

## Vers la création d'un Verbnets du français

Laurence Danlos, Takuya Nakamura, Quentin Pradet

► **To cite this version:**

Laurence Danlos, Takuya Nakamura, Quentin Pradet. Vers la création d'un Verbnets du français. TALN 2014, 21ème conférence sur le Traitement Automatique des Langues Naturelles, Atelier FondamenTAL, Jul 2014, Marseille, France. 2014. <hal-01084681>

**HAL Id: hal-01084681**

**<https://hal.inria.fr/hal-01084681>**

Submitted on 19 Nov 2014

**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.

## Vers la création d'un Verb $\ni$ net du français

Laurence Danlos<sup>1</sup> Takuya Nakamura<sup>2</sup> Quentin Pradet<sup>3</sup>

(1) Université Paris Diderot, Sorbonne Paris Cité, ALPAGE, IUF

(2) IGM-LabInfo Université Paris-Est - 77457 Marne-la-Vallée Cedex 2

(3) CEA, LIST, Laboratoire Vision et Ingénierie des Contenus - F-91191 Gif-sur-Yvette

danlos@inria.fr, takuya.nakamura@univ-mlv.fr, quentin.pradet@cea.fr

**Résumé.** VerbNet est une ressource lexicale pour les verbes anglais qui est bien utile pour le TAL grâce à sa large couverture et sa classification cohérente. Une telle ressource n'existe pas pour le français malgré quelques tentatives. Nous montrons comment adapter semi-automatiquement VerbNet en utilisant deux ressources lexicales existantes, le LVF (Les Verbes Français) et le LG (Lexique-Grammaire).

**Abstract.** VerbNet is an English lexical resource that has proven useful for NLP due to its high coverage and coherent classification. Such a resource doesn't exist for French, despite some (mostly automatic and unsupervised) attempts. We show how to semi-automatically adapt VerbNet using existing lexical resources, namely LVF (Les Verbes Français) and LG (Lexique-Grammaire).

**Mots-clés :** VerbNet, cadres de sous-catégorisations, rôles sémantiques.

**Keywords:** VerbNet, frames, semantic roles.

### 1 Introduction

Le Traitement Automatique des Langues nécessite des lexiques et des gros corpus annotés pour analyser efficacement des textes en domaine ouvert. Obtenir une telle quantité de données est un problème en soi, connu sous le nom de *knowledge acquisition bottleneck* dans la littérature (Gale *et al.*, 1992). Alors qu'annoter de plus en plus de corpus pourra réduire ce goulot d'étranglement pour certains domaines, un travail lexicographique efficace peut apporter des améliorations en décrivant explicitement les similarités et différences entre mots.

Au moment de définir un tel lexique, les deux problèmes principaux sont la granularité et la distinction des sens. Ces deux problèmes ont été traités par les classes de Levin (Levin, 1993). En effet, les classes de Levin regroupent les verbes en utilisant les alternances syntaxiques observables, ainsi que d'autres critères sémantiques et morphologiques clairs. VerbNet (Kipper-Schuler, 2005) est un lexique électronique inspiré des classes de Levin qui encode aussi des rôles thématiques et une décomposition sémantique (section 3). Il a été ajouté au fil des ans à VerbNet de nouvelles constructions et classes, ainsi que des corrections variées, par rapport aux classes de Levin.

Avec VerbNet il est possible d'utiliser une construction syntaxique pour faire correspondre un argument d'un verbe à des rôles sémantiques (Swier & Stevenson, 2005; Pradet *et al.*, 2013). Cette tâche, l'annotation en rôles sémantiques, a pris graduellement de l'importance : elle aide l'extraction d'information (Surdeanu *et al.*, 2003), les systèmes de question-réponse (Shen & Lapata, 2007), l'extraction d'évènements (Exner & Nugues, 2011), la traduction automatique (Bazrafshan & Gildea, 2013) et même la détection de plagiat (Osman *et al.*, 2012), la prédiction des cours de bourse (Xie *et al.*, 2013), l'interprétation de recette de cuisines (Malmaud *et al.*, 2014) ou la génération de scène 3D (Chang *et al.*, 2014). Grâce à sa forte couverture (plus de quatre mille verbes) et son découpage utile des verbes, VerbNet est adapté à l'annotation en rôles sémantiques, en particulier dans les domaines ne disposant pas de corpus annoté.

