

PrepLex: un lexique des prépositions du français pour l'analyse syntaxique

Karen Fort, Bruno Guillaume

► **To cite this version:**

Karen Fort, Bruno Guillaume. PrepLex: un lexique des prépositions du français pour l'analyse syntaxique. Traitement Automatique des Langues Naturelles - TALN 2007, Jun 2007, Toulouse, France. 2007. <inria-00186761>

HAL Id: inria-00186761

<https://hal.inria.fr/inria-00186761>

Submitted on 12 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PrepLex : un lexique des prépositions du français pour l'analyse syntaxique

Karën Fort¹ Bruno Guillaume²

(1) projets Calligramme et TALARIS, (2) projet Calligramme

LORIA/INRIA Lorraine, UMR 7503, Nancy

Karen.Fort@loria.fr, Bruno.Guillaume@loria.fr

Résumé PrepLex est un lexique des prépositions du français. Il contient les informations utiles à des systèmes d'analyse syntaxique. Il a été construit en comparant puis fusionnant différentes sources d'informations lexicales disponibles. Ce lexique met également en évidence les prépositions ou classes de prépositions qui apparaissent dans la définition des cadres de sous-catégorisation des ressources lexicales qui décrivent la valence des verbes.

Abstract PrepLex is a lexicon of French prepositions which provides all the information needed for parsing. It was built by comparing and merging several authoritative lexical sources. This lexicon also shows the prepositions or classes of prepositions that appear in verbs sub-categorization frames.

Mots-clefs : prépositions, lexique, analyse syntaxique

Keywords: prepositions, lexicon, parsing

1 Introduction

Lors de la définition de classes d'entrées lexicales en fonction de leur catégorie, il apparaît naturellement une distinction entre deux types de classes. D'une part, les classes fermées dont on peut énumérer tous les éléments de façon exhaustive ; c'est le cas par exemple des pronoms ou des déterminants. D'autre part, les classes ouvertes pour lesquelles il n'est pas possible de lister tous les éléments (ils peuvent dépendre d'un vocabulaire spécifique à un domaine par exemple) ; les quatre grandes classes ouvertes sont les noms, les verbes, les adjectifs et les adverbes. La méthodologie de construction d'un lexique doit nécessairement être adaptée en fonction de cette notion de classe.

Le statut de la classe des prépositions est plus difficile à établir. A priori, l'ensemble des prépositions peut sembler une classe fermée dont on peut énumérer les éléments ; en pratique, la comparaison de différentes sources disponibles montre qu'il n'est pas facile de déterminer de façon exhaustive la liste des prépositions. Or, celles-ci représentent plus de 14% des lemmes du français¹.

Dans un lexique complet, il est important d'avoir des informations de sous-catégorisation pour les mots prédicatifs (Briscoe & Carroll, 1993; Carroll & Fang, 2004). Ces informations de sous-catégorisation font souvent référence à des prépositions dans la description des arguments. En effet, ces arguments sont souvent contraints à utiliser une préposition particulière (par exemple *compter sur*) ou un ensemble de prépositions qui ont un aspect sémantique commun (par exemple *aller LOC*, où *LOC* peut être remplacé par n'importe quelle préposition locative).

Pour une analyse syntaxique profonde, il est utile de distinguer les compléments indirects requis par le verbe des autres compléments adjoints ne figurant pas dans la valence du verbe. Les deux exemples (1a) et (1b) ont la même structure de surface et seule la sémantique permet de distinguer les deux usages différents de la préposition *sur* : elle introduit un complément oblique dans le premier cas et un complément adjoint dans le second. Ce problème peut être géré par l'utilisation d'informations sémantiques plus fines.

- 1a. *Jean compte sur ses amis*
- 1b. *Jean compte sur ses doigts*

Cette distinction amène à distinguer deux usages différents des prépositions et est donc une source d'ambiguïté lexicale. Pour limiter cette ambiguïté, il est important que le lexique repère les prépositions qui peuvent jouer ces deux rôles (ce sont les prépositions argumentales).

Notre travail est destiné à fournir un lexique utilisable par un analyseur syntaxique. Nous nous sommes restreints aux aspects purement syntaxiques et à quelques éléments sémantiques comme la définition des ensembles de prépositions ayant un aspect sémantique commun (comme *LOC*). Le lexique produit est diffusé sous licence libre et a vocation à être intégré dans des ressources plus larges, existantes ou à venir.

