CAIRN.INFO
MATIÈRES À RÉFLEXION

UNE VERSION ÉLECTRONIQUE DU LVF COMPARÉE AVEC D'AUTRES RESSOURCES LEXICALES

Fadila Hadouche, Guy Lapalme

Armand Colin | « Langages »

2010/3 n° 179-180 | pages 193 à 220

ISSN 0458-726X
ISBN 9782200926403
DOI 10.3917/lang.179.0193

Article disponible en ligne à l'adresse :

https://www.cairn.info/revue-langages-2010-3-page-193.htm

Distribution électronique Cairn.info pour Armand Colin.
© Armand Colin. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

Fadila Hadouche
RALI-DIRO – Université de Montréal

Guy Lapalme
RALI-DIRO – Université de Montréal

Une version électronique du *LVF* comparée avec d'autres ressources lexicales

1. INTRODUCTION

La base de données numérique sur les verbes français réalisée par de J. Dubois et F. Dubois-Charlier, et dont leur ouvrage *Les Verbes français* (désormais *LVF*) est une présentation [1], a été conçue dans une optique de description linguistique sans se soucier d'un codage qui en faciliterait l'utilisation par des programmes de traitement automatique de la langue (TAL). Les informations ont été collectées dans une base de données numérique pour en faciliter la mise à jour, le tri et la consultation. Cette base de données a été confiée à un industriel qui ne l'a ni utilisée ni valorisée ainsi qu'à un éditeur qui a refusé de la publier comme dictionnaire et qui ne l'a diffusée (photocopiée en fait) qu'aux gens qui en faisaient explicitement la demande, si jamais ils en connaissaient l'existence. Malheureusement, ces deux entreprises n'en détenaient pas moins les droits d'utilisation, et une fois qu'il était devenu clair qu'elles ne feraient rien d'autre que d'en bloquer toute exploitation, il a fallu un certain temps afin de les amener à restituer aux auteurs le droit d'utiliser, ou de laisser utiliser par des chercheurs, leur œuvre comme ils le souhaitaient.

Ces circonstances, qui ont différé l'exploitation de l'ouvrage et même seulement sa prise de connaissance par les linguistes, ont retardé des recherches concernant la langue française, car une des raisons expliquant la masse de travail effectué sur l'anglais est la disponibilité de ressources lexicales libres dans cette langue. La plus connue d'entre elles est certainement WordNet *(WN)*, mais

1. Sauf indication contraire, *LVF* désignera ici indifféremment la base de données numérique et l'ouvrage *Les Verbes français*.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

aussi plusieurs autres dont VerbNet *(VN)* et FrameNet *(FN)* pour l'anglais. Le libre accès à ces bases lexicales facilite le développement d'applications basées sur des ressources communes, stables et fiables, ainsi que les collaborations entre les chercheurs qui peuvent les partager. Malheureusement, en français, les ressources de ce type ont été développées majoritairement par des entreprises et des laboratoires privés ou publics qui les gardent jalousement. D'autres ressources, pourtant développées par des organismes publics ou dans le cadre de projets de coopération européenne, sont contrôlées par des organismes qui les commercialisent ou encore sont diffusées dans des formats particuliers ou incompatibles avec les standards d'échange les plus usuels.

J. Dubois et F. Dubois-Charlier, aidés de rédacteurs de multiples dictionnaires et de chercheurs issus d'autres équipes, notamment le LADL, ont produit le *LVF*. Le résultat, trop novateur par rapport aux dictionnaires habituels, n'a pas paru « vendable » aux maisons d'édition classiques, mais a aussi sans doute surpris l'entreprise acquéreuse, qui n'a pas trouvé le moyen de l'exploiter. Lorsqu'enfin il fut possible de rendre ce travail disponible à la communauté des chercheurs, D. Le Pesant en a créé une version sous format Excel qui facilite son exploitation manuelle à l'aide de recherches et de tris selon divers critères. Toutefois, ce mode d'accès n'est pas très commode pour des applications informatiques, d'autant plus que la version Excel utilise un codage assez compact et subtil pour certains types d'information. G. Lapalme a alors proposé une organisation de ces mêmes informations sous la forme d'un fichier XML qui explicite certains codages et facilite le traitement d'un sous-ensemble d'entrées. Le standard XML permet des modes d'accès efficaces par des extractions sous différentes formes et des transformations automatiques en pages web pour une consultation avec un moteur de recherche.

Nous présenterons l'organisation de la base et les modalités d'interrogation disponibles sur le site web de consultation. Nous expliquerons brièvement les étapes de transformation du tableur initial vers le format XML et les pages web en faisant ressortir les éléments les plus prometteurs pour les applications potentielles. Nous donnerons quelques pistes d'utilisation des informations du *LVF* pour des applications de TAL.

Afin de profiter de l'expérience acquise pour le traitement de l'anglais, nous avons cru intéressant de comparer le *LVF* avec d'autres ressources 'équivalentes' : DicoValence en français et FrameNet, VerbNet et WordNet en anglais. Nous sommes conscients que ces ressources traitent deux langues différentes et/ou utilisent des théories linguistiques distinctes, mais elles partagent un même but : décrire la sémantique des unités prédicatives.

Nous ne nous intéressons ici qu'aux verbes car *LVF* ne contient que des verbes. *FN* et *WN* décrivent aussi les noms et les adjectifs qui font l'objet d'autres dictionnaires chez J. Dubois et F. Dubois-Charlier sur lesquels il sera intéressant de travailler un jour, si jamais ils sont rendus disponibles. *FN* et *WN* regroupent les

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

verbes dans des ensembles selon le champ sémantique, déterminé par des relations paradigmatiques ou syntagmatiques. Elles reposent sur des constructions sémantiques conceptuelles (*frame* pour *FN* et *synset* pour *WN*).

VerbNet est basé sur le principe, défendu par B. Levin (1993), que la valence syntaxique et le sens sont liés : les verbes de même valence partagent une composante sémantique. J. Dubois et F. Dubois-Charlier proposent une approche analogue et montrent que la classification des verbes français repose sur une adéquation entre les constructions syntaxiques (et les paradigmes lexicaux qui les actualisent) et l'interprétation sémantique. En termes de ressources anglaises, nous comparons d'abord *LVF* avec *VN*, tous deux basés sur les comportements syntaxiques des verbes ; ensuite, *LVF* avec *FN* et *WN*, qui eux s'appuient plutôt sur des comportements sémantiques. En ce qui concerne le français, la comparaison portera sur *LVF* et Dicovalence, une ressource lexicale exploitée dans des applications en traitement automatique de la langue française comme, par exemple, l'étiquetage sémantique.

Les ressources VerbNet, FrameNet et WordNet sont des ressources essentielles pour le traitement automatique des langues, à savoir la langue anglaise. Cette dernière a tiré profit de ces ressources dans le développement de la sémantique et en réalisant des défis dans différentes applications telles que l'extraction d'informations, la recherche d'informations, le résumé automatique, la compréhension de texte, la désambiguïsation, etc. en tenant compte du sens des unités lexicales.

Nous comparerons *LVF* avec ces ressources non pas pour en identifier les points faibles et les points forts ou pour en tirer la meilleure théorie linguistique. Nous cherchons plutôt à les apparier et à évaluer si *LVF* peut jouer, pour le français, le même rôle que *WN*, *VN* ou *FN* pour l'anglais dans les applications informatiques. En particulier, nous cherchons à déterminer jusqu'à quel point les informations syntaxiques et sémantiques du *LVF* équivalent et sont parfois même supérieures en quantité ou qualité à celles des autres ressources.

2. DESCRIPTION DES RESSOURCES : *LVF*

LVF structure les verbes en classes sémantiques définies par la syntaxe et peut donc être considéré comme un thésaurus de classes syntactico-sémantiques qui forme le cœur de l'ouvrage de J. Dubois et F. Dubois-Charlier (1997). Les principes de la classification reposent sur l'adéquation entre les schèmes syntaxiques et la sélection distributionnelle dans la construction, et l'interprétation sémantique. Les arguments syntaxiques ou distributionnels qui sont la base du classement sémantique des verbes permettent d'associer des sens aux différents emplois d'un verbe et de les distinguer (François, Le Pesant & Leeman 2007). À une forme verbale dans *LVF*, sont associées une ou plusieurs phrases simples qui illustrent le sens et la construction syntaxique. La plupart du temps, une forme verbale a plusieurs emplois et donne donc lieu à

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

plusieurs entrées représentées chacune par un numéro (01, 02, 03, etc.). Par exemple, le verbe *chercher* présente 10 entrées ou emplois différents (notés **chercher 01**, ... **chercher 10**) correspondant à des schèmes syntaxiques différents (notés T1901, N1d, N1j, T1300, etc.), qui indiquent la nature du verbe (transitif direct ou indirect, intransitif, pronominal), le type du constituant sujet et constituant objet (humain, animal ou chose, complétive) et aussi la nature des compléments (locatif, prépositionnel, instrumental, à modalité, etc.). Cette notation correspond aux types et à la nature du sujet et des compléments. Les lettres correspondent aux types transitif, intransitif etc. : A (intransitif), N (transitif indirect), T (transitif direct), P (pronominale). Le sujet correspond au premier caractère après A, N, T, P ; l'objet correspond au deuxième caractère après T et P ; complément prépositionnel et circonstant sont représentés par le deuxième caractère pour N et A et les troisième et quatrième caractères pour T et P. Ainsi, dans T1901, nous avons T : transitif, le premier caractère 1 : sujet humain, le deuxième caractère 9 : objet humain ou chose, le troisième et quatrième caractères 0 et 1 : complément prépositionnel et circonstant. Une entrée d'un verbe peut avoir un schème simple avec objet animal, humain ou non animé. Par exemple, l'entrée **chercher 01**, avec un objet direct humain ou non animé, correspond au schème T1901. Ce verbe peut aussi être employé avec une forme pronominale et avoir des schèmes composés [T1106, P7000]. Par exemple, l'entrée **chercher 09** peut apparaître avec un sujet et un objet humain, ou un complément à modalité qui correspond au schème T1106 ou pronominal P7000 (sujet pluriel humain). Tous les schèmes, simples ou composés, sont regroupés en classes et sous-classes : 248 sous-classes syntaxiques, 54 classes sémantico-syntaxiques et 14 classes génériques. Les emplois des verbes sont répertoriés et identifiés selon certains domaines : Sociologie (SOC), Psychologie (PSY), Mobilier (MOB), etc.

