

Université de Montréal

**Identification et modélisation lexicographique des dépendances syntaxiques régies  
du français: le cas des dépendances nominales**

par  
Évelyne Arbour Barbaud

Département de linguistique et de traduction  
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)  
en linguistique

Avril, 2012

© Évelyne Arbour Barbaud, 2012.

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé:

**Identification et modélisation lexicographique des dépendances syntaxiques régies  
du français: le cas des dépendances nominales**

présenté par:

Évelyne Arbour Barbaud

a été évalué par un jury composé des personnes suivantes:

Marie-Claude L'homme,	président-rapporteur
Alain Polguère,	directeur de recherche
Lyne Da Sylva,	membre du jury

**Mémoire accepté le: 30 juillet 2012**

## RÉSUMÉ

Cette recherche porte sur l'interface entre la sémantique lexicale et la syntaxe, et elle s'inscrit dans le cadre du projet de base lexicale DiCo (acronyme pour *Dictionnaire de combinatoire*) à l'Observatoire de Linguistique Sens-Texte [OLST] de l'Université de Montréal. Le projet découle d'une volonté d'inscrire de façon concise et complète, à même le dictionnaire, le comportement syntaxique typique à chaque unité lexicale. Dans cette optique, nous encodons la cooccurrence des lexies nominales du DiCo avec leurs actants à l'intérieur d'un tableau de régime lexical (aussi connu sous le nom de *schéma valenciel*, *structure argumentale*, *cadre de sous-catégorisation*, *structure prédicats-arguments*, etc.), en notant entre autres les dépendances syntaxiques de surface impliquées. Dans ce mémoire, nous présentons les propriétés syntaxiques d'une dépendance nominale du français, celle que nous avons nommée **attributive adnominale**, de façon à exposer une méthodologie d'identification et de caractérisation des dépendances syntaxiques de surface. Nous donnons également la liste des dépendances nominales régies identifiées au cours de ce travail. Par la suite, nous exposons la création d'une base de données de régimes généralisés du français nommée CARNAVAL. Finalement, nous discutons des applications possibles de notre travail, particulièrement en ce qui a trait à la création d'une typologie des régimes lexicaux du français.

**Mots clés :** lexicologie/lexicographie, syntaxe de dépendance, interface sémantique-syntaxe, régime lexical, Théorie Sens-Texte, Lexicologie Explicative et Combinatoire.

## ABSTRACT

This research focuses on the interface between lexical semantics and syntax, and is part of the lexical database project called DiCo (acronym for *Dictionnaire de combinatoire*) at the Observatoire de Linguistique Sens-Texte [OLST] at the University of Montreal. The project aims to register in a concise and complete way the typical syntactic behavior of each lexical unit directly in the dictionary. To this end, we encode the co-occurrence of the DiCo's nominal lexical units with their actants within a government pattern (also known as *valency structure*, *subcategorization frame*, *predicate-argument structure*, etc.), using, among other things, the surface syntactic dependencies involved. In this dissertation, we present the syntactic properties of a French nominal dependency, the one that we named **attributive adnominale**, so as to expose a methodology for the identification and the characterization of surface syntactic dependencies. We also give the list of valence-controlled nominal dependencies that were identified in our work. Subsequently, we describe the creation of a database of generalized French government patterns named CARNAVAL. Finally, we discuss possible applications of our work, particularly in regard to the creation of a typology of French government patterns.

**Keywords :** lexicology/lexicography, dependency syntax, semantic/syntax interface, lexical valency, Meaning-Text Theory, Explanatory Combinatorial Lexicology.

## TABLE DES MATIÈRES

<b>RÉSUMÉ</b> . . . . .	<b>iii</b>
<b>ABSTRACT</b> . . . . .	<b>iv</b>
<b>TABLE DES MATIÈRES</b> . . . . .	<b>v</b>
<b>LISTE DES TABLEAUX</b> . . . . .	<b>viii</b>
<b>LISTE DES FIGURES</b> . . . . .	<b>ix</b>
<b>LISTE DES ANNEXES</b> . . . . .	<b>x</b>
<b>LISTE DES SIGLES</b> . . . . .	<b>xi</b>
<b>CONVENTIONS D'ÉCRITURE</b> . . . . .	<b>xii</b>
<b>DÉDICACE</b> . . . . .	<b>xiii</b>
<b>REMERCIEMENTS</b> . . . . .	<b>xiv</b>
<b>INTRODUCTION</b> . . . . .	<b>1</b>
<b>CHAPITRE 1 : PROBLÉMATIQUE ET OBJECTIFS DE RECHERCHE</b> . . . . .	<b>4</b>
1.1 Étude préliminaire : ouvrages généraux de lexicographie . . . . .	4
1.2 Analyse : dictionnaires et autres bases de données lexicales . . . . .	7
1.2.1 Produits de la pratique lexicographique anglophone . . . . .	8
1.2.2 Produits de la pratique lexicographique francophone . . . . .	18
1.3 Problème spécifique . . . . .	27
1.4 Objectifs de recherche . . . . .	28
<b>CHAPITRE 2 : LE RÉGIME LEXICAL</b> . . . . .	<b>31</b>
2.1 Un survol de la place du régime lexical en linguistique . . . . .	31

2.1.1	Grammaire générative . . . . .	32
2.1.2	Sémantique conceptuelle . . . . .	34
2.1.3	Grammaire syntagmatique guidée par les têtes . . . . .	35
2.2	Notre cadre théorique : La théorie Sens-Texte . . . . .	38
2.2.1	Présentation générale . . . . .	39
2.2.2	La dépendance syntaxique . . . . .	41
2.2.3	La Lexicologie Explicative et Combinatoire . . . . .	47
2.3	Modélisation des régimes lexicaux en LEC . . . . .	52

### **CHAPITRE 3 : IDENTIFICATION ET CARACTÉRISATION DES DÉPENDANCES SYNTAXIQUES DE SURFACE : LE CAS DE LA DÉPENDANCE ATTRIBUTIVE ADNOMINALE . 56**

3.1	Conceptualisation de la dépendance <b>attributive adnominale</b> .	57
3.1.1	Les différents compléments du nom . . . . .	57
3.1.2	Tesnière et la translation . . . . .	59
3.1.3	Une première description de la construction N1 — <b>attributive adnominale</b> → N2 . . . . .	61
3.2	Méthodologie d'identification/ caractérisation d'une DépSyntS . . . . .	66
3.2.1	Critères de regroupement de différentes constructions syntaxiques sous une même dépendance . . . . .	66
3.2.2	Construction d'un corpus d'exemples . . . . .	67
3.2.3	Propriétés de la dépendance <b>attributive adnominale</b> .	71
3.3	La dépendance <b>attributive adnominale</b> et le « Régime lexicalisé » . . . . .	80
3.3.1	La création lexicale . . . . .	80
3.3.2	Encodage dans les régimes lexicaux . . . . .	81
3.3.3	Encodage dans la zone de combinatoire lexicale . . . . .	84
3.3.4	En résumé . . . . .	89

### **CHAPITRE 4 : MODÉLISATION DES RÉGIMES LEXICAUX DU FRANÇAIS DANS LA BASE DE DONNÉES CARNAVAL 92**

4.1	Tâches effectuées . . . . .	93
4.1.1	Identification et caractérisation des dépendances syntaxiques de surface nominales du français . . . . .	93
4.1.2	Identification et modélisation des différentes composantes d'un régime lexical . . . . .	106
4.1.3	Encodage, dans la base de données, de chacun des régimes associés aux lexies nominales du DiCo . . . . .	109
4.2	Présentation de la base de données CARNAVAL . . . . .	110
4.2.1	Présentation générale . . . . .	110
4.2.2	Le schéma et les balises nécessaires à l'encodage des régimes . . . . .	112
4.2.3	Deux exemples de Tableau de régime . . . . .	115
	<b>CHAPITRE 5 : OUVERTURE ET CONCLUSION . . . . .</b>	<b>120</b>
5.1	Applications futures . . . . .	120
5.2	Vers une typologie des régimes nominaux . . . . .	123
5.3	Limites de la recherche . . . . .	128
5.4	Bilan . . . . .	132
	<b>BIBLIOGRAPHIE . . . . .</b>	<b>135</b>

## LISTE DES TABLEAUX

3.1	Propriétés de la dépendance <b>attributive adnominale</b> . . . . .	78
-----	---	----

## LISTE DES FIGURES

1.1	Extrait de l'entrée lexicale de <code>try.n</code> dans FrameNet . . . . .	16
2.1	Classification des DépSyntP . . . . .	46
2.2	Régime de BLESSURE#I.2 dans DECFC Vol. IV . . . . .	53
4.1	Les DépSyntS dans la Hiérarchie GTN . . . . .	98
4.2	Régime de FRONTIÈRE#I.1 . . . . .	109
5.1	Comparaison des Régimes r.N.2.002 et r.N.2.064 . . . . .	124
5.2	Comparaison des Régimes r.N.2.086 et r.N.2.099 . . . . .	125
5.3	Comparaison des Régimes r.N.2.003 et r.N.2.004 . . . . .	127

## **LISTE DES ANNEXES**

<b>Annexe I :</b>	<b>Liste des lexies étudiées et leur numéro de régime . . . . .</b>	<b>xvi</b>
<b>Annexe II :</b>	<b>Les sous-groupes de la partie du discours Prép. . . . .</b>	<b>xxviii</b>

## LISTE DES SIGLES

### Notions

ASém	Actant Sémantique
ASyntP	Actant Syntaxique Profond
ASyntS	Actant Syntaxique de Surface
D	Dépendant [d'une relation]
DEC	Dictionnaire Explicatif et Combinatoire
DépSém	Dépendance Sémantique
DépSyntP	Dépendance Syntaxique Profonde
DépSyntS	Dépendance Syntaxique de Surface
G	Gouverneur [d'une relation]
L	Lexie
LEC	Lexicologie Explicative et Combinatoire
RSém	Représentation Sémantique
RSyntP	Représentation Syntaxique Profonde
RSyntS	Représentation Syntaxique de Surface
SSém	Structure Sémantique
SSyntP	Structure Syntaxique Profonde
SSyntS	Structure Syntaxique de Surface
TST	Théorie Sens-Texte

### Ouvrages

DECFC	Dictionnaire explicatif et combinatoire du français contemporain
DiCo	Dictionnaire de combinatoire
ILEC	Introduction à la lexicologie explicative et combinatoire
LAF	Lexique actif du français
RLF	Réseau lexical du français
TLFi	Trésor de la langue française informatisé

## CONVENTIONS D'ÉCRITURE

**police typewriter en gras** Noms des dépendances syntaxiques  
**de surface**

police typewriter Extraits du DiCo, de la  
Hiérarchie GTN ou de la base CAR-  
NAVAL

'guillemets simples' 'Sens'

MAJUSCULES LEXIES

À ma nouvelle petite famille.

## REMERCIEMENTS

À tout seigneur tout honneur, je tiens tout d'abord à remercier Lida Iordanskaja qui a bien voulu partager des données qui sont les fruits d'un travail de recherche colossal et qui ont servi de base à ce travail. Merci à Igor Mel'čuk, qui non seulement m'a initiée au monde merveilleux de la linguistique, mais m'a également offert ma première chance. Je crois que tu as eu confiance en mes capacités avant moi-même. Merci à Alain Polguère qui a bien voulu me diriger pour ce travail, malgré tous les autres projets autant professionnels que personnels qu'il avait devant lui. Ces trois personnes et moi-même avons travaillé conjointement pour établir une première liste des dépendances syntaxiques de surface régies du français. Nos réunions de travail ont été extrêmement enrichissantes (et parfois divertissantes) pour moi et je vous remercie de m'avoir incluse dans votre cercle pour ces moments d'études.

Pour ma part, je trouve que le travail de recherche est parfois trop empreint d'isolement. On se sent parfois seule, inutile et on finit par croire qu'à peu près personne ne comprend ce sur quoi on s'acharne à travailler. Je tiens à remercier toutes les personnes qui ont, à un moment ou un autre, brisé ou partagé ce sentiment ; collègues et autres étudiants chercheurs, merci. Plus particulièrement, je tiens à remercier Sébastien Marengo avec qui l'échange d'idées a été particulièrement productif. Merci également d'avoir accepté d'être ma voix en Nouvelle-Orléans. Sylvie, Viviane, merci de m'avoir sortie de ma tête aussi souvent, vous ne pouvez savoir à quel point votre énergie parfois empreinte de folie m'a été nécessaire.

Je tiens également à remercier Jean-Marie Pierrel, directeur de l'ATILF, pour le financement d'un stage au cours des mois d'octobre et novembre 2009. J'ai été accueilli de belle façon par tous les membres du laboratoire et cela m'a permis de bénéficier de belles séances de travail avec mon directeur de recherche et aussi d'une initiation au travail qui est fait là-bas. Je suis fière d'avoir pu ajouter cette expérience à mon parcours.

La rédaction d'un mémoire est un processus long et difficile, d'autant s'il est ponctué d'éléments imprévus de la vie. Je tiens à remercier pour leur patience toutes les personnes qui, dans cette période, ont fait partie de ma vie tant académique que person-

nelle. À Alain, merci d'avoir su être magnanime malgré la longue durée du projet. À Jonathan, merci d'avoir compris que je devais inclure cette rédaction dans nos projets de vie.

## INTRODUCTION

Notre recherche a été réalisée dans le cadre du projet de base lexicale du français DiCo (Jousse et Polguère 2005), dans lequel chaque article lexicographique comporte les dérivations sémantiques qui lient la lexie décrite à d'autres lexies de la langue ainsi que les collocations qu'elle contrôle. Cette description s'accompagne d'une modélisation des structures syntaxiques régies par la lexie et d'une modélisation de son sens, sous forme d'étiquetage sémantique. Nous nous sommes plus précisément intéressée à la section syntaxique de la fiche lexicographique.

Présentement, les fiches du DiCo contiennent une modélisation des structures syntaxiques régies par la lexie. Toutefois, on n'y indique pas explicitement les relations syntaxiques liant les actants à la lexie-vedette. On utilise présentement ce qu'on pourrait appeler une notation abrégée. L'objectif général de cette recherche est donc de vérifier les informations actuellement compilées et d'affiner les descriptions des régimes lexicaux des unités lexicales nominales. Il s'agit d'une amorce d'un travail qui pourra éventuellement être complété pour les autres principales parties du discours (verbe, adjectif et adverbe).

Notre travail s'inscrit dans le cadre théorique de la théorie Sens-Texte [dorénavant TST]. Comme cette théorie a une approche de la langue qui accorde une place primordiale au lexique ainsi qu'à divers types de dépendances, elle permet, notamment grâce à sa conception des actants et à sa notion de relation syntaxique de surface, une description rigoureuse du comportement syntaxique des unités lexicales.

Cette recherche se déroule en plusieurs étapes. Dans un premier temps, elle vise à proposer et appliquer une méthode de description des régimes lexicaux. Après avoir examiné la façon d'encoder ces informations (à la fois lexicales et syntaxiques) en lexicographie traditionnelle, nous tentons de cerner les avantages et inconvénients des pratiques observées et ainsi de proposer une méthode de description rigoureuse et uniforme, que nous appliquons aux lexies nominales du DiCo. Puisque nous proposons d'ajouter le nom des dépendances syntaxiques de surface aux régimes, cette recherche vise dans un deuxième temps l'identification des différentes dépendances

syntaxiques de surface (et plus spécifiquement des dépendances régies par des noms) qui existent en français. Afin de démontrer la méthodologie que nous avons utilisée pour la caractérisation des dépendances nominales régies, la dépendance que nous avons nommée **attributive adnominale** fait d'ailleurs l'objet d'une étude plus approfondie. Notre liste de dépendances nominales établie, nous avons amorcé la construction d'une base de données de régimes généralisés nommée CARNAVAL, qui se veut un outil pour vérifier jusqu'à quel point il existe des corrélations entre les régimes lexicaux et les classes sémantiques (présentement établies grâce aux étiquettes sémantiques) auxquelles appartiennent les lexies. La généralisation de l'information sur le régime lexical pourra éventuellement servir à la création d'une typologie des régimes du français. Nous croyons, en effet, que la généralisation de ce type d'information peut notamment avoir des applications en traitement automatique de la langue et en didactique.

Le présent mémoire comprend 5 chapitres. Tout d'abord, dans le premier chapitre, un regard porté sur différents ouvrages généraux de lexicographie nous permettra de voir l'étendue des données syntaxiques que les lexicologues suggèrent d'inclure dans les dictionnaires. Nous présentons également quelques dictionnaires d'usage et autres bases de données lexicales et les informations syntaxiques qu'ils contiennent. Le but de ce survol est de nous familiariser avec les pratiques lexicographiques courantes et de nous en inspirer dans notre propre travail. C'est également au cours de ce chapitre que nous cernons nos objectifs de recherche.

Dans le chapitre suivant, nous examinerons la place qui est accordée au lexique dans plusieurs théories linguistiques. Par la suite, nous nous consacrons à l'introduction de notions théoriques importantes pour la compréhension du travail effectué. En ce qui a trait à la présentation du cadre théorique, nous souhaitons nous pencher plus sérieusement sur deux aspects prédominants de la TST, soit la place centrale accordée au lexique et la stratification de la modélisation.

Le chapitre subséquent se consacre à l'étude d'une dépendance nominale régie en particulier, soit celle que nous avons nommée **attributive adnominale**. Ce fai-

sant, nous souhaitons proposer une méthodologie qui servira à une identification et caractérisation efficace des autres dépendances syntaxiques de surface du français, autant pour notre travail sur les données du DiCo que pour d'autres projets s'intéressant à l'interaction entre le lexique et la syntaxe.

Le quatrième chapitre traite de la modélisation des régimes nominaux du français. Nous y expliquons la création de la base de données CARNAVAL ; d'un ensemble de régimes dont chacun est propre à une unité lexicale, vers la mise en commun des régimes identiques dans le but de les regrouper de façon ontologique. Il est tout d'abord question des différentes tâches accomplies, dont l'établissement d'une liste des dépendances syntaxiques qui sont gouvernées par la valence des noms en français, pour mener à bien la création de la base de données. Par la suite, en guise de résultats, une présentation sommaire de celle-ci est donnée.

Finalement, en guise d'ouverture, nous proposons quelques projets futurs qui pourraient suivre notre travail. Plus particulièrement, nous tentons de proposer quelques étapes visant la création d'une typologie des régimes nominaux selon les résultats obtenus. Sans être ni définitive ni entièrement satisfaisante, cette esquisse de typologie permet d'effectuer des regroupements au sein des régimes lexicaux. En conclusion, nous mentionnons quelques limites de notre travail et nous terminons le mémoire par un bilan.

## CHAPITRE 1

### PROBLÉMATIQUE ET OBJECTIFS DE RECHERCHE

Ce premier chapitre, en s'appuyant sur un survol d'ouvrages généraux de lexicographie ainsi que sur une brève analyse des informations syntaxiques contenues dans différentes bases de données lexicales, expose notre problématique et précise nos objectifs de recherche.

Nous y verrons que lorsque les lexicologues réfléchissent sur la création des articles de dictionnaires, ils soulignent toujours l'importance de ne pas seulement décrire le sens des unités lexicales, mais aussi leur contexte typique d'utilisation et leur comportement syntaxique dans la phrase. Nous verrons également comment cela se reflète dans la pratique lexicographique en examinant comment les lexicographes s'y prennent pour inscrire différentes informations syntaxiques à l'intérieur des articles de dictionnaire. Bien que nous ayons trouvé des résultats intéressants, nos observations suggèrent qu'il reste du travail à faire, surtout en ce qui concerne la description du comportement syntaxique des lexies non verbales dans les dictionnaires d'usage de la langue française.

#### 1.1 Étude préliminaire : ouvrages généraux de lexicographie

En règle générale, les ouvrages de lexicographie accordent toujours une place aux informations syntaxiques en lien avec le lexique. Par exemple, dans *Lexicography Then and Now* de Zgusta (2006), la sous-section 2.2.1.3 est consacrée à la syntaxe. On y mentionne que, d'un point de vue lexicographique, la syntaxe est en grand développement depuis la fin des années 60. On y affirme que la syntaxe n'est plus seulement étudiée en tant que patron abstrait des phrases, mais aussi en connexion avec le lexique (ce qu'on distingue parfois en termes de macro-syntaxe et de micro-syntaxe). De plus, l'auteur soutient que l'idée de sous-catégorisation des unités lexicales est actuellement d'une grande importance dans toute théorie syntaxique. (Zgusta 2006, 101) La section 2.2.2.4 est quant à elle consacrée aux valences et collocations. On y mentionne que la

théorie syntaxique de Lucien Tesnière (1959) est utile en lexicographie ; que, en effet, les études utilisant sa notion de valence sont au centre des patrons syntaxiques dans les dictionnaires modernes (Zgusta 2006, 105-106).

Il en est de même dans la collection de textes *A Practical Guide to Lexicography* édité par van Sterkenburg (2003). Cet ouvrage regroupe des textes où différents experts traitent du lexique et de sa description dans des dictionnaires de divers types (monolingues, bilingues, spécialisés, électroniques, etc.).

On passe entre autres en revue ce qui fait généralement partie de la microstructure (structure du contenu des articles) de ces dictionnaires. On examine tour à tour les diverses données qui sont communément incluses dans les descriptions lexicographiques des dictionnaires d'usage. Parmi ces informations, on retrouve les informations synchroniques d'identification, celles qui permettent de bien identifier la lexie vedette. Ces critères d'identification sont : 1) phonologiques (l'orthographe et la prononciation) ; 2) syntaxiques (la classe du mot et la structure de ses arguments) ; et 3) morphologiques (la flexion et l'aspect, lorsque nécessaire).

Somme toute, les différents auteurs jugent que les dictionnaires doivent inclure les contraintes qu'une lexie impose sur son environnement syntaxique. Les informations syntaxiques qui sont pertinentes pour les descriptions lexicographiques sont de divers types. Tout d'abord, chacune des entrées de dictionnaire est généralement suivie par une référence à la partie du discours de la lexie vedette. Bien que très concises, ces références sont riches en informations. La seule mention « v. », par exemple, active toute la connaissance que nous avons sur les verbes (par ex., peut constituer le noyau d'une phrase, peut impliquer la présence d'arguments spécifiques, doit s'accorder en genre et en nombre avec son sujet, etc.). De plus, traditionnellement, des étiquettes du type « transitif » et « intransitif » pour les verbes ainsi que « comptable » et « non comptable » pour les noms, apportent des précisions supplémentaires. Ces spécifications indiquent ainsi dans quels contextes syntaxiques la lexie doit ou peut apparaître. Ces références permettent donc d'inclure dans l'article de façon concise ce que nous appelons la valence passive de la lexie vedette. Il est toutefois précisé que l'usage strict des étiquettes pour noter

la sous-catégorisation n'est pas souhaitable, puisque la grammaire d'une langue offre la possibilité de changer la sous-catégorie syntaxique (ex. un verbe transitif peut être utilisé en tant que verbe intransitif dans certaines constructions — *Paul lit [un livre]*). Il importe donc que l'utilisateur soit conscient des propriétés de tels raccourcis méthodologiques. Ces étiquettes permettent de percevoir rapidement, sans demander trop d'espace dans le texte du dictionnaire, certaines informations très pertinentes certes, mais assez générales.

Chaque lexie a aussi des caractéristiques syntaxiques qui lui sont spécifiques. Par exemple, dans le cas des verbes, la structure argumentale spécifie la nature et le nombre des arguments requis pour un verbe en particulier. Les verbes intransitifs sont des prédicats à une place (ex. *marcher, courir*), les verbes transitifs sont des prédicats à deux places qui sélectionnent deux arguments, un sujet et un objet (ex. *savoir, se souvenir*), les verbes ditransitifs sont des prédicats à trois places (ex. *Paul a donné un livre à Marie*). Dans les cas où un verbe requiert un complément prépositionnel (ex. *s'en prendre à qqn, être fâché contre qqn*, etc.), le dictionnaire doit toujours préciser la préposition adéquate. Finalement, on mentionne brièvement que certains noms ou adjectifs qui prennent des arguments doivent être décrits de la même façon. On donne en exemple les cas de : *relation (entre deux personnes ou groupes, d'une chose à une autre)*, *bord (de qqch.)*, *fier-adj. (de qqch.)*.

Ces informations, contrairement aux informations qui concernent la structure des phrases, doivent apparaître à l'intérieur de l'article lexicographique plutôt que dans une grammaire. Pour décrire les informations syntaxiques propres à chaque lexie, les lexicographes doivent trouver un juste milieu entre l'utilisation de nombreux codes et l'utilisation de formules plus longues et détaillées se rapprochant de ce qui est fait dans les grammaires. Ces deux façons de spécifier les différentes possibilités et restrictions se disputent l'avantage d'être plus concise et donc d'économiser de l'espace d'une part et d'être plus facilement déchiffrable de l'autre. Dans le même ordre d'idée, on peut envisager d'utiliser les exemples dans le but de permettre à l'utilisateur de voir ce qu'il peut et ne peut pas faire. En effet, les exemples sont pratiques pour voir ce qui existe, mais l'utilisateur peut seulement déduire ce qu'il ne peut pas faire. Même en indiquant des exemples agrammaticaux, on réglerait certains problèmes, mais on ne pourra jamais

inclure l'ensemble des énoncés possibles et impossibles à l'intérieur d'un dictionnaire avec des exemples d'utilisation.

Bref, dans les ouvrages traitant de lexicographie, on reconnaît l'importance des informations syntaxiques qui doivent être lexicalement spécifiées. On ne propose pas de critères précis quant au nombre ou au type d'informations qui sont nécessaires ; on laisse plutôt ce choix aux lexicographes. Le choix des informations à inclure ainsi que la méthode d'encodage dépendra de l'objectif de chaque dictionnaire, de son public cible, etc.

## **1.2 Analyse : dictionnaires et autres bases de données lexicales**

Pour dresser un panorama de la situation, voyons maintenant certains ouvrages, les informations qu'ils contiennent et la méthode utilisée pour les encoder. Cela nous permettra de vérifier si les dictionnaires et autres bases de données lexicales font état de la micro-syntaxe dans les articles lexicographiques, comme le suggèrent les ouvrages de lexicographie.

Il va sans dire que les concepteurs des dictionnaires d'usage peuvent toujours faire le choix des informations de ce type à inclure dans leurs descriptions. Le public cible et la fonction du dictionnaire auront évidemment un grand rôle à jouer dans cette sélection. Contrairement aux concepteurs de dictionnaires d'usage, les chercheurs qui mettent en branle des projets de recherche ont des contraintes de temps et d'exhaustivité moins rigides. Ceci découle du fait que les bases de données lexicales résultant de ces projets n'ont pas comme unique but d'être des objets commerciaux. Ils visent avant tout à explorer et améliorer les techniques et pratiques existantes. Ainsi, les chercheurs peuvent réfléchir à certains aspects précis, plus ou moins problématiques, de la pratique lexicographique auxquels les concepteurs de dictionnaires d'usage ne peuvent pas nécessairement accorder autant d'attention qu'ils le mériteraient. Comme nous ne pouvons évidemment pas analyser l'ensemble des dictionnaires existants, les ouvrages que nous présentons dans cette section ont été sélectionnés parce que leurs créateurs en sont arrivés, d'après nous, à des résultats intéressants en ce qui concerne l'encodage des in-

formations syntaxiques (LDOCE, COBUILD, BBI, FrameNet, Dictionnaire du français usuel, Tables du LADL, DICOVALENCE) ou qu'ils sont typiques de ce qui est pratique courante en lexicographie française (Petit Robert, TLFi).

## 1.2.1 Produits de la pratique lexicographique anglophone

### 1.2.1.1 LDOCE

Le *Longman Dictionary of Contemporary English* (LDOCE) est produit par Pearson Longman, une division de Pearson Education. En général, tous les dictionnaires de Pearson Longman ont un point commun, ils sont principalement conçus pour favoriser l'apprentissage de l'anglais. Les concepteurs ont utilisé le *Longman Learner's Corpus* (des séquences écrites d'essais et d'examens d'étudiants totalisant 12 millions de mots)<sup>1</sup> pour les aider dans la sélection des informations à inclure dans leurs dictionnaires. Ainsi, les dictionnaires de cette compagnie sont créés de façon à aider les apprenants à ne pas commettre les erreurs les plus typiques relevées dans leur corpus. En ce qui concerne le LDOCE, les concepteurs énoncent clairement son utilité dans la présentation de l'ouvrage en statuant : « Each time we produce a new edition of the *Longman Dictionary of Contemporary English*, our aim is always the same : what can we do to make the dictionary more helpful for advanced level students of English » (Summers 2005, x).

Un effort est donc fait pour fournir aux étudiants des exemples et des listes de collocations, ce qui est décrit comme étant des « words that are typically used together » (Summers 2005, x). La raison première de cette démarche est que leur expérience leur a appris que les étudiants trouvent que « it is easier both to remember and understand the word within a context. » (Summers 2005, x) Dans une entrée, les collocations sont soit soulignées en gras dans les exemples ou avant les exemples, soit incluses dans une case particulière lorsqu'elles sont très nombreuses. La raison pour inclure systématiquement les collocations (ou cooccurrences) les plus courantes retrouvées dans leur corpus, est que cela permet aux étudiants d'arriver à lier les mots ensemble lors de la production orale ou écrite.

---

1. <http://www.pearsonlongman.com/dictionaries/corpus/index.html>

Toujours dans l'optique d'aider les étudiants à utiliser adéquatement les mots en contexte, la dernière édition du LDOCE contient plusieurs nouvelles informations dont les « Grammar boxes ». Ces encadrés fournissent diverses informations grammaticales (partie du discours, nom comptable ~ non comptable, verbe transitif ~ intransitif, etc.). Parmi celles-ci se trouvent les patrons syntaxiques les plus réguliers qui contiennent la lexis vedette. Les prépositions les plus couramment utilisées sont également données.

Bien évidemment, comme ces informations sont données dans un encadré, des abréviations sont nécessaires pour cet encodage. Une liste des « Grammar codes and patterns » utilisés dans le LDOCE est donc fournie en page ii pour rendre l'usage du dictionnaire accessible à tous. On y mentionne entre autres que la section « Grammar » fournit des informations telles que :

« [T] *transitive* a verb that is followed by an object, which can be either a noun phrase or a clause : *I love chocolate.* | *She said she was to busy.* | *I remember going on holiday there.*

[always + adv/prep] shows that a verb must be followed by an adverb or a preposition : *She went upstairs.* | *Robert put the letter in his briefcase.* » (Summers 2005, ii)

Tandis que la section « Patterns » fournit des informations telles que :

« [about] [along] shows that a word can be followed immediately by a particular preposition or adverb : *I'm worried about Rachel.* | *The children skipped along.*

*try to do sth or order sb to do sth* shows that a word can be followed by an infinitive : *Try to forget about it.* | *He ordered them to leave.* | *He's sure to win.* » (Summers 2005, ii)

### 1.2.1.2 COBUILD

Selon les créateurs du COBUILD, ce dictionnaire « offers more, and more accurate, information on the way modern English is used than any previous dictionary. It is a useful guide to writing and speaking English as well as an aid to reading and understanding » (Sinclair 1987, xv). Ces derniers affirment également que, « [t]he quality of information in this dictionary is different from others. With our textual evidence it is

possible to be precise about the shape of phrases and the extent of their variation ; the relative importance of different senses of a word ; and the typical environment in which a word or phrase is used » (Sinclair 1987, xv).

D'après ce qui est indiqué dans l'introduction de l'ouvrage, les exemples et les explications au sein du dictionnaire fournissent une illustration réelle des unités lexicales décrites dans leur contexte grammatical typique et permettent ainsi de montrer comment elles peuvent être utilisées. Évidemment, de telles informations sont particulièrement pertinentes en ce qui nous concerne. Mais, les concepteurs du COBUILD ne se sont pas arrêtés là. En effet, ils ont décidé d'inclure une colonne supplémentaire [=extra column] qui permet d'augmenter le nombre d'informations systématiques données aux utilisateurs. Voici une description partielle de cette colonne :

« The main aim of the Extra Column is to provide the more advanced student with structural information, extending the statements of the main entry. It gives a word class, and indicates the range of syntactic uses. It points out typical syntactic patterns and choices available to the user of English, often going beyond what it is possible to give in examples. [...] So the Extra Column contains a large number of grammatical notes about individual words and senses. More than that, it draws attention to grammatical patterns which are common for that word in that sense. It is rare for grammar books to separate the patterns which actually occur from the ones which can in theory occur ; it is also rare for grammars to be careful about relating the grammar to an individual sense of a word. The Extra Column entries thus provide very substantial information which is otherwise difficult to record in a dictionary, but which is immensely useful for composing speech and writing in English. It is important, therefore to understand its function properly. » (Sinclair 1987, xvi)

Il est important de savoir que ce dictionnaire a été conçu sur la base d'un corpus de données textuelles informatisées de l'Université de Birmingham au Royaume-Uni, la Bank of English. Cette base a été constituée à partir de textes littéraires, journalistiques, techniques, etc. et de corpus de la langue orale. Ainsi, ses informations sont celles que l'on retrouve effectivement plus régulièrement dans la langue. Donc, cette co-

lonne, « [f]rom all of the virtually endless combinations and permutations it picks out what is typical of the use of the word as found in the text collection. Hence the entries for apparently similar words or senses will only be identical if they are used in just the same way as each other » (Sinclair 1987, xvii).