Cependant, un VerbNet de qualité n'existe actuellement que pour l'anglais. Une telle ressource serait pourtant encore plus utile pour des langues moins bien dotées en corpus annotés en rôles sémantiques. VerbNet a un potentiel inter-linguistique, comme cela a été montré avec le Portugais (Kipper-Schuler, 2005, section 2.2.2). Avec l'objectif de développer une version française de VerbNet (nommée Verb $\ni$ net), nous utilisons deux ressources lexicales françaises existantes encodant

le comportement syntaxique et sémantique des verbes (section 3). Nous avons mis en correspondance les classes VerbNet avec ces ressources pour restreindre les traductions françaises des membres de telles classes (section 4.1). La deuxième étape, en cours, adapte les constructions syntaxiques de VerbNet au français. Cette étape soulève différents problèmes (section 4.2). La troisième étape corrigera les verbes obtenus dans la première étape pour s'assurer qu'ils correspondent bien aux constructions retenues dans la deuxième étape.

## 2 Travaux antérieurs

Pour des langues autres que l'anglais, Merlo *et al.* (2002) ont utilisé la similarité entre l'anglais et l'italien pour traduire 20 classes de Levin. Des acquisitions automatiques ont été menées en espagnol (Ferrer, 2004), allemand (Im Walde, 2006) et japonais (Suzuki & Fukumoto, 2009). Les seules traductions directes dont nous avons la connaissance sont les VerbNet estoniens (Jentson, 2014) et portugais brésilien (Scarton & Aluisio, 2012).

Pour le français, Saint-Dizier (1996) a produit en premier une ressource de type VerbNet à partir des classes de Levin. À notre connaissance, cette ressource n'est plus développée et le résultat n'est pas disponible. D'autres travaux se sont concentrés sur l'acquisition automatique de cadres de sous-catégorisation qui ont été regroupés avec des critères syntaxiques et sémantiques. Sun *et al.* (2010) ont utilisé un tel lexique de cadres de sous-catégorisation (Messiant *et al.*, 2010) pour regrouper les verbes avec des traits syntaxiques et sémantiques (collocations et préférence lexicale des verbes). Une évaluation sur une vérité-terrain construite manuellement a mené à un F-score de 55.1 %. Falk *et al.* (2012) ont appliqué un algorithme de regroupement différent qui a amélioré le F-Score à 70 % sur la même vérité-terrain mais un peu plus facile à traiter car plusieurs rôles proches ont été regroupés. Ces ressources mettent en avant de nouvelles manières de regrouper des verbes français, mais comportent trop d'erreurs. Même si leurs résultats peuvent être améliorés, nous pensons qu'il y a aussi un besoin d'un VerbNet du français validé manuellement. Notre traduction sera liée au VerbNet anglais et aux deux ressources linguistiques françaises que nous utilisons, le LVF et le LG. Cette ressource sera aussi ouverte : nous voulons encourager les contributions externes.

## 3 Présentation de VerbNet et des ressources lexicales françaises

**VerbNet** Le premier niveau de la hiérarchie VerbNet est composé de 270 classes. Chacune de ces classes est potentiellement divisée en sous-classes, ce qui forme une hiérarchie arborescente. Pour chaque (sous-)classe, ce lexique électronique liste :

- la liste des verbes,
- les rôles thématiques, éventuellement associés à des restrictions de sélection de type humain ou organisation,
- et une liste de constructions syntaxiques que chaque verbe accepte.