La section 2 décrit les sources utilisées et la méthodologie de comparaison, la section 3 décrit les résultats de cette comparaison. La section 4 décrit comment le lexique est constitué à partir des résultats précédents. Enfin, la section 5 présente un exemple d'utilisation de ce lexique en analyse syntaxique.

¹voir par exemple sur un corpus journalistique : https://www.kuleuven.be/ilt/blf/rechbaselex_kul.php\#freq (Selva *et al.*, 2002)

2 Méthodologie

L'utilisation de prépositions dans le cadre d'une analyse syntaxique nécessite une liste large, inventoriant à la fois les prépositions non argumentales et celles susceptibles d'apparaître dans les cadres de sous-catégorisation des verbes.

2.1 Utilisation de lexiques syntaxiques

Bien entendu, il existe déjà des lexiques syntaxiques, qui proposent un ensemble intéressant de prépositions. Ainsi, le *Lefff* (Sagot *et al.*, 2006) fournit une liste conséquente de prépositions, mais la partie syntaxique du lexique est encore en cours de développement, il présente donc peu de prépositions dans les cadres de sous-catégorisation des verbes. En outre, certaines prépositions du *Lefff* semblent obsolètes ou rares. Le dictionnaire français-UNL (Sérasset & Boitet, 2000) en propose également, mais sa couverture reste limitée et la qualité des entrées est encore inégale. D'autres sources proposent des prépositions dans les cadres de sous-catégorisation des verbes, mais les listes ne sont pas tout à fait cohérentes d'une source à l'autre.

Nous avons donc effectué, dans un premier temps, un travail d'inventaire des prépositions présentes dans un certain nombre de ressources, des lexiques et/ou dictionnaires d'une part, pour la liste générale, des lexiques syntaxiques d'autre part, pour la liste de prépositions argumentales. Deux ressources se classent cependant dans les deux catégories, le *Lefff* et le dictionnaire UNL :

- Le *Lefff* (Lexique des Formes Fléchies du Français (Sagot *et al.*, 2006)) est un lexique morphologique et syntaxique du français à large couverture (plus de 110 000 lemmes). Dans sa version 2.2.1, ce lexique contient 48 prépositions simples et 164 prépositions complexes. Il présente également des informations de sous-catégorisation des verbes, qui font apparaître 14 prépositions que nous qualifierons d'argumentales.
- UNL (Universal Networking Language (Sérasset & Boitet, 2000)), est un dictionnaire du français vers un anglais désambiguïsé conçu pour la traduction automatique, qui comprend des informations syntaxiques dans sa partie française. UNL n'a qu'une couverture assez faible (moins de 27 000 lemmes), mais il propose dans sa partie anglaise des informations de type sémantique que nous envisageons d'utiliser par la suite. UNL contient 48 prépositions simples dont 10 apparaissant dans les cadres de sous-catégorisation des verbes.

2.2 Utilisation de sources de référence

Nous avons ensuite complété la liste de prépositions en utilisant différentes sources construites manuellement, lexique ou dictionnaire, voire grammaire :

- Le Grevisse (Grevisse, 1997), dans sa version papier, nous a permis de vérifier certaines intuitions concernant l'obsolescence ou l'usage de certaines prépositions.
- Le TLFi (Trésor de la langue française informatisé), que nous avons consulté via l'interface du CNRTL², offre une liste de prépositions un peu différente des autres. Elle comporte notamment les formes *voici* et *voilà*, rarement citées dans les autres sources à notre disposition.

²voir : <http://www.cnrtl.fr>

- Enfin, la base de prépositions PrepNet (Saint-Dizier, 2006) nous a permis de vérifier à la fois la complétude de notre liste et les informations sémantiques présentes dans certaines sources.

2.3 Utilisation de dictionnaires de valences verbales

Nous avons ensuite cherché à enrichir la liste des prépositions apparaissant dans les cadres de sous-catégorisation des verbes de *Lefff* et UNL en nous référant à deux sources traitant plus particulièrement des verbes :

- Le dictionnaire de valences des verbes du français DICOVALENCE, héritier de PROTON (van den Eynde & Mertens, 2002), dont la démarche est fondée sur l’approche pronominale. Dans sa version 1.1, ce dictionnaire donne les cadres valenciels de plus de 3700 verbes. Nous avons extrait les prépositions simples et multi-mots qu’il contient (soit plus de 40), ainsi que leurs traits sémantiques associés.
- Nous avons complété cette liste de prépositions argumentales par celle de SynLex (Gardent *et al.*, 2006), lexique syntaxique créé à partir des tables du lexique-grammaire du LADL (Gross, 1975).