Afin de voir comment cette colonne fonctionne plus exactement et ce qu'elle permet de savoir à propos de l'usage des unités lexicales, nous fournissons des exemples de ce que l'on peut y trouver. Nous ne transcrivons que les informations qui nous intéressent (c'est-à-dire, les informations grammaticales et les patrons de constructions syntaxiques). Nous omettons donc les informations sur la fréquence, les informations pragmatiques, les synonymes et antonymes ainsi que les dérivations.

likely 1 ADJ : oft <i>it</i> v-link ADJ that	probable ADJ : oft <i>it</i> v-link ADJ that
2 ADJ : v-link ADJ to-inf	
try VERB V to-inf V adv V -ing V	attempt VERB V to-inf V n
try N-COUNT	attempt N-COUNT : usu with supp. oft N to-inf

Nous pouvons donc voir que des lexies sémantiquement proches (LIKELY~PROBABLE; TRY-v~ATTEMPT-v; TRY-n~TRY-n) n'ont pas exactement le même comportement syntaxique. De ce fait, il est approprié de décrire ces comportements spécifiques à l'intérieur du dictionnaire.

Ce que les « grammar boxes » du LDOCE et les « extra columns » du COBUILD permettent, c'est d'inscrire ce supplément d'informations sans trop alourdir le texte même de l'article. Ainsi, les utilisateurs intéressés pourront s'y référer et ceux qui veulent plus simplement consulter l'article de dictionnaire ne seront pas distraits par un surplus d'information syntaxique.

### 1.2.1.3 BBI

Nous présentons maintenant le *BBI Dictionary of English Word Combinations*. Comme son titre l'indique, l'accent est mis sur la combinaison des mots de l'anglais. Ceci découle du fait que pour acquérir une maîtrise active d'une langue donnée, en plus de maîtriser la prononciation, les formes et le sens des mots, il est primordial d'apprendre à les combiner adéquatement en syntagmes, phrases et textes (Benson et al. 1986, ix). Les créateurs de ce dictionnaire formulent l'idée que, dans les langues, certaines lexies vont naturellement avec d'autres ou s'utilisent dans des constructions grammaticales spécifiques. Ces combinaisons récurrentes, semi-figées se nomment collocations. Les auteurs ont distingué deux types de collocations : les **collocations grammaticales** et les **collocations lexicales**. Les collocations lexicales correspondent aux collocations de la LEC, que l'on encode dans le dictionnaire par les fonctions lexicales. Quant aux collocations grammaticales, elles sont encodées dans le régime lexical. C'est donc cette partie du travail d'encodage du BBI qui nous intéresse particulièrement.

Comme le BBI est un dictionnaire de collocations de l'anglais, on a choisi de ne pas encoder les combinaisons libres. Alors, contrairement au régime lexical du DiCo, on ne retrouvera pas systématiquement la façon d'encoder tous les actants de la lexie vedette. Par exemple, dans le type de collocation G1 (nom + préposition) les combinaisons nom + *of* ne sont généralement pas incluses dans les entrées lexicales d'un nom. C'est qu'on a jugé que ces combinaisons étaient construites de façon libre selon les fonctions de cette préposition, en suivant les règles de l'anglais. Ainsi, dans le BBI, on inclut les collocations *blockade against* et *apathy towards*, mais pas *blockade of* et *apathy of*. Cela en raison du fait que la préposition *of* s'utilise avec presque tous les noms de l'anglais pour introduire des dépendants syntaxiques de nature variée (correspondant à l'objet

direct ou au sujet d'un verbe, dénotant un concept de 'possession', etc.). Donc, seulement les patrons de combinaison impliquant une préposition particulière (ex. *against, towards*) ou une structure syntaxique spécifique (ex. *to + infinitive, that + clause*, etc.) on été inclus.

Les différentes collocations grammaticales identifiées par les lexicographes ont été regroupées en types (= *collocational types*). Bien entendu, les patrons de description utilisés ici sont assez généraux. Ainsi, une unité lexicale peut facilement appartenir à plus d'un groupe. Pour une description complète de ces types de collocations et pour plus d'exemples correspondants, voir les pages xv à xxix de l'introduction du BBI.

G1 nom + préposition (sauf *of* et *by*)

G2 nom + infinitif

nom + verbe en *-ing*

G3 nom + *that* (*that* qui ne pourrait pas être *which*)

G4 préposition + syntagme nominal

G5 adjectif + préposition

adjectif + *of* (le sujet est animé)

G6 adjectif + *to* + infinitive

adjectif + verbe en *-ing*

G7 adjectif + *that*...

G8 Ce dernier groupe comprend 19 patrons verbaux de l'anglais. Chacun de ces patrons est désigné par une lettre majuscule (de A à S) que l'on retrouve dans les entrées lexicales verbales du BBI. Voici à titre d'exemple la description du groupe R.

[R] = *s(it)vo to inf.* (or) *s(it)vo that-clause*

Ces verbes transitifs (qui expriment souvent des émotions) sont précédés du pronom vide *it* et suivis par *to*+infinitives ou par *that*+proposition. Ex. *it puzzled me that they never answered the telephone, it surprised me to learn of her decision.*

Afin de voir concrètement comment cela se transpose dans la pratique lexicographique, nous présentons un exemple de la section sur les collocations grammaticales d'une entrée lexicale du BBI.

**attempt** *n.* [...] 10. an ~ against, on (an ~ on smb.'s life) 11. an ~ at (an ~ at being funny) 12. an ~ to + inf. (we made an ~ to get in touch with them) 13. in an ~ (in our ~ to finish the job quickly, we ran out of supplies)

Ce qui est intéressant dans la pratique lexicographique observée ici, c'est d'une part d'avoir pris en considération les collocations grammaticales des verbes, noms, adjectifs

et prépositions. Contrairement à certains autres ouvrages, on met donc en évidence le fait que les verbes ne sont pas les seuls à devoir être utilisés dans des constructions syntaxiques particulières. D'autre part, nous percevons une volonté de créer des classes de noms, adjectifs et, surtout, de verbes. Ceci a permis aux auteurs de tirer des conclusions telles que : les verbes du groupe R expriment souvent des émotions. Sans que le désir de tirer des conclusions de ce type soit un des buts évidents de la démarche des créateurs du BBI ou que ce soit fait de façon systématique, nous croyons que cela est une piste de réflexion intéressante. Nous croyons en effet que cela démontre que de regrouper les lexies selon leur patron syntaxique afin de voir s'il est possible de percevoir certaines généralités sur les lexies qui composent ces groupes est envisageable.

Toutefois, il est de mise de mentionner que l'encodage des collocations grammaticales n'est pas suffisant afin de décrire le régime syntaxique des unités lexicales. En effet, les actants des lexies vedettes ne sont pas encodés dans le BBI s'ils s'expriment de façon régulière, ils ne le sont que si le patron syntaxique dans lequel ils apparaissent est jugé assez particulier par les lexicographes. De plus, on regroupe dans un même groupe certaines lexies qui ont des caractéristiques grammaticales, ou exigences syntaxiques, assez différentes. Par exemple, certains verbes du groupe P sont transitifs, tandis que d'autres sont intransitifs. De plus, les compléments directs des groupes de verbes où le patron décrit l'expression d'un autre complément (par un groupe prépositionnel, une proposition, etc.) peuvent être soit obligatoires, soit optionnels. Bien entendu, on précise toujours ces distinctions au sein des entrées lexicales le cas échéant. Toutefois, nous désirons encoder ces informations en suivant une méthode d'encodage plus systématique, tant au niveau de l'exhaustivité que de la rigueur. Logiquement, nous aurons besoin d'un nombre plus élevé de patrons de description pour décrire le lexique.

#### 1.2.1.4 FrameNet

Après avoir examiné des dictionnaires, nous souhaitons nous consacrer à la base de données de langue anglaise FrameNet<sup>2</sup>. Ce projet, fondé sur la sémantique des *frames*

---

2. FrameNet est disponible à l'adresse suivante : <http://framenet.icsi.berkeley.edu/>. Pour de plus amples informations, le lecteur intéressé par ce projet pourra également consulter le numéro spécial de

de Fillmore (1977), se construit principalement grâce à l'analyse de phrases d'exemples trouvées en corpus. Dans le chapitre 1 de *FrameNet II : Extended Theory and Practice* ainsi que sur le site internet de FrameNet, les concepteurs résument leur projet ainsi :

« The Berkeley FrameNet project is creating an on-line lexical resource for English, based on **frame semantics** and supported by corpus evidence. The aim is to document the range of semantic and syntactic combinatory possibilities — **valences** — of each word in each of its senses, through computer-assisted annotation of example sentences and automatic tabulation and display of the annotation results. »  
(Ruppenhofer et al. 2010, 5)

Donc, pour créer les annotations de FrameNet, une unité lexicale prédicative est d'abord identifiée comme cible (=target) dans les phrases d'exemples. La première étape du travail d'annotation, est d'identifier le *semantic frame* évoqué par la lexie. Un *semantic frame* est « a script-like conceptual structure that describes a particular type of situation, object, or event along with its participants and props » (Ruppenhofer et al. 2010, 5). Ensuite, on annote les dépendants de cette lexie sur trois niveaux. On assigne d'abord au dépendant un *frame element* ou FE. Puis, on identifie automatiquement sa fonction grammaticale (par exemple **Objet**) et le type de constituant dont il s'agit (par exemple **SN**). Dans le cas le plus simple, la lexie qui évoque un cadre est un verbe et les FEs sont ses dépendants syntaxiques. Par contre, on précise également que certains noms et adjectifs peuvent aussi être évocateurs de *frame* (Ruppenhofer et al. 2010, 5-6). C'est à l'aide de ce travail d'annotation que sont construites les entrées lexicales de FrameNet. Comme dans la pratique traditionnelle en lexicographie, les entrées lexicales des lexies décrites dans FrameNet viennent avec des définitions. De plus, une multitude d'exemples annotés pour chaque lexie sont fournis (environ 20 par lexie). La particularité de FrameNet est que ceux-ci illustrent toutes les possibilités combinatoires de la lexie en question. Ces exemples permettent de spécifier comment les FEs sont réalisés dans les constructions syntaxiques qui ont la lexie décrite comme tête.

Afin de rendre le tout plus clair, voici l'exemple de la lexie `try.n`.

## try.n

### Frame: Attempt

#### Definition:

COD: an effort to accomplish something; an attempt.

Support(s): give, have, make

### Frame Elements and Their Syntactic Realizations

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

Frame Element	Number Annotated	Realization(s)
Agent	(14)	CNI.-- (2) DNI.-- (1) NP.Ext (5) PP[by].Dep (1) PP[from].Dep (2) Poss.Gen (3)
Goal	(12)	AJP.Dep (1) VPto.Dep (2) PPing[of].Dep (1) PPing[at].Dep (1) NP.Obj (1) DNI.-- (2) CNI.-- (1) PP[at].Dep (3)

Figure 1.1 – Extrait de l'entrée lexicale de try.n dans FrameNet

D'après l'entrée lexicale, le *semantic frame* de `try.n` est `Attempt` ; sa définition est : « an effort to accomplish something; an attempt » ; elle a comme verbes supports « give, have, make ». Ses FEs et leurs réalisations syntaxiques sont indiqués dans un tableau. Bien entendu, il faut être familier avec les codes utilisés pour déchiffrer l'ensemble de ces informations. Ce que le tableau indique, c'est que dans le cas de la lexie `try.n`, son `Agent` a été annoté 14 fois et son `Goal` 12. Dans la colonne des réalisations, on retrouve tour à tour le type du constituant (par ex., `PP [at]` est utilisé pour les syntagmes prépositionnels en *at* ; `POSS` est utilisé pour les syntagmes nominaux possessifs incluant les déterminants possessifs et les syntagmes nominaux comprenant le clitique génitif anglais 's), sa fonction grammaticale (par ex., `Dep` est la fonction grammaticale générale qui est assignée aux dépendants, *arguments* et *ajouts*<sup>3</sup> ; `Gen` est la fonction assignée à tout syntagme nominal possessif qui fonctionne comme un déterminant), ainsi que le nombre d'annotations.

On peut donc ici comprendre que la lexie `try.n` peut notamment être utilisée avec un déterminant qui exprime son premier actant (`Agent`) et un complément prépositionnel en *at* qui exprime son deuxième actant (`Goal`) ; comme dans : « His only previous TRY at the marathon was 2 :15 :27[...] ». Pour l'ensemble des codes utilisés dans les annotations ou pour de plus amples explications et exemples, voir Ruppenhofer et al. (2010).

Dans notre optique, ce projet est particulièrement intéressant parce qu'un de ses buts premiers est d'arriver à dériver les valences des unités lexicales en annotant automatiquement des phrases réelles (les phrases d'exemples retenues proviennent du British National Corpus). Une autre de ses tâches principales est de relier indirectement les unités lexicales par le biais de *frame-to-frame relations*. Le fait d'établir un réseau de relations qui lient les unités lexicales entre elles rapproche FrameNet des thésaurus et ontologies. Nous n'entrerons pas en détail dans la description de ces relations, bien que ce soit une partie centrale du travail fait par FrameNet. D'ailleurs, une comparaison avec les fonctions lexicales paradigmatiques de la LEC pourrait être faite pour certaines de

3. On mentionne que l'on rend compte de cette distinction grâce à la notion de « coreness status assigned to a frame element » (Ruppenhofer et al. 2010, 66) et que cela n'est pas repris par des distinctions de fonctions grammaticales.

ces relations (à ce sujet, voir Coyne et Rambow (2009)). De plus, l'étendue du lexique anglais couvert par ce projet est importante, ce qui fait que cette méthode d'encodage a été éprouvée et améliorée avec le temps. Ceci nous laisse croire qu'il nous serait possible de développer une méthode de description rigoureuse et efficace qui puisse être applicable à la création d'une base de données lexicales couvrant une part importante du lexique français, même si notre méthode ne serait pas basée sur la sémantique des *frames*.

## **1.2.2 Produits de la pratique lexicographique francophone**

### **1.2.2.1 Petit Robert**

Outre « le recensement et l'analyse des significations » qui est le « projet fondamental d'un dictionnaire de langue » (Rey-Debove et Rey 2008), l'équipe du Petit Robert accorde une importance particulière à l'histoire et au patrimoine de la langue française, aux variétés du français (régionalismes, niveaux de langue, domaines du savoir), aux relations de sens entre les lexies (arborescence des significations et système analogique) ainsi qu'au caractère vivant de la langue française. C'est dans cette dernière optique qu'elle a procédé à une mise au point de l'orthographe dans la version 2008 du dictionnaire. En ce qui concerne le comportement syntaxique des unités lexicales, les deux responsables du projet, dans Rey-Debove et Rey (2008), déclarent au sujet de leur travail :

« On voit que la différence fondamentale entre le vrai dictionnaire de langue et les ouvrages apparentés réside dans le programme d'information sur le signe. Ce programme n'est pas seulement lié à la reconnaissance et à la compréhension du mot ; il doit permettre la production des phrases en montrant comment le mot s'emploie à l'écrit comme à l'oral. La prononciation est fondamentale, mais aussi l'entourage du mot, les collocations (mots qui apparaissent souvent en même temps), les locutions (expressions figées), et aussi les constructions syntaxiques, les difficultés d'emploi. En somme, ce qui est absolument nécessaire, après la définition, c'est une large exemplification où le mot se trouve dans des contextes attendus, ce qu'on appelle la phraséologie. »

Ainsi, on peut constater que les créateurs du Petit Robert se fient surtout sur la présence d'exemples dans les articles lexicographiques pour encoder les informations syntaxiques spécifiques à chaque unité lexicale. Ils utilisent à la fois les exemples (créés par le lexicographe) et les citations (empruntés à des écrivains) pour « montrer le mot en action, sa place dans la phrase ». (Rey-Debove et Rey 2008) Comme nous l'avons affirmé auparavant, les exemples sont utiles, mais ne peuvent pas constituer une méthode d'encodage systématique et exhaustive.

### 1.2.2.2 Trésor de la langue française informatisé

Dans le *Trésor de la langue française informatisé* [TLFi] (Dendien et Pierrel 2003)<sup>4</sup>, le texte des articles est balisé en plusieurs objets textuels, ce qui permet d'effectuer des recherches pointues sur divers éléments. Plusieurs de ces objets sont susceptibles de contenir de l'information syntaxique — notamment, les « codes grammaticaux » (par ex., verbe trans.), les « indicateurs » (par ex., *Emploi trans.*, *Emploi pronom.*), les « crochets » (par ex., [L'agent, exprimé ou non, est un être vivant], [L'objet désigne un arbre, une plante, etc.], [L'agent est une chose]), les « syntagmes » (par ex., *Blessier la vue, les regards de qqn.*, *Blessier l'amour-propre de qqn.*), et surtout les « constructions » (par ex., **Blessier qqn.**). Ces exemples sont tous tirés de l'entrée de BLESSER.

Au moins un projet vise à extraire les informations de sous-catégorisation à partir du TLFi (Jacquey et Naels 2005). Les deux auteurs rappellent : « Concernant le TLFi, il faut rappeler qu'il n'a pas été conçu dans une perspective d'exhaustivité syntaxique et que les objets « construction » analysés ne décrivent donc pas l'ensemble des constructions syntaxiques dans lesquelles un verbe peut apparaître. D'autre part, l'utilisation de ces objets « construction » a correspondu à une époque de rédaction et n'est donc pas également représentée sur l'ensemble du dictionnaire. » Elles mentionnent aussi que « les informations utiles pour un cadre de sous-catégorisation peuvent être éclatées sur plusieurs niveaux hiérarchiques. Le cas du verbe provoquer est éloquent. Dans son emploi B, on trouve la construction qqn/qqc provoque qqn à, puis en B.1. + subst et + inf. en B.2. Le croisement de B et B.1 permettrait d'obtenir le cadre plus précis qqn/qqc provoque qqn à

---

4. <http://atilf.atilf.fr/>

+ subst, et cela de même en croisant B et B.2 ». C'est donc dire que, même lorsque présentes, les informations d'ordre syntaxique n'ont pas de place spécifique dans l'article de dictionnaire et aucune méthodologie particulière de description n'est prévue. Ces informations se retrouvent à plusieurs endroits différents au sein d'un article et ne sont en aucun cas données systématiquement sur l'ensemble des unités lexicales décrites dans le dictionnaire. Ceci n'est pas unique au TLFi, on voit également ce phénomène dans la plupart des dictionnaires de langue française.

### 1.2.2.3 Dictionnaire du français usuel

Dans le *Dictionnaire du français usuel* de Picoche et Rolland (2002), on a décidé de prioriser les informations concernant les actants sémantiques. Dans le texte présentant la structure des articles, les auteurs mentionnent que « [l]a phrase-titre est toujours accompagnée de sa structure actancielle abstraite, car il est évident qu'un exemple est toujours trop particulier pour être entièrement significatif » (Picoche et Rolland 2002, 14). C'est ainsi que la définition de chacune des lexies est composée de variables numérotées. Lors de transformations syntaxiques d'une phrase (lors de la formation du passif, par exemple) ou d'une nominalisation, les numéros attribués aux actants demeurent les mêmes (Picoche et Rolland 2002, 14). En général, la structure actancielle utilisée ici est assez semblable à la forme propositionnelle utilisée dans le DiCo. On donne parfois aux actants des « déterminants ». Ex. « **Luc porte sa valise à la gare** : A1 humain *porte* A2 concret à A3 spatial. - **Luc porte un blouson noir** : A1 humain *porte* A2 vêtement. » (Picoche et Rolland 2002, 15)

Sans toutefois, accorder une place primordiale à la réalisation syntaxique des actants sémantiques d'une lexie, les concepteurs du dictionnaires ont trouvé nécessaire dans certains cas de faire quelques précisions. Ils soutiennent que, « [d]ans l'ensemble, les structures actanciennes suffisent à préciser la syntaxe particulière aux mots étudiés. Néanmoins, il nous a paru parfois nécessaire de signaler par abréviation GR. une sous-partie consacrée à certaines remarques grammaticales : certaines règles d'accord, certaines particularités morphologiques, les constructions d'un verbe à la syntaxe un peu compliquée » (Picoche et Rolland 2002, 14). En résumé, les auteurs considèrent

aussi que certaines informations syntaxiques doivent figurer dans le dictionnaire plutôt que dans une grammaire. Leur façon d'encoder ces informations se rapproche quelque peu de la méthode du Longman et du COBUILD dont nous avons parlé précédemment. Tout en étant intéressant, ce dictionnaire ne couvre qu'une partie relativement restreinte du lexique français, et ne peut donc pas être utilisé comme dictionnaire d'usage. La méthode utilisée n'est pas non plus systématique, les créateurs du dictionnaire se sont donné un outil assez général pour inscrire des informations de divers types lorsqu'ils le jugeaient nécessaire. Par contre, parmi les dictionnaires de langue française, il est l'un des seuls à porter un intérêt particulier aux actants sémantiques des lexies décrites et à leur expression dans la phrase.

Pour clore le sujet de la pratique lexicographique française traditionnelle, il est à mentionner que d'autres dictionnaires de langue française de moins grande envergure décrivant les valences des verbes ont déjà été conçus. Nous pouvons mentionner par exemple : Bogacki et al. (1983), Frînculescu (1978), Caput et Caput (1969). Toutefois, ces dictionnaires ne sont que très peu utilisés parce que peu connus et même parfois difficiles à obtenir. En ce qui concerne les lexies des autres parties du discours, aucun dictionnaire d'usage grand public ne consigne, à notre connaissance, les informations sur leur valence sur l'ensemble du dictionnaire. Bien que ne provenant pas d'une visée lexicographique, les tentatives les plus notoires à cet égard sont les tables du LADL (Laboratoire d'Automatique Documentaire et Linguistique).

#### **1.2.2.4 Tables du LADL**

Nous souhaitons aborder ici le travail de Maurice Gross tel que présenté dans *Méthode en Syntaxe : régime des constructions complétives* (Gross 1975). Malgré l'année de publication, nous considérons ce texte intéressant dans la mesure où il est à la base d'une méthode de description formelle des langues qui est encore utilisée dans la pratique aujourd'hui. Au départ, M. Gross soulignait :

« [I]'approche très répandue que nous critiquons et que nous proposons de remplacer, consiste essentiellement à effectuer des observations isolées dans des ré-

gions différentes de la structure linguistique. Ces observations ne sont jamais systématiques, et les « trous » qui les séparent ne sont pas explorés empiriquement. Les constructeurs de modèles laissent à l'induction le soin de combler ces trous. » (Gross 1975, 9)

Il souhaitait donc changer la pratique de description alors courante en tâchant d'abord de compiler des données exhaustives pour ensuite faire des généralisations qui peuvent être importantes du point de vue théorique.

Sa démarche se fonde donc sur l'étude systématique des propriétés syntaxiques (ici des verbes). Ces propriétés syntaxiques très variées peuvent être soit distributionnelles, soit transformationnelles. Parmi les distributionnelles on retrouve :

- les compléments nominaux sont soit hum, N-hum ou Nnr (« non restreints ») ;
- la distribution du groupe nominal le fait Qu P ;
- l'énumération des sous-structures acceptables (présence obligatoire des compléments, seul ou en combinaison avec d'autres) ;
- la présence de déterminants spécifiques, adverbes de manière obligatoires, négation, etc. ;
- les entrées pronominales (s'évanouir, l'emporter) ;
- l'auxiliaire adéquat ;
- la nature de la préposition dans les compléments prépositionnels.

Et parmi les transformationnelles :

- la possibilité de changer le temps grammatical dans les complétives ;
- les effacements de préposition ;
- les réductions de complétives ;
- les pronominalisations ;
- la formation du passif ;
- l'extraposition ;
- les permutations ;
- l'interprétation des formes réflexives ;
- l'introduction du discours direct ;
- la possibilité de remplacer un verbe par une séquence verbe opérateur (ou verbe support) + N (ex. *Paul adore Marie* ~ *Paul a de l'adoration pour Marie*) ;
- le comportement des verbes si on leur ajoute un préfixe (*associer* ~ *désassocier*).

La démarche de M. Gross mène ainsi à la construction d'un lexique-grammaire d'environ trois mille verbes. Pour chacun de ces verbes, il présente sous forme de table, sa

distribution selon une centaine de propriétés syntaxiques. Chacun des verbes décrits est donc inclus dans un tableau rectangulaire où les lignes correspondent aux entrées verbales et les colonnes aux diverses propriétés étudiées. À l'intersection de chacune des lignes et colonnes se trouve soit le signe « + » ou le signe « - », ce qui indique si le verbe en question possède ou non la propriété donnée. Chacune de ces tables correspond à une classe de structures. Ces regroupements sont faits selon la (ou les deux) structure principale des verbes de la table. Par exemple, la table 3 inclut les verbes qui peuvent entrer dans la construction « N0 V N1 V1-inf  $\Omega$  » (le verbe en question précédé d'un nom, suivi d'un complément nominal, d'un complément verbal et le complément de ce deuxième verbe. Comme dans l'exemple *Paul envoie la chaise s'écraser contre le mur.*). Ce sont des verbes dits « causatifs de mouvement ».

Comme généralisation possible découlant de sa méthode de description exhaustive et rigoureuse, M. Gross mentionne entre autres la notion de classe grammaticale. Il affirme que les classes grammaticales sont en fait un ensemble de propriétés que partagent un certain nombre de verbes. Il soutient que, suite à son travail, le dénombrement des verbes faisant partie d'une classe était fait de façon plus pertinente. Il affirme aussi que les tables « ont également permis de cerner une limite dans la notion de classe » (Gross 1975, 214). En effet, avec leur centaine de propriétés, les chercheurs du LADL en sont arrivés à séparer les 3000 entrées en 2000 classes (comprenant en moyenne 1,5 verbe). M. Gross soutient toutefois qu'il n'existe pas de verbes qui partageraient toutes leurs propriétés syntaxiques ; si des critères étaient ajoutés à ceux déjà utilisés, tous les verbes seraient dans des classes distinctes. Il faut donc en conclure que plus on examine de propriétés, plus le nombre de classes sera grand et moins chacune d'elles comportera d'éléments constitutifs.

Avant de terminer ce survol du travail de M. Gross, il est important de mentionner que de nombreux travaux ont été réalisés dans l'héritage de M. Gross au LADL. D'ailleurs, des travaux ont été réalisés et publiés dans le cadre du lexique-grammaire sur d'autres éléments prédicatifs (adjectifs, noms, adverbes, expressions figées) depuis les années 1970. On peut télécharger ou visualiser les tables des verbes et autres éléments prédicatifs à l'adresse suivante : <http://infolingu.univ-mlv.fr/> (dans le menu à gauche,

sélectionner : Données Linguistiques, et ensuite Lexique-Grammaire).

Nous retenons de la démarche de M. Gross que la séparation entre la grammaire et le lexique dans la description linguistique n'est pas souhaitable. Nous adhérons à son idée selon laquelle il est plus productif de s'attarder à la description du lexique pour formuler des généralisations ensuite que de donner la priorité absolue à la grammaire, comme c'est trop souvent le cas en linguistique. Bien entendu, les propriétés syntaxiques utilisées ici vont beaucoup plus loin que la simple description de la valence des unités lexicales prédicatives.

### 1.2.2.5 DICOVALENCE

La base de données lexicale DICOVALENCE (anciennement PROTON) (den Eynde et Mertens 2006) est un dictionnaire de valence verbale du français développé par Karel Van den Eynde et Piet Mertens à la K.U. Leuven (Université catholique de Louvain) en Belgique. Ce dictionnaire est élaboré dans le cadre théorique de l'approche pronominale (Blanche-Benveniste 1987; Van den Eynde et Blanche-Benveniste 1978; Van den Eynde et Mertens 2003). Comme son nom le suggère, « [l']*approche pronominale* aborde l'étude de la valence à partir des paradigmes de pronoms qu'acceptent le verbe (ou tout autre constructeur, tel l'adjectif prédicatif). Le but est de caractériser le statut et les propriétés syntaxiques de chaque élément régi. » (Van den Eynde et Mertens 2003, 1) Le projet de construire le dictionnaire de valence verbale est né en 1986. En plus de permettre de vérifier l'utilité de l'approche pronominale, on percevait également l'importance qu'il pourrait prendre dans l'enseignement et dans le traitement automatique de la langue. La première phase de son développement s'est déroulée de 1986 à 1992. Ensuite, la base PROTON a entièrement été revue en 2006 donnant naissance au DICOVALENCE dans son état actuel.

La particularité de ce dictionnaire est donc d'indiquer pour chaque position actancielle d'un verbe le paradigme de pronoms qui peuvent y figurer ; ce qui rend compte des différentes lexicalisations possibles de ses actants. On préconise l'utilisation des pronoms plutôt que celle des entités syntagmatiques (GN, GV, etc.), des fonctions syntaxiques (SUJET, OBJET, etc.), ou des rôles (AGENT, PATIENT, etc.). Pour les auteurs,

l'avantage que présentent les pronoms est qu'ils sont des éléments propres à la langue qui existent en nombre limité et fermé. Ainsi, le

« nombre restreint de pronoms permet de vérifier de façon systématique et exhaustive leurs combinaisons avec les prédicateurs, sans recourir à l'interprétation de certaines propriétés (les traits sémantiques utilisés en syntaxe) posées par le linguiste. La possibilité ou l'impossibilité d'employer tel ou tel pronom est en effet significative : les pronoms révèlent les propriétés fondamentales qu'un prédicateur impose à ses dépendants. » (Van den Eynde et Mertens 2003, 3-4)

Le dictionnaire contient la valence des 3700 formes verbales (ou vedettes de dictionnaire, vocables) les plus courantes du français (les verbes du *Petit Robert* moins ceux appartenant à un vocabulaire spécialisé et les usages archaïques), pour un total d'environ 8600 entrées (ou acceptions, lexies). Par exemple, pour une forme verbale ACHETER, quatre entrées distinctes ont été identifiées.

VAL\$ acheter: P0 P1 (P3) VTYPE\$ predicator simple  
 VERB\$ ACHETER/acheter NUM\$ 1850 EG\$ je lui ai acheté  
 quatre autocollants TR\_DU\$ kopen, afkopen TR\_EN\$ buy  
 (from s.o.) 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

VAL\$ acheter: P0 P1 (P2) VTYPE\$ predicator simple  
 VERB\$ ACHETER/acheter NUM\$ 1860 EG\$ je me suis acheté  
 une nouvelle voiture TR\_DU\$ kopen, anschaffen TR\_EN\$  
 buy s.o. sth 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 P2\$ 0, lui, leur, se réfl., se réc.,  
 l'un l'autre

VAL\$ acheter: P0 P1 (PP<pour>) VTYPE\$ predicator  
 simple VERB\$ ACHETER/acheter NUM\$ 1865 EG\$ j'ai  
 acheté une nouvelle voiture pour Paul TR\_DU\$ kopen,  
 anschaffen TR\_EN\$ buy sth (for) 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 PP\_PR\$ pour PP\$  
 0, qui, lui\_ton, eux, celui-ci, ceux-ci, l'un l'autre

VAL\$ acheter: P0 P1 VTYPE\$ predicator simple VERB\$  
 ACHETER/acheter NUM\$ 1870 EG\$ il a été acheté par ses  
 supérieurs TR\_DU\$ omkopen TR\_EN\$ bribe, buy (over) P0\$  
 qui, je, nous, elle, il, ils, on, celui-ci, ceux-ci P1\$  
 qui, te, vous, la, le, les, se réc., en Q, celui-ci,  
 ceux-ci, l'un l'autre RP\$ passif être, se faire passif

Dans DICOVALENCE, chaque entrée se divise en trois types d'information sur le verbe. La première section que les auteurs nomment « champs génériques » inclut l'identification du verbe et une annotation concise de son cadre valenciens (VAL), le type du verbe (prédicateur, copule, auxiliaire, etc.) (VTYPE), son numéro d'identification (NUM), un ou des exemples d'utilisation (EG) ainsi que la traduction du verbe en néerlandais (TR\_DU) et en anglais (TR\_EN). La deuxième section détaille le schéma valenciens du verbe. On y indique pour chaque paradigme (par exemple, P0, qui correspond au sujet, P1, qui correspond à l'objet direct, etc.), les pronoms qui peuvent être utilisés dans cette position valencienne. Enfin, on consigne dans une troisième section, lorsque nécessaire, d'autres propriétés telles que les reformulations passives et les relations de causation.