Une construction syntaxique (*frame* dans la terminologie VerbNet) inclut :

- un exemple illustratif,
- une formule syntaxique liant les arguments syntaxiques aux rôles thématiques, e.g. *Agent V Patient*,
- et une formule sémantique proche de la logique du premier ordre décrivant l'action et l'état des participants avant, pendant et après le procès.

**Ressources lexicales françaises** Depuis les années 1970, deux ressources lexicales à large couverture ont été développées manuellement pour les verbes français :

- LVF (Les Verbes Français, (Dubois & Dubois-Charlier, 1997)) inclut environ 25 000 entrées classifiées dans 14 classes sémantiques, 54 sous-classes syntaxico-sémantiques et 248 sous-sous-classes.
- LG (Lexique-Grammaire, (Gross, 1975; Boons *et al.*, 1976)) inclut environ 14 000 entrées classifiées en 67 tables, chaque table groupant des verbes partageant les mêmes constructions syntaxiques de base et parfois les mêmes comportements sémantiques. Une table comporte des colonnes qui encodent des restrictions additionnelles (restrictions de sélection, constructions syntaxiques associées, etc.) et qui s'appliquent à un sous-ensemble des verbes de la table.

Les classes LVF et les tables LG peuvent être comparées aux classes VerbNet. Cependant, ces ressources lexicales n'encodent ni rôles thématiques ni formules sémantiques, et ne sont donc pas directement utilisables pour l'annotation en rôles sémantiques. De plus, elles encodent des usages parfois trop techniques, argotiques ou métaphoriques, or nous préférons nous concentrer sur les quelques milliers d'emplois de verbes les plus fréquents. La nouvelle ressource lexicale française,

Verb $\ni$ net, s'appuie à la fois sur l'encodage syntaxique et/ou sémantique riche de ces deux ressources et sur l'information sémantique contenue dans le VerbNet anglais, qui est une langue relativement proche du français.

## 4 Construction de Verb $\ni$ net

Notre principe de base est que le premier niveau de la hiérarchie de Verb $\ni$ net doit être aussi proche que possible de celui de VerbNet et ses 270 classes. Néanmoins, certaines classes doivent disparaître. La raison peut être purement morphologique : toute classe VerbNet composée uniquement de verbes morphologiquement identiques à des noms n'a pas d'équivalent français, ce qui est le cas de la classe [pit-10.7](#) composée de verbes comme *bark* et *bone*, et de la classe [weekend-56](#) avec des verbes comme *weekend* ou *december*. Par contre, la classe [debone-10.8](#) composée de verbes formés avec le préfixe *de-* précédant un nom (*debark*, *debone*) a bien un équivalent français avec les verbes formés par les préfixes *dé-* ou *é-* (*désosser*, *déveiner*, *équeuter*). Étant donné ce principe de base, la construction de Verb $\ni$ net se fait en deux étapes.

### 4.1 Première étape

La première étape de la construction de Verb $\ni$ net a consisté à déterminer les verbes français appartenant à chacune des 270 classes de VerbNet. Cette étape a été réalisée de la façon suivante :

1. Pour une classe VerbNet donnée  $C_e$ , nous avons assigné manuellement les classe(s) LVF  $C_{lvf}$  et les table(s) LG  $C_{lg}$  correspondant à la définition sémantique de  $C_e$ , par exemple : [put-9.1](#)  $\mapsto$  [L3b](#) et [38LD](#), [body\\_internal\\_motion-49](#)  $\mapsto$  [M1a](#) et [32CL](#) ou [32R3](#) ou [32C](#)
2. Nous avons utilisé deux dictionnaires bilingues (SCI-FRAN-EURADIC et le Wiktionnaire) qui nous donnent la liste  $L_{trad}$  des traductions françaises des verbes anglais appartenant à la classe  $C_e$ .
3. Nous avons enfin obtenu la liste des verbes français : ce sont les verbes de  $L_{trad}$  appartenant à l'intersection de  $C_{lvf}$  et  $C_{lg}$  (e.g. *mettre*, *poser* ou *installer* pour [put-9.1](#)).