Des propriétés sémantiques sont également attribuées à chaque entrée par le biais d'un 'opérateur' (OP), sorte d'étiquette interprétative de l'emploi concerné. Par exemple, **acheter 01**, c'est 'obtenir quelque chose de quelqu'un pour soi-même' *(acheter des fruits à l'épicier)*, tandis que **acheter 02**, c'est 'obtenir quelque chose de quelqu'un pour quelqu'un d'autre' *(acheter un collier à sa femme chez le bijoutier)*. Le parallélisme entre 'opérateur' et 'schème' fait ressortir leur forte correspondance. Par exemple, le schème T1901 (sujet humain, objet humain ou chose, complément locatif) de **chercher 01** est interprété sémantiquement par l'opérateur 'lc.qp qc, qn qp' qui signifie 'mettre quelqu'un ou quelque chose quelque part'. *lc* est l'élément primaire de cet opérateur qui signifie la localisation. L'opérateur est l'élément de base de *LVF*. Il permet d'interpréter sémantiquement les schèmes syntaxiques et de définir les classes syntaxico-sémantiques des verbes.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

2.1. Consultation du *LVF* par un site web

Le site web du *LVF* [2] permet de consulter le contenu du *LVF* par classe (organisation choisie dans la version imprimée) en pointant à l'aide de la souris. Les figures 1 et 2 (*infra*) illustrent la navigation pour afficher le contenu de la classe D2c où se trouvent deux emplois du verbe *acheter*.

Il est également possible d'accéder à l'information relative à un verbe en consultant un index alphabétique et en tapant une forme ou une expression régulière sur la forme. La figure 3 montre le début de l'affichage résultant de la requête <acheter> et la figure 4 celui de la requête <a...ter> (*i.e.* les verbes de 7 lettres débutant par 'a' et terminant par 'ter'). L'information sur chaque entrée est alors explicitée sur plusieurs lignes sans utilisation de codes (qui permettent de concentrer l'information sur une seule ligne dans les autres modes de présentation). On dispose également d'un index par domaine et par schème de construction syntaxique qui ne sont pas illustrés ici.

Toutes ces pages ont été générées automatiquement à partir des informations présentes dans le tableur initial fourni par D. Le Pesant qui a été transformé sous un format XML validable avec un schéma Relax NG. Les titres et les noms des classes ont été extraits du document PDF correspondant à la version imprimée du document. La mise à jour de ce fichier initial permet de recréer rapidement cet ensemble de pages. Ces processus de mise en page et de validation ont d'ailleurs permis de détecter quelques erreurs de frappe dans le fichier initial.

Le site web permet de télécharger des versions électroniques des données brutes en format XML pour leur utilisation dans des applications informatiques. Il y a également les descriptions des classes et des opérateurs qui sont disponibles en format XML avec les fichiers de schéma de validation. Les fichiers qui servent à créer les pages web de consultation du site constituent autant d'exemples d'utilisation de ces données pour d'éventuels usagers.

2. Le site web du *LVF* est consultable à l'adresse suivante : http://rali.iro.umontreal.ca/Dubois/

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

14 classes génériques
54 classes sémantico-syntaxiques
244 sous-classes syntaxiques

C: communication (2039 entrées)
D: don, privation (953 entrées)
D1: donner qn à qn/qc, donner aide à qn/qc (157 entrées)
D2: donner qc à qn/qc ; obtenir qc de qn/qc (399 entrées)
D2a: fournir qc à qn (235 entrées)
D2b: donner appui à qc avec qc (22 entrées)
D2c: prendre, supprimer qc à qn/qc (91 entrées)
D2d: enlever qc de soi, enlever partie de qc à qn (23 entrées)
D2e: obtenir, acquérir qc de qn (28 entrées)
D3: figuré de D2 (397 entrées)
E: entrée, sortie (2444 entrées)
F: frapper, toucher (1727 entrées)
H: états physiques et comportements (2079 entrées)
L: locatif (1523 entrées)
M: mouvement sur place (1671 entrées)
N: munir, démunir (3175 entrées)
P: verbes psychologiques (2074 entrées)
R: réalisation, mise en état (2289 entrées)
S: saisir, serrer, posséder (1671 entrées)
T: transformation, changement (2345 entrées)
U: union, réunion (1407 entrées)
X: verbes auxiliaires (212 entrées)

Figure 1 : Sélection du contenu d'une classe générique et d'une classe sémantico-syntaxique dans *LVF* [a]

a. Chaque ligne peut être pliée ou dépliée par un simple clic. L'accès aux entrées d'une sous-classe syntaxique est obtenu en cliquant sur le lien qui en décrit le contenu. La figure 2 montre l'état de l'écran après avoir sélectionné la classe D2c.

Sous-classe syntaxique D2c : *prendre, supprimer qc à qn/qc* (91 entrées)

[Toutes les classes]

Verbe	Opération	Sens	Phrases
acenser 02	abda qc A qn loué	louer	Le fermier **a~** une terre au seigneur. Cette terre peut s'**a~**.
acheter 01	abda qc A qn pr soi	payer	On **a~** cher des fruits à un marchand. On s'**a~** une voiture.
acheter 02	abda qc A qn cadeau	payer à	On **a~** un cadeau au bijoutier pour sa femme.
affréter	abda transp A qn	louer navire pour fret	On **a~** un avion,un navire à l'affréteur.
arracher 02	abda qc A qn forcé	prendre,enlever	On **a~** sa cravate à P. On s'**a~** un cheveu blanc.
arracher 12	abda qc uns A autres	se disputer	On s'**a~** les billets à l'entrée. On s'**a~** P dans les soirées.
asphyxier 04	abda qc A qn p vol	voler,dérober	Le voleur **a~** son portefeuille à P.
attriquer	abda qc A qn	acheter,se procurer	On **a~** une montre à un voleur.
barboter 04	abda qc A qn p vol	voler,chiper	On **b~** son portefeuille à P.
calotter 03	abda qc A qn p vol	dérober,voler	On **c~** à P sa valoche et ses fringues.
carmer	abda qc A qn	payer,acheter	Le fourgue **c~** le fade à P.

Figure 2 : Début de l'affichage des 91 entrées de la sous-classe syntaxique D2c [a]

a. Début de l'affichage des 91 entrées de la sous-classe syntaxique D2c dans lequel on retrouve les emplois 01 et 02 de *acheter*. Les entrées soulignées permettent d'accéder directement à la description des entrées dans la liste alphabétique (voir figure 3).

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

Lettre A

http://rali.iro.umontreal.ca/Dubois/alphabetique/A.html#acheter

acheter	**01**	COM	D2c abda qc A qn pr soi	payer	On **a~** cher des fruits à un marchand. On s'**a~** une voiture.	T13a6, P3006
	02	ECN	D2c abda qc A qn cadeau	payer à	On **a~** un cadeau au bijoutier pour sa femme.	T13a6, P3006
	03	SOC	H2h ger qn p arg	corrompre,payer	On **a~** un fonctionnaire,son silence. On s'**a~** une complicité.	T1106
	04	SOC	D3c abda pr soi abs arg	payer	On **a~** cher sa liberté. Le bonheur s'**a~** à ce prix.	T1308, P3000

Figure 3 : Les quatre emplois du verbe *acheter* répertoriés dans le *LVF* [a]

a. Les quatre emplois du verbe *acheter* répertoriés dans le *LVF* obtenus en cliquant sur le lien **acheter 01** dans la figure 2 ou en tapant <acheter> dans le champ de recherche alphabétique. La deuxième colonne indique le numéro de l'emploi. La troisième comporte l'abréviation du domaine : ici COM (commerce), ECN (économie) et SOC (sociologie). La quatrième décrit de manière compacte le schème sémantique à l'aide d'une expression utilisant des opérateurs. La cinquième donne le sens avec des synonymes. La sixième illustre cet emploi avec des exemples courts et évocateurs. La dernière décrit les schèmes syntaxiques. Les mots soulignés dans cette figure sont des liens permettant d'accéder directement à la description des domaines, des classes ou des schèmes syntaxiques.

78 entrées (27 distinctes) correspondant à *a...ter*

ablater abouter abriter (5) accoter (4) acheter (4) adapter (3) adenter adopter (5) affûter (2) agioter ajouter (4) ajuster (6) alerter (3) alester alloter ameuter (4) amputer (3) annoter (2) anuiter aposter appâter (4) appéter argoter arrêter (15) assoter auditer avorter (2)

[Page d'accueil]

ablater

mot	pronominal:*true* = ablater
domaine	nom:*géologie* = GEL
classe	construction-syntaxique:*d*, generique:*munir*, *démunir*, semantico-syntaxique:*non-animé propre* = N3d
operation	(lc)dmu soi D qc
sens	s'éroder
phrases	**phrase:** Le relief s'a~ . **phrase:** La dissolution a~ le relief.
conjugaison	auxiliaire:*avoir (sauf si pronominal ou entrée en être)*, groupe:*1*, sous-groupe:*chanter* = 1aZ
construction	**scheme:** circonstant:*de*, sujet:*chose*, type:*pronominal* = P30b0 **scheme:** circonstant:*de*, objet:*chose*, sujet:*chose*, type:*transitif direct* = T33b0
derivation	code:--- -- *RA*-- -- **der-ion:** eur, ion
nom	-
lexique	4

Figure 4 : Résultat d'une requête comportant une expression régulière [a]

a. Résultat d'une requête comportant une expression régulière où un <.> correspond à un caractère quelconque. Il s'agit de chercher tous les verbes de 7 lettres débutant par 'a' et terminant par 'ter'. On a retrouvé 78 emplois de 27 verbes, dont *acheter*. En cliquant sur un verbe dans la liste au début de la page, on est amené directement à l'information sur ce verbe.