Un des avantages au fait de fournir l'ensemble des pronoms possibles pour une position valencienne est que cela permet de dériver certains traits sémantiques d'un dépendant. Par exemple, si l'on trouve la configuration de pronom « le, me, te, \*ceci », alors le dépendant sera +HUMAIN ; comme on peut le voir avec l'exemple du verbe « féliciter » (*je me/te/le félicite*, mais *\*je félicite ceci*). (Van den Eynde et Mertens 2003, 18-19)

Dans le but de démontrer la validité de la démarche, on a procédé en 2003 à une comparaison entre les entrées « dégager » / « se dégager » de PROTON et celles du dictionnaire de valence de Busse et Dubost (1983). Ainsi, le dictionnaire de Busse et Dubost compte six formulations différentes là où le lexique PROTON en distingue onze. Ceci mène les auteurs à la conclusion suivante.

« La confrontation des deux dictionnaires met en lumière des différences majeures. Le caractère rudimentaire des étiquettes morphologiques (V, se V, N, de N) employées par Busse et Dubost explique la confusion de formulations différenciées dans PROTON. [...] L'approche pronominale permet ainsi une caractérisation plus

nuancée des traits syntaxiques et sémantiques des actants. » (Van den Eynde et Mertens 2003, 30)

On retiendra de cette méthode que la distinction des différentes structures valenciennes ne peut être faite seulement avec les étiquettes morphologiques. Celles-ci ne sont pas suffisantes pour distinguer des usages différents. Les dépendants superficiellement identiques auront des propriétés syntaxiques différentes, notamment au niveau des pronoms correspondants. Il est donc important de bien distinguer les différentes constructions régies à l'aide d'autres critères que la simple identification de la catégorie grammaticale des syntagmes apparaissant comme compléments. Les tenants de l'approche pronominale le font à l'aide des paradigmes de pronoms ; nous envisageons plutôt le « nommage » précis de la relation syntaxique impliquée. Notons que nous retenons l'usage des critères pronominaux comme étant efficace pour participer à la caractérisation de ces relations.

Finalement, un des objectifs de cette base était la création d'une typologie des schémas valenciens. Comme l'avait perçu les créateurs du DICOVALENCE, « [c]e travail typologique jouerait un rôle important dans la théorie grammaticale et dans l'enseignement des langues. » Nous sommes également d'avis qu'un inventaire exhaustif des constructions attestées peut permettre de ramener la longue liste de constructions à un inventaire relativement petit des régimes possibles. Nous visons également l'intégration éventuelle d'une telle typologie dans une ontologie.

### 1.3 Problème spécifique

Dans l'optique de créer des classes sémantiques à l'intérieur du vocabulaire, J. Picoche mentionne que l'on utilise souvent la syntaxe comme outil révélateur du sens des mots. Selon elle, on procède très souvent à l'observation des propriétés externes d'une lexie, « ce qu'on appelle une « analyse distributionnelle », la « distribution » d'une forme verbale étant l'ensemble des constructions dans laquelle elle peut entrer. » (Picoche 1993, 29) Elle ajoute toutefois que :

« [l]'étude de la construction seule est cependant insuffisante. Deux constructions semblables peuvent produire des différences de sens importantes ; autrement dit une

même **structure de surface** peut correspondre à différentes **structures profondes**. [...] Il arrive aussi que deux constructions différentes produisent le même effet de sens. Cela ne veut pas dire qu'il faille renoncer à l'analyse syntaxique du sens, mais qu'il faut chercher, au-delà des propriétés externes, à l'intérieur de chaque construction, des **propriétés cachées**, que les dictionnaires ne mentionnent jamais, et qui peuvent être relevées par divers **tests syntaxiques**, qui en sont les techniques de découverte. » (Picoche 1993, 31)

C'est dire que d'indiquer la complémentation seulement en terme d'étiquettes morphologiques (syntagme nominal, syntagme prépositionnel en *de*, etc.) n'est pas suffisant. Comme nous le verrons plus loin, c'est dans l'optique d'inclure ces « propriétés cachées » aux descriptions lexicales que nous ajoutons aux régimes les noms des dépendances syntaxiques de surface qui lient la lexie vedette à ses compléments.

À la lumière de ce que nous avons observé ici, il semble d'une grande importance de trouver une méthode de description des informations syntaxiques particulières aux unités lexicales efficace qui pourrait être incluse dans un dictionnaire grand public. Nous souhaitons trouver une méthode qui, tout en permettant au lexicographe d'être précis et exhaustif, demeure assez claire pour être consultée par un grand nombre de gens et assez simple pour que l'on considère son utilisation dans un vaste travail d'édition d'articles lexicographiques.

#### 1.4 Objectifs de recherche

Comme nous avons pu le voir, les informations syntaxiques présentes systématiquement dans les dictionnaires sont plutôt disponibles pour les verbes, par la présence de codes grammaticaux comme « verbe trans. » et d'étiquettes d'emplois comme « Emploi intransitif », « Emploi pronominal », etc. Il en résulte que, d'un point de vue lexicographique, beaucoup moins d'attention a été portée aux dépendants d'autres parties du discours. Bien qu'on en reconnaisse l'importance théorique, la description du comportement syntaxique des unités lexicales prédicatives (de surcroît si elles ne sont pas des verbes) ne semble pas être pratique courante en lexicographie de langue française.

Même lorsque présentes, les informations d'ordre syntaxique sont souvent dispersées dans le texte d'un article de dictionnaire puisque les articles sont souvent divisés

en emplois (pour un verbe, on montre les différents usages intransitif, transitif, pronominal, plus parfois quelques constructions particulières qui revêtent une interprétation sémantique différente).

Comme Marengo (2009) l'a défendu au sujet de l'adjectif, nous croyons qu'il est important de détailler au sein du lexique les dépendants régis des lexies prédicatives de toutes les parties du discours à l'aide des relations syntaxiques de surface. À l'instar de S. Marengo, notre méthodologie se fonde également sur la démarche de Iordanskaja et Mel'čuk (2009), qui ont exposé les principes guides servant à établir les relations syntaxiques de surface. Suite à leur travail, ils ont établi une liste des relations syntaxiques de surface contrôlées par la valence des verbes en français. S. Marengo a quant à lui travaillé à faire l'inventaire de ces relations contrôlées par les adjectifs en français.

Dans un autre ordre d'idée, la création d'une liste des relations syntaxiques de surface du français (qu'elles soient régies ou non) est envisagée dans le cadre de la création d'une ontologie des concepts de lexicologie/lexicographie — nommée *Gros Tas de Notions* [GTN] — (Tremblay 2009, chap.4)<sup>5</sup>.

C'est dans cette optique, et dans la lignée des travaux mentionnés précédemment, que nous avons travaillé à établir un inventaire des dépendances syntaxiques de surface contrôlées par la valence des noms en français. Dans l'ontologie GTN, nous avons pris la décision d'utiliser le terme *dépendance syntaxique de surface* [DépSyntS] au lieu de *relation syntaxique de surface* [RelSyntS]. Comme une dépendance syntaxique est un type particulier de relation syntaxique, nous jugeons que ce terme est plus précis. Ceci permet entre autres de tenir compte du fait que certaines relations présentes dans les représentations syntaxiques de la TST (la Représentation Syntaxique Profonde et la Représentation Syntaxique de surface, voir section 2.2) ne sont pas des dépendances, notamment les relations anaphoriques (la relation entre un pronom et son antécédent par exemple).

Finalement, nous souhaitons en arriver éventuellement à la création d'une base de données qui contienne l'ensemble des tableaux de régime de la langue française et qui permettrait, entre autres, de comparer les unités lexicales qui partagent le même régime.

---

5. <http://olst.ling.umontreal.ca/gtn/>

Ainsi, le vaste éventail de données d'une telle base permettrait la factorisation d'informations sur le régime de lexies apparentées. Comme l'a mentionné J. Milićević, «[d]ans le cadre de la TST, la factorisation de l'information sur le régime a été mentionnée dans le contexte de la création d'articles de dictionnaire généralisés, proposée dans Mel'čuk et Wanner (1996) et Milićević (1997, 90*ssq*)), mais n'a pas encore fait l'objet d'études approfondies.» (Milićević 2009, 6) Or, nous considérons que notre projet constitue une première étape vers cet objectif ambitieux.

Nos objectifs de recherche sont donc, dans l'ordre :

- 1- de contribuer à la création d'une liste des relations syntaxiques de surface du français. Notre apport : les DépSyntS contrôlées par la valence des noms en français.
- 2- de préciser les régimes lexicaux des lexies nominales du DiCo en leur ajoutant le nom des DépSyntS.
- 3- de débiter la création d'une base de données lexicales qui contienne éventuellement l'ensemble des régimes lexicaux du français.

## CHAPITRE 2

### LE RÉGIME LEXICAL

Dans ce chapitre, nous examinons la notion de régime lexical d'un point de vue théorique après l'avoir abordée d'un point de vue plus pratique au chapitre précédent. Nous y traitons de la place qui est accordée au lexique et plus précisément au régime lexical dans des théories linguistiques qui s'intéressent particulièrement à la syntaxe. Avant d'introduire notre propre cadre théorique dans la deuxième section, nous présentons d'abord d'autres théories linguistiques afin qu'il soit possible de saisir ce que la TST a de particulier.

#### 2.1 Un survol de la place du régime lexical en linguistique

Selon l'approche adoptée, les différentes informations concernant la cooccurrence d'une lexie avec ses compléments dans la phrase (ou ses actants syntaxiques) qui doivent figurer dans le lexique sont plus ou moins nombreuses. La présence de ces informations, à la frontière entre la sémantique lexicale et la syntaxe, est rendue nécessaire, parce que la correspondance entre les arguments d'une lexie (ou ses actants sémantiques) et leurs moyens d'expression en surface n'est pas entièrement prévisible à partir de sa définition. Sans prétendre faire un inventaire détaillé des propositions qui ont été faites à travers les diverses théories linguistiques, nous présentons trois positions qui sont plus ou moins divergentes de l'approche adoptée en TST. Nous irons voir le type d'information que l'on encode en théorie dans le lexique dans trois approches différentes, soit la grammaire générative, la sémantique conceptuelle et la grammaire syntagmatique guidée par les têtes. Afin d'exposer les idées de ces trois théories, nous nous baserons sur trois textes célèbres soit : *Knowledge of Language : Its Nature, Origin, and Use*, Chomsky (1986) ; *Semantic Structures*, Jackendoff (1990) ; *Subcategorization and head-driven phrase structure*, Sag et Pollard (1989). Les références utilisées ne sont pas très récentes, mais elles sont encore très importantes dans le domaine. Nous ne prétendons pas que les idées expri-

mées n'ont pas évolué depuis le début des années 90, nous avons choisi ces références parce qu'elles apportent à la réflexion des concepts de la notion de régime lexical qui sont différents.

### 2.1.1 Grammaire générative

En premier lieu, revoyons rapidement les informations qui peuvent être incluses dans la description d'une unité lexicale à l'intérieur du lexique. Tout d'abord, selon Chomsky (1986), le lexique présente pour chaque unité lexicale sa forme phonologique abstraite et toutes les propriétés sémantiques associées à cette forme. Parmi celles-ci (outre la définition), figureront les propriétés de sélection des têtes de constructions : les noms, verbes, adjectifs et particules (prépositions ou postpositions). Ces propriétés de sélection sont de deux types. Premièrement, il y a la sélection sémantique (*s-sélection*), qui précise les rôles sémantiques et le nombre des compléments d'une unité lexicale ainsi que la sélection catégorielle (*c-sélection*), qui précise la nature syntaxique de ces compléments Chomsky (1986, 86). Toujours selon Chomsky, cette dernière spécification semble être redondante. Selon cette approche, si un verbe ou une autre tête *s-sélectionne* une catégorie sémantique donnée, il *c-sélectionnera* automatiquement la catégorie syntaxique qui est sa « canonical structure realization » (CSR). Pour démontrer ceci, il prend l'exemple du verbe *to persuade*. Nous reprenons cet exemple (dont la présentation a été légèrement modifiée) en français.

persuader (x, y, z)

ex. [Jules] a persuadé [Jean] [de l'importance d'aller au collège].

[qu'il était important d'aller au collège].

Ici, le rôle sémantique de Z est *proposition*. Le CSR de *proposition* est SC ou SN (à interprétation propositionnelle). Il est nécessaire de mentionner qu'ici on considère que *de* est une préposition sémantiquement vide utilisée comme marqueur de cas, et que *de l'importance d'aller au collège* n'est donc pas considéré comme un véritable syntagme prépositionnel.

Toutefois, les seules informations sur les rôles sémantiques (ou rôles thématiques) ne sont pas suffisantes. Nous reprenons ici les exemples 62, 63, et 64 (Chomsky 1986, 88-

89) :

- (i) I asked [what time it is]
- (ii) I wondered [what time it is]
- (iii) I (don't) care [what time it is]

- (i) I asked the time
- (ii) \*I wondered the time
- (iii) \*I cared the time

Dans ces exemples, seulement *ask* est transitif et assigne le cas que Chomsky nomme objectif. En fait, les phrases deviennent possibles si une préposition est insérée assignant le cas.

- (i) I wondered about the time (of his arrival)
- (ii) I cared about the time (of his arrival)

Ces exemples montrent que les informations sur la transitivité des verbes doivent également être présentes dans les descriptions lexicales. Notons que cela semble être de l'information qu'il est possible de préciser au sujet des verbes et prépositions, mais difficilement au sujet des têtes nominales et adjectivales. En général, selon cette approche, les problèmes de correspondance sont donc reliés à des interactions entre la théorie des cas et la théorie thématique. Ainsi, les informations catégorielles ne sont pas pertinentes dans le lexique, puisque les informations casuelles suffisent à expliquer les problèmes apparents. En cas de complication, c'est dans cette direction qu'il faut chercher des explications. Notons que cette approche vise la simplification maximale des entrées lexicales, puisque l'on cherche à minimiser la partie apprise des langues naturelles au profit de la grammaire universelle. De ce fait, la c-sélection n'est pas requise dans les informations lexicales, mais on doit tout de même préciser les informations casuelles des têtes lexicales verbales (et prépositionnelles). De plus, on doit aussi formuler des règles pour l'insertion de prépositions vides qui assignent les cas où (et seulement où) cela est nécessaire. On doit également préciser, pour une langue donnée, quelles sont les CSR correspondantes à chacun des rôles sémantiques.

### 2.1.2 Sémantique conceptuelle

Dans cette section, nous exposons brièvement l'approche sémantique de Ray Jackendoff. Son objectif de départ était de mettre au point une théorie du langage générale concernée par la sémantique et la syntaxe et leurs points de connexion. Tout d'abord, il conçoit évidemment une façon qui lui est propre de représenter le sens à l'aide de catégories conceptuelles. Par la suite, il a pu étudier entre autres ce qu'il appelle le problème de correspondance, qui est de caractériser la relation entre le traitement formel du sens et la structure formelle de la syntaxe (Jackendoff 1990, 44). Un de ses présupposés méthodologiques est que les règles de correspondances (incluant les entrées lexicales du lexique) doivent être les plus simples possible pourvu qu'elles mettent en correspondance (*map*) une structure conceptuelle universelle vers une syntaxe particulière à la langue. En ce sens, il arrive à cette conclusion :

« In terms of the simplification of lexical entries, an ideal situation would be one in which conceptual relations were reflected directly and uniformly in syntactic relations. This would permit us to formulate lexical entries just in terms of their syntactic category and their conceptual structure, since both the subcategorization feature and the syntactic position corresponding to each conceptual argument would be predictable on general principles. [...] Unfortunately, the true story is not so simple. There are many apparent mismatches between conceptual arguments - even *expressed* conceptual arguments - and syntactic positions. » (Jackendoff 1990, 155)

Tout comme les tenants de l'hypothèse précédente, il recherche donc la simplicité des entrées lexicales, mais pas la sur-simplification. Bien entendu, les modificateurs potentiels d'une unité lexicale ne sont pas encodés dans son entrée lexicale, mais certaines modifications de la structure argumentale doivent être encodées. Par exemple, il exprime le caractère optionnel de l'expression d'un argument en le plaçant entre < > dans la structure conceptuelle de l'entrée lexicale. Toujours concernant le problème de correspondance, il affirme qu'il y a plusieurs façons de rendre compte des cas où il existe deux possibilités syntaxiques pour un même « mot ». Il mentionne les stratégies syntaxique, lexicale et des règles de correspondances. Pour sa part, il opte pour la dernière des trois, où il n'y a pas de *mapping* entre structures conceptuelle et syntaxique dans l'entrée lexicale, mais où des règles de correspondances vont l'établir. C'est ce qu'on

appelle couramment le *argument linking* en terme de hiérarchie des rôles thématiques. Malgré ce choix, il convient que ces règles ne permettent pas la généralisation complète, en ce sens que certaines informations syntaxiques doivent tout de même figurer dans les descriptions des entrées lexicales (voir section 11.4, p.255-257). Un de ces résidus « concerns semantically unpredictable restrictions on the syntactic category of a complement » (Jackendoff 1990, 257). Ce problème est parallèle à la notion de restrictions de transitivité de N. Chomsky. Toutefois, selon R. Jackendoff, il y en a d'autres ; soit tous les cas où un verbe sélectionne une préposition idiosyncrasique (par ex., *believe in NP*, *convince NP of NP*, *depend/ rely/ count on NP* (Jackendoff 1990, 256)), les adjonctions qui expriment des arguments conceptuels (ce qu'il nomme les *obligatory adjuncts* ; par ex., *rid the room of insects*, *encrusted the throne with jewels*, *provided the students with some books*, *deprived Harry of his money*) ainsi que les verbes à particules (ex. *throw up* 'vomit').

### 2.1.3 Grammaire syntagmatique guidée par les têtes

En ce qui concerne la dernière approche considérée, celle de la HPSG (Head-driven Phrase Structure Grammar), elle est plus près de la vision Sens-Texte en ce qu'elle accorde également une place privilégiée aux propriétés des unités lexicales. Dans cette approche, les objets fondamentaux de l'analyse linguistique sont les signes. Ceux-ci sont représentés par des structures de propriétés représentées à leur tour par des « *attribute-value matrices* » (AVM) (Sag et Pollard 1989, 141). Dans leur AVM, les signes linguistiques lexicaux ont des spécifications pour les valeurs des attributs PHONOLOGIE, SYNTAXE et SÉMANTIQUE. Dans la section SYNTAXE des descriptions lexicales en HPSG, on divise les informations en « LOCAL features » et en « BINDING features » (Sag et Pollard 1989, 142). Les caractéristiques locales spécifient les propriétés syntaxiques inhérentes du signe. D'abord, on y trouve les caractéristiques de tête (=HEAD) qui spécifient les propriétés syntaxiques qu'un signe linguistique partage avec ses projections, c'est-à-dire avec les signes syntagmatiques dont la tête est cette unité lexicale (ce qui correspond à la valence passive dans notre terminologie). On y retrouve également la caractéristique SUBCAT qui donne de l'information à propos de la valence [active] d'un

signe. (Sag et Pollard 1989, 143) Le cadre de sous-catégorisation est utilisé pour encoder les différentes dépendances qui existent entre une tête lexicale et ses compléments. On y précise le nombre et la sorte des signes avec lesquels l'unité lexicale « characteristically combines with in order to become complete. » (Sag et Pollard 1989, 149) La position d'un élément à l'intérieur de la liste contenue dans la section SUBCAT d'une description lexicale correspond au caractère plus ou moins oblique du signe qui occupe cette position, le plus à droite correspondant au dépendant le moins oblique (=le sujet) de l'unité lexicale décrite (Sag et Pollard 1989, 163). En plus d'exercer des restrictions syntaxiques sur ses dépendants, les signes lexicaux peuvent également exercer des restrictions sémantiques (Sag et Pollard 1989, 163). En conséquence, une des propriétés essentielles de la section SUBCAT est d'établir la correspondance entre les relations grammaticales (sujets, objets, etc.) et les rôles décrits dans la situation. Ici aussi, contrairement à ce qui est fait dans le cadre de la grammaire générative, on considère que les rôles thématiques ne sont pas des entités syntaxiques, mais plutôt des constituants de contenu sémantique. On souligne que « It is important to understand the nature of this correspondence clearly, for out it arises the whole communicative power of lexical signs. The point of role assignment is not to make syntactic objects well-formed ; it is to establish a connection between the constituents of an utterance and the constituents of the thing the utterance is about. » (Sag et Pollard 1989, 164) Nous reprenons ici l'exemple (40) qui illustre l'entrée lexicale du verbe ditransitif anglais *give* :

PHON	<i>give</i>								
SYN	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;">HEAD</td> <td style="padding: 5px;">[ MAJ V ]</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">SUBCAT</td> <td style="padding: 5px;">&lt; NP <span style="border: 1px solid black; padding: 0 2px;">1</span>, NP <span style="border: 1px solid black; padding: 0 2px;">2</span>, NP <span style="border: 1px solid black; padding: 0 2px;">3</span> &gt;</td> </tr> </table>	HEAD	[ MAJ V ]	SUBCAT	< NP <span style="border: 1px solid black; padding: 0 2px;">1</span> , NP <span style="border: 1px solid black; padding: 0 2px;">2</span> , NP <span style="border: 1px solid black; padding: 0 2px;">3</span> >				
HEAD	[ MAJ V ]								
SUBCAT	< NP <span style="border: 1px solid black; padding: 0 2px;">1</span> , NP <span style="border: 1px solid black; padding: 0 2px;">2</span> , NP <span style="border: 1px solid black; padding: 0 2px;">3</span> >								
SEM   CONT	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;">RELN</td> <td style="padding: 5px;">GIVE</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">GIVER</td> <td style="padding: 5px;"><span style="border: 1px solid black; padding: 0 2px;">3</span></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">RECEIVER</td> <td style="padding: 5px;"><span style="border: 1px solid black; padding: 0 2px;">2</span></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">GIVEN</td> <td style="padding: 5px;"><span style="border: 1px solid black; padding: 0 2px;">1</span></td> </tr> </table>	RELN	GIVE	GIVER	<span style="border: 1px solid black; padding: 0 2px;">3</span>	RECEIVER	<span style="border: 1px solid black; padding: 0 2px;">2</span>	GIVEN	<span style="border: 1px solid black; padding: 0 2px;">1</span>
RELN	GIVE								
GIVER	<span style="border: 1px solid black; padding: 0 2px;">3</span>								
RECEIVER	<span style="border: 1px solid black; padding: 0 2px;">2</span>								
GIVEN	<span style="border: 1px solid black; padding: 0 2px;">1</span>								

Ce que cette entrée spécifie au niveau syntaxique, c'est que cette tête est de caté-

gorie verbale et qu'elle sélectionne trois compléments de type syntagme nominal. Elle précise également que le verbe *give* dénote : 1) ce qui est donné (= given), 2) celui qui reçoit (= receiver) et 3) celui qui donne (= giver). Ces trois arguments sémantiques sont respectivement réalisés syntaxiquement comme le second objet, l'objet direct et le sujet du verbe. Ainsi, une combinaison syntaxique (création d'un syntagme ou d'une phrase) qui se ferait à partir du signe *give* et avec l'ajout des compléments appropriés assurerait la création d'une construction sémantiquement et syntaxiquement adéquate. Finalement, la section 3.3 de l'article argumente que la sous-catégorisation ne peut être réduite ni à la seule sélection sémantique ni à la seule sélection catégorielle. Les auteurs fournissent certains des arguments qui ont motivé le choix de préférer un traitement uniforme de la sélection fonctionnelle, formelle et sémantique à l'intérieur du cadre de sous-catégorisation en HPSG.

En somme, cette position est plus lexicaliste que les précédentes en ce qu'elle prône l'inclusion d'une plus grande part d'informations syntaxiques à l'intérieur même du lexique. Ici, on ne tente pas d'épurer les entrées lexicales de toutes informations syntaxiques, on considère plutôt le signe linguistique comme une association d'une forme et d'un sens qui est aussi doté de propriétés de combinaison particulières. Bien que la HPSG diffère à plusieurs points de vue de la TST, sa conception théorique du lexique est sans doute celle des trois qui se rapproche le plus de celle adoptée en Lexicologie Explicative et Combinatoire [LEC], la branche lexicographique de la TST. En effet, comme en LEC, l'unité lexicale est décrite comme étant un signe linguistique saussurien (association d'une forme et d'un sens) qui comprend une troisième composante de combinatoire restreinte du signe (les « LOCAL features » vs. le syntactique). Pour ces deux théories, il appartient au lexique de décrire cette dernière composante du signe linguistique.

Dans le présent travail, nous adopterons donc ce que R. Jackendoff a nommé la *stratégie lexicale*. C'est-à-dire que nous désirons décrire le lexique de la façon la plus exhaustive possible afin de rendre compte de la correspondance (ou du manque de correspondance) entre actants sémantiques et actants syntaxiques à l'intérieur même du

lexique<sup>1</sup>. Nous détaillerons ci-après les diverses informations qui sont incluses dans les régimes lexicaux et la façon de les encoder en LEC. Nous ne prétendons pas que des généralisations sont impossibles, au contraire. Nous croyons, tout comme M. Gross, que la formulation de généralisations fiables passe d’abord par un travail d’analyse poussé qui ne peut être fait qu’avec un vaste éventail de données. En d’autres termes, ce n’est pas que la recherche de généralisations n’est pas intéressante, mais plutôt qu’un immense travail de description du lexique français doit être fait pour arriver à une meilleure compréhension de l’interaction entre les contenus sémantiques véhiculés par les lexies prédictives et les constructions syntaxiques qui les expriment. De plus, comme les généralisations parfaites semblent hors d’atteinte, il est sans doute préférable et même inévitable de décrire le comportement syntaxique d’une unité lexicale dans le dictionnaire (que ce soit au sujet de sa sélection catégorielle, de contraintes sur l’ordonnement ou sur la présence obligatoire de ses arguments, des prépositions (véritables ou vides) qu’elle sélectionne, etc.). Comme nous l’avons vu au chapitre précédent, ces informations ne sont généralement pas détaillées de façon constante et ne sont généralement pas disponibles pour l’ensemble des lexies prédictives dans les dictionnaires d’usage.

## 2.2 Notre cadre théorique : La théorie Sens-Texte

Comme notre travail cherche à élaborer une méthode de description pour les informations syntaxiques qui doivent être incluses dans le dictionnaire, le choix de la TST a semblé particulièrement souhaitable. Nous avons opté pour la TST comme cadre théorique parce qu’elle est particulièrement lexico-centrique, et qu’elle accorde une importance aux dépendances dans la modélisation de la langue. Elle offre donc déjà des outils qui seront utiles pour atteindre nos objectifs. Par conséquent, nous débuterons par une présentation sommaire de certains aspects de la TST qui sont indispensables pour bien situer notre travail. Nous souhaitons par la suite nous pencher sur deux aspects prédominants

---

1. En TST, il y a d’ailleurs un troisième type d’actant, à mi-chemin entre les deux. La TST distingue donc trois types d’actants : les Actants Sémantiques, les Actants Syntaxiques Profonds et les Actants Syntaxiques de Surface. Sur la notion d’actant en TST, voir Mel’čuk (2004a, b). Nous y reviendrons également dans la section suivante.

de la TST, soit la LEC et l'approche de syntaxe de dépendance postulée par la théorie. Il est important de mentionner que nous n'expliquerons pas en détail de nombreux aspects de la TST, par ailleurs très intéressants. Puisque notre recherche s'intéresse à la partie du dictionnaire qui décrit le comportement syntaxique d'une unité lexicale donnée, toutes les implications de la TST liées aux sens ou aux formes d'énoncés complexes n'influencent pas directement notre travail. Le lecteur intéressé à une présentation générale de cette théorie ou à d'autres facettes de celle-ci pourra consulter les textes d'introduction suivants : Kahane (2003); Mel'čuk (1997); Milićević (2006); Polguère (1998).

### 2.2.1 Présentation générale

La TST est une théorie linguistique qui a pris naissance en Russie dans les années 1960 (Žolkovskij et Mel'čuk 1967) et qui s'est développée depuis en Europe et au Canada<sup>2</sup>. Dans la TST, le fonctionnement de la langue est décrit au moyen d'un modèle Sens-Texte [MST]. Ce modèle est un ensemble de règles qui simule le comportement linguistique des humains. Plus spécifiquement, un MST est conçu pour opérer la transition entre les sens (toute information ou contenu qu'un locuteur voudra transmettre au moyen du langage) et les textes (manifestations physiques de la parole), et vice-versa (Mel'čuk 2001, 1). Bien qu'en principe les correspondances soient bidirectionnelles, cette modélisation se fait en priorité du sens vers le texte, comme le nom de la théorie l'indique. C'est donc dire que cette approche définit la langue du point de vue du locuteur. Dans nos descriptions lexicales, nous devons donc toujours nous situer au niveau de la production ; toutes les informations pertinentes du point de vue du locuteur qui souhaite formuler un énoncé doivent être incluses dans la description du lexique.

Le MST décrit la langue à l'aide de modules. Il y a évidemment les modules Sémantique [Sém] et Phonologique [Phon] et également deux modules intermédiaires : le Syntaxique [Synt] et le Morphologique [Morph]. Chacun de ces modules correspond à une des divisions reconnues de la linguistique moderne (la sémantique, la syntaxe, la morphologie et la phonologie). Ce qui est particulier, c'est qu'en TST on peut représenter

---

2. Une grande partie des recherches théoriques et appliquées en TST sont menées à l'Observatoire de linguistique Sens-Texte de l'Université de Montréal ([www.olst.umontreal.ca](http://www.olst.umontreal.ca)).

une phrase (ou une partie de phrase : proposition, syntagme, unité lexicale) à chacun de ces niveaux. Chacun d'eux est responsable de la représentation d'un ensemble particulier de propriétés du texte. De plus, tous les niveaux de représentation [R-] linguistique des énoncés (sauf le sémantique) se subdivisent en niveau Profond [-P] et niveau de Surface [-S]. Le MST se résume donc formellement comme suit :

$$\begin{aligned} & \text{RSém [=sens]} \leftarrow \mathbf{sémantique} \Rightarrow \text{RSyntP} \leftarrow \mathbf{syntaxique profond} \Rightarrow \text{RSyntS} \\ & \leftarrow \mathbf{syntaxique de surface} \Rightarrow \text{RMorphP} \leftarrow \mathbf{morphologique profond} \Rightarrow \text{RMorphS} \\ & \leftarrow \mathbf{morphologique de surface} \Rightarrow \text{RPhonP} \leftarrow \mathbf{phonologique} \Rightarrow \text{RPhonS [=textes]} \end{aligned}$$

Chacune des composantes du modèle (en gras) est un ensemble de règles (particulières à chaque langue) qui permettent de faire la transition entre les différentes représentations linguistiques d'un énoncé. Il est à noter que la distinction entre les niveaux Profond et de Surface est particulière à la TST. Le premier sert plutôt à indiquer des informations qui sont sémantiquement pertinentes tandis que le deuxième est surtout important pour la forme des énoncés. Pour de plus amples informations sur la distinction entre Profond et de Surface, voir Mel'čuk (1988, 59-72).

Comme on peut le voir, une des particularités de cette théorie est d'être stratifiée à l'extrême. Ce qui est pertinent pour nous, ce sont les niveaux Sémantiques, Syntaxique Profond et Syntaxique de Surface. Nous désirons précisément encoder dans notre base de données des informations lexicales relevant de ces trois niveaux.

Voici donc une brève description des structures qui constituent la partie centrale des représentations de chacun des trois niveaux de la TST qui nous concernent plus particulièrement (nous laisserons de côté les RMorph — et donc également les dépendances morphologiques qui gèrent notamment les accords — et les RPhon). Notons que les Structures Sémantique [SSém], Syntaxique Profonde [SSyntP] et Syntaxique de Surface [SSyntS] ne sont pas les seules composantes des représentations linguistiques de leur niveau respectif. Afin de ne pas alourdir nos explications inutilement, nous omettons volontairement d'exposer ces composantes additionnelles (les structures communicatives, structures anaphoriques, structures prosodiques, etc.).

La SSém d'une phrase est un réseau sémantique où les noeuds représentent les sens

lexicaux d'un énoncé, les arcs représentent les relations prédicats-arguments et sont numérotés. Ces arcs correspondent à des dépendances sémantiques.

La SSyntP d'une phrase est un arbre où les noeuds sont étiquetés des lexies pleines de la phrase et les branches sont étiquetées avec les noms des Relations Syntaxiques Profondes universelles. Ces relations sont, plus particulièrement, des dépendances syntaxiques.