Cette étape a été réalisée rapidement et a donné de bons résultats : en ne conservant que les verbes à l'intersection de  $L_{trad}$ ,  $C_{lvf}$  et  $C_{lg}$ <sup>1</sup>, les résultats sont précis et cohérents syntaxiquement et sémantiquement. Par exemple, la classe [scribble-25.2](#) contient 18 verbes en anglais ; elle est associée à la classe LVF [R3a.1](#) et la table LG [32A](#), ce qui produit une liste de 16 verbes français : *composer*, *couper*, *donner*, *exécuter*, *fabriquer*, *faire*, *forger*, *former*, *imprimer*, *lever*, *produire*, *reproduire*, *sculpter*, *tailler*, *tirer* et *tracer*. Tous ces verbes sont valides.

Cette première étape a produit un lexique contenant 3888 emplois de verbes correspondant à 2160 lemmes.

### 4.2 Deuxième étape

La deuxième étape de la construction de Verb $\ni$ net est plus laborieuse que la première. Pour chacun des 270 classes françaises  $C_f$ , nous devons déterminer :

- les sous-classes éventuelles en suivant si possible les sous-classes existantes pour l'anglais,
- les constructions valides en français en ajustant, si besoin, les rôles thématiques et les restrictions de sélection.

Cette étape a d'abord demandé le développement d'un outil d'édition (section 4.2.1) afin d'assister le travail lexicographique. Ensuite, il a fallu définir des principes de base pour les frames françaises (section 4.2.2). Enfin, l'étude de chacune des 270 classes  $C_f$  demande de faire un travail syntaxique et sémantique minutieux qui peut aboutir à une réorganisation des classes et sous-classes (Section 4.2.3).

#### 4.2.1 Outil d'édition de Verb $\ni$ net

Nous avons développé un outil en ligne pour pouvoir éditer collaborativement des classes et des frames VerbNet en manipulant directement leur représentation sur le site web, ce qui évite d'avoir à modifier des fichiers XML et permet de se concentrer sur les problèmes linguistiques. Cette interface a été développée à l'aide du framework web Django qui manipule une base de données PostgreSQL qui stocke la ressource et conserve l'historique entier des modifications.

1. Quand cette intersection est vide, la liste non-vide ( $C_{lvf} \vee L_{trad}$  ou  $C_{lg} \vee L_{trad}$ ) a été choisie.

## remove-10.1 <sup>↗</sup>

Classe 10.1 **E3c** <sup>↗</sup> **38LS** <sup>↗</sup>

- Paragon : enlever
- Membres : abolish abstract cull deduct delete depose discharge disengage disgorge dislodge dismiss draw eject eliminate eradicate evict excise excommunicate expel extinguish extirpate extract extrude lop omit ostracize oust partition prise pry ream reap remove retract roust separate sever shoo subtract uproot winkle withdraw wrench
- Traductions : arracher chasser couper distraire dégager dégainer déloger déménager déraciner déterrer effacer enlever exclure expulser extirper extraire lever libérer prélever puiser rejeter retirer soustraire soutirer supprimer tirer traire vider éliminer évacuer ôter barrer casser cueillir dissiper débarquer débloquer décompter déduire défalquer détacher escamoter exciser liquider rabattre rayer retrancher récolter sectionner tailler trancher éjecter éradiquer bannir cloisonner déblayer déboîter décharger décocher défourailler déplacer déposer omettre repousser sélectionner trier écarter éloigner [\[+\]](#)
- Roles : Agent [+int\_control | +organization], Theme, Source [+location]

NP V NP PP.source <span style="float:right">✕</span>	
Exemple	Luc a enlevé les dossiers du bureau.
Syntaxe	Agent V Theme {de} Source
Sémantique	cause(Agent, E) location(start(E), Theme, Source) not(location(end(E), Theme, Source))

Frames supprimées :

- NP V NP (Luc a enlevé les dossiers.)