2.2. Détail d'une entrée du *LVF*

Afin de comparer les informations présentes dans le *LVF* et dans les autres ressources lexicales, nous détaillons les 4 entrées correspondant au lemme *acheter* affichées dans la figure 3. Selon la quatrième colonne, les deux premières sont répertoriées dans la classe D2C, la troisième dans la classe H2h et la dernière dans la classe D3c.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

2.2.1. Sous-classe D2

Selon la figure 1, cette classe sémantico-syntaxique D2 (399 entrées) a pour sens général 'donner qc à qn/qc' ; 'obtenir qc de qn/qc' ('donner quelque chose à quelqu'un/quelque chose' ; 'obtenir quelque chose de quelqu'un – quelque chose'). Les emplois de **acheter 01** et **acheter 02** sont rangés dans la sous-classe syntaxique D2c 'prendre, supprimer qc à qn/qc'. La dernière colonne de la figure 3 indique que les emplois de ces verbes partagent les mêmes schèmes syntaxiques T13a6 (transitif direct, sujet humain, objet chose, circonstant à modalité) et P3006 (pronominal, avec sujet chose et circonstant à modalité).

Le premier (**acheter 01**) a le sens de *payer*. Il consiste à 'payer pour soi-même' qui correspond au schème sémantique 'abda [3] qc A qn pr soi' ('enlever quelque chose à quelqu'un pour soi'). Quand au deuxième, il a le sens de *payer à*. Il consiste à 'payer à quelqu'un d'autre que soi', par exemple un cadeau, qui correspond au schème sémantique 'abda qc A qn cadeau'. Ces deux sens du verbe *acheter* correspondent à deux domaines de classification différents (troisième colonne de la figure 3) : Commerce (COM) pour **acheter 01** et Économie (ECN) pour **acheter 02**.

Ex. d'emplois de **acheter 01** : On achète cher des fruits à un marchand ; On s'achète une voiture
Ex. d'emploi de **acheter 02** : On achète un cadeau au bijoutier pour sa femme

2.2.2. Classe H

Acheter 03 relève de la classe sémantico-syntaxique H2 'avoir tel comportement, telle activité'. Il est rangé en H2h 'diriger quelqu'un dans tel but' qui rassemble les verbes de schème T1106 (verbes transitifs directs, à sujet humain, avec objet humain et circonstant de modalité).

Ce verbe porte le même sens que les verbes *corrompre* et *payer*. Il correspond au schème sémantique 'ger qn p arg' ('diriger quelqu'un par argent'). **Acheter 03** appartient au domaine de la sociologie.

Ex. d'emploi de **acheter 03** : On achète un fonctionnaire, son silence ; On s'achète une complicité

2.2.3. Sous-classe D3

D3 est une classe sémantico-syntaxique qui indique le sens de 'figuré de D2'. **Acheter 04** relève de la sous-classe syntaxique D3e. Il appartient aux schèmes syntaxiques T1308 (transitifs direct, sujet humain, objet chose et circonstant instrumental moyen) et P3000 (pronominal, sujet chose).

Acheter 04, appartenant au domaine de la sociologie, a le sens de 'payer pour soi mais avec un argent abstrait', soit le schème sémantique 'abda pr soi abs arg'.

Ex. d'emplois de **acheter 04** : On achète cher sa liberté ; Le bonheur s'achète à ce prix

3. 'abda' est un opérateur qui signifie *enlever à*, *obtenir de*.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

3. COMPARAISON DU *LVF* AVEC LES AUTRES RESSOURCES

3.1. Dicovalence

3.1.1. Description

Le dictionnaire de valences Dicovalence [4] (Van den Eynde & Mertens 2006) est un dictionnaire syntaxique de plus de 3 700 verbes français simples. Il décrit environ 8 000 entrées. Il répertorie les cadres de valence de ces verbes simples. Ces cadres décrivent une configuration de compléments d'un verbe (le nombre, la nature et la fonction de ses compléments, y compris le sujet). La description des cadres valenciels repose sur les principes de l'approche pronominale : chaque position actantielle – appelée paradigme – est décrite par le paradigme des pronoms tels que P0, P1, P2, P3, PL, etc. qui peuvent l'occuper (les cadres valenciels sont définis sur une base de critères distributionnels plutôt que sur une base sémantique). Ce dictionnaire est constitué d'entrées correspondant aux infinitifs associés à leurs différents cadres de valence. Un verbe peut avoir plusieurs cadres valenciels distincts. Une particularité de Dicovalence est le fait que les cadres valenciels reposent sur les paradigmes pronominaux et que les propriétés de construction, comme les reformulations passives, sont précisées.

Au fil des années, plusieurs modifications mineures ont été apportées au dictionnaire. En 2010, une deuxième version (Dicovalence 2) a été publiée par P. Mertens dans laquelle plusieurs éléments nouveaux ont été ajoutés (Mertens 2010) :

- le champ FRAME qui fournit le cadre de sous-catégorisation : sont indiquées la fonction syntaxique, les réalisations syntagmatiques (pronom, groupe nominal, complétive, infinitive, adverbe, etc.), les restrictions de sélection formulées par des traits sémantiques (humain, concret non humain, abstrait) ;
- le champ AUX qui indique les auxiliaires ;
- les sujets discontinus : par exemple, dans *il en arrive plusieurs ici chaque jour*, les éléments discontinus *il*, *en* et *plusieurs* sont analysés en un seul terme sujet (P0 sujet), trouvé par comparaison avec *ils arrivent*, *plusieurs arrivent*. Dans l'exemple *il en prend plusieurs*, les éléments *il*, *en* et *plusieurs* ne sont pas des sujets discontinus. Ils sont analysés en deux parties : *il* qui correspond à P0 (sujet impersonnel) et *en plusieurs* qui correspond à P1 (complément).

Chaque entrée de ce dictionnaire est décrite par des champs. On en distingue trois catégories (Van den Eynde & Mertens 2006) :

1. champs génériques qui décrivent l'identité de l'entrée ;
2. champs indiquant le schéma valenciel à l'aide de propriétés de la valence du verbe (ou actants du verbe). Celle-ci est composée d'un frame et d'un ensemble de paradigmes P0, P1, P2, etc. ;
3. champs additionnels, comme la reformulation passive, auxiliaire, etc.

4. Dicovalence est consultable en ligne : http://bach.arts.kuleuven.be/dicovalence

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

La figure 5 présente une entrée du verbe *acheter*.

VAL$ acheter: P0 P1 (P3)
VTYPE$ predicator simple
VERB$ ACHETER/acheter
NUM$ 1850
EG$ j'ai acheté ce terrain au voisin d'en face
TR_DU$ kopen, afkopen
TR_EN$ buy (from s.o.)
FRAME$ subj:pron|n:[hum], obj:pron|n:[nhum,?abs], ?objde:pron|n:[hum]
P0$ qui, je, nous, elle, il, ils, on, celui-ci, ceux-ci
P1$ que, la, le, les, en Q, ça, ceci, celui-ci, ceux-ci
P3$ 0, qui, lui, leur, se réfl., se réc., celui-ci, ceux-ci, l'un l'autre
RP$ passif être, se passif
AUX avoir

Figure 5 : L'emploi 1 du verbe *acheter* dans Dicovalence [a]

a. L'emploi 1 du verbe *acheter* dans Dicovalence est décrite par : **sa valence** (VAL) composée d'un lemme (acheter) suivi des positions de paradigmes pronominaux P0, P1 et P3 qui peut être absent ; **le type de verbe** (adjoint, predicator simple, etc.) – VTYPE ; **le verbe lui-même** (VERB) ; **un numéro** qui identifie cette entrée correspondant à un sens du verbe (NUM) ; **un exemple** d'emploi (EG) et **les traductions en néerlandais et en anglais** (TR_DU et TR_EN). Ensuite, le champ FRAME dans lequel on précise la fonction syntaxique (subj) ; la réalisation syntagmatique, par exemple pronom ou bien nom (pron|n) ; le caractère obligatoire ou facultatif (?) – dans cet exemple, le deuxième objet *au voisin d'en face* est facultatif (?objde) – et les restrictions de sélection : humain pour le sujet (hum), non humain ou abstrait pour le 1er objet (nhum, ?abs). L'ensemble des paradigmes pronominaux (P0, P1, P3) avec leurs réalisations possibles en pronoms. Dans cet exemple, la liste de pronoms correspondant à chacune des positions de la valence est donnée. Finalement, la catégorie additionnelle formée de la reformulation passive (RP) et auxiliaire (AUX).

Le verbe *acheter* comporte quatre entrées dans Dicovalence. Toutes ces entrées admettent deux paradigmes obligatoires – P0 (sujet) et P1 (objet direct) –, et la différence entre elles est le troisième paradigme qui peut être P3 (des formes non clitiques présentant la préposition de [Objde]) pour la première entrée, ou P2 (objet indirect ou formes non clitiques présentant la préposition à [objà]) pour la deuxième entrée, ou PP (paradigme prépositionnel [objp<pour>]) pour la troisième entrée, ou rien pour la quatrième entrée.