La SSyntS d'une phrase est un arbre où les noeuds sont étiquetés avec toutes les lexies de la phrase. Les branches de cet arbre sont étiquetées avec les noms des DépSyntS qui représentent chacune une construction particulière d'une langue donnée. Pour un inventaire de ces relations pour l'anglais voir : Mel'čuk et Pertsov (1987), Mel'čuk (2009, 52-58) ; voir Iordanskaja et Mel'čuk (2009) pour un inventaire de celles contrôlées par les verbes en français, et Marengo (2009) pour un travail visant à faire l'inventaire de celles régies par les adjectifs en français.

En somme, la TST se caractérise entre autres par : 1) le recours aux dépendances à tous les niveaux de la représentation linguistique ; et 2) la position centrale qu'y occupe la sémantique lexicale.

### **2.2.2 La dépendance syntaxique**

Deux points de vue divergents existent en ce qui concerne les représentations syntaxiques des phrases : les arbres de dépendances et les arbres syntagmatiques. La grammaire syntagmatique est née d'une vision surtout orientée vers l'analyse des textes à l'aide du distributionnalisme. Comme l'objectif ultime était de rendre compte de certains phénomènes d'acquisition des langues, notamment du fait que certains éléments des langues sont innés, ils ont visé à décrire une grammaire universelle et ont négligé beaucoup de questions de sémantique (propres à chaque langue) et de différences lexicales (au sein d'une langue). M. Gross a d'ailleurs pointé les lacunes de cette approche (Gross 1975, 1979).

La représentation syntaxique d'une phrase par un arbre de dépendance est certainement plus ancienne. L'usage des dépendances remonte à l'antiquité. Kahane et Mel'čuk mentionnent tous deux d'après Owens (1988, 79-81) que les grammairiens arabes du 8<sup>e</sup>

siècle, comme Sibawaih, distinguaient déjà gouverneur et gouverné en syntaxe et utilisaient cette distinction pour formuler des règles d'ordre des mots ou de rection (Kahane 2001; Mel'čuk 2009, 1, 22). Ces deux derniers textes donnent d'ailleurs un bon résumé de l'histoire des grammaires de dépendances, de ses avantages en comparaison avec la grammaire syntagmatique ainsi que de nombreuses autres références bibliographiques.

### 2.2.2.1 Lucien Tesnière

Un des premiers à avoir présenté une théorie syntaxique basée sur une approche de dépendance est Lucien Tesnière. Nous souhaitons donc présenter brièvement sa théorie syntaxique. De plus, c'est une étape nécessaire puisque nous utiliserons plus loin la notion de translation qui en est issue.

Tout d'abord, en guise de présentation, voici reproduits quelques-uns des premiers points de *Syntaxe structurale* (Tesnière 1959, 11-12).

1. — L'objet de la syntaxe structurale est l'étude de la phrase. [...]
- 2.— La phrase est un **ensemble organisé** dont les éléments constituants sont les **mots**.
3. — Tout mot qui fait partie d'une phrase cesse par lui-même d'être isolé comme dans le dictionnaire. Entre lui et ses voisins, l'esprit aperçoit des **connexions**, dont l'ensemble forme la charpente de la phrase.
5. — Il résulte de ce qui précède qu'une phrase du type *Alfred parle* n'est pas composée de **deux** éléments 1° *Alfred*, 2° *parle*, mais bien de *trois* éléments, 1° *Alfred*, 2° *parle* et 3° la connexion qui les unit et sans laquelle il n'y aurait pas de phrase. [...]
7. — La connexion est **indispensable** à l'expression de la pensée. Sans la connexion, nous ne saurions exprimer aucune pensée continue et nous ne pourrions qu'énoncer une succession d'images et d'idées isolées les unes des autres et sans lien entre elles.
11. — La notion de connexion est ainsi la **base** de la syntaxe structurale. On ne saurait donc trop insister sur son importance.

On aura ainsi compris que la syntaxe structurale de L. Tesnière en est une de dépendance et que les « mots » et surtout les relations syntaxiques qui les unissent revêtent une importance cruciale. Ainsi, pour représenter les phrases, il utilise les stemmas, qui représentent l'ensemble des traits de connexion entre les termes régissants (= les gouverneurs) et les termes subordonnés (= les dépendants). Un stemma est donc un arbre de dépendances. Celui-ci ne contient qu'un seul régissant, qui lui peut avoir plus d'un subordonné. Cette pluralité est due au fait que les différents subordonnés assument des fonctions différentes vis-à-vis du régissant. Tesnière appelle « *fonction* des mots le rôle qui leur est assigné dans le mécanisme de l'expression de la pensée. » (Tesnière 1959, 39). « Savoir une langue, c'est donc savoir les différentes fonctions qui doivent y être assurées, et savoir la parler, c'est connaître le **maniement** des outils qui assurent ces fonctions. » (Tesnière 1959, 39)

Dès lors, il avait conclu que l'étude de ces fonctions et de ces outils était primordiale dans l'étude et l'enseignement des langues. Il s'est donc efforcé de concevoir une méthode d'analyse des langues qui permettrait d'en faire comprendre le fonctionnement efficacement. Ainsi, il a travaillé à répertorier les fonctions et les outils qui existent pour exprimer la pensée dans les langues naturelles. Le travail que nous cherchons à faire est d'un point de vue général assez proche, puisque nous tentons d'une part d'établir la liste des relations syntaxiques du français utilisées pour exprimer ces fonctions. Nous croyons qu'un inventaire de ces relations pourrait servir non seulement à l'enseignement de la syntaxe d'une langue, mais également à la description (et éventuellement à l'enseignement) de son lexique.

### 2.2.2.2 La dépendance en TST

#### Dépendances sémantiques

Comme nous l'avons vu dans la présentation générale, la SSém d'un énoncé ou d'une partie d'énoncé est un réseau sémantique où les noeuds représentent les sens lexicaux, et les arcs représentent les relations prédicat-arguments.

On peut considérer que les sens lexicaux se divisent en deux grandes classes, les

prédicats sémantiques et les noms sémantiques.

Les prédicats sémantiques sont des sens liants impliquant au moins un « participant » appelé Actant Sémantique [ASém]. Les ASém sont désignés par des variables du type X, Y, Z, etc. — ‘X regarde Y’, ‘X envoie Y à Z’, ‘[X est] grand’, ‘[X se produit] souvent’, ‘regard de X vers Z’, ‘envoi par X de Y à Z’. On associe généralement les sens liants à des faits (actions, événements, états, etc.).

Cependant, les sens lexicaux prédictifs peuvent aussi dénoter des entités. Par exemple, la lexie FRÈRE [*Alex est le frère de Jonathan.*], comme toutes les lexies dénotant un individu qui entretient une relation particulière avec un autre individu, est un exemple de lexie prédictive dénotant une entité (et non un fait). Un frère est nécessairement le frère de quelqu’un et l’expression *frère de Jonathan* permet d’exprimer un des ASém de ce prédicat. On nommera ces lexies des quasi-prédicats. L’entité dénotée par un quasi-prédicat est donc présentée par la langue de façon particulière, puisqu’elle est définie par son implication dans un fait (un frère n’est pas simplement un individu, mais un individu impliqué dans une relation). C’est pourquoi un quasi-prédicat contrôle aussi des positions actanciennes sémantiques. Les ASém d’un quasi-prédicat sont les actants d’un prédicat « interne » (ici la relation), qui fait partie du sens en question sans en être la composante centrale. Bref, une lexie prédictive (qui contrôle des positions actanciennes ; et qui a donc des ASém) est soit un prédicat, soit un quasi-prédicat.

Finalement, les noms sémantiques sont des sens de lexies qui dénotent des entités n’impliquant intrinsèquement aucun participant — ‘roche’, ‘lion’, ‘cerise’, ‘Évelyne’.

Les prédicats sémantiques sont lexicalisés par des verbes, des adjectifs, des adverbes ou des noms. Les quasi-prédicats et les noms sémantiques sont, quant à eux, toujours lexicalisés par des noms. Pour en savoir plus sur la notion de prédicats vs quasi-prédicats sémantiques, voir Mel’čuk et Polguère (2008).

Une dépendance sémantique est la relation qui existe entre une lexie prédictive (le gouverneur [G] de la relation) et un de ses ASém (le dépendant [D] de la relation). Nous représenterons une dépendance sémantique par un chiffre arabe (1, 2, 3, 4, etc.) identifiant l’actant correspondant. En TST, ces chiffres sont purement distinctifs, ils n’ont aucun sens à proprement parler. Par exemple, le chiffre 1 ne garantit pas que le pre-

mier ASém soit un Agent. Ce sens faisant partie de la définition de la lexie, il n'est pas nécessaire de l'ajouter comme étiquette sur les flèches des dépendances sémantiques.

Les sens lexicaux, ceux que nous retrouverons également dans la description du lexique, sont représentés dans le réseau sémantique des énoncés. Ainsi, le sens d'une lexie correspond à une partie de ce réseau. Le sens d'une lexie prédicative doit donc être représenté de manière formelle à l'aide d'un graphe de ce type également. Dans son article lexicographique, la façon de le représenter sera évidemment différente (formule linéaire au lieu d'un graphe), mais les informations présentes demeurent les mêmes. Comme dans le DiCo, dans le cadre du présent travail, nous utilisons la forme propositionnelle pour représenter les différents ASém des unités lexicales. Par exemple, la forme propositionnelle de la lexie HÉSITATION#1 est : `~ DE LA personne X AU SUJET DU fait Y`. La forme propositionnelle permet donc d'indiquer quels sont les participants impliqués dans la situation que dénote l'unité lexicale HÉSITATION#1 (ses ASém), et une caractérisation sommaire de ces actants grâce à des étiquettes, ici : `personne X et fait Y`.

### **Dépendances Syntaxiques Profondes**

Les Dépendances Syntaxiques Profondes [DépSyntP] sont un concept particulier de la TST puisqu'elles présupposent l'existence de la SSyntP. La SSyntP (grâce à la présence des DépSyntP) permet de visualiser les éléments introduits soit par coordination, soit par subordination. Les relations de subordination peuvent effectuer de la complémentation ou de la modification. Les DépSyntP peuvent donc être actancielle et introduire un Actant Syntaxique Profond [ASyntP] ou être non actancielle et introduire un modificateur adjectival ou adverbial, une coordination ou un élément extra-structural comme une interjection, un complément de phrase, un marqueur de relation, etc.

Les DépSyntP sont universelles (elles sont les mêmes dans toutes les langues) et elles existent en nombre limité. Nous avons vu qu'elles sont soit actancielle (I, II, III, IV, V, VI, et une variante pour l'introduction du discours direct II-disc.dir.), soit non actancielle (ATTR, COORD, APPEND). En voici une classification reprise de Mel'čuk et Milićević (2011).

Subordinatives			Coordinative
Fortes		Faible	
Actanciennes	Modificative		
I – VI	ATTR	APPEND	COORD

Figure 2.1 – Classification des DépSyntP

Comme nous nous intéressons aux dépendances syntaxiques régies, il est évident que nous aurons recours uniquement aux DépSyntP actanciennes, qui sont numérotées par des chiffres romains. Le dépendant d’une telle relation est un ASyntP, également numéroté par un chiffre romain.

Les DépSyntP sont en quelque sorte des généralisations des rôles syntaxiques des éléments de la phrase. Typiquement, I correspond au sujet d’un verbe (LE MINISTÈRE←I– DÉCIDE) ou à un transformé du sujet (LA DÉCISION–I→DU MINISTÈRE, LA DÉCISION–I→MINISTÉRIELLE), II correspond au premier complément d’un verbe — direct ou indirect — (ENVOYER–II→ UNE LETTRE) ou à un transformé (ENVOI–II→ D’UNE LETTRE), etc. Il est à noter que ces distinctions sont importantes puisque deux expressions superficiellement identiques peuvent avoir des structures différentes et exprimer des sens différents. Par exemple, voici deux SSyntP partielles et les sens correspondants pour la forme ambiguë *la crainte des ennemis* où *des ennemis* est le premier complément du nom *crainte* :

1. LA CRAINTE –I→ DES ENNEMIS : ‘la crainte est éprouvée par les ennemis’
2. LA CRAINTE –II→ DES ENNEMIS : ‘la crainte est inspirée par les ennemis’

### Dépendances syntaxiques de surface

Comme nous l’avons vu précédemment, les arcs de la SSyntS sont étiquetées avec les noms des DépSyntS. Ces dépendances représentent chacune une construction particulière d’une langue donnée. En français par exemple, un verbe et son sujet sont liés

par la DépSyntS **subjectale**, un nom et son déterminant sont liés par la DépSyntS **déterminative**, etc. Dans l'exemple donné plus tôt, les DépSyntS correspondantes sont :

1. LA CRAINTE –I→ DES ENNEMIS : *la crainte* –subjectale adnominale→ *des ennemis* : ‘la crainte est éprouvée par les ennemis’
2. LA CRAINTE –II→ DES ENNEMIS : *la crainte* –objectale adnominale→ *des ennemis* : ‘la crainte est inspirée par les ennemis’

Le dépendant d'une telle relation (lorsqu'elle permet d'exprimer un des ASém du gouverneur) est un Actant Syntaxique de Surface [ASyntS]. Les ASyntS correspondent donc aux sujets, compléments d'objet direct, compléments obliques, compléments déterminatifs, etc. de la grammaire traditionnelle.

### 2.2.3 La Lexicologie Explicative et Combinatoire

Comme nous l'avons déjà mentionné, la LEC est la branche de la TST qui s'intéresse à la description du lexique. Elle a permis la création de plusieurs ouvrages lexicographiques qui tentent d'établir la méthode idéale pour la production de dictionnaires, dont les quatre volumes du *Dictionnaire explicatif et combinatoire du français contemporain* [DECFC] (Mel'čuk et coll. 1984-1999). Plus généralement, les Dictionnaires Explicatifs et Combinatoires [DEC], dictionnaires sous-tendus par la LEC, se distinguent des autres dictionnaires entre autres parce qu'ils sont conçus en tenant compte de la production des énoncés plutôt que de la compréhension des énoncés. Ils sont également rigoureusement conçus, chacun des articles lexicographiques étant divisé selon une microstructure qui est constante sur l'ensemble du dictionnaire. Les données que nous utilisons dans cette recherche sont celles du DiCo. Le DiCo est une base de données informatisée, dérivée du DEC, qui se concentre sur la description des propriétés sémantiques et combinatoires des lexies, et est consultable via l'interface Dicouebe (<http://olst.ling.umontreal.ca/dicouebe/index.php>). Cette base de données, a elle-même permis la création de certains outils pédagogiques, notamment le *Lexique actif du français* [LAF] (Mel'čuk et Polguère 2007) ainsi qu'un « site lexical » associé, le DiCoPop

(<http://olst.ling.umontreal.ca/dicopop/>).

La lexie est l'unité de base du lexique, et donc des descriptions lexicales d'un DEC. « Chaque lexie de la langue est associée à un sens particulier, unique, définissable et à un ensemble de formes correspondant aux variantes flexionnelles du signifiant de la lexie » (Polguère 2002, 39). Une lexie est donc un ensemble de signes lexicaux (les mots-formes) qui partagent le même sens lexical. On décrit donc la lexie en la considérant comme un signe linguistique ; elle est composée d'un sens lexical (le signifié), d'une forme (le signifiant) et d'un ensemble de traits de combinatoires (le syntactique). L'article de dictionnaire est donc logiquement divisé en trois parties qui servent à la description de chacune de ces composantes. Le patron de description (repris ici de *Introduction à la lexicologie explicative et combinatoire* [ILEC] (Mel'čuk et al. 1995, 70)) qui permet de décrire une lexie est structuré selon ce qui suit :

- zone phonologique (= le signifiant de la lexie vedette L) ;
- zone sémantique (= le signifié de L) ;
- zone de combinatoire (= le syntactique de L).

Cette dernière zone est celle qui fait l'objet d'une description particulièrement développée dans le DiCo. Elle se divise en six parties :

- la sous-zone de combinatoire morphologique (partie du discours, types de déclinaison/conjugaison, formes non réalisables, formes irrégulières, etc.) ;
- la sous-zone de combinatoire stylistique (marques d'usage du type spéc., litt., fam., arch., etc., qui indiquent le type de contexte textuel dans lequel la lexie peut s'insérer) ;
- la sous-zone de combinatoire syntaxique (classe et sous-classes syntaxiques, constructions typiques, régime, etc.) ;
- la sous-zone de combinatoire lexicale restreinte ;
- zone d'illustrations ;
- zone phraséologique.

Nous présentons ici ces différentes zones et sous-zones en faisant ressortir les points pertinents dans le cadre de notre travail. Mentionnons brièvement que la zone d'illustrations est celle où sont données les phrases d'exemples illustrant l'utilisation de la

lexie vedette. La zone phraséologique donne les locutions ou expressions figées qui contiennent la lexie vedette. Chacune des locutions étant considérée comme une unité lexicale à part entière, elle fait elle-même l'objet d'une description et est donc associée à une fiche complète.

### **Zone sémantique**

La partie centrale de la zone sémantique est la définition lexicographique de L (= la lexie vedette). Notons que dans la base DiCo, on n'inclut pas de définition lexicographique à proprement parler. En ce qui concerne la zone sémantique, on ne fait usage que du défini (c'est-à-dire la lexie L incluse dans une forme propositionnelle) et d'une étiquette sémantique attribuée à L. L'étiquette sémantique est une forme normalisée du genre prochain de L (sa paraphrase minimale). En ce sens, le DiCo est moins riche que le DEC. Toutefois, la présence de la forme propositionnelle permet l'encodage de la valence sémantique d'une lexie prédicative (c'est-à-dire des Asém inclus dans sa définition), ce qui est primordial en lexicologie et plus particulièrement en LEC. L'étiquette sémantique, quant à elle, permet d'établir la classe sémantique de L, ce qui nous permet entre autres de déterminer si L dénote un fait ou une entité. Ceci nous permet également de classer les lexies ayant un sens liant en terme de prédicat ou de quasi-prédicat sémantique. Il est à noter que le projet RLF qui est en cours à l'Université de Nancy projette d'intégrer les définitions analytiques (avec genre prochain et différences spécifiques) dans les articles lexicographiques (Lux-Pogodalla et Polguère 2011).

### **Zone de combinatoire lexicale**

Il est évident que pour maîtriser une lexie, il ne suffit pas d'en connaître la forme et le sens. Pour être apte à bien comprendre ou formuler des énoncés d'une langue donnée, il faut également connaître les liens que les lexies entretiennent entre elles au sein du réseau complexe qu'est le lexique d'une langue. On appelle ce phénomène la connexion lexicale. Ces liens lexicaux sont de deux types, les liens paradigmatiques (correspondent au phénomène de dérivation sémantique – la synonymie, l'hyponymie/hyperonymie, la no-

minimalisation/verbalisation, etc.) et les liens syntagmatiques (les propriétés individuelles d'une lexie qui contrôlent la façon dont elle se combine aux autres unités lexicales dans la phrase — modificateurs adjectivaux, verbes supports, etc.). Les concepteurs de dictionnaires ont conscience de ce phénomène et souhaitent inclure ces renseignements dans les articles. On voit souvent à la fin des articles un ou des synonymes. Dans *Le Petit Robert* (Rey-Debove et Rey 2008), CLOCHE 7 a comme synonyme AMPOULE. Dans ce même article (CLOCHE), on voit aussi des renvois à d'autres lexies du dictionnaire : BOURDON, CLOCHETTE sont des types de CLOCHE 1. On choisit également certains exemples d'utilisation précis qui mettent en scène des expressions semi-idiomatiques (collocations) construites avec la lexie décrite : CLOCHE 1 *cloche qui tinte, cloches qui sonnent l'angélus, le glas, le tocsin.* ; CLOCHE 2 *cloche à oxygène, cloche à fromage, cloche de métal.*

En LEC, ces deux types de liens lexicaux font systématiquement et explicitement partie d'un article lexicographique. Pour ce faire, elles sont modélisées au moyen du système formel des fonctions lexicales. Chaque fonction lexicale possède un sens général et permet d'encoder une relation récurrente en langue. Une fonction lexicale se comporte un peu comme une fonction mathématique, en ce sens qu'elle s'applique à un argument (une unité lexicale de base, choisie pour son sens) et produit une valeur (un ensemble de dérivés sémantiques ou de collocatifs). Une application de fonction lexicale se décrit par une formule du type  $f(x)=y$ , où  $x$  est l'argument de la fonction lexicale  $f$  et  $y$  est la valeur « retournée ». Par exemple, la fonction lexicale **Magn** (du latin *magnus* 'grand') associe à une lexie l'ensemble des lexies ou expressions linguistiques qui expriment auprès d'elle l'intensification, c'est-à-dire le sens général 'intense', 'très', 'beaucoup', etc. Ainsi, la lexie PLEUVOIR, qui est choisit librement pour son sens par un locuteur qui souhaite parler du fait qu'il pleut, a comme « intensificateurs » potentiels : *des cordes, abondamment, à verse, à boire debout*, etc. Selon le formalisme des fonctions lexicales, ceci est représenté par la formule suivante qui doit être présente dans la fiche lexicographique de la lexie PLEUVOIR : **Magn(pleuvoir) = des cordes, abondamment, à verse, à boire debout**

### Zone de combinatoire syntaxique

Cette zone contient d'abord des informations qui servent principalement à décrire la valence passive de L, c'est-à-dire l'ensemble des contextes dans lesquels elle peut apparaître en tant que dépendant. La simple étiquette « nom », « verbe », etc. permet de noter de façon concise une multitude d'informations au sujet de ces différentes classes de mots. On sait donc en voyant « nom » que la lexie en question peut (potentiellement) apparaître comme sujet d'un verbe conjugué, comme objet direct d'un verbe transitif, comme complément d'une préposition, etc. Toutefois, la valence passive ne peut être réduite à la seule dénomination en tant que partie du discours, deux noms pouvant évidemment apparaître (ou ne pas apparaître) dans tous les mêmes contextes. On peut alors ajouter de nombreuses autres précisions (par ex., « peut être employé performativement », « apposition (avec ou sans trait d'union, collocatif de son gouverneur syntaxique) », « locution prépositionnelle à emploi adverbial ou adjectival », etc.)

De plus, il est assez généralement admis que la transitivité ou l'intransitivité d'un verbe, le caractère obligatoire ou non des compléments, la sélection de prépositions régies, etc. sont des éléments qui ne peuvent être entièrement déduits du sens d'une lexie. Effectivement, des lexies sémantiquement proches (grosso modo des synonymes) n'expriment pas toujours leurs ASém de la même façon. On peut le voir par exemple en comparant :

1. *Michel est capable de se sacrifier. ~ Michel est apte à se sacrifier.*
2. *Il demande/ réclame un bec à sa copine. ~ Il exige/ revendique/ sollicite un bec de sa copine.*
3. *Julie permet à Luc de lire cette revue. ~ Julie autorise la lecture de cette revue à Luc.*
4. *Jules informe Ève de son intention de partir. ~ Jules apprend à Ève qu'il a l'intention de partir. ~ Jules renseigne Ève au sujet de son intention de partir. ~ Jules communique son intention de partir à Ève.*
5. *son [= le tueur] homicide ~ son [= la victime] assassinat ~ son [= tueur ou victime] meurtre*

La cooccurrence syntaxique d'une lexie avec ses actants relève plutôt de propriétés inhérentes qui, bien que concernant la syntaxe, ne relèvent pas de la syntaxe grammaticale d'une langue. Il s'agit plutôt d'une syntaxe lexicale ou micro-syntaxe. Cette

micro-syntaxe étant soumise aux caprices de chaque lexie, elle doit par conséquent être considérée comme une des facettes du lexique et faire partie des descriptions lexicales au sein du dictionnaire. Bien que ce soit généralement admis comme étant vrai, ces descriptions ne sont pratiquement jamais complètes ni systématiques dans les dictionnaires d'usage, particulièrement en français.

Toutefois, la LEC accorde à ce type d'informations une place privilégiée et c'est pourquoi la zone de combinatoire syntaxique d'un DEC permet de noter les propriétés « micro-syntaxiques » d'une lexie prédicative à l'intérieur d'un Tableau de régime ou, plus simplement, Régime (la principale composante de la zone de combinatoire syntaxique). Le Tableau de régime de L décrit donc sa valence active, les dépendants syntaxiques régis qu'elle contrôle.

### 2.3 Modélisation des régimes lexicaux en LEC

Dans le Tableau de régime, chaque ASém de L correspond à une colonne particulière, où l'on inscrit également l'ASyntP correspondant (ex. X = I, Y = II). Dans chacune d'elle, une ligne est consacrée à chacune des expressions possibles d'un ASyntP donné, les constructions syntaxiques (*de* N, *que* PROP, etc.) avec éventuellement un moyen morphologique (le cas nominal dans les langues où il existe, le mode verbal, etc.).

Les données sur l'expression des actants présentes dans le Tableau de régime sont donc en lien avec trois types de régimes différents, distingués dans Mel'čuk (En préparation, Section 1). Le régime<sub>1</sub> se manifeste lorsqu'un G impose certaines valeurs flexionnelles à son D, sans que ce soit nécessairement lié à l'accord (cas grammatical d'un D nominal, mode verbal d'un complément, etc.). Il s'agit donc d'un cas particulier de dépendance morphologique. Le régime<sub>2</sub> se manifeste lorsqu'un G sélectionne une lexie en particulier (le plus souvent une préposition ou une conjonction sémantiquement déficiente, DIFFÉRENCE *entre* N et N, STRATAGÈME *pour que* PROP-subj.) pour introduire son D. Il s'agit donc d'un cas particulier de dépendance syntagmatique entre deux lexies au niveau de la syntaxe de surface. Finalement, le régime<sub>3</sub> se manifeste lorsqu'un G — de par son sens — impose au D l'appartenance à une partie du discours

particulière (N, V, ADJ, PROP, etc.). Il s'agit donc d'un cas particulier de dépendance sémantique.

Le Tableau de régime est également accompagné d'une liste de contraintes additionnelles et d'exemples qui en illustrent l'application.

Prenons en exemple le Régime de BLESSURE#I.2 (*Blessure de X [par Y] à Z au moyen de W*), tiré du DECFC Vol. IV.<sup>3</sup>

**Régime**

X = 1	Y = 2	Z = 3	W = 4
1. <i>de</i> N	_____	1. <i>à</i> N	1. <i>à</i> N
2. <i>A<sub>poss</sub></i>			2. <i>par</i> N

1) C <sub>4.1</sub>	: N désigne une arme blanche
2) C <sub>4.2</sub>	: N = <i>balle</i> ...
C <sub>1</sub>	: <i>la blessure de Jean &lt;du soldat, du cheval&gt;, sa blessure</i>
C <sub>3</sub>	: <i>une blessure à l'épaule &lt;au cœur, à l'abdomen&gt;, des blessures au corps</i>
C <sub>4</sub>	: <i>une blessure à l'arme blanche &lt;au couteau&gt;, une blessure par balle</i>
C <sub>1</sub> + C <sub>3</sub>	: <i>les blessures de l'enfant aux bras, sa blessure au poignet droit</i>
C <sub>1</sub> + C <sub>3</sub> + C <sub>4</sub>	: <i>sa blessure par balle à la jambe</i>

Figure 2.2 – Régime de BLESSURE#I.2 dans DECFC Vol. IV

Comme le DiCo est une version informatisée du DEC, la modélisation est faite différemment. Le Régime n'est pas encodé dans un véritable tableau, mais plutôt linéaire-

3. Les créateurs du DECFC ne faisaient pas l'usage des chiffres romains pour les ASynP. Ici, les chiffres arabes sont utilisés au niveau de la syntaxe profonde ; voir les arbres syntaxiques lorsqu'il y a des modifications de régimes.

ment ; chacune des expressions possibles pour chacun des ASém étant donnée sur une ligne différente. Par exemple, le Tableau de régime précédent correspondrait au texte suivant :

X = I = de N, A-poss  
 Y = II = --  
 Z = III = à N | N désigne une arme blanche  
 W = IV = à N, par N | N = balle...

Les concepteurs du DiCo ont choisi d'intégrer les contraintes additionnelles à la suite de chacune des lignes au lieu d'en faire une liste à la suite des autres informations (évitant ainsi d'avoir à utiliser des codes parfois assez complexes à déchiffrer rapidement, C4.1, C1.1 + C2, etc.). La zone syntaxique du DiCo contient également moins d'exemples que les articles du DECFC. Comme notre base de données vise la factorisation d'informations sur le régime, nous avons omis les exemples d'utilisation, puisque trop lexicalemement spécifiques. Nous considérons que ces exemples peuvent faire partie de la zone d'illustrations des articles du dictionnaire, et ne sont donc pas nécessaires dans notre base de données sur les régimes.

La LEC et les dictionnaires qui en sont issus ont développé une méthodologie stricte qui prévoit : 1) la description de la diathèse d'une lexie, c'est-à-dire la correspondance entre ses ASém et ses ASyntP ; et 2) les constructions qui réalisent les ASyntP à la surface. C'est donc dire que le Régime de L spécifie les correspondances entre les ASém, les ASyntP et les ASyntS de L.

Toutefois, en ce qui concerne les ASyntS, il n'est pas suffisant d'écrire qu'un ASyntP donné s'exprime avec « de N », par exemple. C'est que le syntagme en « de N » peut correspondre à différents syntagmes (en fonction de la DépSyntS impliquée) comme nous avons pu le voir avec l'exemple de *crainte des ennemis*. Pour que l'information soit complète, le Régime devrait donc comprendre explicitement le nom des dépendances particulières qui sont présentes entre L et ses différents ASyntS et pas seulement les moyens d'expression de surface (*de N, à N, par N, etc.*). Le Tableau de régime tel qu'actuellement utilisé se sert donc d'un raccourci méthodologique en ne mentionnant

pas les dépendances, mais seulement les actants (c'est le cas aussi pour les ASém et les ASyntP). Nous voulons donc que nos Régimes soient construits de façon à inclure les dépendances en plus des actants. Nous verrons la méthodologie retenue au chapitre 4.

## CHAPITRE 3

### IDENTIFICATION ET CARACTÉRISATION DES DÉPENDANCES SYNTAXIQUES DE SURFACE : LE CAS DE LA DÉPENDANCE ATTRIBUTIVE ADNOMINALE

« En l’occurrence, l’anglais l’emporte sur le français : *fire dancer*, c’est tellement mieux que danseur de feu. Pauvre français de besogneux analytique, qui doit établir un constat d’accident – un complément déterminatif – est-ce un génitif objectif ou subjectif ? Qu’est-ce que la grammaire vient faire entre deux divinités ? C’est l’anglais qui a raison, il faut jeter les deux mots l’un contre l’autre – et qu’ils se débrouillent – et aussitôt ils crépitent ensemble. »

Amélie Nothomb, *Tuer le père*

Dans ce chapitre, nous proposons une méthodologie d’identification/ caractérisation des DépSyntS en l’appliquant à une dépendance nominale, la dépendance **attributive adnominale**. Lors de la première section, nous présentons d’abord différentes constructions syntaxiques qui contiennent potentiellement la DépSyntS que nous souhaitons étudier. Ce faisant, nous examinerons comment ces constructions sont généralement analysées et ce qui les distingue des autres compléments du nom.

Notons que nous avons choisi cette dépendance précise parce que nous considérons qu’elle représente un défi intéressant, à la fois pour son identification et pour son encodage. Premièrement, au niveau de l’identification, nous nous demandions si des constructions où le D est régi et des constructions où le D est non régi pouvaient être rassemblées sous cette même dépendance. En 3.2, nous visons l’élaboration de critères qui permettent d’identifier adéquatement la dépendance **attributive adnominale**. Ces critères, rassemblés dans le Tableau 3.1, pourront également fournir une réponse à la question précédente. Au cours de cette réflexion, nous proposons une marche à suivre qui pourra être applicable à l’identification et à la caractérisation d’autres DépSyntS.

Puis, en ce qui concerne l’encodage, il faut être conscient du fait que dans les cas où l’expression d’un D ne se fait pas de façon totalement libre, ceci doit apparaître dans

la description lexicographique de la lexie gouvernante. Nous souhaitons réfléchir sur la meilleure méthode qui permette cet encodage lorsque le D (qui semble correspondre à un ASém du G) est exprimé par la dépendance **attributive adnominale**. Cela nous semble soulever un problème intéressant en ce qui concerne notre travail de recherche, puisqu'il n'est pas évident de prime abord si un dépendant de ce type sature toujours la position actantielle de son G. Conséquemment, dans la section 3.3, nous réfléchirons sur ce qui doit être modélisé à l'aide des régimes lexicaux, en exposant deux solutions de modélisation.

### 3.1 Conceptualisation de la dépendance **attributive adnominale**

Avant d'exposer notre méthodologie d'identification/ caractérisation des DépSyntS en 3.2, il est important de bien conceptualiser la dépendance qui servira lors de cette démonstration. Nous avons choisi de traiter des constructions où un élément nominal (précédé d'une préposition sémantiquement déficiente) est utilisé pour effectuer une modification syntaxique analogue à celle de l'adjectif (comme dans le *danseur de feu* de la citation mise en exergue) tout comme le fait la construction appositive « N2 N1 » de l'anglais (par ex., *fire dancer*). Avant de bien cerner cette dépendance particulière, revoyons rapidement les différents compléments du nom qui existent en français.