FIGURE 1 – Interface web pour analyser et éditer VerbNet. Chaque frame peut être entièrement modifiée en cliquant dessus et la structure peut être réorganisée. Les traductions en violet appartiennent à l’intersection de  $C_{lvf}$  et  $C_{lg}$  ; les traductions rouges (respectivement vertes) appartiennent uniquement à  $C_{lvf}$  (respectivement  $C_{lg}$ ).

Cet outil a d’abord été rempli automatiquement avec les frames VerbNet et les traductions identifiées lors de la première étape, qui sont mises à jour dès que l’on change la correspondance avec LVF ou LG. Il permet d’éditer, ajouter et supprimer les frames et les classes. À l’aide de cet outil, le travail de la deuxième étape est parfois très facile. Par exemple, la classe [coloring-24](#) (qui n’a pas de sous-classe) a des équivalents directs en français : il suffit simplement d’entrer les exemples français avec les prépositions correctes, e.g. *with* doit être remplacé par *de*.

### 4.2.2 Principes sur les frames

Toutes les frames impliquant une alternance conative, bénéfactive ou dative peut être systématiquement supprimée étant donné que ces alternances n’existent pas en Français.

De plus, nous avons laissé de côté pour l’instant toutes les frames qui correspondent à des sous-structures, c’est-à-dire à des frames avec de compléments manquants, par exemple la sous-structure *NP V* de 25.1 illustrée par *Smith was inscribing = Smith gravait*. Le codage des sous-structures est assez dépendant du genre du corpus et donc demande des études de corpus, qui devront être faites ultérieurement.

### 4.2.3 Analyse au cas par cas

La seconde étape peut demander un travail délicat pour au moins deux raisons. La première est qu’il existe des différences sémantiques qui sont prises en compte dans VerbNet mais pas dans le LVF ni dans le LG. Par exemple, parmi les verbes de

Sending and Carrying (11), les verbes des classes 11.3, 11.4 et 11.5 décrivent un mouvement où l'Agent accompagne le Thème dans son changement de lieu, voir *Pamela drove packages to NY* où Pamela est allée à NY avec les paquets, tandis que les verbes de 11.1 et 11.2 décrivent un mouvement où seul le Thème change de lieu, voir *Pamela sent packages to NY* où seuls les paquets sont allés à NY. Dans les ressources françaises, il existe des classes de verbes pour un changement de lieu du Thème causé par un Agent, mais l'éventuel déplacement de l'Agent n'est pas codé. Face à cette difficulté, deux solutions sont possibles : soit faire le codage du déplacement de l'Agent soit ignorer cette différence sémantique. Nous sommes plutôt pour la seconde solution, ce qui nous amène à adopter dans Verb $\exists$ net une hiérarchie différente de celle de VerbNet pour les verbes concernés.

Faisons à ce propos une remarque sur la complémentarité du LVF et du LG. Dans le cas précédent, ni le LVF ni le LG ne codent une certaine propriété sémantique. Mais il arrive qu'une propriété sémantique soit codée dans le LVF et pas dans le LG, ce que nous pouvons illustrer par les verbes de Combining and Attaching (22). Les classes mix-22.1 et *almagamate-22.2* correspondent toutes les deux à 36S dans le LG avec le frame *Agent V Patient avec/et Co-Patient* mais différent par l'aspect résultatif, ce qui est marqué par l'adverbe *ensemble* qui est possible en mix-22.1 (*Luc a mélangé le sucre et l'eau ensemble*) et impossible en *almagamate-22.2* (*\*Luc a alterné le rouge et le noir ensemble*). Cette propriété sémantique n'est pas codée en 36S. Par contre, il semble que les verbes de 22.1 appartiennent à la classe U3.b du LVF et ceux de 22.2 à T4e.3<sup>2</sup>. Le LVF permet donc de différencier des verbes selon une propriété sémantique non codée dans le LG. À rebours, les nombreuses propriétés codées dans les colonnes du LG permettent facilement d'identifier des sous-classes. Ainsi, la classe cut-21.1 est associée à 38PL ou 32CL (*Agent V Patient en Result<+plural> = Luc a débité le veau en morceaux*), la sous-classe cut-21.1-1 à 38PL[+N0 lui V N1pc W] ou 32CL[+N0 lui V N1pc W], c'est-à-dire aux verbes de 38PL ou 32CL qui ont un '+' dans la colonne intitulée *NO lui V N1pc W* (*Luc a coupé le veau en morceaux, Luc lui a coupé le doigt*).

