

Extension et couplage de ressources syntaxiques et sémantiques sur les adverbes du français

Benoît Sagot, Karën Fort, Fabienne Venant
INRIA Paris-Rocquencourt / Paris 7, INIST, Loria / Université Nancy2

Introduction

La question de la disponibilité des ressources pour le Traitement Automatique des Langues (TAL) reste, encore aujourd'hui, une question cruciale, notamment pour le français. Si les disparités avec l'anglais s'estompent un peu en ce qui concerne les ressources morphologiques et syntaxiques (Sagot *et al.*, 2006), l'accès à des ressources sémantiques de qualité reste difficile. Des efforts sont néanmoins faits pour mettre à la disposition de la communauté un WordNet du français, librement disponible, le WOLF (cf. section 1.2). Enfin, le couplage entre ressources syntaxiques et sémantiques reste à faire.

Dans cet article, nous présentons un premier pas dans cette direction. En restreignant notre champ d'investigation aux adverbes, nous avons cherché à compléter le WOLF puis à en coupler les entrées à celles du lexique syntaxique *Lefff* (cf. section 1.1). Ces dernières ont été construites en grande partie à partir des tables du lexique-grammaire des adverbes en *-ment* de (Molinier & Levrier, 2000). Pour atteindre ces objectifs, nous nous sommes également appuyés sur la base de synonymes DicoSyn (Ploux & Victorri, 1998).

Cet article est organisé comme suit. La section 1 décrit les trois ressources sur lesquelles repose ce travail. La section 2 décrit la façon dont nous avons étendu le WOLF à l'aide de deux techniques complémentaires, et la section 3 décrit les résultats de l'évaluation manuelle exhaustive des entrées obtenues. Enfin, la section 4 décrit la façon dont nous avons associé à certaines entrées adverbiales du *Lefff* un ou plusieurs identifiants de synsets du WOLF, en exploitant les classes d'adverbes de (Molinier & Levrier, 2000).

1. Ressources

1.1. Le *Lefff* et les tables du Lexique-Grammaire

Le *Lefff* (Lexique des Formes Fléchies du Français) (Sagot *et al.*, 2006), est un lexique morphologique et syntaxique du français à large couverture, librement

disponible¹. Le *Lefff* a pour objectif de concilier la pertinence linguistique et l'utilisabilité dans des applications de TAL. Il est notamment utilisé dans divers analyseurs syntaxiques reposant sur différents formalismes linguistiques (LFG, TAG). Le *Lefff*, aujourd'hui en version 3, couvre toutes les catégories et est progressivement enrichi d'informations syntaxiques et sémantiques, notamment par comparaison avec d'autres ressources syntaxiques (Sagot & Danlos, 2007; Danlos & Sagot, 2007). Ainsi, les entrées adverbiales du *Lefff* ont pu être complétées (Sagot & Fort, 2007) grâce aux tables du Lexique-Grammaire des adverbes en *-ment*, dites tables de Molinier (Molinier & Levrier, 2000). En particulier, chaque entrée du *Lefff* correspondant à un adverbe en *-ment* est associée à une classe syntaxico-sémantique de Molinier. Naturellement, un même lemme adverbial peut correspondre à plusieurs entrées, chacune appartenant alors nécessairement à une classe différente.

Les adverbes en *-ment* forment une classe numériquement importante d'adverbes, qui, contrairement aux autres adverbes, est une classe ouverte. Ces adverbes forment une classe morphologiquement homogène, puisqu'ils sont majoritairement construits sur un schéma du type adjectif+*ment*. Toutefois, nombre d'autres adverbes existent, et notamment une quantité importante de locutions adverbiales. Un grand nombre d'entre elles ont été décrites dans les tables d'adverbes figés de Maurice Gross. Ces tables comportent également de nombreuses collocations à fonction adverbiale mais dont le statut d'adverbe composé fait débat. Nous n'avons donc pas cherché à en tirer parti pour l'instant.