Nous avons, à partir de ces différentes sources, effectué une étude systématique de la présence de chaque préposition, de leur appartenance éventuelle à des cadres valenciels, ainsi que de certains traits sémantiques qui leur sont associés. Nous avons ensuite regroupé les prépositions qui apparaissaient à la fois en tant qu’entrée lexicale et dans les cadres de sous-catégorisation des verbes.

Il est à noter que nous avons dû restreindre notre analyse des cadres de sous-catégorisation à ceux des verbes, du fait qu’il n’existe encore à notre connaissance aucun lexique présentant une information syntaxique suffisamment riche sur les adjectifs ou les noms.

Les prépositions multi-mots présentant des caractéristiques (nombre) et des difficultés spécifiques (segmentation), nous en avons fait un inventaire séparé selon les mêmes méthodes.

3 Résultat de la comparaison des sources

3.1 Prépositions simples

Nous avons ainsi listé 85 prépositions simples, dont 24 apparaissent dans des cadres de sous-catégorisation de verbes (cf. tableau 1).

Il est à noter que les 4 sources qui décrivent des cadres de sous-catégorisation utilisent des formats très différents. *Lefff* propose une vision condensée des verbes, les cadres valenciels étant regroupés dans une seule entrée ; à l’opposé, SynLex et DICOVALENCE décrivent de nombreux cadres en distinguant par exemple systématiquement les différentes réalisations syntaxiques des arguments. Les valeurs relatives sur une ligne ne reflètent donc pas du tout la couverture lexicale des sources.

	Lexiques					Cadres de sous-catégorisation			
	<i>Lefff</i>	TLFi	Grevisse	PrepNet	UNL	<i>Lefff</i>	DV ^a	SynLex	UNL
à	X	X	X	loc		319	895	887	246
après	X	X	X	loc	X	2	12	1	
aussi					X				
avec	X	X	X	X	X	35	193	611	49
chez	X	X	X	loc	X		9		1
comme	X				X	14	11	10	3
de	X	X	X	deloc	X	310	888	1980	282
depuis	X	X	X	deloc	X		2	1	
derrière	X	X	X	loc	X		3		
devers	X	X	X						
dixit	X								
emmi		X							
entre	X	X	X	loc	X		19	4	
hormis	X	X	X	X	X				
jusque	X	X	X		X		7		
lès	X	X	X						
moyennant	X	X	X	X	X				
par	X	X	X	loc	X	3	38	73	8
parmi	X	X	X	loc	X		7	7	
passé		X			X				
selon	X	X	X	X	X		1	1	
voici		X			X				

TAB. 1 – Quelques prépositions simples dans les différentes sources

^aDICOVALENCE

3.2 Prépositions multi-mots

Nous avons obtenu une liste de 222 prépositions multi-mots, dont 18 apparaissent dans des cadres de sous-catégorisation de verbes (cf. tableau 2). Il est intéressant de noter que seuls DICOVALENCE et SynLex proposent des prépositions multi-mots dans leurs cadres de sous-catégorisation. Le *Lefff* fournit quant à lui une liste impressionnante de prépositions multi-mots (plus de 150) qui représente une excellente base de travail.

4 Construction du lexique

Le premier critère de sélection que nous avons choisi d'appliquer pour construire notre lexique est qu'une préposition doit être listée dans au moins une source parmi celles citées. Par ailleurs, nous considérons qu'une préposition est argumentale si elle apparaît dans au moins un cadre de sous-catégorisation de verbe.

	Lexiques					Cadres de sous-catégorisation			
	Lefff ^a	TLFi	Grevisse	PrepNet	UNL	Lefff	DV ^a	SynLex	UNL
à cause de	X		X	X					
à la faveur de			X	X					
à partir de	X		X	deloc				1	
afin de	X	X	X	X					
au nord de				loc					
au vu de	X								
auprès de	X	X	X	loc			27	35	
comme de							1		
conformément à	X			X					
d'avec			X				1	6	
en faveur de	X		X	X			13		
il y a	X								
jusqu'à	X			loc	X		10		
jusqu'en	X								
jusqu'où	X								
loin de	X		X	loc					
par suite de			X						
pour comble de	X								
près de	X		X	loc					
quant à	X	X	X						
tout au long de	X			X					
vis-à-vis de	X		X	X				1	