La seconde étape se heurte aussi à des difficultés qui viennent de différences basiques entre le français et l'anglais. Nous laissons de côté les problèmes de traduction archi-connus<sup>3</sup> pour nous concentrer sur des différences plus subtiles. Ainsi dans VerbNet, les verbes de Change of Possession (13) sont organisés en 10 classes 13.i avec  $1 \leq i \leq 10$ . Une telle hiérarchie ne peut pas être gardée pour le français pour les raisons suivantes :

- L'absence d'alternances dative et bénéfactive en français fait que la différence entre les classes 13.1 et 13.2 ne peut pas être gardée.
- La différence sémantique entre les classes 13.1 and 13.3 (à savoir HAS-POSSESSION versus FUTURE-POSSESSION) est peut-être trop subtile (même pour l'anglais) et peut être ignorée.
- La préposition *with* dans le frame *Agent V Recipient with Theme* de 13.4-1 et 13.4-2 correspond en français à *en et/ou de* selon le verbe (e.g. *Luc livre Max en/\*de lait, Luc équipe Max en/de téléviseurs, Luc dote Max \*en/de téléviseurs*), ce qui demande une réorganisation en sous-classes.

Au total, entrer dans le détail des frames syntaxico-sémantiques est un travail parfois laborieux qui peut amener à adopter pour Verb $\exists$ net une hiérarchie différente de celle de l'anglais, même si nous essayons d'éviter des écarts trop importants. Ajoutons qu'après avoir établi les (sous-)classes en fonction des frames, il faut revoir la répartition des verbes dans les sous-classes.

## 5 Conclusion

Nous avons présenté une méthode pour adapter la ressource syntaxique et sémantique VerbNet vers le français en se servant de deux ressources lexicales existantes, le LVF et le LG. Le travail est en cours et Verb $\exists$ net sera mis librement à la disposition de la communauté avec l'outil associé permettant une édition collaborative.

La structure de Verb $\exists$ net ne suit pas exactement celle de VerbNet mais nous documentons les différences pour qu'elles soient explicites et connues. À terme, des correspondances pourront être établies entre Verb $\exists$ net et d'autres ressources françaises comme WOLF (Sagot & Fišer, 2012) et le Framenet du français développé dans le cadre du projet ASFALDA<sup>4</sup>.

2. Les auteurs linguistes de cet article sont spécialistes du LG mais pas du LVF et donc ne sauraient affirmer leurs jugements sur le LVF.

3. Par exemple, la traduction des verbes de mouvement : *John swam across the river* → *Jean a traversé la rivière à la nage* (lit. John crossed the river with a swim).