En revanche, le *Lefff* ne dispose pas encore d'informations sémantiques, pas même d'identifiants sémantiques à associer à chaque entrée. Un des objectifs de ce travail est de remédier à ce manque, en ce qui concerne les entrées adverbiales, en établissant des liens entre les entrées adverbiales du *Lefff* et celles du lexique sémantique WOLF.

1.2. WOLF

Le WOLF (WOrdnet Libre du Français) est une ressource lexicale sémantique pour le français, librement disponible² (Sagot & Fišer, 2008). Il s'agit d'un *wordnet*, à l'image du Princeton WordNet (PWN), le premier à avoir été développé, et qui traite de l'anglais (Fellbaum, 1998). Comme tout *wordnet*, le WOLF est une base de données lexicales dans laquelle les mots (lexèmes) sont répartis en catégories et organisés en une hiérarchie de nœuds. Chaque nœud a un identifiant unique, et représente un *concept*, ou *synset* (ensemble de synonymes). Il regroupe un certain nombre de lexèmes synonymes dénotant ce concept. Ainsi, dans la version 2.0 du PWN, le *synset* ENG20-02853224-n comprend les lexèmes {*car, auto, automobile,*

1. <http://gforge.inria.fr/projects/alexina/>

2. <http://wolf.gforge.inria.fr/>

machine, motorcar}. Les lexèmes peuvent être simples ou composés, et les usages métaphoriques et idiomatiques sont pris en compte. Les synsets sont précisés par une brève définition et sont reliés à d'autres synsets. Ainsi, le synset ci-dessus est relié au synset {*motor vehicle, automotive vehicle*} par un lien d'hyperonymie, et au synset {*cab, hack, taxi, taxicab*} par un lien d'hyponymie.

Le WOLF a été construit à partir de la version 2.0 du PWN et de diverses ressources multilingues, au moyen de deux approches complémentaires. Les lexèmes polysémiques ont été traités au moyen d'une approche reposant sur l'alignement en mots d'un corpus parallèle en cinq langues, dont le français. Différents lexiques multilingues ont été extraits de ces corpus alignés, en prenant en compte trois à cinq des langues disponibles (précision et rappel de ces lexiques variant en fonction du nombre de langues conservées). Les lexiques multilingues extraits ont été désambiguïsés sémantiquement à l'aide de wordnets des langues concernées. Par ailleurs, les lexèmes monosémiques du PWN n'ont nécessité que des ressources bilingues anglais-français pour construire de nouvelles entrées. Pour cela, des lexiques bilingues ont été extraits à partir de ressources wiki (Wikipedia, Wiktionary) et de thésaurus. Les sous-wordnets nominaux et verbaux du WOLF ont été évalués par rapport au wordnet français issu du projet EuroWordNet³.

Le WOLF contient tous les synsets du PWN 2.0, y compris ceux pour lesquels aucun lexème français n'est connu. En ce qui concerne les adverbes, la dernière version du WOLF avant ce travail, la version 0.1.4, ne dispose de lexèmes français que pour 676 des 3 664 synsets adverbiaux. Ceci ne représente que 18,4% des synsets adverbiaux, et seulement 983 couples lexème–synset ne mettant en jeu que 665 lemmes adverbiaux distincts. C'est pourquoi, avant de coupler les entrées adverbiales du *Lefff* et du WOLF, nous avons mis en œuvre deux techniques complémentaires pour en augmenter la couverture. L'une de ces techniques repose sur la relation de dérivation morphologique et sémantique qui existe souvent entre un synset adverbial et un synset adjectival, tant en anglais qu'en français. L'autre de ces techniques repose sur l'exploitation de la base de synonymes DicoSyn.

1.3. DicoSyn et les cliques de synonymes

DicoSyn est un dictionnaire électronique de synonymes, dont les versions les plus récentes sont consultables en ligne⁴. La base initiale (Ploux & Victorri, 1998) est issue de la fusion de sept dictionnaires classiques (Bailly, Benac,

3. Le wordnet développé dans le cadre du projet EuroWordNet (Vossen, P., 1999) est le seul autre wordnet du français. Il ne comporte que des synsets nominaux et verbaux, mais aucun synset adjectival ni adverbial. De plus, d'importants problèmes de licence en font une ressource très peu utilisée par la communauté. Enfin, et en partie pour cette même raison, elle n'a pas évolué depuis sa création. Ce sont là les trois principales motivations du projet de développement du WOLF.