La correspondance entre ces entrées et celles de *LVF* est partielle :

– L'entrée de la figure 5 qui fait intervenir le paradigme P3 (objde) correspond à **acheter 01** de *LVF*. L'exemple *j'ai acheté ce terrain au voisin d'en face* est classé dans cette entrée. *Voisin* qui correspond au P3 avec le clitique *de* ; *terrain du voisin* est considéré ici le vendeur du terrain. Dans cet exemple, il pourrait y avoir une ambiguïté : *voisin* pouvant être interprété comme le vendeur ou comme le bénéficiaire du cadeau, car la pronominalisation *je lui achète un terrain* est la même. Quoique, dans cet exemple précis, on utiliserait probablement *pour le voisin* s'il était bénéficiaire.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

– La deuxième entrée fait intervenir le paradigme P2 (objà), ex. *je me suis acheté une nouvelle voiture* (ici P2 est égal 'à moi-même'). Dans cette entrée, le paradigme P2 est considéré comme bénéficiaire du cadeau. Dans Dicovalence, un emploi du verbe avec P2 égal 'à soi-même' ou avec P2 égal 'à quelqu'un' est considéré dans une même entrée. Cette deuxième entrée avec P2 signifie 'à soi-même' correspond à l'entrée **acheter 01** dans *LVF*. Ces deux premières entrées du Dicovalence correspondent à la même entrée **acheter 01** de *LVF*. La raison étant que, dans *LVF*, en plus du comportement syntaxique (sans deuxième complément, l'interprétation est que 'l'achat est pour l'acheteur lui-même' : *j'ai acheté un terrain = je me suis acheté un terrain*), le comportement sémantique désigné par l'opérateur est aussi pris en compte, de même que la commutation distinctive : *payer* dans le premier cas, *payer à* dans le deuxième. Ces deux entrées de Dicovalence sont désignées dans *LVF* par un seul opérateur qui est 'abda qc à qn pr soi' qui signifie 'enlever quelque chose à qqn pour soi en le payant'.
– La troisième entrée correspond à **acheter 02** dans *LVF*. Ex. *j'ai acheté une nouvelle voiture pour Paul*, ce qui signifie 'le fait d'acheter quelque chose pour quelqu'un d'autre'. Dans *LVF*, ce cas est considéré comme un cadeau. Par contre, dans Dicovalence, la raison de cette entrée est la présence du paradigme prépositionnel PP<pour> (objp<pour>) qui la distingue des deux autres entrées.
– La quatrième entrée correspond au fait d''acheter une personne, dans le sens de la corruption'. Dans ce cas, le cadre valenciel est composé uniquement du paradigme P0 (sujet) et du paradigme P1 (objet direct). Cette entrée correspond à **acheter 03** dans *LVF*. L'emploi de **acheter 04** dans *LVF*, qui signifie 'le fait d'acquérir pour soi un argument abstrait' – ex. *on achète cher sa liberté* –, n'a pas une entrée spécifique dans Dicovalence. Mais, ces emplois peuvent être classés dans la première ou la deuxième entrée parce qu'il y a la possibilité d'avoir dans ces entrées un objet de type abstrait. Notons cependant que, en dehors du paradigme lexical de l'objet, les propriétés des deux constructions diffèrent : ainsi, on peut acheter quelqu'un, le silence ou la complicité de quelqu'un, mais non (dans le même sens) le bonheur ou la liberté de quelqu'un ; ou encore, on peut dire d'un fonctionnaire, mais non du bonheur, qu'il se laisse acheter (son silence) ou qu'il se fait acheter (sa complicité), etc.

L'insuffisance soulevée au dernier point montre que Dicovalence ne présente que la construction syntaxique et les restrictions de sélection qui, à eux seuls, ne permettent pas de séparer le sens 'acheter un fonctionnaire' de celui 'acheter le bonheur'. Par contre, dans le *LVF*, ces deux exemples sont considérés comme des emplois différents et classés dans des classes syntaxico-sémantiques différentes à l'aide des propriétés sémantiques données par l'opérateur.

3.1.2. Comparaison avec le LVF

Dans Dicovalence, le comportement syntaxique des verbes est décrit, mais non leur comportement sémantique. Seules les restrictions de sélection telles

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

que humain, animé, concret, abstrait, etc. sont précisées, contrairement au *LVF* où les comportements syntaxique et sémantique sont tous les deux décrits. Dans *LVF*, les opérateurs tels que *abda* 'enlever', *loq* 'parler', etc. avec leurs compléments tels que *A qn* 'à quelqu'un', *qc* 'quelque chose', *pr soi* 'pour soi', etc. interprètent sémantiquement les schèmes syntaxiques. Dans Dicovalence, *Frame* décrit les restrictions de sélection pour les paradigmes définis pour une entrée. Dans la description de l'entrée *acheter* de la figure 5, *Frame* définit 'humain' pour le paradigme P0 (sujet), 'non humain' pour le paradigme P1 (objet) et 'humain' pour le troisième paradigme P3 (objde). L'interprétation sémantique dans *LVF* est plus riche et significative qu'une simple déclaration de restriction de sélection dans Dicovalence. Dans *LVF*, la rubrique « Sens » contient, par ailleurs, les parasynonymes ou les formes abrégées de définition qui éclairent éventuellement le sens de l'entrée *LVF*.

Dicovalence repose sur une approche pronominale. La valence est caractérisée par les paradigmes pronominaux qu'accepte le verbe. Chaque position de valence appelée « paradigme » est associée à un paradigme de pronoms. Dans d'autres approches syntaxiques, on utilise généralement la notion de « syntagme » (syntagme nominal, syntagme verbal, syntagme prépositionnel, etc.) plutôt que celle de paradigme pronominal. L'avantage de cette approche est que le nombre de pronoms est restreint, ce qui permet de vérifier facilement leurs combinaisons avec le prédicat. L'inconvénient de cette approche pronominale est que le pronom ne désambiguïse pas la position syntaxique, comme le montre le cas du verbe *acheter* dans *J'achète un champ à mon voisin, Je lui achète un champ*.

Dans *LVF*, cette ambiguïté est levée par la prise en considération non pas de la pronominalisation, mais de la construction syntaxique et des opérations distributionnelles pour rendre compte que l'énoncé cache deux interprétations (la suppression : *j'achète un champ* fait nécessairement du voisin le vendeur dans *J'achète un champ à mon voisin* ; la commutation : *j'achète un champ à/pour mon voisin* fait nécessairement du voisin le destinataire du cadeau). L'approche de *LVF* est plutôt distributionnelle.

Dans *LVF*, les verbes partageant les mêmes comportements syntaxiques et sémantiques sont regroupés dans des classes sémantiques. Dans Dicovalence, il n'y a pas de regroupement de verbes dans des classes syntaxiques ni sémantiques. La seule spécification de variétés de classes de verbes est faite par le champ VTYPE, dans la description d'une entrée d'un verbe, indiquant le type du verbe : verbes pleins, les verbes composés, les verbes adjoints, etc.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

3.2. VerbNet

3.2.1. Description

Le lexique VerbNet [5] *(A Class-Based Verb Lexicon)* est une hiérarchie de classes de verbes anglais (Kipper, Dang & Palmer 2000). *VN* englobe des informations à la fois syntaxiques et sémantiques, compte environ 5 200 verbes distribués dans 237 classes et 194 sous-classes en utilisant les classes de verbes de B. Levin (Levin 1993). Ce lexique regroupe des verbes anglais dans des classes en se basant sur les liens de leur comportement syntaxique et sémantique. Les verbes de la même classe partagent les mêmes comportements syntaxiques et sémantiques, mais ne constituent pas nécessairement des synonymes. Une classe de verbes dans *VN* regroupe plusieurs verbes, identifie des rôles thématiques avec des restrictions de sélection et décrit des structures syntaxiques et des prédicats sémantiques (*Frame* dans la figure 6) (Kipper-Schuler 2005). Une classe est donc composée :

- de verbes membres décrits par un lien vers le sens correspondant dans FrameNet (voir section 3.3) et un lien vers le sens correspondant dans WordNet (voir section 3.4) ;
- de rôles thématiques qui sont plus ou moins génériques tels que *Agent*, *Theme*, *Source*, *Beneficiary*, etc. et des restrictions de sélection telles que *animate*, *organization*, *concrete*, *region*, *location*, etc. ;
- des structures syntaxiques et des prédicats sémantiques accompagnés d'un ou de plusieurs exemples ;
- d'éventuelles sous-classes, comportant d'autres verbes que leur classe parente, permettent de raffiner cette dernière. Elles ajoutent des informations comme, par exemple, de nouveaux rôles thématiques et restriction de sélection, et créent de nouveaux *frames* syntaxiques ou sémantiques. Elles héritent des informations de leur classe parente.

La syntaxe dans VerbNet est indiquée par des patrons syntaxiques accompagnés de leurs rôles (les arguments du verbe sont associés à des rôles thématiques). Par exemple, le premier emploi de *buy* de la figure 6 a pour patron syntaxique [NP V NP] et la rubrique 'syntax' donne '*Agent* V *Theme*' qui indique que le groupe nominal (NP=Noun Phrase) qui précède le verbe (V) est l'*Agent* et que le groupe nominal qui suit le verbe est le *Theme*. Ces rôles sont définis d'une manière sémantique et sont associés à chaque argument du verbe.

3.2.2. Exemple la classe de verbe 'buy' de la classe get-13.5.1

Le verbe *buy* a 8 emplois différents selon les constructions syntaxiques et sémantiques (seules les 3 premières sont affichées dans la figure 6) illustrés avec des exemples simples.

- Agent V Theme *(Carmen bought a dress)*

5. VerbNet est consultable en ligne : http://verbs.colorado.edu/∼mpalmer/projects/verbnet.html

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

– Agent V Theme from Source *(Carmen bought a dress from Diana)*
– Agent V Theme for Beneficiary *(Carmen bought a dress for Mary)*
– Agent V Beneficiary Theme *(Carmen bought Mary a dress)*
– Agent V Theme for capital (avoir ou biens) *(Carmen bought a dress for 50$)*
– Capital V Theme *(50$ won't even buy a dress)*
– Agent V Theme from Source at for Asset *(FMC has bought 565,000 shares from Nortek Inc at 23.50$ a share)*
– Agent V Theme {from for on} Source *(John earned 10$ million on proceeds from the sale)*

Les rôles thématiques sont accompagnés par des restrictions de sélection pour chaque description de classes de verbes. Par exemple, l'agent de *buy* doit avoir un trait [animé] ou [organisation], la source doit être de type [concret] et le bénéficiaire doit avoir un trait [animé] ou [organisation].