TAB. 2 – Quelques prépositions multi-mots dans les différentes sources

^aDICOVALENCE

4.1 Filtrage manuel

Nous avons ensuite trié ces prépositions en fonction de critères simples et avons identifié en particulier celles qui étaient à éliminer parce que :

- visiblement erronées, c'est le cas par exemple de *aussi*, présent dans le dictionnaire UNL en tant que préposition,
- obsolètes ou d'un emploi extrêmement rare, comme *emmi* (TLFi), *devers* (Lefff, TLFi, Grevisse) ou encore *comme de* (DICOVALENCE).

Nous avons également effectué un tri dans les traits sémantiques et avons supprimé les entrées erronées, telles que *avec* comme locatif dans SynLex et dans DICOVALENCE.

4.2 Quelques remarques

Certaines sources font apparaître comme des prépositions des formes qui ne sont pas considérées comme telles en linguistique. C'est le cas en particulier de :

- *comme*, qui n'est pas cité dans les trois sources de référence que sont le Grevisse, le TLFi et PrepNet, il est en effet ambigu et peut également être considéré comme une conjonction,
- *il y a* ou *y compris*, qui ne sont citées que dans le Lefff,

Lexiques						Cadres de sous-catégorisation				
Lefff	TLFi	Grevisse	PrepNet	UNL	PrepLex	Lefff	DV	SynLex	UNL	PrepLex
44	69	55	36	46	49	14	24	18	11	23

TAB. 3 – Total des prépositions simples par source

Lexiques						Cadres de sous-catégorisation				
Lefff	TLFi	Grevisse	PrepNet	UNL	PrepLex	Lefff	DV	SynLex	UNL	PrepLex
166	11	77	89	2	206	0	16	4	0	15

TAB. 4 – Total des prépositions multi-mots par source

– *d’avec*, qui, bien qu’il apparaisse dans le Grevisse, n’est présent que dans les cadres de sous-catégorisation de DICOVALENCE et SynLex.

Nous avons choisi de conserver ces formes dans notre lexique, pour des raisons pratiques liées à l’application visée, l’analyse syntaxique.

Par ailleurs, même si sa couverture est large, notre lexique n’est évidemment pas exhaustif. Il serait intéressant d’y ajouter certaines formes manquantes, notamment :

- des prépositions présentes dans le DAFLES (Selva *et al.*, 2002), comme par exemple la forme *au détriment de*,
- des prépositions citées dans des grammaires de référence, comme *question*, dans la Grammaire méthodique du français (Riegel *et al.*, 1997),
- les multiples prépositions locatives (et, par métonymie, temporelles) que peut préfixer la forme *jusqu’*, par exemple *jusqu’auprès de*. Cette forme élidée de *jusque* pourrait sans doute faire l’objet d’un traitement particulier en tant que *modifieur* de préposition. Il en va d’ailleurs de même de *dès*, suivi d’un temporel (ou d’un locatif, par métonymie).

Ce tri a également permis de mettre en évidence certaines difficultés, en particulier les élisions dans les formes multi-mots, telle *afin de*, *afin d’*, ou les contractions telles *face à*, *face au* ou *à partir de*, *à partir du*, qui seront traitées lors de la segmentation.

D’autres, comme *lès*, qui n’est utilisé qu’en toponymie dans des formes avec tirets (comme *Bathelémont-lès-Bauzemont*), seront traitées en amont, lors de la segmentation.

4.3 Résultats

Au final, nous obtenons une liste de 49 prépositions simples, dont 23 apparaissent dans des cadres de sous-catégorisation des verbes, dans au moins une source, et sont donc considérées comme argumentales (cf. tableau 3).

Nous obtenons également une liste de plus de 200 prépositions multi-mots, dont 15 apparaissent dans des cadres de sous-catégorisation des verbes, dans au moins une source, et sont donc considérées comme argumentales (cf. tableau 4).

Nous avons pour l’instant limité les informations sémantiques utilisées dans le lexique à *loc* (locatif) et *deloc* (délocatif), mais nous avons l’intention d’étendre ces catégories à celles retenues dans DICOVALENCE (temps, quantité, manière).

En outre, nous nous sommes également référés aux sources pour renseigner les catégories des arguments introduits par les prépositions argumentales.