4. <http://dico.isc.cnrs.fr/> et <http://elsap1.unicaen.fr/dicosyn.html>

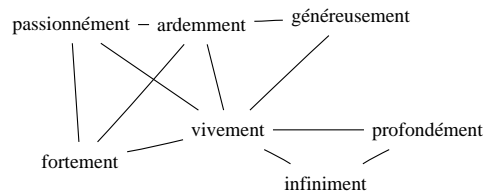


Figure 1. *Un extrait du graphe adverbial de synonymie*

Du Chazaud, Guizot, Lafaye, Larousse et Robert) dont ont été extraites les relations synonymiques. L'intérêt principal de ce dictionnaire est de donner de manière explicite le graphe de la relation de synonymie⁵. Ploux et Victorri ont mis au point un logiciel, Visusyn, permettant d'explorer ce graphe. On peut ainsi visualiser et caractériser automatiquement les propriétés sémantiques d'une unité grâce au sous-graphe qu'elle constitue avec ses synonymes (François *et al.*, 2002; Venant, 2004), ou étudier plus globalement les caractéristiques sémantiques de tout un paradigme lexical (Venant, 2007). Nous avons ainsi pu exploiter un graphe de synonymes adverbiaux. DicoSyn ne comportant aucune indication de catégorie, ce graphe adverbial a été construit en croisant DicoSyn avec les entrées adverbiales en *-ment* du *Lefff*. Ce graphe comporte 1 597 sommets (les adverbes) et 4 344 liens (de synonymie). Parmi les sommets, certains ne sont pas des adverbes en *-ment*, mais des synonymes de ces adverbes (par exemple *bien* est un sommet du graphe car DicoSyn le signale comme étant synonyme de *amplement* ou *copieusement*). La figure 1 présente un extrait de ce graphe.

L'exploitation de ce graphe repose sur la notion de clique. Une clique est un ensemble, le plus grand possible, de sommets deux à deux connectés. Ainsi le graphe de la figure 1 comporte 3 cliques : la clique {*ardemment*, *fortement*, *passionnément*, *vivement*} (on ne peut pas ajouter *généreusement* qui n'est pas synonyme de *fortement*, ni de *passionnément*), la clique {*ardemment*, *généreusement*, *vivement*} et la clique {*infiniment*, *profondément*, *vivement*}. Le graphe adverbial que nous avons construit comporte 2 247 cliques. L'idée est qu'en première approximation, une clique correspond à un emploi adverbial possible. Ce sont en effet des ensembles de synonymes, c'est-à-dire approximativement des synsets au sens d'un wordnet comme le WOLF. Ce sont donc les cliques qui constituent l'unité structurelle d'analyse sémantique du graphe.

5. Il s'agit bien sûr d'une relation de synonymie partielle, que Ploux et Victorri définissent de la façon suivante : « Deux unités lexicales sont en relation de synonymie si toute occurrence de l'une peut être remplacée par une occurrence de l'autre dans un certain nombre d'environnements sans modifier notablement le sens de l'énoncé dans lequel elle se trouve. » Il s'agit d'une relation symétrique.

2. Extension du WOLF

Comme indiqué précédemment, la première étape de notre travail a été d'étendre le WOLF, afin d'augmenter le nombre de synsets adverbiaux non vides (pour lesquels au moins un lexème français est connu), et d'augmenter le nombre de lexèmes dans chaque synset non vide. Pour cela, nous nous sommes appuyés sur deux types de relations entre lexèmes : la relation de dérivation entre un adverbe en *-ment* et l'adjectif sur lequel il est construit, et la relation de synonymie entre adverbes telle que définie par les cliques extraites de DicoSyn.

