80 >	<BASE>	<TERM>*chromatographie en phase gazeuze*</TERM>
	</BASE>	
	<MODIF>	<TERM>*chromatographie capillaire en phase gazeuze*</TERM>
	</MODIF>	
</NPNA>		
<NA freq=14>	<BASE>	<TERM>*chromatographie gazeuze*</TERM>
	</BASE>	
</NA>		
</CAND>		

Table 1: Output of the first step

we deduce that *alimentaire* is a relational adjective;

2. From the coordinational variant *produit agricole et alimentaire* "farm and food product", we deduce that *agricole* is a relational adjective.

This classic learning algorithm that is normally bound by the number of adjectives in the corpus converges in five steps. It allows to extend the set of relational adjectives from 143 to 239. The following are some examples of acquired relational adjectives:

Relational adjective	Number of iterations	Coordinated relational adjectives
gazeux	1	(*microbien solide liquide organoleptique*)
ferme	2	(*élastique*)
productif	1	( )
global	2	(*micro-économique spécifique local*)
peroxydasique	1	(*polyphénoloxydasique lipoxygénasique catalasique*)
hydrodynamique	3	(*thermique*)

Using this extended list of relational adjectives, we automatically check all the modification variants of collected base-terms:

- if a relational adjective is present, we infer an hyperonymy link between the variant and the base term as for *contraction isométrique* "isometric contraction" and *contraction musculaire isométrique* "isometric muscular contraction", but not for *organisation ordonnée des molécules* "ordered molecule organization" that remains a syntactic variation of *organisation moléculaire* "molecule organization";
- if an adverb of negation is present, we infer an antonymy link between the variant and the base term as for *brunissement non enzymatique* "non enzymatic browning" and *brunissement enzymatique* "enzymatic browning".

4.2.2 Morphological conflating

To identify the conceptual relationships denoted by derivational links, we perform a morphological analysis using the same method as in section 4.1: we wrote stripping-recoding morphological rules for each conceptual relationship, we apply the overgeneration method and the filtering based on the presence or not of the generated base term candidates. In order to browse the list of candidate terms, we apply to each candidate terms successively all the possible derivations.

The output of the conceptual structuring program is a list of candidate terms ranked, each of them representing a set of conceptually linked candidate terms. An example of such structure is given in Table 2.

5 Results and Evaluation

We apply our program on a technical corpus in the field of agriculture which consists of 2,702 scientific abstracts for a total of 427,482 tokens and an average size of a record of 316 tokens.
Table 3 gives the results of the collecting phase and Table 4 shows the percentages of the different types of variations for candidate terms appearing at least two times and the number of synonymic conflations. This conflating has a linguistic precision of 99 %.

Number of occurrences base structures	1	≥ 2	Total
Nom1 Prep (Det) Nom2	17 232	5 949	23 181
Nom Adj	12 344	4 778	17 122
Nom à Vinf	203	16	219
Total	29 912	10 895	40 807

Table 3: Number of candidate base terms

Syntactic variation		
Coor + Modif	Coor	Modif
61 (0,5 %)	458 (4 %)	1651 (15,1 %)
19,2 %		
Morphological variation N1 (Prep (Det)) N2 / N1 AdjR with AdjR derived from N2 343		

Table 4: Number of base-term variations

Conceptual Link	Syntactic	Morphological
Spec	731	
Anti	183	106
N_{Res}		132
MICRO		59
AGAIN		36
BEFORE		29
Mult		23
INTER		20
...		...
Total	914	558

Table 5: Number of major conceptual links

Table 5 gives the number of the most frequent automatically identified conceptual relationships.
Concerning morphological links, we note that two non-standard functions that have not been presented yet obtain a consequent representativity: **MICRO** induces from the suffixe *micro(-)*: **MICRO**(*film perforé* "perforated film") = *film micro-perforé* "pinhole film"; **INTER** infers from the suffixe *inter* expressing a reprocivity relationship: **INTER**(*échelle nationale* "national scale") = *échelle internationale* "international scale". The average precision of the morphological links is 95 %. The wrong links are 75 % due to the prefixes *re, ré* refering to the function **AGAIN**. Examples of false drop are: *action/réaction* "reaction" in several candidate base-terms: *action/réaction enzimatique* "enzimatic action/reaction", *action/réaction acide* "acide action/reaction", ..., *production/reproduction*, *solution/résolution* "resolution", etc.