Il n'existe pas encore de format normalisé pour les lexiques syntaxiques, même si un effort d'homogénéisation est en cours (Projet Lexsynt). Actuellement, PrepLex est donc présenté dans un format texte qui permet à la fois l'édition manuelle et l'utilisation dans un analyseur ou dans d'autres outils de traitement de la langue. Dans ce format, les informations syntaxiques sont décrites à l'aide structures de traits récursives de profondeur 2. Le niveau externe décrit la structure en termes d'“argument” : ce niveau contient toujours un trait “head” et un trait pour chacun des “arguments” de l'entrée. Le niveau interne décrit alors plus finement chaque argument. De plus, ce format permet de définir des informations syntaxiques de façon modulaire en factorisant les parties redondantes. Dans le cas des prépositions, toutes les entrées partagent le même squelette :

```
Prep : [
head [cat=prep, prep=#, funct=#]
comp [cat=#, cpl=@]
]
```

Lorsque ce squelette est instancié pour une préposition particulière, les 3 valeurs de traits (notées “#”) doivent impérativement être renseignées et la valeur de trait notée “@” est optionnelle. Ces traits sont repérés par leur nom (*prep*, *funct*) pour la tête, par une notation pointée (*comp.cat*, *comp.cpl*) pour l'argument.

```
à          Prep [prep=a|LOC; funct=aobj|loc|adj; comp.cat=np|sinf;
                comp.cpl=void|ceque]
après      Prep [prep=apres|LOC; funct=obj|loc|adj; comp.cat=np]
avec       Prep [prep=avec; funct=obj|adj; comp.cat=np]
à_travers Prep [prep=a_travers;funct=obj|adj; comp.cat=np]
```

Techniquement, la seule difficulté est de choisir comment représenter l'appartenance à une classe sémantique de prépositions comme *loc*. Ici, nous avons choisi de définir comme valeurs atomiques possibles pour le trait “prep”, l'ensemble des prépositions argumentales et l'ensemble des classes sémantiques (notées en majuscules). On utilise alors la disjonction *a|LOC* pour indiquer que la préposition *à* peut être utilisée soit comme une préposition particulière, soit comme une préposition locative.

Nous avons en outre décidé d'indiquer dans le lexique les sources dans lesquelles la préposition apparaît, afin de permettre un éventuel filtrage pour des utilisations particulières. Dans le cas des prépositions argumentales, nous avons ajouté un champ comportant la fréquence d'apparition et un exemple pris dans l'une des sources.

5 Un exemple d'utilisation dans un système TAL

Nous exposons ici de manière succincte quelques problèmes posés par les prépositions lors de l'analyse syntaxique.

5.1 Spécificités liées à la segmentation

La première difficulté pour l'intégration des prépositions dans un analyseur se situe au niveau de la segmentation en lexèmes. Il faut gérer les phénomènes d'élision : *au* doit être traité comme *à*

le, le *de* finissant certaines prépositions multi-mots peut être élide en *d'*. Cependant, ces phénomènes ne sont pas spécifiques aux prépositions, il sont traités soit dans le lexique (par exemple *Lefff* distingue les deux formes *au cours de* et *au cours d'*), soit lors de la segmentation. Nous avons choisi de les traiter dans le segmenteur, afin de simplifier la maintenance du lexique.

Une difficulté plus directement liée aux prépositions multi-mots, est la possibilité d'avoir une ambiguïté de segmentation. Par exemple, dans les deux phrases (2a) et (2b) la suite de mots *au cours de* est une préposition multi-mots dans le premier cas, mais elle doit être décomposée dans le deuxième. D'autres prépositions multi-mots ne nécessitent jamais de découpage, par exemple *y compris*.

2a. *Il a beaucoup travaillé au cours de cette année*

2b. *Il a beaucoup travaillé au cours de M. Durand*

5.2 Prépositions compléments vs prépositions argumentales

Lors de l'analyse syntaxique, on doit nécessairement distinguer l'usage d'une préposition pour introduire un argument du verbe de celui d'une préposition introduisant un complément. Comme on l'a vu (exemples (1a) et (1b)), cette distinction est souvent difficile à établir et repose sur des considérations sémantiques. L'analyse syntaxique doit alors maintenir l'ambiguïté de rattachement. Le fait d'avoir des informations précises sur les prépositions argumentales permet de contrôler ces ambiguïtés.