2.1. Extension par la relation de dérivation

La technique qui repose sur la relation de dérivation part du double constat suivant :

- Le Princeton WordNet inclut une relation de dérivation (*derived*) qui relie certains synsets adverbiaux à un ou plusieurs synsets adjectivaux. Ce lien indique que certains lexèmes adjectivaux présents dans le synset adjectival permettent de construire, par dérivation morphologique (suffixe *-ly*), certains lexèmes adverbiaux présents dans le synset adverbial. Naturellement, ce lien indique également une parenté sémantique entre les deux synsets.
- Le mécanisme de dérivation morphologique et sémantique entre adjectifs et adverbes est souvent parallèle entre l'anglais (adjectif + *ly*) et le français (adjectif_{fem,sing} + *ment*⁶).

Nous avons donc récupéré, pour chaque synset adverbial, les adjectifs (français) associés par le WOLF au synset adjectival qui lui est relié par la relation *derived*. Nous avons appliqué l'algorithme de dérivation morphologique à ces adjectifs⁷. Ceux des adverbes ainsi construits qui sont présents dans le *Lefff* ont été conservés, et attribués au synset adverbial de départ (en notant que ces liens lexème–synset ont été construits par dérivation morphologique).

Par exemple, soit le synset ENG20-00115661-b. Dans le WOLF 0.1.4, seuls les lexèmes *toujours* et *invariablement* y sont présents (et corrects). Or, ce synset est lié par la relation *derived* au synset adjectival ENG20-02417249-a, qui, dans le WOLF, comporte les lexèmes *permanent*, *invariable* et *perpétuel*. Les adverbiaux hypothétiques *permanemment*, *invariablement* et *perpétuellement* sont donc construits. Le premier est éliminé car il n'est pas présent dans le *Lefff*, le second confirme un lexème déjà présent dans le synset adverbial, et le dernier permet la création d'un nouveau lien lexème–synset. Au final, le synset ENG20-00115661-b devient donc {*toujours*, *invariablement*, *perpétuellement*}.

6. Ceci n'est naturellement pas toujours exact (cf. courante/couramment et bien d'autres), mais constitue une heuristique raisonnable.

7. Le féminin singulier de l'adjectif a été récupéré dans le *Lefff*.

Grâce à l'application de cette technique, le nombre de relations lexème–synset adverbial dans le WOLF est passé de 983 à 1 536 (56% d'augmentation). Le nombre de synsets adverbiaux non vides est quant à lui passé de 676 à 969 (43% d'augmentation). Le nombre de lexèmes adverbiaux présents dans le WOLF passe de 665 à 889 (23% d'augmentation).

2.2. Extension par la relation de synonymie

Une fois les synsets adverbiaux du WOLF complétés par l'exploitation de la relation de dérivation entre adverbes en *-ment* et adjectifs, nous avons mis en œuvre une technique qui repose sur la relation de synonymie telle que définie par les cliques extraites de DicoSyn. Pour ce faire, nous avons procédé en trois temps.

1. Nous avons tout d'abord attribué heuristiquement un poids à chaque lien lexème–synset en fonction de leur origine (cf. section 1.2), de la façon suivante. Si un lien a été construit (entre autres) à partir de ressources bilingues (extraites de ressources wiki), il reçoit un poids de 5. Si un lien a été extrait à partir de corpus multilingues alignés, il reçoit un poids de 4 si l'un des corpus comportait au moins 4 langues, et de 3 si tous ne comportaient que 3 langues. Dans les autres cas, y compris pour les liens construits à l'aide de la relation de dérivation, un poids de 2 est attribué au lien.
2. Pour chaque synset adverbial, on lui associe la clique DicoSyn qui maximise la somme des poids des lexèmes communs entre la clique et le synset.
3. On complète chaque synset par les lexèmes (adverbes) présents dans la clique qui lui est associée.