Go To Comments | **get-13.5.1** | Post Comment
Members: 25, Frames: 7

Class Hierarchy
get-13.5.1*
get-13.5.1-1

Members (KEY)

ATTAIN	CONSERVE (WN 4; G 1)	PICK (FN 1; WN 2; G 2)	SHOOT (FN 1; WN 2; G 1)
BOOK (WN 2; G 1)	FIND (WN 3, 7; G 1)	PLUCK (WN 1, 6; G 1)	SLAUGHTER (WN 1; G 1)
BUY (FN 1; WN 1, 3; G 1)	GATHER (WN 1, 6; G 1, 2)	PROCURE (FN 1; WN 1, 2)	VOTE (WN 5; G 5)
CALL (FN 1; WN 5, 23; G 1, 7)	HIRE (FN 1; WN 1, 2, 3; G 1)	PULL (FN 1; WN 2, 6, 17; G 1, 3)	WIN (WN 2; G 2)
CATCH (WN 4, 5, 8; G 1, 2)	LEASE (FN 1; WN 2, 4; G 1)	REACH (WN 8; G 3)	
CHARTER (FN 1; WN 3)	ORDER (FN 1; WN 2; G 2)	RENT (FN 1; WN 3, 4; G 2)	
CHOOSE (WN 1, 2; G 1)	PHONE (FN 1; WN 1; G 1)	RESERVE (WN 3, 4; G 2)	

Roles (REF)

- Agent [+animate | +organization]
- Theme
- Source [+concrete]
- Beneficiary [+animate | +organization]
- Asset [-location & -region]

Frames (REF KEY)

NP V NP
EXAMPLE "Carmen bought a dress."
SYNTAX Agent V Theme
SEMANTICS has_possession(start(E), ?Source, Theme) transfer(during(E), Theme) has_possession(end(E), Agent, Theme) cause(Agent, E)

NP V NP PP.source
EXAMPLE "Carmen bought a dress from Diana."
SYNTAX Agent V Theme {from} Source
SEMANTICS has_possession(start(E), Source, Theme) transfer(during(E), Theme) has_possession(end(E), Agent, Theme) cause(Agent, E)

NP V NP PP.beneficiary
EXAMPLE "Carmen bought a dress for Mary."
SYNTAX Agent V Theme {for} Beneficiary
SEMANTICS has_possession(start(E), ?Source, Theme) transfer(during(E), Theme) has_possession(end(E), Agent, Theme) cause(Agent, E) benefit(E, Beneficiary)

Figure 6 : Début de l'information à propos du verbe *buy* dans VerbNet [a]

a. *Buy* se trouve dans la classe ***get*** qui comprend d'autres verbes tels que *book*, *choose*, *pick*, *win*, etc.

Les rôles thématiques définis dans VerbNet peuvent aussi servir à décrire le prédicat sémantique, par exemple le premier exemple de *buy* indique que :

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

– *has-possession (start(E), ?source, Theme)* : au démarrage de l'événement, *Theme* sans source n'appartient pas à *Agent* ;
– *transfer (during(E), Theme)* : durée de transfert de *Theme* à *Agent* ;
– à la fin de l'événement, *has_Possession (END(E), Agent, Theme)* : *Theme* appartient à l'*Agent* ;
– donc *cause (Agent, E)* : l'*Agent* cause l'événement.

3.2.3. Comparaison avec LVF

Dans VerbNet, comme dans *LVF*, les verbes partageant les mêmes comportements syntaxiques sont regroupés en classes sémantiques cohérentes. Or, dans *VN*, les *frames* syntaxiques d'un verbe sont accompagnés de rôles thématiques définis de manière sémantique. Par exemple, au patron [NP V NP] du premier emploi de *buy*, est associé le rôle *Agent* à NP qui précède le verbe *buy* et *Theme* à NP qui suit ce verbe. Le prédicat sémantique définit le verbe en appliquant les fonctions sémantiques telles que *has_possession*, *duration*, *transfer*, *cause*, etc. sur les différents arguments du verbe représenté par leurs rôles sémantiques. Par exemple, le prédicat sémantique

has_possession (Start(E),?Source, Theme)
transfer (During(E), Theme)
has_possession (END (E), Agent,Theme)

définit le premier emploi de *buy*.

Le cas d''acheter un cadeau à quelqu'un' est représenté dans le troisième emploi de *buy* de la figure 6. Cet emploi est décrit par le patron syntaxique [NP V NP PP.Beneficiary], accompagné par des rôles sémantiques 'Agent V Theme {for} Beneficiary' et du prédicat sémantique

has_possession (Start(E),?Source, Theme)
transfer (During(E), Theme)
has_possession (END (E), Agent,Theme)
Cause (Agent, E) Benefit (E, Beneficiary).

Dans *LVF*, des informations syntaxiques et sémantiques définissent le verbe. Ce sont les opérateurs qui interprètent sémantiquement les schèmes syntaxiques du verbe et la rubrique « sens » clarifie la définition du verbe. Par exemple, l'emploi **acheter 01** est décrit par son schème syntaxique [T13a6, P3006], l'opérateur 'abda' et le sens de *payer*. Tandis que **acheter 02** – qui signifie 'l'achat d'un cadeau' – a pour schème syntaxique [T13a6, P3006], l'opérateur 'abda' et le sens de *payer à*. Nous constatons une certaine correspondance entre la description du sens du verbe dans *LVF* par le schème syntaxique, l'opérateur et le sens, et celle dans *VN* par les patrons syntaxiques, les rôles thématiques et le prédicat sémantique. *LVF* ne définit pas les rôles thématiques à associer aux différents arguments du verbe alors qu'ils sont indiqués dans *VN*. Ces informations sur les rôles thématiques sont importantes pour désambiguïser certains sens. Par exemple, si un argument objet du verbe ayant le rôle de *Theme* se trouve en position sujet dans la phrase, cet argument ne prend pas le rôle *Agent*, mais il gardera son rôle thématique même s'il change de position. Cela est important en traduction automatique pour les langues dont les positions des arguments sont

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

différentes, par exemple entre les langues de type « sujet verbe objet » (SVO) et « sujet objet verbe » (SOV) pour lesquelles il est primordial de détecter les rôles sémantiques.

Dans *VN*, une classe de verbes est constituée par un ensemble d'emplois de verbes qui partagent les mêmes composants (structures syntaxiques, rôles thématiques avec les restrictions de sélection et prédicats sémantiques). Ces regroupements des emplois de verbes en classes ressemblent à celui du *LVF* qui regroupe les emplois de verbes dans des classes sémantico-syntaxiques selon les schèmes syntaxiques et leurs interprétations sémantiques.

Dans *LVF*, chaque classe est constituée par les schèmes syntaxiques et les propriétés sémantiques représentées par un opérateur qui donne l'interprétation sémantique du verbe en utilisant des fonctions telles que *abda*, *lc*, *dat*, etc. Dans *VN*, le prédicat sémantique interprète sémantiquement un verbe en utilisant des fonctions telles que *start*, *has-possession*, *duration*, *transfer*, etc. Des restrictions de sélection (*animate*, *organization*, *concrete*, *location*, etc.) sont associées aux rôles thématiques (*Agent*, *Theme*, *Beneficiary*, etc.). Ces restrictions sont semblables aux traits du *LVF* (*humain*, *animal*, *chose*, *animé*, *non animé*, etc.) définis dans le schème ou aux traits (*abstrait* (abs), *concret*, *quelqu'un* (qn), *quelque chose* (qc), *quelque part* (qp), etc.) définis dans l'opérateur.

LVF et *VN* décrivent essentiellement les mêmes informations syntaxico-sémantiques associées aux verbes. Dans les deux ressources, les constructions syntaxiques sont interprétées sémantiquement. Dans *LVF*, les opérateurs et leurs compléments interprètent les schèmes syntaxiques ; dans *VN*, le prédicat sémantique et les rôles thématiques avec les restrictions de sélection interprètent sémantiquement les constructions syntaxiques. Cela diffère de Dicovalence où les constructions syntaxiques sont définies, mais non interprétées sémantiquement. Dans Dicovalence, les arguments gouvernés par un verbe et leurs fonctions syntaxiques sont détectés en utilisant les paradigmes pronominaux P0 (sujet), P1 (objet), P2 (objet à), P3 (objet de), PL (locatif), PT (temporel), etc., mais il n'y a aucune interprétation sémantique des verbes contrairement à *LVF*, où l'opérateur et le sens définissent le verbe. Par exemple, dans *LVF* **acheter 01** et **louer 02** partagent les mêmes schèmes (T13a6, P3006) ; ce sont plutôt les champs 'opérateur' et 'sens' qui permettent de les distinguer. C'est également le cas pour *VN* qui, à l'aide de son prédicat sémantique, interprète différemment *louer* et *acheter*. Dans Dicovalence, cette distinction est impossible au moyen des seuls paradigmes syntaxiques. Dicovalence a été utilisé dans plusieurs applications de TAL (pour l'étiquetage sémantique ou la catégorisation des verbes) malgré cette absence dans l'interprétation sémantique. *LVF*, qui comble cette lacune, pourrait donc jouer un rôle très intéressant dans les applications de TAL du français. *LVF* peut donc être considéré comme équivalent à VerbNet en termes de constructions syntaxiques et d'interprétation sémantique, et même supérieur au Dicovalence en termes d'interprétation sémantique.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

3.3. FrameNet

3.3.1. Description

FrameNet [6], un projet mené à Berkeley à l'initiative de C. Fillmore (Baker, Fillmore & Lowe 1998), est fondé sur la théorie *frame semantics* (Fillmore 1982) dans laquelle les *frames* sont considérés comme des descriptions linguistiques de scénarios conceptuels. L'auteur de cette théorie suppose que, pour comprendre le sens d'une unité lexicale, nous devons d'abord connaître tous les *Semantic Frames* (SF) – la structure conceptuelle à laquelle elle appartient (Fillmore 1982). Le modèle SF caractérise les propriétés sémantiques et syntaxiques des unités lexicales en les rattachant à des *frames*. Ce sont des représentations de situations impliquant des participants qui, dans ce cadre théorique, sont appelés *Frame Elements* (FE). Dans cette approche, les sens des unités lexicales ne sont pas liés directement unité à unité, mais seulement selon leur relation aux *Frames* (Fillmore & Atkins 1992). *FN* a pour objectif d'incorporer des informations syntaxiques et sémantiques pour chacun des sens d'une entrée lexicale à travers une annotation en rôles sémantiques d'exemples d'emplois des unités prédicatives. Chaque sens d'un mot est décrit par un *frame* qui comporte les actants sémantiques ou *Frame Elements* (*FE*). Ces FE comportent dans leur description les réalisations syntaxiques et les patrons syntaxiques où ils apparaissent (voir l'exemple de la figure 8). *FN* est une base de données lexicale pour l'anglais qui définit des *Semantic Frames* d'unités prédicatives telles que des verbes, noms et adjectifs.