#### 3.1.1 Les différents compléments du nom

La grammaire traditionnelle (aussi nommée troisième grammaire scolaire) a introduit au début du 20<sup>e</sup> siècle la notion de complément déterminatif ou de complément du nom (*Antidote Prisme* 2004, Guide : Historique, Article : XX<sup>e</sup> siècle : la troisième grammaire scolaire, ou grammaire traditionnelle). Le complément déterminatif se subordonne au nom pour en limiter le sens de façons très variées. Il peut notamment indiquer :

- l'espèce : *un cor de chasse*.
- l'instrument : *un coup de lance*.
- le lieu : *la bataille de Waterloo*.
- la matière : *une statue de bronze*.
- la mesure : *un trajet de dix kilomètres*.
- l'origine : *un jambon d'Ardenne*.

- la possession : *la maison de mon père.*
- la qualité : *un homme de coeur.*
- le temps : *les institutions du moyen âge.*
- la totalité : *une partie de cette somme.*
- la destination : *une salle d'escrime.*
- le contenu : *une tasse de lait.* (Grevisse 1995, 51-52)

Grevisse n'utilise pas seulement les sens impliqués pour classifier les différents compléments déterminatifs. En effet, il distingue plusieurs types de compléments déterminatifs selon la nature du nom gouverneur. Nous résumons ici les classifications de Grevisse et Goose (2007, 429-432) et celles d'*Antidote Prisme* (2004, Guide : Grammaire, Article : Compl. du nom). Il y a les noms qui correspondent à des verbes (*acceptation, amour, remerciement, transmission, volonté*, etc.). Ces noms peuvent avoir comme compléments les mêmes compléments que le verbe correspondant (sujet, objet direct, objet indirect, complément adverbial, etc.). On nomme souvent le complément correspondant au sujet d'un verbe « complément subjectif » (*la survivance du latin classique, la création de/par l'artiste*) et celui correspondant à l'objet du verbe « complément objectif » (*la distribution des vivres, la création d'une sculpture*). Certains noms correspondent plutôt à des adjectifs (*adresse, facilité, fierté, gentillesse*, etc.). Ces noms peuvent accepter les mêmes compléments que leurs analogues adjectivaux. Par exemple, *adresse* accepte un complément en *à* comme l'adjectif *adroit* : *adroit au tir à l'arc* ~ *adresse au tir à l'arc*. Finalement, les noms quelconques (les autres), peuvent aussi avoir des compléments (qui sont donc particuliers aux noms) : le complément de relation, qui concerne autant la possession que d'autres relations (*le nez de Racine, la voiture de mon père, les gants de Marie, la porte de la rue*) et le complément de caractérisation (*un cor de chasse, sa jupe à pois, mon chapeau de paille, le cuir du Maroc, un fruit à noyau, un canon contre avions, la viande de cheval*). Comme son nom l'indique, le complément de caractérisation donne une caractéristique du nom auquel il est subordonné, et en restreint ainsi le sens. Grevisse remarque que « le complément déterminatif sans article joue le rôle d'une locution adjectivale équivalent à un qualificatif : *un coup d'audace* = *un coup audacieux* ; *un soleil de printemps* = *un soleil printanier* » (Grevisse 1995, 51).

Grosso modo, la dépendance pour laquelle nous présentons une méthode d'identification est celle liant un nom à cedit complément de caractérisation.

### 3.1.2 Tesnière et la translation

Dans sa théorie syntaxique, Tesnière a également traité de ce type de complément. Pour bien comprendre l'analyse qu'il en a faite, nous devons d'abord introduire la notion qui fait l'objet de la troisième partie de *Syntaxe structurale* (Tesnière 1959, 359-664), la translation<sup>1</sup>.

Pour Tesnière, il existe une opposition entre les mots pleins et les mots vides. Les mots pleins sont « ceux qui sont chargés d'une **fonction sémantique**, c'est-à-dire ceux dont la forme est associée directement à une idée, qu'elle a pour fonction de représenter et d'évoquer. » (Tesnière 1959, 53) En contrepartie, « [l]es mots vides sont ceux qui ne sont pas chargés d'une fonction sémantique. Ce sont de simples outils grammaticaux dont le rôle est uniquement d'indiquer, de préciser ou de transformer la catégorie des mots pleins et de régler leurs rapports entre eux. » (Tesnière 1959, 53) Il existe quatre espèces de mots pleins, qui sont, par ordre d'importance : le **verbe**, le **substantif**, l'**adjectif** et l'**adverbe**. Il existe aussi deux espèces de mots vides, les **jonctifs** (qui sont les marqueurs morphologiques de la jonction) et les **translatifs** (qui sont les marqueurs morphologiques de la translation).

Dans la théorie de Tesnière, les différentes classes de mots pleins prennent comme dépendants des classes de mots précises. Les dépendants du verbe peuvent être soit des substantifs (correspondent aux actants du verbe), soit des adverbes (correspondent aux circonstants). Les dépendants les plus courants du substantif sont des adjectifs épithètes ou autres subordonnés épithétiques qui servent à qualifier le substantif dont ils dé-

---

1. Pour notre part, nous nous en tiendrons à une notion assez restreinte de la translation par rapport à celle formulée par Tesnière. Nous croyons que certains phénomènes décrits comme étant de la translation sont couverts de façon satisfaisante en TST. Soit, en particulier, la flexion casuelle de certaines langues et la dérivation morphologique (en synchronie) ainsi que la création lexicale par dérivation morphologique (en diachronie). Pour Tesnière, lorsqu'on crée le substantif *chanteur* à partir du verbe *chanter* ou qu'on utilise un nom en apposition, on a également affaire à de la translation. Notre raisonnement n'est pas en désaccord total avec cette idée, seulement, nous désirons nous concentrer sur un type de translation en particulier, celle qui est la source de la DépSyntS que nous nommons **attributive adnominale**.

pendent. L'adjectif n'a comme subordonné possible que l'adverbe et, en ce qui concerne les adverbes, ils ne peuvent prendre comme dépendants que d'autres adverbes. Selon ce raisonnement, un dépendant syntaxique ne se conformant pas à la catégorie de dépendant typique passe donc par un phénomène de translation.

Ainsi, lorsqu'un nom a un groupe prépositionnel comme dépendant, la préposition fait office de translatif pour modifier la catégorie nominale de ce qui suit en adjectif ; afin que le groupe Prép+Nom fasse office de modificateur épithétique. Ainsi, à côté des adjectifs proprement dits, les noms peuvent avoir comme dépendants des quantités de formules adjectivales introduites par des prépositions, dont les plus fréquentes sont *de* (par ex., *un homme de coeur*, *un homme de goût*, *un poète de génie*) et *à* (par ex., *la femme à barbe*). Tesnière souligne d'ailleurs que « [t]outes ces expressions sont de véritables adjectifs [...]. La preuve c'est qu'on peut souvent leur substituer un adjectif sans modifier en rien la structure syntaxique de la phrase : *un poète de génie*, c'est *un poète génial*. » (Tesnière 1959, 439)

Ainsi, pour Tesnière, les cas suivants contiennent tous des exemples de la translation du substantif en adjectif.

- *le livre de Pierre*
- *le train de Paris* ('qui arrive de Paris')
- *le train de Paris* ('qui va vers Paris')
- *le chien du maître*
- *le maître du chien*
- *la gare de Sceaux*

C'est-à-dire que sans être à proprement parler des adjectifs, les expressions *de N* précédentes ont valeur d'adjectifs, puisqu'ils se comportent syntaxiquement comme tels ; ils dépendent d'un gouverneur nominal.

Bien qu'à la base ces constructions soient syntaxiquement semblables, nous ne pouvons les identifier par la même DépSyntS. Effectivement, les dépendants se comportent différemment. Tesnière distingue d'ailleurs plus d'un type de transféré, voir ses classifications en pages 441-446.

Il est donc important de mentionner que même si la notion de translation de Tesnière a servi d'inspiration, nous n'avons pas repris intégralement ses classifications. Ce serait impossible entre autres parce que nous avons une notion d'actant plus large que ce dernier (nous en admettons plus de trois par prédicat, les noms quasi-prédicatifs, les adjectifs et les adverbes peuvent aussi avoir des actants, etc.). Ainsi, nous nous servons de la notion de translation pour l'identification d'une seule dépendance. Notre translation adjectivale est celle où un substantif est transféré en adjectif « qualificatif » à l'aide d'une préposition (comme dans les exemples *un homme de coeur*, *un homme de goût*, *un poète de génie*, *une femme à barbe*). Nous excluons donc de ceci les cas où (dans les termes de Tesnière) le substantif transféré est un substantif subjectif ou objectif ainsi que ceux où le dépendant est un adjectif d'appartenance (*le livre de Paul*), un adjectif circonstanciel (*le train de Paris*) ou un adjectif de quiddité (*la ville de Paris*, *un fripon d'enfant*, *ce chien de Morin*). Ces types de dépendants sont liés à leur gouverneur par des dépendances que nous avons identifiées autrement.

### 3.1.3 Une première description de la construction

#### **N1 —attributive adnominale → N2**

Pour donner une première description de cette construction particulière, nous allons comparer les deux syntagmes suivants :

1. *ceinture de pantalon*
2. *ceinture de mon pantalon*

Au niveau syntaxique, ces deux constructions s'opposent par de nombreuses propriétés. Nous verrons l'ensemble des propriétés de la dépendance **attributive adnominale** (celle présente dans la construction en 1) en 3.2.3. Nous souhaitons d'abord parler tour à tour de la fonction du dépendant, de l'interprétation sémantique de la construction, de l'identité du D prépositionnel et, finalement, de la détermination de N2 ; parce que nous croyons que ce sont ces éléments qui permettent le mieux de bien conceptualiser cette dépendance.

### 3.1.3.1 Fonction du dépendant

En français, le principal rôle d'un adjectif est de qualifier un nom. Toutefois, si aucun adjectif dénotant le sens que l'on souhaite exprimer n'existe, il faudra trouver une façon de pallier ce manque lexical. Plusieurs façons de réussir à exprimer le sens souhaité sont envisageables.

Premièrement, on pourrait dériver morphologiquement un nouvel adjectif, comme cela est possible dans certaines langues. Par exemple, si l'on souhaite modifier le nom *ceinture* en précisant de quel type particulier est la ceinture, on pourrait dériver un adjectif à partir du nom *pantalon* : *pantalonière*, pour former un composé Nom + Adj ; soit *ceinture pantalonière* qui voudrait signifier 'ceinture destinée à être portée avec un pantalon'. Toutefois, une dérivation de ce type en français ne se fait pas de façon totalement libre et est possible presque exclusivement en diachronie, c'est-à-dire qu'un néologisme (par exemple un nouvel adjectif) met du temps à se lexicaliser officiellement.

La deuxième option est d'utiliser une construction appositive comme on le fait en anglais avec les *compound nouns*. En effet, ce sens précis est exprimé en anglais à l'aide d'une construction appositive de type N2 N1 : *pant belt*. Bien que l'on voit parfois une construction appositive (N1 N2) en français (ex. *des camions monstres, une soupe maison, une alarme incendie, un gel douche*), cette option n'est pas aussi systématique qu'en anglais.

Bien sûr, en français, il est possible de modifier un nom analytiquement avec une proposition relative ou une construction participiale : *ceinture qui est destinée à être portée avec un pantalon, ceinture que l'on porte généralement avec un pantalon, ceinture destinée à être portée avec un pantalon, ceinture servant à attacher un pantalon*, etc. Toutefois, cette façon de faire oblige à avoir recours à des constructions parfois très lourdes.

Finalement, la dernière option est d'avoir recours à ce que Tesnière a nommé la translation. Dans ce cas particulier, il s'agit d'utiliser un élément non adjectival en fonction adjectivale. En français, il est donc possible de créer le composé *ceinture de pantalon*, où le nom *pantalon* sert à modifier le nom *ceinture* comme le ferait l'adjectif *pantalonière*,

et la préposition *de* sert de marqueur de la translation.

Ce processus peut être considéré comme étant à la frontière entre la syntaxe et la morphologie. On reconnaît d'ailleurs le statut de composés de certains groupes de ce type. Certains sont lexicalisés, c'est-à-dire qu'ils constituent des locutions : *robe de chambre*, *pomme de terre*, *fruits de mer*. Certains autres sont semi-lexicalisés : *bateau de pirate*, *ceinture de sécurité*. Ces deux types de constructions se distinguent parce que, dans les deuxièmes, N1 garde son statut référentiel (un *bateau de pirate* est un type de *bateau*, une *ceinture de sécurité* est un type particulier de *ceinture*) ; alors que ce n'est pas le cas des premiers (une *robe de chambre* n'est pas une *robe*, une *pomme de terre* n'est pas une *pomme*, un *fruit de mer* n'est pas un *fruit*). Dans son travail sur les compléments du nom en [*de NP*], M. L. Knittel arrive à des conclusions similaires en ce qui a trait à la lexicalisation des constructions de ce type. Suivant le travail de M. Gross sur les noms composés (Gross 1996), elle arrive à la conclusion que « la lexicalisation, dans une langue comme le français, doit être vue comme sujette à gradation. » (Knittel 2009, 268)

### 3.1.3.2 Interprétation sémantique

Revenons à nos deux constructions :

1. *ceinture de pantalon*
2. *ceinture de mon pantalon*

Il est de mise de se demander si dans les deux cas N1 (*ceinture*) est lié à son dépendant prépositionnel par la même dépendance syntaxique. Au niveau du sens, les deux constructions amènent deux interprétations différentes, parce qu'elles expriment effectivement deux sens différents. En (2), N2 réfère à la pièce de vêtement avec lequel la ceinture est portée, dans un moment précis, ce qui n'est pas le cas dans l'exemple (1). Dans ce cas, la précision apportée par le groupe en « *de N2* » ne garantit pas que la ceinture soit utilisée avec un pantalon. Ce groupe sert plutôt à préciser le sous-groupe auquel N1 appartient ; il est utilisé pour sous-catégoriser N1. Ainsi, le groupe « *N1 de N2* » fait référence à une ceinture d'un certain type. Il forme un groupe nominal référentiel contrairement à N2 pris isolément. Par opposition, en (2), N2, tout comme N1, est

pleinement référentiel et le groupe en « *de N2* » est utilisé afin de spécifier N1. En général, les auteurs qui se sont penchés sur les constructions en « *N1 de N2* » parviennent également à des conclusions similaires. À ce sujet, M. L. Knittel mentionne :

« une distinction claire oppose les constructions possessives et les structures en [de NP]. En effet, l'absence de déterminant qui distingue les possesseurs des NP introduits par *de* résulte systématiquement en une réinterprétation du second nom, qui apparaît alors comme une sorte de complément à valeur taxinomique, permettant d'exprimer une sous-classe du référent du nom principal. » (Knittel 2009, 261)

Dans le même ordre d'idées, L. Benetti distingue principalement deux types d'opérations pouvant être supportées par une expansion en « *de N2* » : l'opération de sous-catégorisation (repérage intensionnel) où N1 peut être épithétisé par un sous-catégorisateur comme dans *chien de chasse, maison de campagne, blouson d'aviateur* et l'opération de spécification (repérage extensionnel) où N1 peut être augmenté de ce qu'elle appelle un « circonstant » de type Prép SN comme dans *le blouson → de l'aviateur, le chien de chasse → de Pierre, l'espèce de maison de campagne → sur la colline*. (Benetti 2008, 118-119)

### 3.1.3.3 Identité du D prépositionnel

En général, lorsque l'on traite de ce type de constructions, on précise que l'on s'en tient aux constructions de type « *N1 de N2* ». Il est de mise de se demander s'il est possible qu'une autre préposition puisse introduire un syntagme qui aurait la même fonction, soit introduire un syntagme prépositionnel sous-catégorisateur pour former un « GN complexe ». En fait, il semble que *de* ne soit pas la seule préposition qui puisse introduire un complément de ce type. En effet, il existe des cas où une autre préposition sert à introduire un élément nominal qui prend une valeur de modificateur adjectival. Grevisse donnait comme exemples du complément de caractérisation *un fruit à noyau, un verre à eau, un canon contre avions, une table en chêne*, Tesnière parlait de *femme à barbe*, etc. Ce semble également être le cas dans les constructions suivantes : *gâteaux aux noix, tir à l'arc, prison pour femmes, menu pour enfants, magazine pour femmes*, etc. Il semble que

ce qui caractérise la préposition (qu'elle soit À, DE, POUR, CONTRE, EN, ou autre) est qu'elle perde en partie de son sens. Elle est avant tout un outil qui introduit le N2 ; sa valeur est avant tout syntaxique. Dans ce type de constructions, c'est le N2 ainsi introduit qui est sémantiquement important, il donne une caractéristique du N1.

#### 3.1.3.4 Détermination de N2

De prime abord, il pourrait sembler que le nom suivant la préposition soit non défini, c'est-à-dire utilisé sans déterminant (ce qu'on qualifie parfois de nom nu). Par contre, ce n'est pas toujours le cas. En effet, le nom suivant la préposition (N2) est parfois utilisé avec l'article défini LE et la construction Prép N2 effectue la même modification auprès de N1 que les constructions en « *de N2* » dont nous avons traité plus haut. C'est ce qui se produit dans les exemples suivants (les 3 derniers étant repris de Benetti (2008)) : *tir à l'arc, gâteau au chocolat, tarte aux pommes, fleur des champs, rose des vents, fraise des bois, Faculté des Lettres, les sciences de l'ingénieur.*

Toutefois, dans les rares cas où l'article défini (LE ou LES) est présent, il semble que ce soit le LE-générique. En somme, dans la construction qui nous intéresse, le N2 est soit non déterminé, soit déterminé par un article défini générique. Ceci est compatible avec le fait que N2 est non référentiel puisqu'il ne dénote pas un élément précis de la réalité, seulement un groupe de choses. Les D en « Prép N2-déf. » des exemples précédents servent donc également de classifiants. Ainsi, dans une expression comme *tarte aux pommes*, le dépendant *aux pommes* sert à préciser le type particulier de tarte auquel on fait référence, ce qui distingue la tarte dont on parle des autres tartes. C'est donc dire que l'on observe la même interprétation sémantique que dans une construction comme *ceinture de pantalon*, où l'absence de déterminant est en quelque sorte analogue au défini générique. Il en est probablement ainsi parce que l'ancien français « marquait régulièrement par l'absence de déterminant la référence virtuelle ou générique ». (Riegel et al. 1994, 166)

### 3.2 Méthodologie d'identification/ caractérisation d'une DépSyntS

Le premier objectif de la présente recherche étant de contribuer à la création d'une liste des relations syntaxiques de surface du français, nous proposons dans cette section de recenser les étapes qui permettent une caractérisation complète d'une dépendance donnée. Ces étapes de caractérisation visent à bien cerner les propriétés d'une dépendance, ce qui en permet l'identification efficace. Nous fournissons une démonstration pour la dépendance **attributive adnominale** basée sur une sélection de lexies nominales du DiCo.

#### 3.2.1 Critères de regroupement de différentes constructions syntaxiques sous une même dépendance

Les propriétés linguistiques pertinentes des DépSyntS, ou de leurs dépendants, sont liées au fait que les relations sont sujettes à trois types de contraintes, telles qu'explicitées dans Iordanskaja et Mel'čuk (2009, 155).

- Une DépSyntS est associée au rôle sémantique que son dépendant remplit — via la DépSyntP correspondante (qui est à son tour reliée à un rôle sémantique).
- Les DépSyntS se combinent à l'intérieur de la SSyntS, où elles ont chacune un comportement particulier.
- Les DépSyntS servent à la linéarisation, prosodisation (ou ponctuation à l'écrit) et morphologisation de la SSyntS (vers la SMorphP).

Tous les dépendants possibles d'une DépSyntS donnée doivent donc avoir des propriétés linguistiques identiques ou très similaires en regard à ces trois types de contraintes. Afin d'établir avec certitude les propriétés qui permettent d'identifier adéquatement la dépendance **attributive adnominale**, nous examinerons le comportement des dépendants potentiels selon des propriétés : 1) sémantico-syntaxiques ; 2) syntaxiques ; et 3) morpho-syntaxiques.

Pour une première caractérisation générale d'une DépSyntS, il est utile d'établir quels sont les gouverneurs possibles et les dépendants prototypiques de la dépendance

ciblée. Dans le cas qui nous concerne, la DépSyntS **attributive adnominale** en est une qui est exclusivement nominale, c'est-à-dire que le gouverneur est nécessairement un nom. Quant au dépendant prototypique, c'est un nom non référentiel précédé d'une préposition. Ceci étant fait, nous pouvons rechercher des propriétés générales des trois types mentionnés précédemment. En ce qui concerne la dépendance **attributive adnominale**, nous avons créé cette instance de relation pour rendre compte des constructions où un groupe prépositionnel semblait modifier un nom, comme le ferait ordinairement un adjectif. Nous avons donc nommé ce critère syntaxique : Source de la dépendance : Translation  $\rightarrow$  Adj.

Évidemment, il est important de tester les propriétés à l'aide de phrases d'exemples qui contiennent cette dépendance. Les constructions que nous retenons pour arriver à cerner les critères d'identification de la dépendance **attributive adnominale** sont celles où nous suspectons que cette Translation  $\rightarrow$  Adj a été utilisée. Nous désirons entre autres établir si la DépSyntS est la même dans les constructions où le groupe prépositionnel exprime un ASém du gouverneur et dans celles où ce n'est pas le cas. Évidemment, nous commençons par cerner les propriétés des constructions où le dépendant est un ASém potentiel du gouverneur, et donc que le gouverneur contrôle son expression. Nous pourrions ensuite décider si les D qui ne sont pas régis par la valence du G nominal partagent assez de ces propriétés pour faire partie de la même dépendance.

### 3.2.2 Construction d'un corpus d'exemples

Afin de bien cerner les diverses propriétés des DépSyntS, la construction d'un corpus d'exemples qui contiennent des instances supposées de cette dépendance est primordiale. Voici donc une sélection de lexies nominales où un des ASém est potentiellement exprimable par la DépSyntS **attributive adnominale**. La lexie est suivie de son numéro d'acception dans le DiCo, de son étiquette sémantique (és) et de sa forme propositionnelle (fp). Pour chacune d'entre elles, nous fournissons quatre exemples en identifiant clairement le gouverneur [G] et le dépendant [D] de la dépendance. Potentiellement, la relation syntaxique qui les lie dans la phrase est la DépSyntS **attributive adnominale**. Les exemples sont repris directement du

DiCo ou tirés d'une recherche sur Internet<sup>2</sup>. Il est important de comprendre que ce corpus sert à illustrer notre méthodologie et qu'il est nettement insuffisant, de par sa petite taille, pour tester de façon adéquate l'ensemble des propriétés d'une dépendance. Notre travail de caractérisation s'est effectué sur un corpus plus large, soit l'ensemble des lexies nominales du DiCo qui semblaient pouvoir être modifiées par un D de ce type, ainsi que des exemples construits avec d'autres lexies. Nous avons également construit des exemples où cette dépendance n'exprime pas un des ASém de son gouverneur.

#### CEINTURE#I.1

és: pièce d'habillement

fp: ~ DE L'individu X QU'IL UTILISE AVEC LE vêtement Y

1. *Pour ceux qui souhaitent nouer leur Obi (ceinture[G] de[D] kimono) de façon traditionnelle, voici quelques images qui pourront vous apporter un meilleur confort lors de l'apprentissage de l'Aïkido.*
2. *Suite à une césarienne, vous pouvez accrocher une ceinture[G] de[D] robe de chambre aux barreaux du bout du lit afin de pouvoir vous redresser grâce à la traction des bras.*
3. *À partir de ce jour, il est formellement interdit à tout Écossais de porter le kilt, le plaid, la ceinture[G] de[D] plaid, la ceinture[G] de[D] kilt, le petit kilt ou toute pièce de vêtements ou d'ornement pouvant permettre l'identification d'un Clan.*
4. *À vendre : Ceinture[G] de[D] pantalon qui permet de cacher les billets, parfaite pour mettre en sécurité votre argent durant vos déplacements à l'étranger.*

#### FARINE<sup>3</sup>

és: substance alimentaire

fp: ~ DE plante X

1. *Dès le premier repas à la ferme-auberge locale, le ton est donné : jambon corse et pâté de tête, cannelloni au brocciu et gâteau à la farine[G] de[D] châtaigne.*

---

2. Il est à noter que certains de ces exemples ont été modifiés par volonté de clarté. On aura par exemple enlevé certains éléments non pertinents pour ne pas complexifier les phrases, ou ajouté quelques mots qui aident à leur compréhension. Par exemple, la phrase suivante (trouvée sur le site <http://www.cesarine.org/apres/hopital/>) : *Vous pouvez aussi installer des moyens de fortune comme une ceinture de robe de chambre accrochée aux barreaux du bout du lit afin de pouvoir vous redresser grâce à la traction des bras.* est devenue : *Suite à une césarienne, vous pouvez accrocher une ceinture[G] de[D] robe de chambre aux barreaux du bout du lit afin de pouvoir vous redresser grâce à la traction des bras.*, le deuxième exemple pour la lexie CEINTURE#I.1.

3. Un exemple trouvé suggère qu'il manque au moins une acception dans ce vocable. *Mme de Palacio demande que la réforme distingue mieux les formes de pêche qui alimentent en priorité la consommation des particuliers de la pêche industrielle, surtout destinée à la production de farine de poissons.*

2. *Après l'élection, quel qu'en soit le vainqueur, le prix de la farine[G] de[D] maïs va doubler.*
3. *Les meuniers français font pression pour obtenir la diminution du tarif douanier sur la farine[G] de[D] blé.*
4. *La farine[G] de[D] sarrasin foncée contient environ deux fois plus de protéines que la farine de sarrasin plus pâle.*

## FLEUVE#II

és: grande quantité

fp: ~ DE X

1. *Un fleuve[G] de[D] boue a recouvert les quartiers nord de la ville.*
2. *Le volcan a déversé un fleuve[G] de[D] lave.*
3. *La douleur m'a brisée, la fraternité m'a relevée, de ma blessure a jailli un fleuve[G] de[D] liberté.*
4. *Lorsqu'on lui rend visite à Rio, comme ce fut le cas d'un cortège d'architectes et de journalistes ravis de voir de près ce monument de sympathie, ce fleuve[G] d'[D]élégance naturelle, il joint toujours le geste à la parole.*

## GÂTEAU

és: préparation alimentaire

fp: ~ FAIT AVEC X

1. *Quand j'ai vu ce beau gâteau[G] à[D] l'érable dans le magazine, c'est son originalité qui est venue piquer ma curiosité.*
2. *Avec l'arrivée des self-services dans les cantines, l'enfant peut choisir entre le gâteau[G] à[D] la semoule et la salade de fruits.*
3. *Quant au toponyme «Chazeron», il fournit maintenant une appellation au gâteau[G] à[D] la noisette inventé en 1996 par le maître boulanger Nury.*
4. *Le gâteau[G] à[D] la mayonnaise est le meilleur gâteau[G] au (= à[D]le) chocolat sur terre*

## LAIT#b

és: substance alimentaire

fp: ~ D'animal X

1. *Son auteur relève une contamination par l'iode 131 de plus de 10 000 becquerels par litre dans du lait[G] de[D] brebis en Corse.*
2. *Le lait[G] de[D] vache entier pasteurisé peut être donné au nourrisson à partir de neuf à douze mois et pendant toute la deuxième année.*
3. *La brousse du Rove, ce fromage onctueux, est fabriquée exclusivement avec du lait[G] de[D] chèvre, qu'on fait bouillir.*

4. *Mais recevoir, des mains du chef de famille, un morceau de soie bleue sur lequel il a versé quelques gouttes de lait[G] de[D] yack, est évidemment un honneur.*

#### OUTIL#I

és: instrument

fp: ~ utilisé par l'individu X pour faire l'action Y(X)

1. *La dégauchisseuse est un outil[G] de[D] menuiserie qui permet de dégauchir le bois.*
2. *J'ai en ma possession un outil[G] de[D] cordonnerie pour réparer les chaussures à la main.*
3. *Tous sont costumés d'une manière grotesque ayant de grands tabliers et tenant chacun à la main un outil[G] de[D] maçonnerie.*
4. *Comment choisir le bon outil[G] de[D] jardinage ?*

#### RESTAURANT

és: établissement de restauration

fp: ~ OÙ TRAVAILLE L'ensemble d'individus X POUR SERVIR À L'ensemble d'individus Y LA nourriture Z

1. *Mesquite - un restaurant[G] pour[D] carnivores.*
2. *Elle travaillait le soir dans un restaurant[G] pour[D] allergiques à Montréal.*
3. *Le restaurant[G] pour[D] enfants La Fermette est ouvert le week-end et pendant les vacances scolaires.*
4. *Les restaurants[G] pour[D] ouvriers se caractérisent par un menu relativement fixe et relativement bon marché.*

fp: ~ OÙ TRAVAILLE L'ensemble d'individus X POUR SERVIR À L'ensemble d'individus Y LA nourriture Z

1. *Ce petit restaurant[G] de[D] saucisses traditionnelles a perdu une grande partie de sa clientèle.*
2. *La petite troupe se sépare en se promettant de se faire une bouffe chez Omar, un restaurant[G] de[D] couscous de la rue de Bretagne.*
3. *Cette jeune femme timide dirige Le Pescador, le restaurant[G] de[D] fruits de mer qui jouxte la Playa Minitas.*
4. *Son fils a racheté un hôtel à l'abandon pour en faire un restaurant[G] de[D] cuisine traditionnelle.*

#### TÉMOIN#II.1

és: artefact servant à fournir une information

fp: ~ INDIQUANT L'INFORMATION Y À L'individu X

1. *Il est dommage que ce modèle de voiture soit équipé d'un témoin[G] d'[D]essence et non d'une jauge.*

2. *Si le témoin[G] d'[D]alimentation ne s'allume pas et que l'appareil photo ne fonctionne pas, mettez l'appareil photo sous tension.*
3. *Grâce au témoin[G] d'[D]activité qui mesure vos pas minute par minute, vous pourrez désormais vous fixer des objectifs de marche et mesurer l'évolution de votre activité au quotidien.*
4. *La couleur du témoin[G] de[D] batterie intégré au commutateur indique l'autonomie restante de la batterie.*

TIR#1b

és: sport

fp: ~ [PRATIQUÉ PAR L'individu X] AVEC L'arme Y

1. *Anne Briand a mis fin en 1999 à sa carrière en biathlon — un sport associant ski de fond et tir[G] à[D] la carabine.*
2. *Les droits TV sont les plus lucratifs pour les fédérations importantes : quel contraste entre les 740 heures de diffusion du football sur les chaînes hertziennes et les quelques minutes consacrées à l'haltérophilie ou au tir[G] à[D] l'arc.*
3. *Il est devenu champion de France de tir[G] au(=à[D] le) pistolet à cinquante mètres.*
4. *Ils sont propriétaires d'un stand de tir[G] à[D] la carabine.*

ZONE

és: espace

fp: ~ OÙ A LIEU LE fait X

1. *L'avion, entré dans une zone[G] de[D] turbulence, se mit brusquement à tanguer.*
2. *Au cours des années 1990, la création de nombreuses zones[G] de[D] libre-échange permit d'aplanir les obstacles à la circulation internationale des capitaux et des marchandises.*
3. *Ce n'est pas l'objet de ce rapport que d'analyser les nombreuses zones[G] de[D] conflits politiques.*
4. *Ce bateau aura pour tâche d'intercepter les émissions radio et radar à proximité de la zone[G] de[D] crise où il évolue.*

### 3.2.3 Propriétés de la dépendance attributive adnominale

Étudions maintenant les cas décrits ci-dessus dans l'optique de cerner les propriétés des dépendants selon les critères fixés au préalable. La liste des premières propriétés pertinentes pour l'analyse des constructions G[nominal] → D a été établie suite à un travail fait par Lidija Iordanskaja (inédit). Nous avons donc utilisé ces propriétés pour

tester la DépSyntS qui nous intéresse. Nous présentons ici chacune des propriétés pertinentes pour cette dépendance particulière avec quelques exemples qui permettent de bien illustrer la propriété testée. L'ensemble des propriétés que nous avons retenues pour l'analyse des constructions  $G[\text{nominal}] \rightarrow D$  sont consignées plus loin dans le tableau 3.1., avec des précisions concernant les constructions étudiées ici.