Par exemple, soit le synset ENG20-00115661-b, le même que précédemment. Après extension par la relation de dérivation, il contenait les adverbes *toujours*, *invariablement* et *perpétuellement*. Les deux premiers ayant été construits grâce au Wiktionary français, ils reçoivent un poids de 5. L'adverbe *perpétuellement*, construit par dérivation, reçoit un poids de 2. La clique qui maximise la somme des poids des lexèmes communs est la clique {*éternellement*, *invariablement*, *perpétuellement*, *sans cesse*, *toujours*}. Deux adverbes sont donc ajoutés au synset ENG20-00115661-b, la locution adverbiale *sans cesse* et l'adverbe en *-ment* *perpétuellement*.

En procédant ainsi, nous passons de 1 536 à 2 149 relations lexème–synset adverbial, soit une augmentation de 28,5%.

3. Evaluation du WOLF adverbial étendu

3.1. Méthodologie

Nous avons procédé à une évaluation manuelle de l'ensemble des synsets adverbiaux obtenus, c'est-à-dire de 2 149 couples lexème–synset mettant en jeu

1 025 lexèmes adverbiaux. Chacun des auteurs a validé manuellement les couples concernant un quart des lexèmes ; le dernier quart a été évalué par les trois auteurs, ce qui permet d'évaluer l'accord inter-validateurs.

Valider un couple lexème–synset consiste à lui attribuer un des codes suivants :

- OK : association correcte ;
- SC (Semantically close) : un des sens du lexème est sémantiquement proche du synset (hyponyme, hyperonyme, pseudo-synonyme) ;
- SR (Semantically related) : un des sens du lexème entretient un lien sémantique moins proche avec le sens correspondant au synset ;
- NR (Non Related) : aucun sens du lexème n'a le moindre lien sémantique avec le sens correspondant au synset ;
- CC (Composed Component) : association fautive, mais le lexème est un des composants d'un lexème composé qui, lui, serait à sa place dans le synset ;
- ID (Incorrect Derivation) : association fautive, du fait d'un problème de dérivation telle que l'ambiguïté de l'adjectif intermédiaire ou l'absence de parallèle entre dérivation morphologique et dérivation sémantique (par exemple *absolument* dans le synset défini par *in a royal manner*) ;
- SE (Spelling Error) : lexème mal orthographié, l'association est à rejeter ;
- WC (Wrong Category) : association fautive, du fait d'une mauvaise catégorisation grammaticale du lexème (par exemple *bougonnerie*).

3.2. Accord inter-validateurs

Pour un quart des lexèmes, les trois auteurs ont procédé à l'évaluation de façon indépendante. Si l'on met ensemble tous les codes autres que OK, les trois validateurs sont en accord sur 366 des 530 couples lexème–synset concernés, soit 69%. Parmi ces couples, 292 (55%) ont été validés (code OK) par les trois validateurs. Si l'on regarde la répartition des codes pour chaque validateur, on constate des différences entre degrés de tolérance (cf. figure 2). La frontière entre des codes comme SC, SR et NR étant difficile à définir de façon parfaitement objective, les divergences obtenues à leur propos ne sont pas surprenantes. En revanche, parmi les 456 couples ayant reçu le code OK de la part d'au moins un des validateurs, seuls 292 ont été codés trois fois OK (64%), et 94 deux fois (20,6%). Le désaccord est donc assez élevé, ce qui s'explique principalement par la difficulté de la tâche (certains synsets n'étant pas aisément distinguables) et par la rareté de certains adverbes.

Après analyse de ces résultats, nous avons décidé d'attribuer un code unique aux couples lexème–synset évalués par les trois validateurs de la façon suivante :

- code OK si les trois évaluations sont de la forme OK-OK-OK, OK-OK-SC, OK-OK-SR, ou OK-SC-SC ;
- code SC si elles sont de la forme OK-SC-SR ou SC-SC-SR ;