6 Conclusion

En comparant divers lexiques et dictionnaires, nous avons établi une liste de prépositions utiles pour le TAL. Nous nous sommes concentrés surtout sur les aspects syntaxiques. Un tri manuel a permis d'écarter des prépositions obsolètes ou très rares et quelques cas d'erreur. Le lexique ainsi produit contient plus de 250 prépositions dont 49 sont des prépositions simples.

Dans les lexiques syntaxiques, les cadres de sous-catégorisation décrivent les prépositions introduisant certains arguments. Des prépositions apparaissant dans les entrées verbales d'un lexique syntaxique sont appelées argumentales. Nous avons identifié 40 prépositions argumentales.

Le lexique développé est librement disponible³. Ce lexique va nécessairement évoluer. D'autres sources d'information auraient leur place dans ce travail, notamment les champs *constructions* des verbes du TFLi qui font référence à des prépositions qui sont donc argumentales. Il est prévu d'utiliser prochainement cette source pour faire évoluer le lexique.

La mise en place d'une base de donnée contenant ce lexique est en cours³. Cela devrait permettre de faciliter la maintenance du lexique, mais aussi d'enrichir les données pour chaque entrée, notamment avec des exemples d'usage ou des traits sémantiques plus variés (*loc*, *deloc*, mais aussi *tim*, *man*, *qty*). Nous envisageons d'y ajouter des informations de fréquence sur corpus.

Une tâche de bien plus grande ampleur serait d'enrichir ce lexique avec des informations sémantiques plus fines que la seule référence aux classes *loc*, *deloc*, ... Il existe de nombreux

³<http://loriatat.loria.fr/Resources.html>

travaux de linguistique portant sur les prépositions. Cependant, la plupart s'attache à des descriptions sémantiques fines d'un petit nombre de prépositions ; une exception notable étant le travail réalisé dans PrepNet (Saint-Dizier, 2006). Il conviendrait donc de transformer ces ressources pour les rendre directement utilisables par un système de traitement automatique des langues.

Remerciements

Nous tenons à remercier chaleureusement M. Guy Perrier, Professeur à l'Université Nancy 2, pour ses conseils patients et éclairés.

Références

- BRISCOE T. & CARROLL J. A. (1993). Generalised probabilistic LR parsing for unification-based grammars. *Computational Linguistics*.
- CARROLL J. A. & FANG A. C. (2004). The automatic acquisition of verb subcategorisations and their impact on the performance of an HPSG parser. In *Proceedings of the 1st International Joint Conference on Natural Language Processing (IJCNLP)*, p. 107–114, Sanya City, China.
- GARDENT C., GUILLAUME B., PERRIER G. & FALK I. (2006). Extraction d'information de sous-catégorisation à partir des tables du LADL. In *Proceedings of TALN 06*, p. 139–148, Leuven.
- GREVISSE M. (1997). *Le Bon Usage – Grammaire française, édition refondue par André Goosse*. Paris – Louvain-la-Neuve : DeBoeck-Duculot, 13^e édition.
- GROSS M. (1975). *Méthodes en syntaxe*. Hermann.
- RIEGEL M., PELLAT J.-C. & RIOUL R. (1997). *Grammaire méthodique du français*. PUF, 3^e édition.
- SAGOT B., CLÉMENT L., VILLEMONTÉ DE LA CLERGERIE E. & BOULLIER P. (2006). The Lefff 2 syntactic lexicon for French : architecture, acquisition, use. In *Actes de LREC 06, Gênes, Italie*.
- SAINT-DIZIER P. (2006). PrepNet : a Multilingual Lexical Description of Prepositions. In *LREC, Gênes, 12/05/2006-14/05/2006*, p. 877–885 : European Language Resources Association (ELRA).
- SELVA T., VERLINDE S. & BINON J. (2002). Le DAFLES, un nouveau dictionnaire pour apprenants du français. In *Actes du dixième congrès EURALEX'2002 (European Association for Lexicography)*. Copenhague.
- SÉRASSET G. & BOITET C. (2000). On UNL as the future "html of the linguistic content" and the reuse of existing NLP components in UNL-related applications with the example of a UNL-French deconverter. In *Proceedings of COLING 2000, Saarebruecken, Germany*.
- VAN DEN EYNDE K. & MERTENS P. (2002). *La valence : l'approche pronominale et son application au lexique verbal*, In *Journal of French Language Studies*, p. 63–104. Cambridge University Press, 13^e édition.