### 3.2.3.1 Propriétés formelles de G et D

Tout d'abord, comme nous l'avons mentionné précédemment, le Gouverneur de cette relation est un nom commun et le Dépendant prototypique est nom non référentiel qui est gouverné par une préposition dont le sens est plus ou moins vide. Il s'agit le plus souvent de la préposition DE, dans une moindre mesure de la préposition À et parfois d'une autre préposition (EN, POUR, CONTRE, etc.).

### 3.2.3.2 Propriété(s) définitive(s)

1. Le complément attributif se comporte comme un adjectif qualificatif, il modifie son G nominal. Il est le résultat d'un phénomène de Translation  $\rightarrow$  Adj.

### 3.2.3.3 Autres propriétés descriptives

1. Si un nom suit la préposition, il est non déterminé ou déterminé par un article défini générique.

Comme nous l'avons mentionné précédemment, dans la plupart des cas, le nom suivant la préposition n'est pas défini, c'est-à-dire utilisé sans déterminant. Par contre, dans certains cas, ce nom est utilisé avec l'article défini LE (ex. *tir à l'arc*, *gâteau au* (= à *le*) *chocolat*). Toutefois, dans les quelques cas où l'article LE est présent, il s'agit du LE-générique.

2. Le Dépendant ne peut être paraphrasé par un adjectif possessif (A-poss).
  - *La ceinture de pantalon de Marie est détachée.*  $\sim$  *\*Sa ceinture de Marie est détachée.*
  - *Hier, j'ai acheté une ceinture de pantalon.*  $\sim$  *\*Hier, j'ai acheté sa ceinture.*
3. Le Dépendant ne peut être pronominalisé par *dont*.

Dans des exemples comme : *C'est un pantalon dont j'ai égaré la ceinture* et *Voici le pantalon dont j'ai fabriqué la ceinture*, le *dont* pronominalise plutôt le « de N » de la dépendance **complétive adnominale** (la ceinture du/de ce pantalon).

4. Le Dépendant ne peut être pronominalisé par *en*.

– *Je veux du lait de chèvre.* ~ *\*[De chèvre,] j'en veux du lait.*

5. Le Dépendant ne peut pas être modifié par une relative descriptive.

– *\*une ceinture de kimono que je porte quand je fais du karaté* [dans le sens où je porte le kimono]

– *\*de la viande de cheval qui galope dans le pré*

– *\*J'ai dévoré un gâteau au chocolat avec lequel ma soeur fait ses fondues.*

Ceci découle évidemment de la non-référentialité du N2. Il faut ajouter que la précision quant au statut descriptif de la proposition relative est nécessaire, parce que si la relative modifie la classe d'objet, la relativisation est possible.

– *\*une ceinture de kimono que je porte quand je fais du karaté* [dans le sens où je porte le kimono] ~ *une ceinture de kimono qui sert à faire du karaté*

– *\*de la viande de cheval qui galope dans le pré* ~ *de la viande de cheval qui est nourrie au grain*

Le dépendant peut donc être modifié par une construction qui effectue une modification analogue à celle du complément attributif ; soit une relative classifiante, un autre complément attributif ou un adjectif.

– *\*J'ai dévoré un gâteau au chocolat avec lequel ma soeur fait ses fondues.* ~ *J'ai dévoré un gâteau au chocolat qui vient des Andes.* *J'ai dévoré un gâteau au chocolat au lait.* *J'ai dévoré un gâteau au chocolat biologique.*

6. La dislocation à gauche est impossible.

– *Un restaurant de mets chinois a ouvert ses portes dans le quartier.* ~ *\*De mets chinois, un restaurant a ouvert ses portes dans le quartier.*

– *Ce restaurant de moules offre un très bon rapport qualité-prix.* ~ *\*De moules, ce restaurant offre un très bon rapport qualité-prix.*

7. Le Dépendant tend à se linéariser immédiatement à la suite du gouverneur nominal, avant les autres dépendants.

– *La farine de blé du Canada se vend bien à l'étranger.* ~ *\*La farine du Canada de blé se vend bien à l'étranger.*

– *Le témoin d'essence de ma voiture clignote sans cesse.* ~ *\*Le témoin de ma voiture d'essence clignote sans cesse.*

- *Le bateau de pirate de Paul est arrivé ce matin.* ~ *\*Le bateau de Paul de pirate est arrivé ce matin.*

Mais il s'agit bien là d'une tendance, puisque dans certains cas où l'on souhaite apporter des précisions supplémentaires, un adjectif qualifiant le gouverneur peut précéder le complément attributif (par ex., *gel nettoyant pour bébés*, *ouvrages généraux de mathématique*, *gâteau monté aux fruits*, [Qc] *gâteau renversé à l'ananas*). Par contre, il est tout de même plus difficile d'insérer un adjectif entre le G et le D lorsque la préposition est DE. Même si l'on voit souvent des exemples comme : *station manuelle d'alarme*, *tomate mure de jardin*, *mois gratuit d'entraînement*, *frais cachés d'inscription*, il est souvent plus naturel de préconiser l'ordre inverse : *station d'alarme manuelle*, *tomate de jardin mure*, *mois d'entraînement gratuit*, *frais d'inscription cachés*.

#### 8. Le Dépendant est répétable<sup>4</sup>.

Il est évident qu'il n'est pas possible d'exprimer de façon répétée un D qui correspond à un ASém donné.

- *\*Je pratique le tir à l'arc à la carabine au fusil.*
- *\*Pourrais-tu acheter du lait de vache de chèvre de brebis ?*

Toutefois, il semble que la dépendance **attributive adnominale** en tant que telle soit répétable si elle exprime à chaque fois un contenu sémantique différent, surtout si un des deux compléments est en processus de lexicalisation.

- *Maryse préfère les gâteaux de fête au chocolat.*
- *Mark a acheté un moulin à café à piles.*
- *J'ai une aversion envers les restaurants de hamburgers pour enfants.*
- *Je ne vais jamais là-bas, c'est un bar à tapas d'alcooliques*

#### 9. Paraphrases : le complément attributif peut être remplacé par un adjectif de sens équivalent (si un tel adjectif existe).

---

4. Il est important de porter une attention particulière en testant cette propriété, parce que la linéarisation de plusieurs D prépositionnels est assez complexe en français. On ne peut pas tester, par exemple, avec CEINTURE#I.1. dans des exemples comme *\*une ceinture de pirate[=X] de pantalon[=Y]* ; *une ceinture[G] de pantalon[=Y] de pirate[=X]*, parce qu'on interprète le deuxième complément attributif (*de pirate*) comme dépendant du N inclus dans le premier complément attributif (*pantalon*). Même lorsque c'est possible, certaines contraintes d'ordonnancement doivent être prises en considération (ex. *un gâteau de fête au chocolat* ~ *\*un gâteau au chocolat de fête* ; *un moulin à café à piles* ~ *\*un moulin à piles à café*).

- ‘un restaurant de cuisine mexicaine’ = ‘un restaurant mexicain’
- ‘un magazine pour femmes’ = ‘un magazine féminin’

Il est important de souligner que les propriétés descriptives que nous venons d'énumérer ne font pas partie intégrante de la description de la dépendance **attributive adnominale** dans la hiérarchie GTN (voir p.26), du moins pour l'instant. Bien que n'étant pas explicitement fournies dans la hiérarchie, nous reconnaissons évidemment que ces propriétés sont grandement utiles lors de l'identification de cette dépendance et absolument nécessaires pour une catégorisation complète d'une DépSyntS donnée. On pourra également retrouver ces propriétés dans la grammaire du français. Les règles de grammaire diront, par exemple, qu'il est impossible de remplacer un complément attributif par un adjectif possessif, qu'il doit être réalisé linéairement le plus près possible de son gouverneur, etc.

#### 3.2.3.4 La DépSyntS **attributive adnominale** : une dépendance régie, non régie ou mixte ?

D'abord, comme nous le verrons plus en détail dans le chapitre suivant, les DépSyntS sont classées dans l'ontologie GTN selon deux optiques différentes, soit : vis-à-vis de la partie du discours du gouverneur et vis-à-vis du caractère typiquement régi ou non régi de la dépendance. Évidemment, comme les dépendances pertinentes pour mener à bien notre travail sont celles qui sont gouvernées par des noms, nous avons choisi pour notre démonstration une dépendance faisant partie de la classe des DépSyntS nominales. Toutefois, il n'est pas tout à fait clair d'entrée de jeu si cette dépendance doit être classée comme une dépendance typiquement régie (typiquement gouvernée par la valence du gouverneur), non régie (typiquement combinable librement à un gouverneur) ou mixte (typiquement combinable librement à un gouverneur ou gouvernée par sa valence).

La prochaine étape dans la caractérisation de la DépSyntS **attributive adnominale** est donc de vérifier si cette dépendance est une DépSyntS mixte. Après avoir réfléchi sur le sujet, il semble que la construction syntaxique que nous étudions soit indifférente à la nature d'ASém du dépendant. Il est tout à fait possible de construire

une construction analogue avec un D qui n'est pas un des ASém potentiels de la lexie nominale gouvernante.

Comme nous l'avons vu précédemment, il est possible d'indiquer un type de ceinture en précisant avec quel vêtement elle est typiquement utilisée (actant Y). D'autre part, la lexie CEINTURE#I.1 peut aussi être classifiée par :

– le porteur prototypique (actant X) :

1. *Loin de singer la femme, les tuniques en mousseline, les jabots flasques, les ceintures[G] de[D] pirates à boucle de nacre et les chemisettes en daim parlent de fragilité.*
2. *Il allie le coton d'un tee-shirt à la fluidité de la mousseline et tranche le tout par une large ceinture[G] de[D] cow-boy.*
3. *La collection se distingue par une sobriété soutenue, par des coupes parfaites pour des pantalons de laine à ceinture[G] de[D] geisha, des jupes aux genoux et une multitude de tuniques bicolores.*

– son usage (pas un actant) :

1. *Il est à noter que la ceinture[G] de[D] karaté se distingue (sauf pour la noire) par un liseré rouge central sur toute la longueur.*
2. *La ceinture[G] de[D] judo est un équipement spécifique au judogi, car il permet à la fois de fermer son costume d'entraînement et d'afficher son galon et son niveau.*
3. *En Belgique, le mot apertintaille désigne une ceinture[G] de[D] carnaval avec des grelots.*

– le matériel dont elle est composée (pas un actant) :

1. *Tout de blanc vêtu, hormis une large ceinture[G] de[D] satin jaune, le vieillard a placé un plateau d'offrandes pour les dieux et les démons sur la pointe de l'île.*
2. *«Il a un oeil diabolique», dit Anne-Marie, dont la ceinture[G] de[D] croco portée sur un sweater de cachemire chocolat signait l'allure.*
3. *Le jeune a découvert, dans le couloir de sa maison, une ceinture[G] de[D] toile kaki à laquelle est attaché un poignard.*

On peut également voir cette opposition dans des constructions comme :

- *gardien de prison* [actant Y potentiel] ~ *gardien de nuit* [circonstant de temps où se pratique l'activité] (exemple de Knittel (2009, 262))
- *chien de berger* [actant X potentiel] ~ *chien de chasse, chien de traîneau* [usages potentiels]

– *bateau de pirate* [actant X potentiel] ~ *bateau de pêche* [usage potentiel] ~ *bateau à moteur, à voile* [méthodes de propulsion]

Ainsi, si les constructions où le D n'est pas un ASém de son G et celles où il l'est présentent un nombre suffisant de propriétés communes, nous pourrions considérer que l'ensemble de ces constructions peuvent appartenir à la même DépSyntS. À titre de démonstration, nous présentons ici, à l'aide d'un tableau, les propriétés cernées pour l'identification des constructions G(nominal) → D en précisant si elles sont : 1) syntaxico-sémantiques ; 2) purement syntaxiques ; ou 3) morpho-syntaxiques. Évidemment, ce tableau permet également de comparer les constructions où le D correspond à un ASém du G et celles où ce n'est pas le cas. De plus, le tableau est conçu de façon à pouvoir être utilisé lors de l'étude d'autres constructions G[nominal] → D.

Le — indique la non-pertinence d'une propriété pour la dépendance examinée. Dans les deux cas où il est présent, le +/- sert à indiquer que : 1) il n'est pas possible d'insérer un élément entre le complément attributif et son gouverneur, sauf un autre complément attributif (et parfois un adjectif) ; et 2) le complément attributif se positionne généralement devant les autres dépendants, sauf évidemment lorsqu'il y a plus d'un complément attributif (ou parfois en présence d'un adjectif).

Type de propriétés	Propriété	C. régies	C. non régies
	G	N commun	N commun
	D prototypique	Prép (+ N non réf.)	Prép (+ N non réf.)
Sémantico-syntaxiques	Relation syntaxique profonde correspondante	I, II, ..., VI	ATTR
	Paraphrases	Adjectif	Adjectif
Syntaxiques	Possibilité de dislocation à gauche	-	-
	D paraphrasable par A-poss	-	-
	Pronominalisation du D par <i>dont</i>	-	-
	Pronominalisation du D par <i>en</i>	-	-
	Possibilité d'insertion entre G et D	+/-	+/-
	Possibilité pour D d'être modifié par une relative descriptive	-	-
	Résultat d'un phénomène de translation → Adj.	+	+
	Répétabilité de D	+	+
Morpho-syntaxiques	Dépendance morphologique (D → G ou G → D)	—	—
	Expression de détermination de N2	Non déterm. ou Déf. génér.	Non déterm. ou Déf. génér.
	Flexion indépendante du D	+	+
	Position linéaire du D par rapport à son G	Post-posé	Post-posé
	Ordre des codépendants : Avant les autres D	+/-	+/-
	Prosodie/Ponctuation	—	—

Tableau 3.1 – Propriétés de la dépendance **attributive adnominale**

Bien que le dépendant prototypique de cette dépendance soit un nom non référentiel gouverné par une préposition sémantiquement déficiente, il est possible que cette préposition (le véritable D syntaxique) soit suivie d'un verbe à l'infinitif. Dans des exemples comme *problème à résoudre*, *bicyclette à vendre*, etc. la construction Prép + V-inf effectue une modification analogue à celle présente dans les constructions analysées plus haut. Nous considérons donc qu'elles font également partie de la dépendance **attributive adnominale**.

De plus, nous croyons intuitivement qu'un groupe  $G[N] \rightarrow D[\text{complément attributif}]$  a certaines propriétés prosodiques particulières, soit former une certaine unité prosodique plus « solide » que d'autres syntagmes nominaux. Nous ne l'avons pas indiqué clairement ici, car nous n'avons pas de véritables moyens accessibles pour tester cette intuition.

Suite à cette comparaison, il semble que les propriétés mentionnées précédemment sont également valables pour les dépendants non actanciers. Puisque les D semblent partager un nombre important de propriétés, nous pouvons affirmer qu'il s'agit de la même dépendance dans l'ensemble des cas examinés. La DépSyntS **attributive adnominale** fait donc vraisemblablement partie de la classe des DépSyntS Mixtes. Nous l'avons placée à cet endroit dans la hiérarchie GTN, aux côtés des autres DépSyntS mixtes, soit : **appositive compositive, complétive adnominale et déterminative**.

Maintenant que l'ensemble des propriétés (définitoires et descriptives) de la dépendance **attributive adnominale** ont été cernées, il est temps de traiter de la manière qui soit la plus appropriée pour encoder, dans l'entrée lexicale de la lexie gouvernante, la DépSyntS du français que nous venons d'identifier (dans les cas où elle exprime un ASém de son G bien entendu).

### 3.3 La dépendance **attributive adnominale** et le « Régime lexicalisé »

En débutant cette recherche, nous souhaitions réfléchir sur la modélisation lexicographique des dépendances syntaxiques régies. Nous avons proposé de suivre la méthode de la LEC, soit d'utiliser le régime lexical pour inclure les dépendances syntaxiques qui permettent d'exprimer les ASém des lexies prédicatives de toutes les parties du discours. Toutefois, comme la dépendance **attributive adnominale** est quelque peu atypique du fait de sa mixité, nous présentons dans cette section une réflexion qui s'est imposée d'elle-même lors de l'étude plus approfondie de cette dépendance nominale : la place qui doit être accordée aux compléments attributifs au sein des fiches lexicographiques.

#### 3.3.1 La création lexicale

Comme nous avons pu le voir en 3.1.3.2, la DépSyntS **attributive adnominale** sert à préciser de quel type est le référent du gouverneur nominal. D'ailleurs, ce n'est pas par hasard que l'on se sert abondamment de cette dépendance en terminologie lors de la création de nouvelles unités lexicales spécialisées (ou termes). En effet, dans de nombreux cas, un terme se crée à l'aide d'une unité lexicale de la langue générale à laquelle on ajoute de l'information pertinente dans le domaine de spécialisation. Par exemple, on retrouve sous la lettre C du DiCoInfo (*Dictionnaire fondamental de l'informatique et de l'Internet*), quelques termes formés d'un gouverneur nominal et d'un dépendant lié à celui-ci par ce que nous croyons être la dépendance **attributive adnominale** : *cadence d'horloge, cartouche d'encre, cartouche d'impression, champ de saisie, chemin d'accès, clavier d'ordinateur, clé de chiffrement, clé de cryptage, clé de mémoire, clé de mémoire USB, clé de stockage USB, clic de souris, copie de sauvegarde*. Dans ces constructions, le dépendant peut correspondre à un des actants sémantiques de l'unité lexicale générale qui est le gouverneur ou non (par ex., *clavier d'ordinateur* ~ *copie de sauvegarde*).

Bien évidemment, ce type de construction syntaxique n'existe pas seulement dans les langues spécialisées, mais aussi dans la langue générale. Ce qui est exprimé au côté de

la lexie gouvernante pour en préciser la sous-classe est choisi en fonction de cette lexie ou de sa classe sémantique. En général, pour une *préparation alimentaire*, on précisera un ou plusieurs de ses ingrédients (par ex., *tarte aux pommes*, *gâteau au chocolat*) ; pour un *artefact*, on précisera souvent – directement ou de façon abrégée – son utilité (par ex., *moulin à café*, *dé à coudre*, *témoin d'essence*), ou ce qui le fait fonctionner (par ex., *bateau à voile*, *moulin à vent*, *moteur à essence*). Le choix en question n'est donc pas constant (on ne peut pas dire que c'est toujours l'ASém1 par exemple) et est contraint lexicalement. Ceci veut dire qu'en plus d'exprimer le type d'une unité lexicale donnée à l'aide de ses modificateurs potentiels, il est possible de le faire en précisant un de ses actants sémantiques potentiels. Dans ces cas précis où N2 est un ASém potentiel de N1, tout porte à croire que le dépendant exprimé ne correspond pas à une véritable variable (X, Y, etc. de la forme propositionnelle), mais serait plutôt un classifiant (N2 classe la lexie N1, il n'exprime pas vraiment un de ses ASém.). Il s'agirait en quelque sorte d'un « actant métonymique » qui ne sature pas vraiment la position actancielle de la lexie prédicative, lorsqu'exprimé dans une phrase<sup>5</sup>.

Dans les cas où le complément attributif semble exprimer un ASém du gouverneur, deux possibilités sont envisageables : soit l'encoder dans les régimes lexicaux, soit dans la zone de combinatoire lexicale. Cette incertitude découle du fait qu'il n'est pas évident que la position actancielle soit effectivement saturée par la présence d'un complément attributif. En fait la question cruciale à laquelle on doit répondre dans de tels cas, est de savoir si le dépendant est un ASyntP ou alors un modificateur de type ATTR de L, ce qui n'est pas évident. Nous examinerons donc brièvement tour à tour ces deux options.

### 3.3.2 Encodage dans les régimes lexicaux

Dans les cas où l'expression d'un complément attributif est régie par la valence de L, il est nécessaire d'indiquer cette information dans son régime lexical. Voici hypothé-

---

5. Le terme d'actant métonymique a été utilisé parce que dans des cas comme : *moulin à café*, *prison pour femmes*, *ceinture de pantalon*, *restaurant de hamburgers*, le café est effectivement une partie de ce que peut moulin, les femmes sont une partie de ce que la prison peut enfermer, le pantalon est l'un des vêtements avec lequel on peut porter une ceinture, les hamburgers sont une partie de ce que peut servir le restaurant, etc.

tiquement comment seraient modélisés, dans le DiCo, quelques régimes de trois lexies envisagées jusqu'à maintenant. Nous avons bien évidemment ajouté la DépSyntS.

Il est à noter que l'usage du « = » dans les régimes est peu élégant formellement. Il est évident que les éléments figurant de chaque côté de ce signe ne sont pas des équivalents. Nous avons simplement repris le formalisme qui était utilisé lors de l'édition des fiches du DiCo à l'aide du programme Filemaker. Cette formalisation n'est utilisée ici (et ailleurs dans ce mémoire) que pour présenter ces informations dans le corps du texte, sans devoir inclure à chaque fois une image de l'interface Dicouebe (qui, d'ailleurs, ne permet pas toujours, à notre avis, de visualiser les informations concernant le régime lexical de façon très conviviale).

CEINTURE#I.1

és: pièce d'habillement

fp: ~ DE L'individu X QU'IL UTILISE AVEC LE vêtement Y

X = I = complétive adnominale = de N

= déterminative = A-poss

Y = II = complétive adnominale = de N

= déterminative = A-poss

= attributive adnominale = de+N

OUTIL#I

és: instrument

fp: ~ utilisé par l'individu X pour faire l'action Y(X)

X = I = attributive adnominale = de+N | X est un professionnel

= complétive adnominale = de N

= déterminative = A-poss

Y = II = attributive adnominale = de+N | Y est une activité

professionnelle

= objectale oblique = pour N, à V-inf, pour V-inf

TÉMOIN#II.1

és: artefact servant à fournir une information

fp: ~ INDIQUANT L'INFORMATION Y À L'individu X

X = I = --

Y = II = attributive adnominale = de+N

Cependant, un ASém donné peut parfois être exprimé simultanément par la dépendance **attributive adnominale** et par une autre DépSyntS. Par exemple, avec la lexie CEINTURE#I.1, on peut à la fois exprimer Y par la dépendance **complétive adnominale** (*la ceinture de mon /ce pantalon*) et par la dépendance **attributive adnominale** (*la /une ceinture de pantalon*). Dans le cas d'une complétive adnominale, le groupe « de N » réfère à la pièce de vêtement avec lequel la ceinture est portée, dans un moment précis.

- *Mais, attaché à la ceinture de son pantalon, dissimulé sous son blouson, le revolver de José est prêt à servir, dans son étui de cuir.*
- *Assis au milieu d'un cercle de lumière, Meisho Tisha a le geste lent pour choisir l'une des trois flûtes calées dans la ceinture de son kimono.*

En contrepartie, le complément attributif ne garantit pas que la ceinture soit utilisée avec cette pièce de vêtement en particulier.

- *Pour se donner un style, il porte toujours une ceinture de robe de chambre à son jeans.*
- *À défaut d'avoir trouvé autre chose, j'ai noué mon manteau à l'aide d'une ceinture de pantalon.*

Ce qui est problématique ici, c'est que bien que ce ne soit pas très courant ou particulièrement naturel, on peut même envisager que les deux dépendants soient exprimés simultanément.

- *La ceinture de pantalon de son jeans est détachée.*
- *La ceinture de kimono de son uniforme a été déchirée par les coups de son adversaire.*

Il faut donc en conclure que l'un des deux dépendants ne sature pas la position actancielle ; ou encore que la position actancielle est saturée simultanément par ces deux dépendants. Si deux D saturent la même position actancielle, une nouvelle méthode de modélisation devrait être mise sur pied. En ce sens, on pourrait se demander s'il est possible d'inclure la relation ATTR dans le régime pour l'expression du complément attributif. Dans cette optique, le régime lexical de la lexie CEINTURE#I.1 serait plutôt :

X = I = complétive adnominale = de N  
           = déterminative = A-poss  
 Y = II = complétive adnominale = de N  
           = déterminative = A-poss  
       = ATTR = attributive adnominale = de+N

Ceci permettrait de montrer dans le régime qu'il est possible de qualifier la lexie gouvernante à l'aide de l'ASém2 (la variable Y) via la relation profonde ATTR. Ceci éviterait du même coup d'avoir à affirmer que l'ASyntP II, qui n'est pas répétable, est exprimé deux fois dans une phrase comme *La ceinture de pantalon de mon jeans est détachée*. Toutefois, la relation profonde ATTR n'en est pas une qui est régie, c'est une relation qui est combinable librement aux lexies et l'inclure dans les régimes lexicaux est donc problématique. On pourrait alors poser une nouvelle DépSyntP II, d'un type particulier, par exemple II-ATTR. Toutefois, la création d'une nouvelle DépSyntP, qui existent en nombre très restreint puisqu'elles se veulent très générales, est tout aussi problématique. Nous croyons que l'option la plus heureuse serait plutôt de conclure que l'un des deux D ne sature pas vraiment la position actancielle. Conséquemment, sa présence dans le régime lexical doit être remise en question, car il s'agirait plutôt d'un type de cooccurrence lexicale qu'on n'encode pas systématiquement au moyen des fonctions lexicales.

### 3.3.3 Encodage dans la zone de combinatoire lexicale

Tout d'abord, rappelons que la principale section de la description lexicale du DiCo est celle où sont compilés les liens de dérivations sémantiques et les relations base-collocatifs contrôlés par la lexie vedette («f1»). En LEC, ces liens lexicaux sont encodés grâce aux fonctions lexicales. Grosso modo, les fonctions lexicales paradigmatiques sont utilisées pour rendre compte des dérivations sémantiques et les fonctions lexicales syntagmatiques des relations base-collocatifs. Ce champ «f1», grâce aux fonctions lexicales syntagmatiques, permet donc de noter à même la fiche lexicographique les contraintes lexicales qu'impose la lexie vedette sur son environnement. Il serait donc possible d'envisager que les liens entre une lexie et ses compléments attributifs soient modélisés dans cette section à l'aide d'une fonction lexicale.

Présentement, on peut retrouver dans le DiCo quelques endroits où une relation **attributive adnominale** est encodée sous forme de cooccurrence lexicale dans la zone de combinatoire lexicale à l'aide de fonctions lexicales non standard. C'est

le cas de la lexie FOUR#<sub>a</sub>; és: appareil de cuisson; fp: ~ UTILISÉ PAR L'individu X POUR CUIRE LA nourriture OU L'artefact Y; où l'on retrouve la possibilité d'exprimer Y par le syntagme « à+N » à la fois dans le régime grâce à la ligne Y = II = à+N et dans la zone de combinatoire lexicale :

```
/*Tel que Y est du pain*/
{Y = 'pain'} ~ à pain
/*Tel que Y est des briques*/
{Y = 'briques'} ~ à briques
```

Bien que la méthode utilisée ne soit pas constante sur l'ensemble du DiCo, elle est présente à plusieurs endroits. Il est donc de mise de se demander s'il serait utile d'uniformiser cette pratique en l'étudiant d'abord de façon plus approfondie. Nous nous demandons, entre autres, si la création d'une nouvelle fonction lexicale standard serait pertinente pour remplacer cette modélisation sous forme de X = . . . , Y = . . . , Z = . . . , etc. Puisque les fonctions lexicales standard sont organisées en un système cohérent, il est important de ne pas prendre la création d'une nouvelle fonction à la légère. La complexité d'une telle démarche est d'ailleurs détaillée dans Jousse (2003) et Polguère (2007).

Pour que cette fonction lexicale soit considérée comme standard, le collocatif potentiel doit exprimer vis-à-vis de son gouverneur un sens assez vague et assez général qui est potentiellement exprimable vis-à-vis de nombreuses lexies. En ce qui concerne notre potentielle fonction lexicale standard, le D exprime le sens 'type de' (au lieu de 'beaucoup', 'très' ou 'intense' pour Magn par exemple) et il semble que cette modification soit possible avec la majorité des noms. Les cas que l'on considère actuellement comme étant régis, sont ceux pour lesquels le N de la valeur de la fonction lexicale (ici les constructions à+N, de+N, etc.) correspond à un des actants typiques de la lexie clé.

Par exemple, la lexie CEINTURE#<sub>I.1</sub>, dont la forme propositionnelle est : ~ DE L'individu X QU'IL UTILISE AVEC LE vêtement Y, qui prend potentiellement le sémantème 'pantalon' comme deuxième actant. L'expression du deuxième actant (ici 'pantalon') pourrait être consignée dans son régime. Comme nous l'avons vu, il existe des cas où ces deux actants peuvent être exprimés simultanément (par ex., *la*

*ceinture de pantalon de mon jeans, la ceinture de kimono de son uniforme*). Ceci rend cet encodage problématique (notamment parce que la DépSyntP II n'est théoriquement pas répétable). En contrepartie, nous avons vu que nous avons parfois recours à une fonction lexicale non standard pour encoder une relation de ce type. Cette différence de traitement s'explique par le fait que l'expression du complément attributif tend à se lexicaliser en combinaison avec son gouverneur pour former une locution. Les locutions DÉ À COUDRE et MOULIN À VENT sont probablement de bons exemples de ce processus de lexicalisation. Bien sûr, une étude en diachronie des locutions s'imposerait pour en être absolument certain. Dans les cas où ce processus de lexicalisation est potentiellement en cours, on aura eu tendance à encoder la cooccurrence dans la zone de FL plutôt que dans le régime.

Avant de conclure sur le sujet, nous tenons à souligner que le sémantème général que nous avons utilisé comme sens général de notre fonction lexicale potentielle ('type de') est déjà associé à la relation sémantique d'hyponymie. Dans une fiche lexicographique, Hypo est utilisé pour modéliser le lien entre une lexie et ses hyponymes (grosso modo ses synonymes plus riches). Il serait intéressant de voir si l'on ne pourrait pas élargir ce que couvre cette relation, en permettant par exemple que les hyponymes ne soient pas exclusivement des lexies différentes de L. En forçant un peu, on pourrait peut-être inclure les constructions G —**attributive adnominale**→ D avec les hyponymes de L dans sa fiche lexicographique. Par exemple, on pourrait ajouter à la suite des hyponymes de la lexie GÂTEAU, tous ses « hyponymes analytiques » (*gâteau au chocolat, gâteau à la vanille, gâteau aux fruits, etc.*) grâce à la formule suivante :

Hypo(*gâteau*) = *baba, biscuit, cake, charlotte, chou, clafoutis, éclair, flan, millefeuille, moka, quatre-quarts, religieuse, tarte // ~ à N[=Y]*

Cette dernière option, celle d'encoder les « hyponymes analytiques » d'une lexie L aux côtés de ces hyponymes réguliers (des lexies distinctes), semble d'ailleurs trouver d'autres appuis. S. Marengo dans son travail sur les adjectifs jamais attribués note que pour l'expression des co-hyponymes d'une lexie nominale, les adjectifs qu'il nomme « classifiants » sont concurrencés tantôt par des « syntagmes prépositionnels sans déter-

minant » (nos compléments attributifs), tantôt par des « substantifs épithètes » (nos N2 en apposition, voir p. 59). À ce sujet, il donne les exemples suivants :

- *ski alpin* ~ *ski de fond*, *chien bâtard* ~ *chien de race*, *piano droit* ~ *piano à queue*
- *chapeau mou* ~ *chapeau melon*, *col anglais* ~ *col Mao*, *porte française* ~ *porte patio* (Q) (Marengo 2011, 123)

La solution de modéliser ce type de cooccurrence de L avec ses ASém potentiels par une fonction lexicale (nouvelle fonction lexicale standard, ou une extension de *HYP0*) pourrait être utilisée pour l’encodage de diverses informations qui, au moment d’écrire ces lignes, sont présentes dans les régimes des unités lexicales ou encodées sous forme de « régime lexicalisé », une sorte de fonction lexicale non-standard un peu étrange. Toutefois, il n’est pas dans l’optique du présent mémoire de décider de la validité de cette fonction lexicale potentielle dans le système des fonctions lexicales. La modélisation des régimes lexicaux s’est donc faite sans trop en tenir compte, soit en ajoutant simplement le nom de la dépendance **attributive adnominale** lorsque l’expression d’un ASém via cette dépendance était présente dans les régimes du DiCo. Nous n’avons pas voulu revoir la zone de combinatoire lexicale, ni enlever d’informations présentes dans les fiches lexicographiques. Nous avons tout de même fait l’effort de relever les cas douteux. Un travail futur pourra toujours être consacré à l’uniformisation de l’encodage du phénomène de « régime lexicalisé ».

Pour notre part, nous croyons que le recours à cette fonction lexicale présenterait l’avantage de permettre d’encoder à la fois les D qui semblent être actanciels et des D non-actanciels au même endroit de la fiche lexicographique. Ceci serait pratique parce que cela fournirait un outil pour encoder les expressions en voie de figement, ou qui forment éventuellement des unités lexicales spécialisées. Ce qui ne serait pas possible si l’on encode les D qui correspondent à des ASém potentiels d’une lexie prédicative dans son régime. Par exemple, pour la lexie BATEAU, on pourrait inscrire au même endroit de la fiche une construction base-collocatif où le collocatif correspond à un ASém1 potentiel de BATEAU (*bateau de pirate*), d’autres où il correspond à un des usages potentiels d’un bateau (*bateau de croisière, de pêche, de sauvetage*) et d’autres où il correspond à ce qui permet au bateau de se mouvoir (*bateau à moteur, à rames, à voiles*). Il est perti-

ment d'inscrire ces constructions semi-lexicalisées dans la zone de combinatoire lexicale, parce qu'un usager doit savoir qu'il faut dire *bateau de pêche* et non *\*bateau à pêcher* ou *\*bateau à poissons* par exemple.