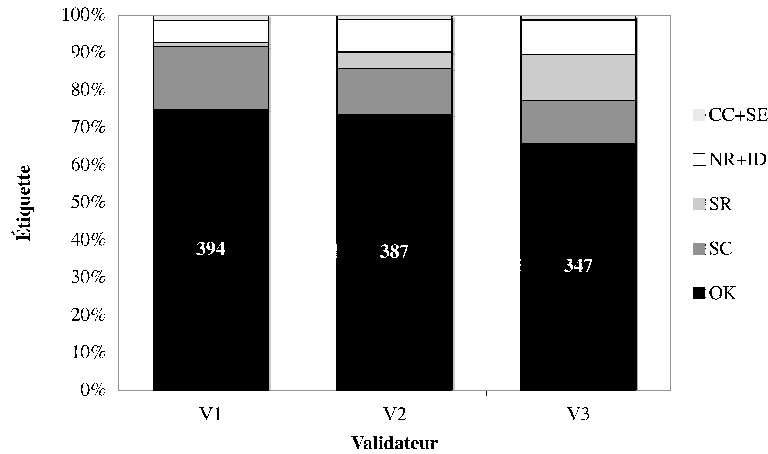


Figure 2. Comparaison entre les trois validateurs des répartitions entre codes d'évaluation pour le même ensemble de 530 couples lexème–synset

- code SR dans les autres cas où il y a un ou deux codes OK parmi les trois, ainsi que dans les cas SC-SR-SR et SR-SR-SR ;
- code SE (ID, CC, WC) dans les autres cas, dès lors qu'un validateur au moins a donné le code SE (ID, CC, WC) ;
- code NR dans les autres cas.

Naturellement, les couples lexème–synset évalués par un seul validateur conservent le code que ce dernier leur a attribué.

3.3. Résultats de l'évaluation et ressource obtenue

Les résultats sont encourageants (voir table 1). On obtient en effet plus de 68% d'associations lexème–synset correctes (elles ont reçu le code unique OK selon l'heuristique expliquée ci-dessus). Sur les 2 149 relations lexème–synset construites automatiquement, nous en avons donc conservé 1 461 (contre 983 avant ce travail, non validées manuellement). Le WOLF contient désormais 871 lexèmes adverbiaux (contre 665 au départ) répartis dans 871 synsets non vides (contre 676 initialement). L'amélioration du WOLF adverbial n'est donc pas que qualitative, grâce à la validation manuelle, elle est également quantitative.

Total	OK	SC	SR	NR	ID	CC	WC	SE
2 145	1 461	296	147	162	41	26	13	3
100%	68,1%	13,8%	6,9%	7,6%	1,9%	1,2%	0,6%	0,1%

Tableau 1. Résultats de la validation manuelle

4. Couplage entre entrées du *Lefff* et entrées du WOLF

Le couplage des ressources lexicales syntaxiques et sémantiques est indispensable pour de nombreuses applications de TAL, en particulier pour la construction d'analyseurs syntaxico-sémantiques performants. Les entrées du *Lefff* pour les adverbes en *-ment* et les synsets adverbiaux du WOLF, après les extensions décrites à la section précédente, constituent de bons candidats à un tel couplage. En effet, si chaque lemme adverbial en *-ment* peut avoir différentes entrées dans les tables de Molinier et donc dans le *Lefff*, il a au plus une entrée dans chacune des classes d'adverbes définies par (Molinier & Levrier, 2000). Associer un ou plusieurs synsets du WOLF à une entrée du *Lefff* pour un adverbe en *-ment* revient donc à associer à chaque synset une classe d'adverbes de Molinier. Pour cela, nous avons simplement attribué à chaque synset adverbial la classe la mieux partagée parmi ses lexèmes. Les couplages obtenus seront validés manuellement dans un proche avenir.

5. Conclusion et perspectives

A l'heure où l'absence de ressource lexicale à grande échelle pour le français se fait cruellement sentir, nous avons montré l'importance de faire collaborer différentes ressources existantes, de façon à enrichir ou diversifier les informations qu'elles contiennent. L'interaction *Lefff*-WOLF, via l'utilisation de DicoSyn, a permis de faire évoluer chacune de ces ressources vers une plus grande complétude qualitative et quantitative. Ce travail nous a en effet permis d'obtenir une augmentation de près de 55% du nombre de relations lexème-synset adverbial dans le WOLF, et d'attribuer des identifiants de synset à de nombreuses entrées adverbiales du *Lefff*.