6 Conclusion

Links between complex terms can be used to assist terminolographers to handle extracted terminological data more conveniently, since several related concepts are clustered. This conceptual structuring relies on term variation and we have stressed the crucial part of this handling. The method can be easily adapted to other romance languages and English for which it suffices to define patterns for base-terms and their variations and to list appropriate affixes reflecting semantic relationships.

References

Steven Abney. 1997. Part-of-Speech Tagging and Partial Parsing. In Steve Young and Gerrit Bloothooft, editors, *Corpus-Based Methods in Language and Speech Processing*, volume 2, chapter 4. Kluwer Academic Publishers.

M. Teresa Cabré, Rosa Estopà Bagot, and Jordi Vivaldi Platresi. 2001. Automatic term detection: A review of current systems. In Didier Bourigault, Christian Jacquemin, and Marie-Claude L'Homme, editors, *Recent Advances in Computational Terminology*, volume 2 of *Natural Language Processing*, pages 53–88. John Benjamins.

Danielle Corbin. 1987. *Morphologie dérivationnelle et structuration du lexique*. Max Niemeyer Verlag Tübingen.

XML structure		
<SETCAND **ident=1613** >		
<LINK type=spec ident1=1643 ident2=15789 >		
<LINK type=anti ident1=1643 ident2=25128 >		
<LINK type=N$_{Res}$ ident1=1643 ident2=16667 >		
<CAND ident=1643 freq=2 >		
<NPN freq=2 >	<BASE>	<TERM> *phosphorylation de la caséine* </TERM>
	</BASE>	
</NPN>		
</CAND>		
<CAND ident=15789 freq=1 >		
<NPN freq=1 >	<MODIF>	<TERM>*phosphorylation chimique de la caséine*</TERM>
	</MODIF>	
</NPN>		
</CAND>		
< CAND ident=1667 freq=1 >		
<NA freq=3 >	<BASE>	<TERM>*caséine phosphorylée*</TERM>
	</BASE>	
</NA>		
</CAND>		
< CAND ident=25128 freq=2 >		
<NPN freq=1 >	<BASE>	<TERM>*déphosphorylation de la caséine*</TERM>
	</BASE>	
< /NPN >		
</CAND>		
</SETCAND>		

Table 2: Example of conceptual structuring

Judit Feliu Cortès and M. Teresa Cabré. 2002. Conceptual relations in specializes texts: new typology and an extraction system proposal. In *Proceeding of the 6th International Conference of Terminology and Knowlegde Engineering (TKE'02)*.

Béatrice Daille. 2001. Qualitative terminology extraction. In Didier Bourigault, Christian Jacquemin, and Marie-Claude L'Homme, editors, *Recent Advances in Computational Terminology*, volume 2 of *Natural Language Processing*, pages 149–166. John Benjamins.

Ted Dunning. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74.

Natalia Grabar and Pierre Zweigenbaum. 2000. Automatic acquisition of domain-specific morphological resources from thesauri. In *In Proceedings of RIAO 2000: Content-Based Multimedia Information Access*, pages 765–784, Paris, France.

J. Grimes. 1990. Inverse lexical functions. In J. Steele, editor, *Meaning-Text Theory: Linguistics, Lexicography and Implications*, pages 350–364. Ottawa University Press, Ottawa.

Thierry Hamon and Adeline Nazarenko. 2001. Detection of synonymy link between terms: Experiment and results. In Didier Bourigault, Christian Jacquemin, and Marie-Claude L'Homme, editors, *Recent Advances in Computational Terminology*, volume 2 of *Natural Language Processing*, pages 185–208. John Benjamins.

Zelig S. Harris. 1968. *Mathematical Structures of Language*. Wiley, New York.

C. Jacquemin. 2001. *Spotting and Discovering Terms through Natural Language Processing*. Cambridge: MIT Press.

Kyo Kageura. 2002. *The Dynamics of Terminology: A Theoretico-Descriptive Study of Term Formation and Terminological Growth*, volume 5 of *Terminology and Lexicography Research and Practice*. John Benjamins.

M.-C. L'Homme. 2002. Fonctions lexicales pour représenter les relations smantiques entre termes. *Traitement automatique des langues (TAL)*, 43(1):19–42.

Emmanuel Morin and Christian Jacquemin. 1999. Projecting corpus-based semantic links on a thesaurus. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99)*, pages 389–396.

J.C. Sager. 1990. *A Practical Course in Terminology Processing*. John Benjamins.

Leo Wanner, editor. 1996. *Lexical Functions in Lexicography and Natural Language Processing*. John Benjamins, Amsterdam/Philadelphia.