Il ne faut toutefois pas oublier le fait que certaines lexies sélectionnent un complément attributif, il est parfois même obligatoire (PLAQUE#I.1 fp : ~ DE matière X). Dans des cas semblables, il est inévitable de conserver les compléments attributifs dans le régime lexical. Cela est nécessaire, parce qu'il serait impossible d'affirmer qu'un complément obligatoire doit être encodé dans la zone de combinatoire lexicale. Dans certains cas donc, la dépendance **attributive adnominale** doit être incluse dans le régime lexical, parce que le D sature la position actancielle de L, ce qui n'est vraisemblablement pas le cas des exemples étudiés jusqu'ici. Par exemple, le régime suivant :

PLAQUE#I.1

fp: ~ DE matière X

X = I = attributive adnominale = de+N | oblig.

permet de voir qu'il est possible de construire des expressions comme *plaque de glace*, *plaque de neige*, *plaque de verglas*, *plaque d'athérome*<sup>6</sup>.

6. Dans l'état actuel de nos recherches, nous considérons ce D comme un complément attributif, parce qu'il présente les propriétés syntaxiques identifiées précédemment. Toutefois, si l'on suit le travail de Marengo (2011, 159-177), ce dernier complément ne correspond pas à un adjectif classifiant, mais est le complément d'un nom de quantité (ou substantif quantificateur). Il serait peut-être intéressant d'ajouter des propriétés à notre liste actuelle (notamment en ce qui a trait à leur comportement vis-à-vis des adjectifs) pour voir si une séparation serait souhaitable. Nous tenons à rappeler que nous n'avons pas voulu créer des instances de DépSyntS uniquement basées sur la nature des noms G, mais bien à l'aide des propriétés syntaxiques des D. Le fait qu'*une plaque#I.1 de glace*, *un morceau#I.2 de sucre*, *un fleuve#I de lave*, etc. soient des sous-catégories de *plaque#I.1*, de *morceau#I.1* ou de *fleuve#I* qui n'ont pas besoin d'être consignées dans la zone de fonction lexicale parce que construit de façon libre par la langue (on voit la façon de faire dans le régime), contrairement aux constructions N1-**attributive adnominale**→N2 typiques, ne devrait pas non plus nous forcer à créer une instance de dépendance distincte si les dépendances partagent tout de même un nombre important de propriétés. Nous convenons qu'il reste sans doute quelque chose à tirer au clair à ce sujet.

### 3.3.4 En résumé

Nous avons d'abord exposé au cours de ce chapitre la méthode qui a été utilisée pour l'identification de la DépSyntS **attributive adnominale**. Les mêmes étapes ont été utiles lors de la dénomination et description d'autres DépSyntS nominales, et elles seraient également adéquates lorsque ce travail sera fait pour les gouverneurs des autres parties du discours du français, et même, en partie du moins, pour les dépendances non régies. La méthodologie que nous avons proposée sera reprise dans le chapitre suivant avec la liste des DépSyntS nominales qu'elle nous a permis de créer. Par la suite, nous avons questionné ce qui devait être modélisé dans les régimes lexicaux en examinant le cas de la DépSyntS **attributive adnominale**. Plus précisément, nous nous sommes demandé si l'expression de tous les ASém d'une lexie prédicative doit être encodée dans son régime lexical ou si certains cas peuvent échapper à cette règle. Cette réflexion est notamment pertinente lorsqu'un syntagme qui semble correspondre à un ASém ne sature vraisemblablement plus la position actancielle correspondante ; lorsqu'une construction présente dans un régime se « lexicalise ».

Nous pouvons conclure de cette dernière réflexion que ce qui doit être inclus dans le régime lexical est l'expression d'un ASém lorsque le syntagme sature effectivement la position actancielle correspondante. La zone de combinatoire lexicale serait alors utilisée pour encoder les D qui semblent correspondre à des ASém d'une lexie, mais qui ne sont pas réellement régis par celle-ci. En conclusion, le régime de la lexie CEINTURE<sub>I.1</sub> doit être :

X = I = complétive adnominale = de N  
           = déterminative = A-poss  
 Y = II = complétive adnominale = de N  
           = déterminative = A-poss

Les constructions *ceinture de pantalon*, *ceinture de kimono*, *ceinture de robe de chambre*, etc. devraient quant à elles être ajoutées d'une façon ou d'une autre dans la zone de combinatoire lexicale. Nous croyons en effet que ces expressions sont semi-phraséologisées. On ne peut pas dire par exemple, *\*ceinture pantalonière*, *\*ceinture à*

*kimono*, \**ceinture pour robe de chambre* ; la façon adéquate d'exprimer ces sens étant *ceinture de+N[=Y]*. Il en est de même pour plusieurs autres lexies du DiCo qui devraient voir leur régime modifié. Par exemple, le régime de la lexie OUTIL#I est présentement le suivant :

és: instrument

fp: ~ UTILISÉ PAR L'individu X POUR FAIRE L'action Y(X)

X = I = attributive adnominale = de+N | X est un professionnel  
 = complétive adnominale = de N  
 = déterminative = A-poss

Y = II = attributive adnominale = de+N | Y est une activité  
 (professionnelle)  
 = objectale oblique = pour N, à V-inf, pour V-inf

Nous croyons toutefois que, comme le complément attributif en « *de+N* » de CEINTURE #I.1, les deux compléments attributifs en « *de+N* » neaturent pas leur position actancielle respective et ne doivent donc pas faire partie du régime. C'est du moins notre opinion puisqu'il est possible de construire des exemples comme : *l'outil de maçon[=X] de Fred[=X] <son[=X] outil de maçon[=X]>, un outil de maçonnerie[=Y] pour travailler le bois[=Y], un outil de cordonnerie[=Y] pour réparer les chaussures[=Y]*.

C'est ce qui nous porte à croire que pour les deux D en « *de+N* », les DépSyntP correspondantes sont des ATTR plutôt que I et II, et ces cooccurrents devraient donc apparaître dans la zone de combinatoire lexicale de OUTIL#I. En ce sens, le régime serait plutôt :

és: instrument

fp: ~ UTILISÉ PAR L'individu X POUR FAIRE L'action Y(X)

X = I = complétive adnominale = de N  
 = déterminative = A-poss

Y = II = objectale oblique = pour N, à V-inf, pour V-inf

Nous croyons que beaucoup de cas similaires ont été inclus dans les régimes étant donné que les compléments attributifs ont tendance à se linéariser immédiatement à la droite de la lexie gouvernante. Il est donc parfois pertinent de remettre en question la décomposition du sens prédicatif de certaines lexies ; certains ASém ayant été considérés comme tels par la seule possibilité d'avoir un complément attributif à la suite d'une lexie.

Par exemple, la description des ASém de la lexie TUYAU#I.1, et par conséquent de son régime, a été modifiée à un certain moment lors de la rédaction de ce mémoire (sans que cette modification soit liée à notre travail).

Avant :

fp: ~ PERMETTANT LE déplacement X DU liquide OU gaz Y  
 X = I = attributive adnominale = de+N [tuyau d'aération]  
 Y = II = attributive adnominale = de+N [tuyau d'eau]

Maintenant :

fp: ~ POUR LA CIRCULATION DU liquide OU gaz X DE Y À Z  
 X = I = attributive adnominale = de+N [tuyau d'eau]  
 Y = II = --  
 Z = III = --

et dans les fonctions lexicales, on retrouve :

Qui sert à la circulation d'air dans une construction:  
*d'aération*

Qui sert à l'écoulement des eaux: *d'écoulement, de descente*  
*//descente*

Qui sert à amener un gaz dans une machine: *d'admission*

Qui sert à évacuer un gaz d'une machine: *d'échappement*

Qui sert à évacuer la fumée d'un poêle: *de poêle*

Qui sert à la production du son dans un orgue: *d'orgue*

Qui est la partie d'une pipe que l'on met à sa bouche: *de pipe#I*

Nous croyons que cet exemple démontre bien la difficulté dont nous avons parlé dans la précédente sous-section. En ce qui nous concerne, nous avons fait l'effort de relever les cas de ce type lors de notre travail d'encodage, mais il est possible que certains problèmes persistent.

## CHAPITRE 4

### MODÉLISATION DES RÉGIMES LEXICAUX DU FRANÇAIS DANS LA BASE DE DONNÉES CARNAVAL

Tout d'abord, mentionnons d'entrée de jeu que le projet de la création d'un *Réseau lexical du français* (RLF) se situant dans la lignée d'autres projets tels que WordNet (Fellbaum 1998) et FrameNet (Ruppenhofer et al. 2010) créait la nécessité d'une base de données des régimes du français pour soutenir leur travail d'édition. Ce besoin est clairement exprimé dans la présentation du projet :

« **3. Government pattern** This section describes how the keyword's semantic actants can be expressed as its syntactic dependents. A database of French government patterns will be included in the FLN data structure and valency tables (roughly, subcategorization frames) appearing in a lexicographic article will ultimately be directly imported from this base rather than manually typed by lexicographers. » (Lux-Pogodalla et Polguère 2011)

Ceci découle de la volonté d'offrir le plus de contenu normalisé possible dans l'édition des articles. Ainsi, les extraits de l'article lexicographique qui sont choisis à partir de menus ne sont pas entrés manuellement par les lexicographes, ce qui peut éviter l'insertion de coquilles ou le manque d'uniformité dans les descriptions.

L'objectif général derrière notre travail était donc d'en arriver éventuellement à la création d'une base de données qui contienne l'ensemble des tableaux de régime de la langue française. Bien évidemment, il est plus réaliste, à court terme, de débiter le développement de cette base par un nombre relativement restreint de régimes. Pour notre part, nous nous sommes fixée comme objectif de départ les régimes des lexies nominales du DiCo. Nous avons choisi cet échantillon, d'abord parce que le DiCo a surtout été développé avec ce type de lexies. Ceci est dû au fait que le DiCo est avant tout un dictionnaire de combinatoire. Comme nous l'avons mentionné précédemment, chacune des lexies décrites dans le DiCo est accompagnée, entre autres, des collocations qu'elle contrôle ; verbes supports, modificateurs, etc. Bien que les lexies de toutes les parties du

discours puissent en principe contrôler ce type de relations syntagmatiques, c'est dans la description des noms prédicatifs et quasi-prédicatifs que cela s'avère le plus pertinent. D'autre part, comme les constructions syntaxiques gouvernées par des lexies nominales n'ont été que peu étudiées dans une perspective lexicographique, elles nous semblaient offrir de nombreuses possibilités de réflexions intéressantes.

Dans la première section du chapitre, nous explicitons les étapes nécessaires à la conception de la base de données. Dans un deuxième temps, nous présentons la base de données en cours de création.

#### 4.1 Tâches effectuées

Pour arriver à créer la partie de la base de données qui nous concerne, les tâches que nous avons dû effectuer sont les suivantes : 1) établir les DépSyntS régies par des noms qui existent en français ainsi que leurs propriétés définitives ; 2) établir quelles sont les différentes composantes d'un régime lexical afin de prévoir les différentes balises de notre encodage ; 3) entamer la conception de la base de données en détaillant chacun des régimes associés aux lexies nominales du DiCo.

##### 4.1.1 Identification et caractérisation des dépendances syntaxiques de surface nominales du français

La première étape de notre travail a été d'établir la liste des DépSyntS dont le gouverneur est potentiellement nominal. Comme notre approche est avant tout lexicographique, les dépendances identifiées sont essentiellement des dépendances potentiellement régies. Toutefois, certaines dépendances typiquement non régies ont également été décrites au cours de ce travail, car elles ont été nécessaires dans certains cas particuliers. Il est important de noter que le travail n'est pas terminé à cet égard et que la liste des dépendances nominales non régies est donnée à titre indicatif seulement, puisqu'incomplète.

Suite au travail effectué au chapitre précédent sur la DépSyntS **attributive adnominale**, le patron de description des DépSyntS que nous proposons, se détaille en cinq étapes (qui diffèrent légèrement de celles explicitées dans Iordanskaja et Mel'čuk

(2009, 169)).

- Identifier les G possibles et les D prototypiques de la dépendance
- Identifier les propriétés définitoires de la dépendance
- Identifier les propriétés descriptives de la dépendance
- Justification (si nécessaire)

Dans les cas les moins évidents, il peut être utile de démontrer que la dépendance **d** ne pourrait pas être considérée comme étant une instance d'une dépendance **d'**

- Commentaires (si nécessaire)

#### 4.1.1.1 Caractérisation des dépendances à l'aide de leurs propriétés définitoires

La première étape de notre travail de description est donc de cerner quels sont les G possibles et les D prototypiques des DépSyntS nominales, ainsi que leurs autres propriétés définitoires, ce qui les distingue fondamentalement les unes des autres. Pour qu'une propriété soit considérée comme définitoire, elle devait être assez générale pour être éventuellement applicable dans d'autres langues que le français. Comme de telles propriétés générales n'étaient pas suffisantes pour distinguer toutes les dépendances que nous avons recensées, certains autres critères typiques du français ont été retenus. Par exemple, lorsqu'une relation a comme dépendant une lexie particulière du français, nous l'avons indiqué. Les propriétés définitoires sont donc celles qui sont nécessaires à la différenciation des dépendances et qu'une dépendance particulière possèdera toujours dans une construction syntaxique donnée.

Pour débiter, il est avant tout important d'observer les lexies nominales du français et leur comportement syntaxique respectif. Cette étape permet de dresser une première liste sommaire des DépSyntS nominales. Un premier travail de repérage effectué par Lidija Iordanskaja sur les données du DiCo (inédit) a permis d'effectuer cette étape préliminaire. Cette ébauche a servi à établir une première liste des propriétés définitoires de ces DépSyntS nominales. Il est à noter que la classe dépendance syntaxique de surface nominale n'est pas exclusive. Ainsi, le fait qu'une dépendance appartienne à cette classe n'exclut pas qu'elle puisse éventuellement appartenir à d'autres classes. Par exemple, la dépendance **objectale oblique** sera considérée comme

une seule dépendance, peu importe la partie du discours du gouverneur. Cette dépendance fera donc à la fois partie des classes dépendance syntaxique de surface verbale, dépendance syntaxique de surface nominale et dépendance syntaxique de surface adjectivale.

Ce travail effectué par Iordanskaja est analogue à celui de Marengo (2009) sur les dépendants syntaxiques des adjectifs. Dans cette recherche, S. Marengo part d'un groupe de lexies adjectivales témoins et regroupe les différentes dépendances faisant partie de leur régime sur la base de leurs propriétés particulières. Il semble cependant que cette méthode de travail n'ait pas un degré d'abstraction suffisant pour rendre compte de certains faits. Par exemple, une même dépendance peut servir à spécifier des constructions syntaxiques qui peuvent être ou non régies lexicalement. Selon notre approche, on aura la même DépSyntS dans les deux cas suivants :

*aide* –**d**→ *américaine*

*meuble* –**d**→ *américain*

Dans le premier exemple, le D correspond à un ASém de son gouverneur, plus précisément le premier (grosso modo : *aide DE LA personne X À LA personne Y EN CE QUI CONCERNE LE problème Z AVEC LE moyen W*), ce qui n'est pas le cas dans le deuxième exemple. Malgré cette différence, ces deux constructions peuvent être associées à la même dépendance (soit la DépSyntS **modificative**). Il en est de même lorsque la partie du discours du gouverneur diffère d'une construction à l'autre. Dans les deux cas suivants, la dépendance syntaxique **d** est la DépSyntS **agentive**.

[*L'homme est souvent*] *exploité* –**d**→ *par l'homme*

*exploitation* –[*de l'homme*]–**d**→ *par l'homme*

Ainsi, les données recueillies par Iordanskaja ont certes servi de base, mais nous avons pris soin d'élargir notre perspective. Par la suite, le travail consistant à encoder les différents régimes du DiCo nous permettra de vérifier la validité des critères d'identification qui auront été établis.

Voici reproduit ici le patron servant à l'identification des DépSyntS tel qu'il est pour l'instant. Il ne contient que les propriétés définitoires qui sont utiles pour distin-

guer les DépSyntS les unes des autres ; les propriétés descriptives n’y sont donc pas incluses. G signifie ‘Gouverneur’, D signifie ‘Dépendant’, PDD signifie ‘partie du discours’. D’autres explications suivront.

- I. Taxinomie des G et D
  - a. G = (<PDD>) [nature sémantique]
  - b. D = <Lexème> (<PDD>)
- II. Source de la dépendance
  - a. <DépSyntP> <DépSém>
  - b. Sémantème <‘s’>
  - c. Translation → Adj
- III. Manifestation de la dépendance
  - a. <Ordre linéaire>
  - b. <Prosodie>
  - c. <Dépendance morphologique>

Lors de la création d’une instance de dépendance dans la hiérarchie GTN, nous sélectionnons les critères nécessaires à son identification dans cette liste de propriétés définitoires. Par exemple, en ce qui concerne la dépendance **attributive adnominale**, les critères pertinents sélectionnés sont : I.a., I.b. et II.c. Ensuite, on apporte les précisions nécessaires (par ex., I.a. G = (Ncomm)) pour chacune des dépendances.

Bien entendu, cette liste de propriétés définitoires devra être élargie au fur et à mesure que de nouvelles DépSyntS seront répertoriées. Celles qui y figurent sont celles qui ont été utiles jusqu’à présent. La « taxinomie des G et D » fait référence à la classification des gouverneurs et des dépendants selon différents critères. Pour l’instant, il semble pertinent d’identifier les gouverneurs et les dépendants d’une dépendance par leur partie du discours accompagnée d’autres précisions — les traits syntaxiques (par ex., V transitif, N commun, etc.), leur nature sémantique (par ex., prédicat, quasi-prédicat) ou par un lexème spécifique (par ex., DE#1, QUE#1, etc.). La « source de la dépendance » fait référence à ce qui donne naissance à cette dépendance aux niveaux plus profonds (SyntP ou Sém). Nous avons conclu que les DépSyntP, les DépSém et éventuellement certains sémantèmes ou certaines réalités syntaxiques telles que la translation peuvent être la source

d'une DépSyntS particulière. Finalement, dans quelques cas, nous avons jugé que la façon générale dont la dépendance se manifeste dans la langue est une propriété définitoire de celle-ci. Par exemple, dans tous les cas où une dépendance **circonstancielle** apparaît dans un texte français, aucune dépendance morphologique n'existera entre le G et le D.

Comme nous l'avons mentionné au chapitre précédent, les informations concernant les DépSyntS sont actuellement compilées dans la hiérarchie des concepts de lexicologie/lexicographie GTN, dont voici un extrait.

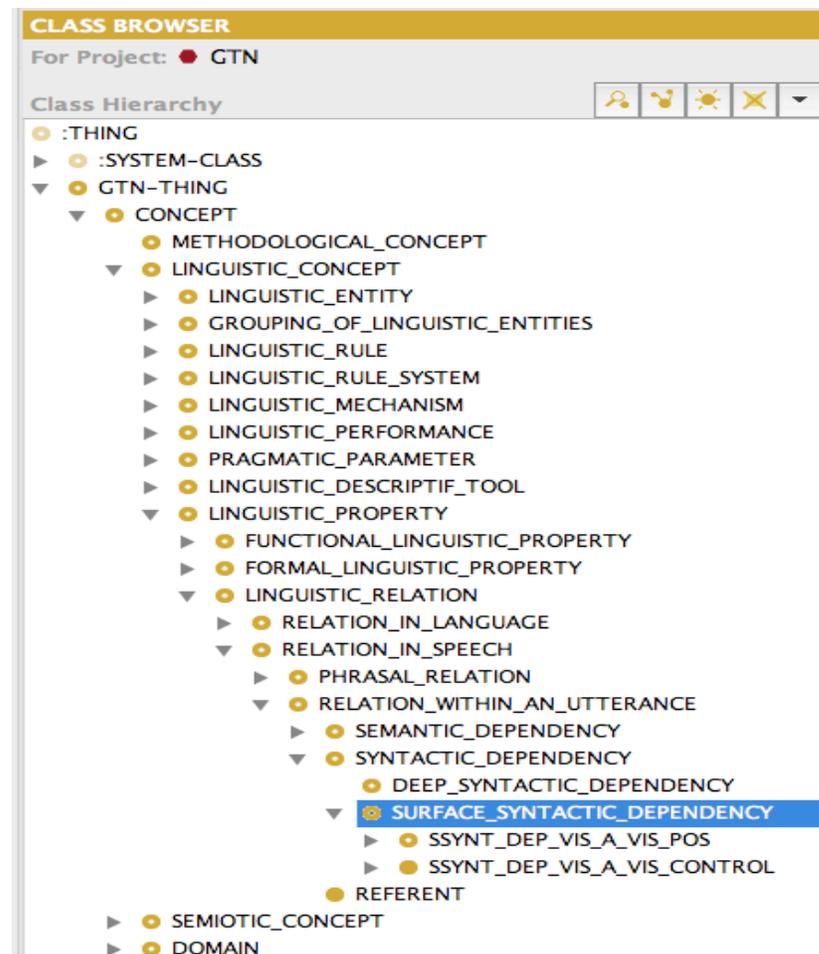


Figure 4.1 – Les DépSyntS dans la Hiérarchie GTN

Comme on peut le voir, elles y sont classées selon deux axes : 1) la partie du discours du gouverneur (POS pour *part of speech*); et 2) leur caractère typiquement régi ou non régi (CONTROL). Nous présentons ici les DépSyntS nominales (pouvant avoir un nom — et éventuellement autre chose — comme G) du français selon le deuxième axe. Les descriptions comprennent les propriétés définitives de chacune des dépendances ainsi que quelques exemples dans lesquels elles sont présentes.

#### 4.1.1.2 Les DépSyntS nominales du français

##### DépSyntS régies (ou *Valence-controlled SSyntDep*)

###### 1. agentive

Propriétés :

I.a. G = (Vtrans, infinitif) ou (Ncomm, prédicat) ou (Adj, ppassé)

b. D = PAR ou DE#3

II.a. RelSém1  $\Leftrightarrow$  Soit RelSyntPI (G est Vtrans, infinitif ou Ncomm, prédicat); soit RelSyntPII avec passif (G est Adj, ppassé)

Exemples :

*Le roman est traduit[G] par[D] Léo  
traduction[G] de ce livre par[D] Léo  
verre cassé[G] par[D] Léo  
tué[G] par[D] une pierre  
aimé[G] de[D] tout le monde*

###### 2. appositive énonciative

Propriétés :

I.a. G = (Ncomm)

II.a. RelSyntP II-disc-dir

Exemples :

*son cri[G] « Au[D] secours »*

###### 3. objectale adnominale

Propriétés :

I.a. G = (Ncomm, prédicat ou S1 de lexie  
prédicative)

b. D = DE#1 ou QUE#1

II.b. RelSém 2

Exemples :

*désir[G] d'[D] un départ immédiat*

*désir[G] de[D] partir immédiatement*

*désir[G] que[D] Marie parte immédiatement*

*assassinat[G] de[D] Jean*

*assassin[G] de[D] Lucien*

*mangeur[G] de[D] pommes de terre [= 'personne qui est  
en train de manger des pommes de terre']*

#### 4. objectale indirecte

Propriétés :

I.a. G = (prédicat ou quasi-prédicat)

b. D = Prép-datif ou Clit-datif

Exemples :

*écrire[G] à[D] Jean*

*donner[G] un cadeau à[D] Marc*

*dessiner[G] un mouton à[D] Lucien*

*abandon[G] de ses biens à[D] ses enfants*

*don[G] aux[D] pauvres*

*lettre[G] à[D] Jean*

*fidèle[G] à[D] sa blonde*

*ça lui[D] est[G] égal (Contrôlé par le régime de  
ÉGAL)*

#### 5. objectale oblique

Propriétés :

I.a. G = (prédicat ou quasi-prédicat)

b. D = (Prép ≠ Prép-datif) ou DE#2 ou QUE#2

II.b. RelSém Alpha≠1

Exemples :

*se rapprocher[G] de[D] Jean*

*parler[G] de[D] politique*

*rêve[G] de[D] partir*

*rapprochement*[G] *avec*[D] *Jean*  
*différent*[G] *de*[D] *Marcel*  
*le même*[G] *que*[D] *le tien*

## 6. subjectale adnominale

Propriétés :

- I.a. G = (Ncomm, prédicat)
- b. D = DE#1 ou "fam" À ou ENTRE
- II.b. RelSém 1 si DE ou RelSém 1+2 si ENTRE

Exemples :

*désir*[G] *de*[D] *Paul*  
*faute*[G] *de*[D] *Marc*  
 "fam" *faute*[G] *à*[D] *Marc*  
*lutte*[G] *entre*[D] *Paul et Pierre*  
*lutte*[G] *de*[D] *Paul contre l'ignorance*  
*différence*[G] *entre*[D] *Igor et Lida*  
*traduction*[G] *de*[D] *Léo de ce livre (cf. traduction*  
*de ce livre par Léo)*

## DépSyntS non régies (ou *Non-valence-controlled SSyntDep*)

### 1. appositive appellative

Propriétés :

- I.a. G = (Ncomm)
- b. D = (Npr) ou DE#4 (régit Npr)
- II.b. sémantème 's'appeler'

Exemples :

*fleuve*[G] *St-Laurent*[D]  
*ville*[G] *de*[D] *Montréal*  
*bibliothèque*[G] *Miterrand*[D]

### 2. appositive copulative

Propriétés :

- I.a. G = (Ncomm)
- b. D = (Ncomm)
- II.b. sémantème 'être'

Exemples :

*enfant*[G] *prodige*[D]  
*moteur*[G] -*alternateur*[D]

### 3. appositive numérative

Propriétés :

- I.a. G = (Ncomm)
- b. D = (Num)
- II.b. sémantème 'numéro'

Exemples :

*place*[G] *de stationnement 18*[D]  
*abattoir*[G] *cinq*[D]  
*paragraphe*[G] *vingt*[D]  
*version*[G] *3.2*[D]

### 4. circonstancielle

Propriétés :

- I.a. G = (V ou N ou Adj)
- b. D = (Adv ou GPrép ou GN)
- II.b. Sémantème 'lieu' ou 'moment' ou 'durée' ou 'manière' ou 'cause' ou 'but'
- III.c. Pas de dépendance morphologique

Exemples :

*partir*[G] *ce matin*[D]  
*se sentir*[G] *bien en*[D] *Italie*  
*marcher*[G] *vite*[D]  
*arriver*[G] *ce matin*[D]  
*maison*[G] *sur*[D] *la falaise*  
*table*[G] *près*[D] *du mur*  
*Paris*[G] *il y a*[D] *trente ans*  
*séjour*[G] *chez*[D] *Jean*  
*exil*[G] *en*[D] *Angleterre*  
*promenade*[G] *avec*[D] *Marie*  
*né*[G] *en*[D] *juin*  
*beau*[G] *de*[D] *façon bizarre*

### 5. modificative

Propriétés :

- I.a. G = (N sauf Pron)
- b. D = (Adj non déterm)
- III.c. Dépendance morphologique G → D

Exemples :

*excellent*[D] *comportement*[G]  
*troisième*[D] *porte*[G] *à gauche*  
*aide*[G] *américaine*[D]

## 6. numérale

Propriétés :

- I.a. G = (N sauf Pron)
- b. D = (Num)
- III.a. Linéarisation prend en compte une éventuelle dépendance déterminative
- c. Dépendance morphologique D → G

Exemples :

*deux*[D] *amis*[G]  
*les deux*[D] *amis*[G]

## DépSyntS Mixtes (ou *Mixed SSyntDep*)

### 1. appositive compositive

Propriétés :

- I.a. G = (Ncomm)
- b. D = (Ncomm ou Npr)
- III.c. Pas de dépendance morphologique G → D

Exemples :

*un ticket*[G] *restaurant*[D]  
*des tickets*[G] *restaurant*[D]  
*espace*[G] *fumeur*[D]  
*espace*[G] *fumeurs*[D]  
*ministère*[G] *Donadieu*[D]  
*(un) vert*[G] *émeraude*[D]

### 2. attributive adnominale

Propriétés :

- I.a. G = (Ncomm)

b. D = (Prép + N-non réf.)

II.c. Translation → Adj

Exemples :

*gâteau*[G] *au*[D] *chocolat*  
*congé*[G] *de*[D] *maladie*  
*viande*[G] *de*[D] *cheval*  
*moulin*[G] *à*[D] *café*  
*prison*[G] *de/pour*[D] *femmes*  
*problème*[G] *à*[D] *résoudre*  
*fleur*[G] *des*[D] *champs*  
*tir*[G] *à*[D] *l'arc*  
*morceau*[G] *de*[D] *sucre*

### 3. complétive adnominale

Propriétés :

I.a. G = (N, quasi-prédictat)

b. D = DE#1 ou "fam" À ou (prép)

Exemples :

*jouet*[G] *de*[D] *Marc*  
 "fam" *jouet*[G] *à*[D] *Marc*  
*visage*[G] *de*[D] *Marc*  
*morceau*[G] *de*[D] *cette météorite*  
*médecin*[G] *de*[D] *Jean*  
*voisine*[G] *de*[D] *Gustave*  
*frontière*[G] *entre*[D] *la France et la Suisse*  
*(Il est la) fierté*[G] *de*[D] *la famille*  
 "fam" *bâtiment*[G] *à*[D] *Pierre* (dans le sens où  
 Pierre a construit ce bâtiment → ne correspond pas  
 à un ASém de G)

### 4. déterminative

Propriétés :

I.a. G = (N sauf Pron)

b. D = (Dét)

III.a. Position extrême dans le syntagme de G

c. Dépendance morphologique G → D

Exemples :

*le/un/ce*[D] *dégoût*  
*son*[D] *dégoût*[G]

#### 4.1.1.3 Caractérisation des dépendances à l'aide de leurs propriétés descriptives

Il est important de mentionner que cette liste a été testée en l'utilisant pour l'encodage des régimes lexicaux du DiCo ; elle a donc été modifiée en cours de travail selon les données rencontrées. Lorsque la liste des dépendances a été jugée suffisamment précise, nous avons pu passer à la deuxième étape du travail d'identification, c'est-à-dire établir une liste des propriétés descriptives spécifiques à chacune des dépendances. Encore ici, la liste des premières propriétés pertinentes pour l'analyse des constructions G[nominal] → D a été établie dans un premier temps suite au travail de Lidija Iordanskaja et enrichie par la suite. Ces propriétés sont les suivantes :

1. Dépendance morphologique (D → G ou G → D)
2. Flexion indépendante du D
3. Expression de détermination
4. Position linéaire du D par rapport à son G
5. Possibilité de dislocation à gauche
6. Prosodie/Ponctuation
7. Paraphrasage du D par A-poss
8. Pronominalisation du D par *dont*
9. Pronominalisation du D par *en*
10. Ordre de codépendants
11. Possibilité d'insertion entre G et D
12. Possibilité pour D d'être modifié par une relative
13. Actant profond de G auquel correspond D
14. Répétabilité de D
15. Paraphrases

À l'heure actuelle, les DépSyntS ne sont pas associées de façon satisfaisante à l'ensemble de leurs propriétés descriptives respectives. Ce travail de caractérisation est donc « a work in progress ». Nous avons déjà donné et expliqué au chapitre précédent celles

qui sont associées à la DépSyntS **attributive adnominale**. Nous espérons que cela sera suffisant à la bonne compréhension du lecteur. Comme nous l'avons mentionné précédemment, ces informations ne feront pas partie intégrante de l'ontologie GTN dans la première étape de sa conception. De plus, pour les principales tâches qui nous occupent, soit ajouter les DépSyntS nominales dans les régimes lexicaux du DiCo et construire une base de données des régimes lexicaux, les propriétés définitoires se sont avérées suffisantes. Nous donnons donc en quelque sorte cette étape à titre indicatif. Pour quelqu'un dont le but serait de véritablement caractériser l'ensemble des DépSyntS et d'en donner la liste exhaustive de propriétés, cette étape serait cruciale. Notons également qu'il existe probablement de nombreuses autres propriétés spécifiques aux dépendants de gouverneurs nominaux. On pourra se référer au travail effectué au LADL pour se rendre compte que de cerner la quasi-totalité de ces propriétés et de les associer aux dépendants correspondants à un grand nombre de lexies est une tâche ardue et complexe qui, bien qu'intéressante, est en dehors de la portée de notre recherche actuelle.