Ces résultats encourageants montrent également la pertinence de l'exploitation d'un lexique sous forme de graphe, au moins en ce qui concerne l'accès automatique aux informations sémantiques qu'il contient. La synonymie et les adverbes en *-ment* ont constitué un terrain d'expérimentation idéal, et nous incitent à l'exploration d'autres relations, paradigmatisées (hyponymie, antonymie, ...) ou syntagmatiques (via l'analyse de corpus), ainsi qu'à d'autres parties du discours, comme par exemple les noms en *-ité* ou les verbes en *-ifier* et *-iser*.

Œuvres citées

- Danlos L. & Sagot B. 2007. Comparaison du Lexique-Grammaire et de Dicovalence : vers une intégration dans le *Lefff*. In *Actes de TALN 07*, Toulouse, France.
- Fellbaum C. 1998. *WordNet : An Electronic Lexical Database*. MIT Press.
- François J., Victorri B. & Manguin J.-L. 2002. Polysémie adjectivale et synonymie : l'éventail des sens de curieux. *La polysémie*.
- Molinier C. & Levrier F. 2000. *Grammaire des adverbes. Description des formes en -ment*. Genève, Suisse : Droz.
- Ploux S. & Victorri B. 1998. Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes. *Traitement Automatique des Langues (T.A.L.)*, 39(1), 161-182.

- Sagot B., Clément L., Villemonte de La Clergerie E. & Boullier P. 2006. The *Lefff* 2 syntactic lexicon for French : architecture, acquisition, use. In *Proc. of LREC'06*.
- Sagot B. & Danlos L. 2007. Améliorer un lexique syntaxique à l'aide des tables du lexique-grammaire – Constructions impersonnelles et expressions verbales figées. *Cahiers du Cental*.
- Sagot B. & Fišer D. 2008. Building a free french wordnet from multilingual resources. In *Actes de Ontolex 2008*, Marrakech, Maroc. (à paraître).
- Sagot B. & Fort K. 2007. Améliorer un lexique syntaxique à l'aide des tables du lexique-grammaire – adverbes en -ment. In *Actes du Colloque Lexique et Grammaire*, Bonifacio, France.
- Venant F. 2004. Polysémie et calcul du sens. In *Actes de JADT 2004*, Louvain-la-Neuve, Belgique.
- Venant F. 2007. Une exploration géométrique de la structure sémantique du lexique adjectival français. *Traitement Automatique des Langues (T.A.L.)*, 47(2).
- Vossen, P. 1999. *EuroWordNet : a multilingual database with lexical semantic networks for European Languages*. Kluwer, Dordrecht.

Résumé

Cet article présente un travail sur l'extension des entrées adverbiales du WOLF, une ressource lexicale sémantique du français, ainsi leur couplage avec celles du lexique syntaxique *Lefff*, principalement extraites des tables du lexique-grammaire de (Molinier & Levrier, 2000). Le travail repose sur l'exploitation des relations de dérivation et de synonymie, ces dernières étant extraites de la base de synonymes DicoSyn. La ressource sémantique adverbiale obtenue, librement disponible, est évaluée et validée manuellement de façon exhaustive.

Summary

Extension and coupling of syntactic and semantic resources for French adverbs

This paper presents a work on extending the adverbial entries of the WOLF, a semantic lexical resource for French, and connecting them with those of the syntactic lexicon *Lefff*, which were mostly extracted from the lexicon-grammar tables from (Molinier & Levrier, 2000). This work relies on the exploitation of the derivation and the synonyms relations; the latter are extracted from the DicoSyn synonyms database. The resulting semantic resource, which is freely available, is manually evaluated and validated in an exhaustive manner.

Adresse de l'auteur:

Benoît Sagot

Alpage, INRIA Paris-Rocquencourt / Paris 7 — benoit.sagot@inria.fr

Karën Fort

SRDI, INIST — karen.fort@inist.fr

Fabienne Venant

Talaris, INRIA Nancy Grand-Est — fabienne.venant@loria.fr