On définit des *frames* auxquels sont liées les unités lexicales en utilisant les FE de cette unité lexicale. Chaque sens d'une unité polysémique appartient à des *Semantic Frames* différents. Les FE sont identifiés et définis selon qu'ils sont obligatoires *(core)* ou optionnels *(non-core)*. Par exemple, le verbe *buy* appartient au frame C*ommerce_buy*. Ce frame est défini par les FE obligatoires *Buyer* et *Goods*. D'autres participants sont considérés optionnels *(non-core)* : on retrouve *Manner*, *Means*, *Money*, etc. Dans *FN*, la description d'une unité lexicale est faite au moyen de la description des frames auxquels cette unité appartient.

Nous décrivons ici l'unité lexicale anglaise verbale *buy* qui a un sens apparenté à celui de l'unité lexicale française *acheter* traitée dans *LVF*. Cette unité lexicale, dans *FN*, appartient à un seul frame *Commerce_buy*. La figure 7 présente la caractérisation du frame *Commerce_buy* (l'unité lexicale *buy* sera définie ensuite).

Dans la définition donnée au haut de la figure 7, on constate la mise en évidence des rôles sémantiques des participants. Le FE *Buyer* indique celui qui réalise l'action C*ommerce_buy* et le FE *Goods* indique l'objet que le *Buyer* paie et qui entre en sa possession. Dans l'exemple donné dans la définition, on voit la correspondance des éléments avec les FE cités dans la définition du frame.

6. FrameNet est consultable à l'adresse suivante : http://framenet.icsi.berkeley.edu/

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

La réalisation du FE *Buyer* est *Abby*, celle du FE *Goods* est *car*, et *Robin* et *$5,000* sont respectivement les réalisations de FE *Seller* et *Money*. Ces types de FE sont décrits dans la partie des *Frame Elements*. Les FE obligatoires sont *Buyer* et *Goods*. *Money* et *Seller* sont optionnels. Les unités lexicales appartenant à ce frame sont *buy.v*, *purchase.v* et *purchase_(act).n*. Ces unités sont synonymes.

Commerce_buy

Definition:

These are words describing a basic commercial transaction involving a buyer and a seller exchanging money and goods, taking the perspecitive of the buyer. The words vary individually in the patterns of frame element realization they allow. For example, the typical pattern for the verb BUY: BUYER buys GOODS from SELLER for MONEY. Abby bought a car from Robin for $5,000.

FEs:

Core:

Buyer [Byr] — The Buyer wants the Goods and offers Money to a Seller in exchange for them.
Jess BOUGHT a coat.
Lee BOUGHT a textbook from Abby.

Goods [Gds] — The FE Goods is anything (including labor or time, for example) which is exchanged for Money in a transaction.
Only one winner PURCHASED the paintings

Non-Core:

Manner [M]
Semantic Type Manner — Any description of the purchasing event which is not covered by more specific FEs, including secondary effects (*quietly, loudly*), and general descriptions comparing events (*the same way*). It may also indicate salient characteristics of the Buyer that affect the action (*presumptuously, coldly, deliberately, eagerly, carefully*).
She gleefully BOUGHT the rock.

Means [Mns]
Semantic Type
State_of_affairs — The means by which a commercial transaction occurs.
Will they allow you to PURCHASE by check?

Money [Mny] — Money is the thing given in exchange for Goods in a transaction.
Sam BOUGHT the car for $12,000.

Figure 7 : Consultation du Frame *Commerce_buy* dans FrameNet [a]

a. Des codes de couleurs différencient les types des éléments. Les éléments essentiels *(core)*, appelés actants dans d'autres théories, sont distingués des éléments optionnels *(non-core)*, souvent appelés circonstants.

Description du verbe 'buy'

La figure 8 présente l'unité lexicale *buy* en précisant le frame auquel elle appartient et en donnant une définition soit extraite du *Concise Oxford Dictionary* (COD), soit rédigée par l'équipe de FrameNet (notée par *FN*). On donne ensuite les *Frame Elements* avec leurs réalisations syntaxiques et, finalement, une valence ou construction syntaxique.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

Frame: Commerce_buy

Definition

COD: obtain in exchange for payment

Frame Elements and Their Syntactic Realizations

The Frame elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realizations(s)
Buyer	(82)	CNI.-- (24) NP.Ext (54) PP[by].Dep (4)
Goods	(82)	DNI.-- (13) NP.Ext (13) NP.Obj (48) Sinterrog.Dep (2) AJP.Dep (2) NP.Dep (4)

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns	
82 TOTAL	Buyer	Goods
(3)	CNI --	DNI --
(9)	CNI --	NP Ext
(11)	CNI --	NP Obj
(1)	CNI --	Sinterrog Dep
(2)	NP Ext	AJP Dep
(10)	NP Ext	DNI --
(4)	NP Ext	NP Dep
(37)	NP Ext	NP Obj
(1)	NP Ext	Sinterrog Dep
(4)	PP[by] Dep	NP Ext

Figure 8 : Frame de l'entrée *buy* dans FrameNet [a]

a. On retrouve, pour chaque type de construction du verbe, le nombre d'exemples en corpus annotés. Ici, on rapporte 82 exemples d'achat de biens par un acheteur.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

buy.v

Frame Elements	Core Type
Buyer	Core
Duration	Peripheral
Goods	Core
Manner	Peripheral
Means	Peripheral
Money	Peripheral
Place	Peripheral
Purpose	Extra-Thematic
Purpose_of_goods	Extra-Thematic
Rate	Peripheral
Reason	Extra-Thematic
Recipient	Extra-Thematic
Seller	Peripheral
Time	Peripheral
Unit	Peripheral

- swh
  1. He advises : ` BUY what is functional and decorative . CNI
  2. ` I 've always BOUGHT when other people are squeezed and have got to sell . DNI
  3. The electorate will BUY what they 're shown is right , though persuading them costs a lot of bucks .
  4. George and Lennie have a dream about their own piece of land which they will be able to BUY when they get enough money .
- np-ap
  1. Four years ago I BOUGHT an old Harmony Sovereign acoustic guitar for £20 from an absolute prat .
  2. Heston , who had director , script and cast approval , BOUGHT Peckinpah extra filming time by returning his salary to the studio , convinced they would n't accept it .
  3. On other occasions , borrowing may be the only way you will ever be able to afford to BUY something expensive like a house .
  4. As a result of your win I can BUY something special for your ma .
- np-np
  1. She and husband Tom have BOUGHT a villa a stone 's throw from the BBC 's Spanish soap set -- so they can get away from tourists at home !

Figure 9 : Rapport d'annotation sur le frame *buy* dans FrameNet [a]

a. Des codes de couleur permettent de connaître le type de *Frame Element* qu'on retrouve dans les exemples du corpus.

Dans le rapport d'annotation (figure 9), on retrouve un tableau de description des FE obligatoires et optionnels ; puis, un ensemble d'exemples de contextes sous forme de phrases instanciant le verbe *buy* où tous les FE sont annotés par des rôles sémantiques. La figure 9 ne présente que quelques exemples et FE. Trois étiquettes (DNI, CNI et INI) sont aussi utilisées dans *FN* pour décrire les FE. L'étiquette DNI *(Define Null Instantiation)* indique que le FE n'apparaît pas explicitement dans la phrase et qu'il est porté dans le discours ou le contexte. Les étiquettes CNI *(Constructional Null Instanciation)* et INI *(Idefinite Null Instantiation)* permettent d'annoter les anaphores et les ellipses.

La représentation (illustrée dans les figures 7 à 9) est utilisée pour faciliter la consultation d'une unité prédicative par un utilisateur humain. Les données de *FN* sont organisées en une base de données exploitable à l'aide de programmes informatiques. *FN* offre aussi un module d'interrogation intégré *(FrameNet SQL)*.

Dans *FN*, les frames sont reliés par des relations sémantiques telles que *héritage*, *sous-frame*, *causative*, *point de vue*, etc. Par exemple, le frame C*ommerce_buy*

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

est le fils de *Commerce_goods-transfer*, et il est le petit fils du frame *Commercial_transaction* défini par les FE *Buyer*, *Goods*, *Seller* et *Money*. Dans le graphe des frames, le frame *Commerce_sell* est le frère du frame *Commerce_buy*. Ils héritent tous les deux du frame principal *Commerce_transaction*. La différence est que les FE obligatoires pour *Commerce_buy* sont *Buyer* et *Goods*, et les optionnels sont *Seller* et *Money*. Quant au frame *Commerce_sell*, les FE obligatoires sont *Seller* et *Goods*, et les optionnels sont *Buyer* et *Money*.

3.3.2. Comparaison avec LVF

LVF et *FN* comportent tous deux la notion de groupements des verbes. Pour *FN*, cette notion est fondée sur le partage du comportement sémantique ou des *Frame Elements* et ne tient pas compte du comportement syntaxique alors que, pour *LVF*, le comportement syntaxique est indispensable. On parle alors de *frames sémantiques* pour *FN* et de schèmes syntaxiques et de constructions sémantiques pour *LVF*.