4. Le travail présenté ici a été en partie financé par le projet ANR ASFALDA ANR-12-CORD-0023.

## Références

- BAZRAFSHAN M. & GILDEA D. (2013). Semantic Roles for String to Tree Machine Translation. In *ACL 2013*.
- BOONS J. P., GUILLET A. & LECLÈRE C. (1976). *La structure des phrases simples en français : constructions intran-sitives*.
- CHANG A. X., SAVVA M. & MANNING C. D. (2014). Semantic parsing for text to 3d scene generation. In *ACL 2014 Workshop on Semantic Parsing*.
- DUBOIS J. & DUBOIS-CHARLIER F. (1997). *Les verbes français*. Larousse.
- EXNER P. & NUGUES P. (2011). Using semantic role labeling to extract events from Wikipedia. In *DeRiVE 2011*.
- FALK I., GARDENT C. & LAMIREL J.-C. (2012). Classifying French Verbs Using French and English Lexical Resources. In *ACL 2012*.
- FERRER E. E. (2004). Towards a Semantic Classification of Spanish Verbs Based on Subcategorisation Information. In *ACL 2004 : Student Research Workshop*, Barcelona, Spain.
- GALE W. A., CHURCH K. W. & YAROWSKY D. (1992). Using bilingual materials to develop word sense disambiguation methods. In *4th International Conference on Theoretical and Methodological Issues in Machine Translation*, p. 101–112.
- GROSS M. (1975). *Méthodes en syntaxe. Régime des constructions complétives*. Hermann.
- IM WALDE S. S. (2006). Experiments on the automatic induction of German semantic verb classes. *Computational Linguistics*, **32**(2), 159–194.
- JENTSON I. (2014). VerbNet Workbench. In *GWC 2014*.
- KIPPER-SCHULER K. (2005). *VerbNet : A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania.
- LEVIN B. (1993). *English verb classes and alternations : a preliminary investigation*. University Of Chicago Press.
- MALMAUD J., WAGNER E. J., CHANG N. & MURPHY K. (2014). Cooking with semantics. In *ACL 2014 Workshop on Semantic Parsing*.
- MERLO P., STEVENSON S., TSANG V. & ALLARIA G. (2002). A Multilingual Paradigm for Automatic Verb Classification. In *Proceedings of 40th Annual Meeting of the Association for Computational Linguistics*, p. 207–214, Philadelphia, Pennsylvania, USA : Association for Computational Linguistics.
- MESSIANT C., GÁBOR K., POIBEAU T. *et al.* (2010). Acquisition de connaissances lexicales à partir de corpus : la sous-catégorisation verbale en français. *Traitement automatique des langues*, **51**(1), 65–96.
- OSMAN A. H., SALIM N., BINWAHLAN M. S., ALTEEB R. & ABUOBIEDA A. (2012). An improved plagiarism detection scheme based on semantic role labeling. *Applied Soft Computing*, **12**(5), 1493–1502.
- PRADET Q., DE CHALENDAR G. & PUJOL G. (2013). Revisiting knowledge-based Semantic Role Labeling. In *LTC'13*.
- SAGOT B. & FIŠER D. (2012). Automatic Extension of WOLF. In *GWC 2012*.
- SAINT-DIZIER P. (1996). Constructing Verb Semantic Classes for French : Methods and Evaluation. In *COLING 1996*.
- SCARTON C. & ALUISIO S. (2012). Towards a cross-linguistic VerbNet-style lexicon for Brazilian Portuguese. In *Workshop on Creating Cross-language Resources for Disconnected Languages and Styles Workshop Programme*, p. 11.
- SHEN D. & LAPATA M. (2007). Using Semantic Roles to Improve Question Answering. In *EMNLP-CoNLL 2007*.
- SUN L., KORHONEN A., POIBEAU T. & MESSIANT C. (2010). Investigating the cross-linguistic potential of VerbNet : style classification. In *COLING 2010*.
- SURDEANU M., HARABAGIU S., WILLIAMS J. & AARSETH P. (2003). Using predicate-argument structures for information extraction. In *Annual Meeting of the ACL 2003*, p. 8–15.
- SUZUKI Y. & FUKUMOTO F. (2009). Classifying Japanese Polysemous Verbs based on Fuzzy C-means Clustering. In *Proceedings of the 2009 Workshop on Graph-based Methods for Natural Language Processing (TextGraphs-4)*, p. 32–40, Suntec, Singapore : Association for Computational Linguistics.
- SWIER R. & STEVENSON S. (2005). Exploiting a Verb Lexicon in Automatic Semantic Role Labelling. In *HLT-EMNLP 2005*.
- XIE B., PASSONNEAU R. J., WU L. & CREAMER G. G. (2013). Semantic Frames to Predict Stock Price Movement. In *ACL 2013*.