Dans *FN*, une unité lexicale (*i.e.* un lemme associé à un sens) évoque un frame unique. Chaque frame peut être associé à plusieurs unités lexicales de plusieurs types (nom, adjectif, verbe) et est décrit par les *Frame Elements* ou bien les rôles sémantiques correspondant aux arguments de l'unité lexicale prédicative. *FN* ne tient pas compte du comportement syntaxique pour décrire les frames et le sens des verbes, il se base principalement sur les *Frame Elements* (rôles sémantiques) pour décrire le sens des unités lexicales. Ces FE sont nécessaires, on doit les annoter dans les contextes, mais ils sont parfois absents de la phrase du contexte ; c'est pourquoi *FN* a introduit trois étiquettes (instanciation nulle : CNI, INI, DNI) pour annoter ces FE omis dans les phrases.

– Exemples de cas de CNI : un sujet omis dans une phrase impérative *(Cook on low heat until done [CNI FOOD]* [7]*)*, un agent omis dans une phrase passive *(Family feuds last for generation and [slurs on honor INJURY] are avenged [by murder PUNISHMENT][CNI AVENGER]).*
– Exemples de cas DNI ou type anaphorique : cas où les éléments omis doivent être les éléments qui sont déjà compris dans le contexte du discours *([The monkey AVENGER] avenged [himself INJURED_PARTY][by growing to the size of a giant and setting fire to the city PUNISHMENT][DNI OFFENDER]).*
– Les cas d'INI sont illustrés par les objets omis des verbes tels que *eat*, *drink*, *sew*, *bake*, etc.

Dans *LVF*, chaque sens associé à un verbe est numéroté. Les emplois ou sens du verbe sont répertoriés dans des classes sémantico-syntaxiques selon la construction lexico-syntaxique (schèmes syntaxiques) et la construction sémantique (opérateur) qui interprète ces schèmes syntaxiques. Une unité lexicale peut appartenir à plusieurs classes sémantico-syntaxiques dans *LVF*, mais à un seul

7. Les rôles sémantiques des *Frame Elements* (les éléments qui sont entre crochets) sont écrits en majuscule.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

frame dans *FN*. Par exemple, *acheter (buy)* n'appartient qu'au frame de la transaction commerciale dans *FN* faisant intervenir le vendeur, l'acheteur et l'argent ; dans *LVF*, cela correspond aux entrées **acheter 01** ou **acheter 02** (*buy something for...* avec le rôle sémantique RECIPIENT pour le bénéficiaire), dans les domaines COM (commerce) ou ECN (économie). *FN* ne prend pas en considération l'emploi de *buy* comme 'corrompre une personne' ou l'emploi de *buy* 'acheter quelque chose avec autre chose que l'argent' comme 'acheter sa liberté en se sacrifiant' ou 'avec son sang' qui appartiennent aux domaines SOC (social) ou PSY (psychologique). *FN* définit d'abord les frames pour y introduire ensuite les unités lexicales. *FN* étant toujours en développement, cette insuffisance pourrait être comblée dans le futur.

3.4. WordNet

3.4.1. Description

WordNet [8] (Miller 1995) est une ressource lexicale de large couverture, développée pour la langue anglaise par des linguistes du laboratoire des sciences cognitives de l'Université Princeton. C'est un réseau sémantique de la langue anglaise qui se fonde sur une théorie psychologique du langage dont le but est de classer et de mettre en relation le contenu sémantique et lexical de cette langue.

Il consiste en un réseau de mots et une hiérarchie de concepts de la réalité perçue. Ces concepts sont abstraits (*motion*, *communication*, *cognition*, etc.) ou concrets (*animal*, *humain*, *plants*, etc.) et reliés par des relations sémantiques (Fellbaum 1998 ; Miller 1998). Il regroupe les noms, les verbes, les adjectifs et adverbes dans des *synsets* (ensemble de synonymes). Ainsi, une liste de *synsets* est associée à un mot correspondant aux différents sens ou acception répertoriés de ce mot. *WN* repose sur deux moyens pour définir le sens de l'unité considérée : les *synsets* et les relations lexicales.

Synset

Un sens d'un mot est représenté par l'ensemble des mots utilisés pour exprimer ce sens les *synsets* (*i.e.* par un ensemble de synonymes) et une définition (Fellbaum 1998 ; Miller 1998).

Un *synset* regroupe des mots synonymes dénotant un sens ou un usage particulier. Dans *WN*, les noms et verbes sont organisés en hiérarchies. Les relations sémantiques permettent de déterminer le sens d'un mot en relation avec d'autres sens. Ces relations sont l'hyperonymie ('est-un') et l'hyponymie qui constituent une hiérarchie d'héritage.

8. WordNet est consultable à l'adresse suivante : http://wordnet.princeton.edu/

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

- S: (n) bargain, **buy**, steal (an advantageous purchase) "*she got a bargain at the auction*"; "*the stock was a real buy at that price*"

Figure 10 : Information sur le nom *buy* dans WordNet [a]

a. Les éléments du synset sont *bargain*, *buy* et *steal*, suivis d'une courte définition et de quelques exemples d'usage.

Les relations lexicales

Dans WordNet, des relations existent entre les *synsets* et entre les mots. Les *synsets* sont rassemblés sur une intuition d'ordre conceptuel, selon le sens que l'on attribue intuitivement à un mot. Deux types de relation sont identifiés :

– relations entre les synsets : hyperonymie/hyponymie (du type *tree/plant*) ;
– relations entre les unités lexicales ou les lemmes : synonymie (du type *buy/purchase*), antonymie (du type *rich/poor*).

Les verbes dans WordNet

Les verbes sont classés par champ sémantique. Trois grandes classes sont considérées : actions, événements, états. Actions et événements sont divisés en 14 classes : verbes de mouvement, de perception, de contact, de communication, de compétition, de consommation, de comportement social, etc. Un même champ sémantique peut avoir plusieurs hiérarchies indépendantes. Par exemple, les verbes de possession sont regroupés au sein de 3 hiérarchies distinctes dont les origines sont {give, transfer}, {take, receive} et {have, hold}. Les verbes partageant la même valence syntaxique ont en commun une composante sémantique. Ainsi, les arbres des verbes créés sur une base sémantique correspondent à un certain comportement syntaxique (Miller 1998) ou à ce à partir de quoi les comportements syntaxiques dans *WN* sont exprimés par des classes sémantiques.

- S: (v) **buy**, purchase (obtain by purchase; acquire by means of a financial transaction) "*The family purchased a new car*"; "*The conglomerate acquired a new company*"; "*She buys for the big department store*"
- S: (v) bribe, corrupt, **buy**, grease one's palms (make illegal payments to in exchange for favors or influence) "*This judge can be bought*"
- S: (v) **buy** (be worth or be capable of buying) "*This sum will buy you a ride on the train*"
- S: (v) **buy** (acquire by trade or sacrifice or exchange) "*She wanted to buy his love with her dedication to him and his work*"
- S: (v) **buy** (accept as true) "*I can't buy this story*"

Figure 11 : Consultation de *buy* dans WordNet (5 sens) [a]

a. En cliquant sur le S, au début des items de la liste, on accède aux relations sémantiques et lexicales. La figure 12 montre celles du premier sens.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

- S: (v) **buy**, purchase (obtain by purchase; acquire by means of a financial transaction) "*The family purchased a new car*"; "*The conglomerate acquired a new company*"; "*She buys for the big department store*"
  - *direct troponym* / *full troponym*
  - *verb group*
    - S: (v) **buy** (be worth or be capable of buying) "*This sum will buy you a ride on the train*"
  - *domain category*
    - S: (n) commerce, commercialism, mercantilism (transactions (sales and purchases) having the objective of supplying commodities (goods and services))
  - *direct hypernym* / *inherited hypernym* / *sister term*
    - S: (v) get, acquire (come into the possession of something concrete or abstract) "*She got a lot of paintings from her uncle*"; "*They acquired a new pet*"; "*Get your results the next day*"; "*Get permission to take a few days off from work*"
  - *entailment*
    - S: (v) pay (give money, usually in exchange for goods or services) "*I paid four dollars for this sandwich*"; "*Pay the waitress, please*"
    - S: (v) choose, take, select, pick out (pick out, select, or choose from a number of alternatives) "*Take any one of these cards*"; "*Choose a good husband for your daughter*"; "*She selected a pair of shoes from among the dozen the salesgirl had shown her*"
  - *phrasal verb*
  - ***antonym***
    - W: (v) sell [Opposed to: buy] (exchange or deliver for money or its equivalent) "*He sold his house in January*"; "*She sells her body to survive and support her drug habit*"
  - *derivationally related form*
  - *sentence frame*

Figure 12 : Relations sémantiques et lexicales associées au premier sens de *buy*

3.4.2. Comparaison avec LVF

La hiérarchie des verbes dans WordNet est composée de verbes d'états et de verbes qui dénotent des actions et des évènements. Les verbes d'états sont divisés en *be*, *belong*, *auxiliaries* et *control verb* (comme *want*, *fail*, *prevent*). Les verbes d'actions et d'évènement sont divisés en 14 classes sémantiques : *motion*, *perception*, *contact*, *communication*, *competition*, *change*, *cognition*, *consumption*, *creation*, *emotion*, *perception*, *possession*, *bodily care* and *functions*, *social behavoir and interaction*. Certaines classes correspondent aux 14 classes génériques de *LVF* (communication ; don, privation ; entrée, sortie ; frapper, toucher ; états physiques et comportements ; locatif ; mouvement sur place ; munir, démunir ; verbes psychologiques ; réalisation, mise en état ; saisir, serrer, posséder ; transformation, changement ; union, réunion ; verbes auxiliaires). La classe des verbes auxiliaires peut correspondre avec les verbes d'états de *WN*. *WN* et *LVF* divisent d'abord l'ensemble des verbes de la langue selon les champs sémantiques conceptuels définis plus haut. *WN* a répertorié plusieurs sens de *buy* correspondant à *acheter* dans *LVF*. *WN*, contrairement à *FN*, répertorie le sens de corruption pour le verbe *buy* (deuxième sens de la figure 11 : *This judge can be bought*) correspondant à l'entrée **acheter 03** dans *LVF*. Il a aussi identifié le sens d''acheter un argument abstrait' comme le bonheur, la liberté, l'amour, etc. (quatrième sens de la figure 11 : *She wanted to by his love with her dedication to him and his work*) correspondant à l'entrée **acheter 04** dans *LVF*.

WordNet organise les verbes dans un ensemble de synonymes appelé *synsets*. Il inclut un ensemble riche de relations sémantiques entre les *synsets* et entre

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

les mots ou unités lexicales (hyponymie, hyperonymie, antonymie, synonymie, etc.), mais il ne fournit pas la structure argumentale du prédicat ni les structures syntaxiques associées aux sens. Dans *WN*, les verbes sont regroupés en champs sémantiques, mais les relations utilisées sont toutes paradigmatiques et on ne retrouve aucune information syntagmatique (*i.e.* sur les contextes dans lesquels les mots apparaissent). Alors que dans *LVF*, les informations syntaxiques et sémantiques sont riches, *WN* ne fournit pas les structures syntaxiques pour les différents sens des verbes. Il suppose que l'utilisateur connaît déjà le contexte dans lequel le mot est employé. Malgré cette lacune de *WN*, cette ressource reste très utilisée pour les développements sémantiques ou dans le cadre du web sémantique (ce qui est sûrement en partie dû à sa libre disponibilité dans une formation informatique pratique pour les applications).

En s'inspirant de ce modèle, *LVF* pourrait donc devenir une ressource très intéressante pour ces applications de TAL sachant qu'elle offre une combinaison d'informations syntaxiques et sémantiques. Cette combinaison de syntaxe et sémantique est plus explicite que la seule conceptualisation des significations de *WN*.

4. TABLEAU RÉCAPITULATIF

Même si les théories linguistiques fondant ces ressources diffèrent (approche pronominale du Dicovalence, approche distributionnelle et transformationnelle de *LVF*, approche de *Semantic Frames* de *FN*, approche de *synsets* de *WN*), celles-ci présentent plusieurs points communs dans leurs descriptions des unités lexicales : le sens d'un mot est toujours déterminé par des relations paradigmatiques et syntagmatiques (sauf pour *WN* dans ce dernier cas). Elles utilisent, toutefois, des éléments propres à leurs théories dans le but de former un réseau ou une hiérarchie de relations sémantiques entre les mots. Dans cette section, nous mettons en évidence les points communs entre *LVF* et les autres ressources en le comparant à Dicovalence, VerbNet, FrameNet et WordNet. Le tableau 1 compare les cinq ressources sur certains points syntaxiques ou sémantiques de classification.

Tableau 1 : Comparaison des caractéristiques de *LVF* avec d'autres ressources lexicales : une pour le français et trois pour l'anglais

	LVF	Dicovalencee	VerbNet	FrameNet	WordNet
Regroupement en classes	Hiérarchie, sans héritage Verbes partageant les mêmes propriétés sémantico-syntaxique	Pas de hiérarchie Verbes partageant les mêmes propriétés syntaxiques ou frame	Hiérarchie héritage par la sous-classe Verbes partageant les informations syntaxiques et sémantiques	Graphe de scénarios conceptuels ou frames Les unités lexicales partageant les mêmes Frame Elements (rôles sémantiques)	Réseau de mots et de concepts Classification selon les synsets, et les relations séman-tiques et lexicales

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

Tableau 1 : (suite)

Types de liens	Description de propriétés séman-tiques et syntaxi-ques	Description des cadres de valence et du champ frame	Description de frame syntaxique et prédicat sémantique, ainsi que les rôles thématiques	Héritage, causative, inchoative, point de vue, etc.	Synonymie, antonymie hyperonymie, hyponymie
Définition	Donnée dans la rubrique Opérateur (formule syntaxico-sémantique) à l'aide d'opérateurs	Pas de définition	Pas de définition	Tirée d'un dictionnaire ou donnée par l'équipe de *FN*	Courte
Description	Construction syntaxique (schème) Opérateur (séman-tique) Classe sémantico-syntaxique Numéro, domaine, sens (les verbes syno-nymes) Une simple phrase d'emploi sans aucune annotation	Informations syntaxiques suivies de trait sémantique cadre valenciel Type du verbe Numéro, traduction en néerlandais et en anglais Paradigmes pronominaux Reformulation passive et auxiliaire Exemple d'utilisation	Frames syntaxiques Restriction Rôles thématiques Prédicat sémantique Exemple sans annotation par type de construction	Rôles sémantiques Frame Définition Construction syntaxique Phrases d'emplois annotées par des rôles sémantiques	Pauvre
Rôles sémantiques	Absents	Absents	Essentiels	Essentiels	Absents
Synonymes	Donnés par la colonne sens	Difficile à distinguer	Difficile à distinguer	Dans un même frame	Par les synsets
Sens différents	Représentés par des numéros et regroupés dans des classes différentes	Représentés par des numéros différents et donnés selon des cadres valenciels différents	Classes différentes ; un verbe d'une classe est sémanti-quement non ambigu	Regroupés dans des frames différents	Numérotation des sens différents
Domaine	Explicite	Verbes simples de la langue les plus utilisés et pas de verbes de spécialité	Non explicite	Non explicite	Non explicite
Exemples d'emploi	Un pour chaque sens, sans annotation	Un par cadre valenciel du verbe représentant un sens du verbe	Un par construction	Plusieurs exemples d'emplois annotés par des rôles sémantiques	Exemples simples
Sens figurés	Par des sous-classes	Non représentés explicitement	Non représentés	Non représentés	Non explicites

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

5. CONCLUSION

Nous avons présenté une correspondance entre certains éléments du *LVF* et d'autres ressources lexicales françaises et anglaises, riches en informations sémantiques, à grande couverture et qui ont été souvent utilisées pour plusieurs applications en TAL.

Malgré des différences d'approches et de théories telles que l'approche pronominale pour Dicovalence, les classes de B. Levin pour VerbNet, les *frames* pour FrameNet ou les *synsets* pour WordNet, nous avons établi plusieurs correspondances entre les informations décrites dans ces ressources et *LVF*.

LVF, fournissant des informations syntaxiques et sémantiques riches, gagnerait à être mieux connu de la communauté TAL. Les schèmes syntaxiques nous informant sur l'organisation et les relations syntaxiques des différents éléments entourant un verbe. L'opérateur et ses compléments, qui interprètent sémantiquement ces schèmes syntaxiques ainsi que la rubrique sens qui définit le verbe, permettent d'organiser et de classer les différents emplois ou sens d'un verbe donnant ainsi naissance à plusieurs entrées d'un même verbe. Son organisation par domaine est un atout intéressant, car il facilite la consultation du vocabulaire d'un domaine en particulier. C'est pourquoi, nous pensons que la version XML du *LVF* en facilitera la diffusion et son utilisation éventuelle dans des applications de TAL.

Références

BAKER C. F., FILLMORE C. J. & LOWE J. B. (1998), "The Berkeley FrameNet project", *Proceedings of the COLING-ACL '98 conference*, Montréal: Université de Montréal, 86-90.

DUBOIS J. & DUBOIS-CHARLIER F. (1997), *Les Verbes français*, Paris : Larousse-Bordas.

FELLBAUM C. (ed.) (1998), *WordNet: An Electronic Lexical Database*, Cambridge (MA): MIT Press.

FILLMORE C. J. (1982), "Frame semantics", *Linguistics in the Morning Calm*, Seoul: Hanshin Publishing Co, 111-137.

FILLMORE C. J. & ATKINS B. T. (1992), "Towards a Frame-Based Organization of the Lexicon: The Semantics of RISK and its Neighbours", *in* A. Lehrer & E. Kittay (eds), *Frames, Fields, and Contrast : New Essays in Semantics and Lexical Organization*, Hillsdale: Lawrence Erlbaum Associates, 75-102.

FRANÇOIS J., LE PESANT D. & LEEMAN D. (2007), « Présentation de la classification des *Verbes français* de Jean Dubois et Françoise Dubois-Charlier », *Langue française* 153, 3-19.

KIPPER K., DANG H. & PALMER M. (2000), "Class-based construction of a verb lexicon", *Proceedings of the Seventh National Conference on Artificial Intelligence*, Austin (TX): AAAI Press, 691-696.

KIPPER-SCHULER K. (2005), *VerbNet: a broad-coverage, comprehensive verb lexicon*, Computer and Information Science PhD, Thesis, University of Pennsylvania.

LEVIN B. (1993), *English Verb Classes and Alternations: A Preliminary Investigation*, Chicago: The University of Chicago Press.

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)

MERTENS P. (2010), « Restrictions de sélection et réalisations syntagmatiques dans DICOVALENCE. Conversion vers un format utilisable en TAL », *Actes de la 17e Conférence sur le Traitement Automatique des Langues Naturelles* (TALN 2010), Montréal. [http://bach.arts.kuleuven.be/pmertens/papers/PMertens_taln2010.pdf]

MILLER G. A. (1995), "WordNet: A Lexical Database for English", *Communications of the ACM* 38, 39-41.

MILLER G. A. (1998), "Foreword", *in* C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database*, Cambridge (MA): MIT Press, xv-xxii.

VAN DEN EYNDE K. & MERTENS P. (2006), *Le dictionnaire de valence* Dicovalence : *Manuel d'utilisation*, Leuven : Université de Leuven. [http://bach.arts.kuleuven.be / dicovalence / manuel_100625.pdf]

© Armand Colin | Téléchargé le 02/01/2022 sur www.cairn.info (IP: 90.126.175.170)