#### 4.1.2 Identification et modélisation des différentes composantes d'un régime lexical

En vue de la création de la base de données, il a été nécessaire de bien identifier les différentes informations présentes dans les régimes lexicaux. Précédemment, les régimes contenaient les différents actants de la lexie vedette (les ASém identifiés avec des variables — X, Y, Z, W —, les ASyntP —I, II, III, IV —, et les syntagmes correspondants en syntaxe de surface — *de* N, V-inf, *que* PROP-subj., etc). Voici, par exemple, le régime de la lexie FRONTIÈRE#I.1 tel qu'il apparaît dans sa fiche du DiCo (nous avons ajouté un exemple pour chacune des constructions).

és: lieu qui sépare

fp: ~ DU territoire X AVEC LE territoire Y [SURVEILLÉE PAR LES individus Z POUR CONTRÔLER LE PASSAGE DES individus W]

Rég.1

X+Y = I = de N et de N, entre N et N, A-poss [*la frontière de l'Allemagne et de la France, entre l'Allemagne et la France,*

*leur frontière]*

Z = II = --

W = III = --

Rég.2

X = I = de N, A-poss

Y = II = avec N [*la frontière de l'Allemagne avec la France, sa frontière avec la France*]

Z = III = --

W = IV = --

Comme nous souhaitons préciser les régimes, nous avons ajouté à ce régime le nom de la DépSyntS qui correspond aux différents syntagmes. Afin d'illustrer ceci, voici le régime modifié de la lexie FRONTIÈRE#I.1. Nous indiquons en gras des informations que nous avons jugées manquantes dans le régime précédent.

Rég.1

X+Y = I = complétive adnominale = de N et de N, entre N et N

= déterminative = A-poss

**= modificative = Adj.** [*la frontière franco-allemande*]

**= appositive compositive = N-N** [*la frontière France-Allemagne*]

Z = II = --

W = III = --

Rég.2

X = I = complétive adnominale = de N

= déterminative = A-poss

Y = II = objectale oblique = avec N

Z = III = --

W = IV = --

Ceci a fait en sorte que les régimes décrivent maintenant les correspondances entre les différentes dépendances (sémantique, syntaxique profonde et syntaxique de surface) — en rouge dans CARNAVAL — en plus des actants correspondants — en noir dans CARNAVAL. Premièrement, le réseau sémantique minimal de la lexie décrite est modélisé en indiquant dans le régime la dépendance sémantique — 1, 2, 3, 4 — ainsi que la variable — X, Y, Z, W — attribuée à la position actancielle correspondante ; ce qui permet de modéliser sa valence sémantique. Ensuite, la deuxième partie du régime sert à

décrire la diathèse de la lexie vedette. On y retrouve, pour chacun des ASém : la DépSém — 1, 2, 3, 4 —, la DépSyntP — I, II, III, IV — et implicitement l'ASyntP correspondant (I, II, III ou IV selon le cas), et finalement, la (ou les) DépSyntS qui permet l'expression de cet ASém ainsi que les syntagmes correspondants en syntaxe de surface.

Évidemment, dans le cas des lexies qui permettent les modifications de régime (pour une explication de cette notion voir entre autres Mel'čuk et coll. (1984-1999, Vol.IV :22) ou Milićević (2009, 101-102)), comme la lexie FRONTIÈRE#I.1, nous retrouvons une seule fois la description de la valence sémantique et ensuite les deux (et éventuellement plus) options de diathèses possibles (Rég. 1, Rég. 2). Ceci fait en sorte que la structure d'un régime multiple est légèrement différente de la structure d'un régime simple (de loin les plus courants). Voici donc le nouveau régime de FRONTIÈRE#I.1 tel que présent dans la base CARNAVAL.

<b>r.N.4.015</b> 1 → X 2 → Y 3 → Z 4 → W	
<b>Rég.1</b>	
1+2 ↔ I	
complétive adnominale	de N-coord   entre N-coord
déterminative	A-poss
modificative	Adj
appositive compositive	N-N
3 ↔ II	—
4 ↔ III	—
<b>Rég.2</b>	
1 ↔ I	
complétive adnominale	de N
déterminative	A-poss
2 ↔ II	
objectale oblique	avec N
3 ↔ III	—
4 ↔ IV	—

Figure 4.2 – Régime de FRONTIÈRE#I.1

### 4.1.3 Encodage, dans la base de données, de chacun des régimes associés aux lexies nominales du DiCo

Finalement, la troisième tâche consistait en la création de la base de données des régimes généralisés. Il s'agissait donc d'encoder tous les régimes différents rencontrés dans le DiCo et de leur attribuer un code. Bien entendu, cette étape implique de vérifier et de modifier au besoin les régimes lexicaux existants et de leur ajouter le code approprié. Ce code pourra éventuellement être inclus dans la section des caractéristiques grammaticales de la fiche du DiCo ou du RLF. Il sera ainsi aisé de retrouver les lexies qui partagent le même régime en interrogeant l'interface appropriée à l'aide d'un numéro de régime. De plus, on pourrait envisager d'établir un hyperlien dans chacune des fiches

lexicographiques du DiCo vers son régime particulier. Ceci permettrait d’avoir accès à l’information sur le régime via l’interface DiCoPop, ce qui n’est pas le cas présentement. Cette étape a permis la création de 326 modèles de régimes différents qui sont tous associés à au moins une lexie nominale du DiCo. Les régimes de 710 lexies nominales sont donc décrits dans la base de données CARNAVAL, que nous présentons dans la section suivante. De plus, nous fournissons en annexe du présent mémoire la liste de ces lexies nominales ainsi que leur numéro de régime respectif.

## **4.2 Présentation de la base de données CARNAVAL**

Notre troisième objectif de recherche était d’entamer la création d’une base de données des régimes lexicaux du français. Bien que notre base pourrait éventuellement contenir l’ensemble des régimes de la langue française, elle ne contient pour l’instant que les régimes des lexies nominales du DiCo, puisque ce sont ces régimes que nous nous proposons d’encoder. Nous exposons nos résultats en regard de cet objectif en commençant par une présentation générale de la base, nous poursuivons avec le schéma et les balises nécessaires à l’encodage des régimes et nous terminons avec deux Tableaux de régime détaillés.

### **4.2.1 Présentation générale**

La base CARNAVAL est développée sous forme de fichier XML grâce au logiciel oXygen. Comme nous venons de le préciser, avant d’amorcer les modélisations, nous avons analysé les régimes afin de bien identifier leurs différentes composantes. Nous avons ensuite hiérarchisé ces composantes en leur attribuant chacune un nom qui est devenu le nom d’une balise dans notre fichier XML. Avant de débiter la conception de nos fiches, un schéma .rnc (RELAX NG (REGular LAnguage for XML, New Generation), Compact Syntax) a été créé. Ce schéma permet la vérification automatique de la structure des fiches de la base de données. Le nom de chacune des balises se retrouve en majuscules dans nos descriptions, ce qui permet de les distinguer facilement du texte de nos fichiers. De plus, leur nom est en anglais, ce qui permet de différencier le métalangage

de la langue sur laquelle portent les descriptions. Cela procure également l'avantage de permettre à un plus grand nombre de personnes de comprendre la structure de nos fichiers, même si la langue décrite ne leur est pas accessible. Ce schéma détaille donc la modélisation des régimes lexicaux que nous avons conçue.

La base de données est une suite de GP et de MULTIPLE\_GP (GP est utilisé pour government pattern); un régime multiple (MULTIPLE\_GP) est utilisé pour la description des lexies qui permettent les modifications de régimes comme c'est le cas pour FRONTIÈRE#I.1. Chaque régime (GP ou MULTIPLE\_GP) a un numéro d'identification unique. Un régime simple GP est constitué d'un élément VALENCY et d'un élément DIATHESIS. L'élément VALENCY permet d'associer les variables sémantiques (SEM\_VAR) de la forme propositionnelle de la lexie (X, Y, Z, etc.) à la relation sémantique (SEM\_REL) impliquée (1, 2, 3, etc.). À l'intérieur de l'élément DIATHESIS, l'expression de chacun des actants sémantiques est décrite dans une balise ACTANT\_EXPRESSION. L'élément DIATHESIS est donc ce que l'on pourrait considérer comme le régime en tant que tel. On y retrouve toutes les correspondances entre les différentes relations du régime, soit la SEM\_REL, DSYNT\_REL (relation syntaxique profonde) et la/les SSYNT\_DEP (dépendance syntaxique de surface). On y retrouve également l'ensemble des syntagmes qui servent à exprimer chacun des actants. Dans les cas où il est impossible d'exprimer un actant directement vis-à-vis de la lexie décrite, on retrouve la balise NO\_REALIZATION. Il est également possible d'ajouter des conditions à divers endroits du régime, ce qui permet notamment d'indiquer la présence obligatoire d'un élément, ou d'autres contraintes sur le régime.

La plupart des éléments qui composent le texte de nos fichiers sont également contraints par ce schéma. Les valeurs possibles ont été listées de façon à limiter l'introduction d'erreurs dans la base de données. En cours d'annotation, lorsque nécessaire, il est possible et relativement aisé d'introduire de nouvelles possibilités à ce schéma. Par exemple, cet extrait du schéma indique le texte qui peut être inséré à l'intérieur de la balise SSYNT\_DEP :

```
SSYNT_DEP-ELEM = element SSYNT_DEP {"agentive"
    | "appositive appellative" | "appositive compositive"}
```

```

| "appositive copulative" | "appositive numérative"
| "appositive énonciative" | "attributive adnominale"
| "circonstancielle" | "complétive adnominale"
| "déterminative" | "modificative" | "numérale"
| "objectale adnominale" | "objectale indirecte"
| "objectale oblique" | "subjectale adnominale"}

```

#### 4.2.2 Le schéma et les balises nécessaires à l'encodage des régimes

Afin de permettre de voir l'ensemble des balises utilisées, voici le schéma .rnc qui a été conçu. Contrairement à nos explications précédentes, il contient l'ensemble des noms de balises et permet de rendre compte plus efficacement de la construction hiérarchique des régimes.

```

datatypes xsd = "http://www.w3.org/2001/XMLSchema-datatypes"

start = REG-ELEM

REG-ELEM = GOV_PATTERNS-ELEM

GOV_PATTERNS-ELEM = element GOV_PATTERNS {GP-ELEM+ & MULTIPLE_GP-ELEM+}

GP-ELEM = element GP {
  attribute id {text},
  VALENCY-ELEM,
  DIATHESIS-ELEM+
}

DIATHESIS-ELEM = element DIATHESIS {
  ACTANT_EXPRESSION-ELEM*
}

VALENCY-ELEM = element VALENCY {
  SEM_ACTANT_NUMBER-ELEM+
}

SEM_ACTANT_NUMBER-ELEM = element SEM_ACTANT_NUMBER {
  SEM_REL-ELEM,
  SEM_VAR-ELEM
}

MULTIPLE_GP-ELEM = element MULTIPLE_GP {
  attribute id {text},

```

```

VALENCY-ELEM,
GP_OPTION-ELEM+
}

GP_OPTION-ELEM = element GP_OPTION {
  attribute name {"Rég.1" | "Rég.2"},
  DIATHESIS-ELEM*
}

ACTANT_EXPRESSION-ELEM = element ACTANT_EXPRESSION {
  SEM_CORRESPONDANCE-ELEM,
  SSYNT_REALIZATION-ELEM
}

SEM_REL-ELEM = element SEM_REL {"1" | "2" | "3" | "4" | "5" | "6" |
  "1+2" | "2+3"}

SEM_VAR-ELEM = element SEM_VAR {"U" | "V" | "W" | "X" | "X1" | "X2" |
  "Y" | "Z"}

SEM_CORRESPONDANCE-ELEM = element SEM_CORRESPONDANCE {
  SEM_REL-ELEM,
  DSYNT_REL-ELEM,
  SEM_CORR_COND-ELEM}

DSYNT_REL-ELEM = element DSYNT_REL {"I" | "II" | "III" | "IV" | "V" |
  "VI"}

SEM_CORR_COND-ELEM = element SEM_CORR_COND {"oblig." | "~ est réciproque"}?

SSYNT_REALIZATION-ELEM = element SSYNT_REALIZATION {
  NO_REALIZATION-ELEM |
  SSYNT_DEPENDENCY-ELEM*}

NO_REALIZATION-ELEM = element NO_REALIZATION {"---"}

SSYNT_DEPENDENCY-ELEM = element SSYNT_DEPENDENCY {
  SSYNT_DEP-ELEM,
  PHRASES-ELEM,
  SSYNT_COND-ELEM}

SSYNT_DEP-ELEM = element SSYNT_DEP {"agentive" | "appositive appellative" |
  "appositive compositive" | "appositive copulative" |
  "appositive numérative" | "appositive énonciative" |
  "attributive adnominale" | "circonstancielle" |
  "complétive adnominale" | "déterminative" | "modificative" |

```

"numérale" | "objectale adnominale" | "objectale indirecte" |  
 "objectale oblique" | "subjectale adnominale"}

PHRASES-ELEM = element PHRASES {  
 PHRASE-ELEM+}

PHRASE-ELEM = element PHRASE {  
 USAGE\_NOTE-ELEM?,  
 (PHRASE\_GOV\_LEX-ELEM | PHRASE\_GOV\_POS-ELEM | PHRASE\_GOV\_LF-ELEM)?,  
 ACTANT\_SSYNT\_EXPR-ELEM,  
 SSYNT\_COND-ELEM?}

USAGE\_NOTE-ELEM = element USAGE\_NOTE { "familier" | "littéraire" | "soutenu" }

PHRASE\_GOV\_LEX-ELEM = element PHRASE\_GOV\_LEX { "à" | "à cause de" |  
 "à l'adresse de" | "à l'effet que" | "à l'égard de" |  
 "à l'endroit de" | "à l'encontre de" | "à l'intention de" |  
 "à l'usage de" | "à la vue de" | "adressé à" | "après" |  
 "à propos de" | "auprès de" | "au bénéfice de" | "au détriment de" |  
 "au moyen de" | "au nom de" | "au profit de" |  
 "au sein de" | "au service de" | "au sujet de" | "au visage de" |  
 "aux mains de" | "avec" | "concernant" | "contre" |  
 "d'entre" | "dans" | "(dans A-poss[X] combat/lutte) contre" |  
 "(de)" | "de" | "de cuisine" | "de la part de" |  
 "de ne pas" | "de soins" | "destiné à" | "devant" | "du fait de" |  
 "dû à" | "due à" | "en" | "en direction de" | "en échange de" |  
 "en face de" | "en faveur de" | "en l'honneur de" |  
 "en réaction à" | "entre" | "envers" | "face à" | "par" |  
 "par suite de" | "pour" | "pour ce qui a trait à" | "pour que" |  
 "pour soins" | "quant à" | "que" | "que (ne)" |  
 "sous la direction de" | "suite à" | "suite à la disparition de" |  
 "suite au décès de" | "sur" | "sur la joue de" | "sur la personne de" |  
 "vers" | "visant" | "vis-à-vis de" }

PHRASE\_GOV\_POS-ELEM = element PHRASE\_GOV\_POS { "Prép-cause" | "Prép-datif" |  
 "Prép-devant" | "Prép-envers" | "Prép-percep" | "Prép-propos" }

PHRASE\_GOV\_LF-ELEM = element PHRASE\_GOV\_LF { "Loc-ad" | "Loc-in" }

ACTANT\_SSYNT\_EXPR-ELEM = element ACTANT\_SSYNT\_EXPR {"A-num" | "A-poss" |  
 "Adj" | "N" | "N-déf" | "N-pl" | "N-propre" | "+N" | "+N-sing" |  
 "+N-pl" | "N-coord" | "N-N" | "N À BORD" | "Num" | "Num (N)" | "Num %" |  
 "Num COPIE/EXEMPLAIRE" | "Num N" | "PRON-pl" | "PROP" |  
 "« PROP »" | "PROP-inter" | "PROP-subj" | "V-inf" | "V-inf-passé" }

SSYNT\_COND-ELEM = element SSYNT\_COND {text}?

### 4.2.3 Deux exemples de Tableau de régime

Pour terminer la présentation de la base CARNAVAL, voici donc deux Tableaux de régime en faisant partie. Le r.N.2.002 (pour : régime Nominal à 2 actants numéro 002) et le r.N.2.020 (pour : régime Nominal à 2 actants numéro 020) qui correspondent respectivement à plusieurs lexies dont les suivantes : CRIME#I.b, FAUTE#I.1a d'une part et ANIMOSITÉ, AVERSION d'autre part. Nous ajoutons également le texte du fichier XML des régimes r.N.2.002 et r.N.2.020 incluant l'ensemble de leurs balises. Ceci permettra au lecteur (initié à ce langage) de bien voir les différentes composantes de ces régimes.

r.N.2.002 1→X 2→Y	
1 ⇔ I	
subjectale adnominale	de N
déterminative	A-poss
2 ⇔ II	
—	

```
<GP id="r.N.2.002">
  <VALENCY>
    <SEM_ACTANT_NUMBER>
      <SEM_REL>1</SEM_REL>
      <SEM_VAR>X</SEM_VAR>
    </SEM_ACTANT_NUMBER>
    <SEM_ACTANT_NUMBER>
      <SEM_REL>2</SEM_REL>
      <SEM_VAR>Y</SEM_VAR>
    </SEM_ACTANT_NUMBER>
  </VALENCY>
  <DIATHESIS>
    <ACTANT_EXPRESSION>
      <SEM_CORRESPONDANCE>
        <SEM_REL>1</SEM_REL>
```

```

    <DSYNT_REL>I</DSYNT_REL>
  </SEM_CORRESPONDANCE>
  <SSYNT_REALIZATION>
    <SSYNT_DEPENDENCY>
      <SSYNT_DEP>subjectale adnominale</SSYNT_DEP>
      <PHRASES>
        <PHRASE>
          <PHRASE_GOV_LEX>de</PHRASE_GOV_LEX>
          <ACTANT_SSYNT_EXPR>N</ACTANT_SSYNT_EXPR>
        </PHRASE>
      </PHRASES>
    </SSYNT_DEPENDENCY>
    <SSYNT_DEPENDENCY>
      <SSYNT_DEP>déterminative</SSYNT_DEP>
      <PHRASES>
        <PHRASE>
          <ACTANT_SSYNT_EXPR>A-poss</ACTANT_SSYNT_EXPR>
        </PHRASE>
      </PHRASES>
    </SSYNT_DEPENDENCY>
  </SSYNT_REALIZATION>
</ACTANT_EXPRESSION>
<ACTANT_EXPRESSION>
  <SEM_CORRESPONDANCE>
    <SEM_REL>2</SEM_REL>
    <DSYNT_REL>II</DSYNT_REL>
  </SEM_CORRESPONDANCE>
  <SSYNT_REALIZATION>
    <NO_REALIZATION>---</NO_REALIZATION>
  </SSYNT_REALIZATION>
</ACTANT_EXPRESSION>
</DIATHESIS>
</GP>

```

<b>r.N.2.020</b> 1 → X 2 → Y	
Rég.1	
1+2 ↔ I	
subjectale adnominale	entre N-coord
déterminative	A-poss
Rég.2	
1 ↔ I	
subjectale adnominale	de N
déterminative	A-poss
2 ↔ II	
objectale oblique	Prép-envers N

```

<MULTIPLE_GP id="r.N.2.020">
  <VALENCY>
    <SEM_ACTANT_NUMBER>
      <SEM_REL>1</SEM_REL>
      <SEM_VAR>X</SEM_VAR>
    </SEM_ACTANT_NUMBER>
    <SEM_ACTANT_NUMBER>
      <SEM_REL>2</SEM_REL>
      <SEM_VAR>Y</SEM_VAR>
    </SEM_ACTANT_NUMBER>
  </VALENCY>
  <GP_OPTION name="Rég.1">
    <DIATHESIS>
      <ACTANT_EXPRESSION>
        <SEM_CORRESPONDANCE>
          <SEM_REL>1+2</SEM_REL>
          <DSYNT_REL>I</DSYNT_REL>
        </SEM_CORRESPONDANCE>
        <SSYNT_REALIZATION>
          <SSYNT_DEPENDENCY>
            <SSYNT_DEP>subjectale adnominale</SSYNT_DEP>
          <PHRASES>
            <PHRASE>
              <PHRASE_GOV_LEX>entre</PHRASE_GOV_LEX>
              <ACTANT_SSYNT_EXPR>N-coord</ACTANT_SSYNT_EXPR>
            </PHRASE>
          </PHRASES>
        </SSYNT_REALIZATION>
      </ACTANT_EXPRESSION>
    </DIATHESIS>
  </GP_OPTION>
</MULTIPLE_GP>

```

```

        </PHRASES>
    </SSYNT_DEPENDENCY>
    <SSYNT_DEPENDENCY>
        <SSYNT_DEP>déterminative</SSYNT_DEP>
        <PHRASES>
            <PHRASE>
                <ACTANT_SSYNT_EXPR>A-poss</ACTANT_SSYNT_EXPR>
            </PHRASE>
        </PHRASES>
    </SSYNT_DEPENDENCY>
</SSYNT_REALIZATION>
</ACTANT_EXPRESSION>
</DIATHESIS>
</GP_OPTION>
<GP_OPTION name="Rég.2">
    <DIATHESIS>
        <ACTANT_EXPRESSION>
            <SEM_CORRESPONDANCE>
                <SEM_REL>1</SEM_REL>
                <DSYNT_REL>I</DSYNT_REL>
            </SEM_CORRESPONDANCE>
            <SSYNT_REALIZATION>
                <SSYNT_DEPENDENCY>
                    <SSYNT_DEP>subjectale adnominale</SSYNT_DEP>
                    <PHRASES>
                        <PHRASE>
                            <PHRASE_GOV_LEX>de</PHRASE_GOV_LEX>
                            <ACTANT_SSYNT_EXPR>N</ACTANT_SSYNT_EXPR>
                        </PHRASE>
                    </PHRASES>
                </SSYNT_DEPENDENCY>
            </SSYNT_REALIZATION>
            <SSYNT_DEPENDENCY>
                <SSYNT_DEP>déterminative</SSYNT_DEP>
                <PHRASES>
                    <PHRASE>
                        <ACTANT_SSYNT_EXPR>A-poss</ACTANT_SSYNT_EXPR>
                    </PHRASE>
                </PHRASES>
            </SSYNT_DEPENDENCY>
        </ACTANT_EXPRESSION>
    <ACTANT_EXPRESSION>
        <SEM_CORRESPONDANCE>
            <SEM_REL>2</SEM_REL>
            <DSYNT_REL>II</DSYNT_REL>
        </SEM_CORRESPONDANCE>
    </ACTANT_EXPRESSION>

```

```
<SSYNT_REALIZATION>
  <SSYNT_DEPENDENCY>
    <SSYNT_DEP>objectale oblique</SSYNT_DEP>
    <PHRASES>
      <PHRASE>
        <PHRASE_GOV_POS>Prép-envers</PHRASE_GOV_POS>
        <ACTANT_SSYNT_EXPR>N</ACTANT_SSYNT_EXPR>
      </PHRASE>
    </PHRASES>
  </SSYNT_DEPENDENCY>
</SSYNT_REALIZATION>
</ACTANT_EXPRESSION>
</DIATHESIS>
</GP_OPTION>
</MULTIPLE_GP>
```

## CHAPITRE 5

### OUVERTURE ET CONCLUSION

La base de données CARNAVAL présentée au chapitre précédent représentait évidemment le coeur de ce projet de recherche. Les efforts que nous avons mis pour la créer nous incitent à compléter ce mémoire en proposant quelques applications possibles d'une telle base de données. Plus particulièrement, nous nous servons de quelques régimes de CARNAVAL afin de fournir différentes étapes visant une ontologisation des régimes du français. De plus, la complexité du sujet traité nous pousse également à discuter des limites de la présente recherche. Nous terminons le mémoire par un bilan des conclusions générales tirées de notre étude et par un retour sur le travail accompli.

#### 5.1 Applications futures

##### **Le régime lexical comme outil en traitement automatique des langues et en lexicographie**

Bien que ce ne soit pas notre champ d'expertise, ni le sujet du présent mémoire, nous croyons qu'une base de données lexicales couvrant l'ensemble du lexique français et décrivant les régimes selon la méthode que nous avons proposée pourrait trouver de nombreuses applications en TAL. Puisque nos régimes lexicaux sont détaillés de façon très précise et qu'ils contiennent diverses composantes, ils pourraient être une source d'informations précieuses pour diverses tâches en TAL. D'une part en analyse de textes, puisque la grande quantité d'informations stockées dans les régimes peuvent aider à la désambiguïsation de constructions superficiellement identiques ; d'autre part en production automatique de textes. Comme le régime contient des informations de tous les niveaux de la représentation linguistique, ils permettent de faire la transition entre la représentation d'un sens que l'on souhaite exprimer et les constructions de surface qui expriment ce sens. Pour deux exemples d'utilisation concrets des tableaux de régime en TAL, voir Milićević (2009, 105-107).

Dans ce même article, Milicević mentionne également certaines applications lexicographiques des tableaux de régimes. Ils peuvent tout d'abord servir « d'outil diagnostique » pour postuler certaines composantes du sens d'une lexie ou pour discriminer des acceptions au sein d'un vocable polysémique.

De plus, nous avons également mentionné précédemment que notre base pourrait servir à l'édition des régimes lexicaux d'un projet lexicographique de plus grande envergure comme le RLF. Nous croyons en fait que l'outil le plus approprié pour cette tâche serait une version ontologique de notre base. En effet, si la base CARNAVAL devait être utilisée à vaste échelle pour l'édition d'articles de dictionnaire comme nous l'avons cru possible pour le RLF, il faudrait vraisemblablement la modifier. Contrairement à ce que nous avons d'abord cru, il semble que même si la partie du lexique couverte par la base est de plus en plus grande au fur et à mesure de nos encodages, il est toujours nécessaire d'ajouter des régimes de façon régulière. Tout porte à croire que ce sera toujours le cas puisque nos régimes comprennent des éléments qui sont très spécifiques (par ex., certains commentaires comme *doit être exprimé sauf si Y signifie 'boisson alcoolisée'* ou *si ~ n'est pas modifié par un Magn*; certaines prépositions régies comme *sur la personne de*; etc.). Ceci ne serait évidemment pas très pratique pour l'édition d'articles si les lexicographes doivent sans cesse ajouter de nouveaux régimes à la base parce que le régime particulier qu'ils cherchent ne s'y trouve pas. Nous croyons donc que leur outil d'édition devrait prendre en considération certaines des étapes d'ontologisation présentées plus bas en section 3. En d'autres termes, la sélection qu'ils auraient à faire se ferait dans une base de régimes simplifiés, les différentes informations omises pouvant être ajoutées « à la mitaine » avant ou à la suite du Tableau de régime (un peu comme c'était fait dans le DECFC pour les contraintes).

### **La création d'articles de dictionnaire généralisés**

Évidemment, notre base de régimes permet de comparer facilement les unités lexicales qui partagent exactement le même régime, en trouvant tout simplement les lexies qui ont le même numéro de régime. Toutefois, ce qui nous semble le plus intéressant,

c'est que la base de données permette la généralisation des informations sur les régimes, ces généralisations pouvant à leur tour contribuer à la création d'articles de dictionnaire généralisés (également d'après Milićević (2009), voir aussi Milićević (1997, 90*ssq*) et Mel'čuk et Wanner (1996)).

Comme nous l'avons vu précédemment, une étiquette sémantique définit une classe de lexies qui ont toutes la même composante centrale dans leur définition. Par exemple, l'étiquette sémantique *instrument* regroupe différentes lexies : ANCRE, CLÉ#I.1, CLOCHE#I, MARTEAU#2, OUTIL#I. Cette étiquette est placée assez haut dans la hiérarchie et englobe donc les étiquettes suivantes (qui définissent des sous-classes) *instrument agricole*, *instrument d'aiguisage*, *instrument d'optique*, *instrument de mesure*, *instrument servant à éclairer*, *instrument servant à fumer*, *instrument servant à produire un son*, *instrument servant à se protéger*, *instrument tranchant*, *outil*, *ustensile*.

On peut donc supposer que les lexies dans la descendance de *instrument* (BALAI, CLÉ#I.2, COUTEAU, FUSIL(2), HACHE, MARTEAU#1, MICROSCOPE, PARAPLUIE, PELLE(1), PIPE#I, RASOIR, SIFFLET#1, TAMBOUR#I.1, THERMOMÈTRE#I.1) héritent de celle-ci certaines propriétés sémantiques. Par exemple, il est légitime de croire que les lexies dans la descendance de *instrument* auront au moins deux ASém, le premier correspondant au potentiel utilisateur de l'instrument et le deuxième à l'usage pour lequel l'instrument est utilisé. Leur forme propositionnelle pouvant être résumée comme suit : ~ DESTINÉ À ÊTRE UTILISÉ PAR L'individu X POUR FAIRE Y.

Effectivement, on voit que la majorité des lexies mentionnées précédemment sont décrites de cette façon (par ex., COUTEAU ~ UTILISÉ PAR L'individu X [POUR COUPER / PERCER Y], MARTEAU#2 ~ UTILISÉ PAR L'individu X [POUR VENDRE L'objet Y], OUTIL#I ~ UTILISÉ PAR L'individu X POUR FAIRE L'action Y(X), etc.). Pour les cas résiduels, il serait peut-être à propos de réviser leur forme propositionnelle. Par exemple, pour les lexies ANCRE et SIFFLET#1, nous croyons que leur forme pro-

positionnelle actuelle : ~ DU bateau X, ~ UTILISÉ PAR L'individu X pourraient probablement être révisées en : ~ DESTINÉE À ÊTRE UTILISÉE PAR L'individu X POUR ARRÊTER LE DÉPLACEMENT DU bateau Y, ~ DESTINÉ À ÊTRE UTILISÉ PAR L'individu X POUR PRODUIRE LE son expressif Y.

De plus, comme les lexies qui portent la même étiquette en héritent certaines propriétés de cooccurrence lexicales (verbes supports, modificateurs, etc.) et certaines propriétés de cooccurrence avec les actants, il est possible de faire des prévisions concernant les modifications probables de leur diathèse. Milićević donne l'exemple des lexies dénotant des sentiments potentiellement réciproques qui partagent la possibilité d'une modification de régime permettant l'expression des deux ASém (celui qui expérimente le sentiment et celui qui est l'objet de ce sentiment) par le même ASyntP. Par exemple, pour les lexies COUP DE FOUDRE#1 et ANIMOSITÉ, il est possible d'exprimer les deux ASém via les DépSyntP I et II (par ex., *le coup de foudre#1/ l'animosité de Sylvie pour Alex*) ou, si le sentiment est réciproque, via la même DépSyntP I (par ex., *le coup de foudre/ l'animosité entre Sylvie et Alex*).

## 5.2 Vers une typologie des régimes nominaux

Une grande partie des travaux futurs qui pourraient être menés dans la lignée de notre projet, notamment l'utilisation d'une base de régimes pour l'édition d'un grand nombre de fiches lexicographiques et la création d'articles de dictionnaire généralisés, présuppose la généralisation des informations sur les régimes. Il va sans dire que les régimes présents dans notre base de données sont trop « complexes » pour être parfaitement utilisables dans le cadre de tels projets. Nous devons donc prévoir des étapes servant à factoriser les informations sur les régimes et permettant éventuellement de créer une typologie des régimes du français. Nous utilisons donc ici quelques régimes nominaux de la base CARNAVAL afin de présenter certaines étapes de mise en commun des régimes semblables.

### Mise en commun de la valence

Présentement, une section de nos régimes sert à indiquer à quels ASém correspondent les variables de la forme propositionnelle. Bien que l'utilisation de ces variables soit très utile, entre autres pour percevoir les liens qu'entretiennent les acceptions d'un vocable polysémique, elles ne relèvent pas de la syntaxe. La première étape pour la création de la typologie des régimes, est donc de regrouper les lexies qui ont le même régime, mais dont les noms de variables de la forme propositionnelle sont différentes. À titre d'exemple, nous pouvons regrouper les régimes r.N.2.002 et r.N.2.064.

<b>r.N.2.002</b> 1 → X 2 → Y <hr/> 1 ⇔ I <hr/> subjectale adnominal <span style="border: 1px solid black; padding: 2px;">de N</span> déterminative <span style="border: 1px solid black; padding: 2px;">A-poss</span> <hr/> 2 ⇔ II <hr/> —	<b>r.N.2.064</b> 1 → Y 2 → Z <hr/> 1 ⇔ I <hr/> subjectale adnominal <span style="border: 1px solid black; padding: 2px;">de N</span> déterminative <span style="border: 1px solid black; padding: 2px;">A-poss</span> <hr/> 2 ⇔ II <hr/> —
---	---

Figure 5.1 – Comparaison des Régimes r.N.2.002 et r.N.2.064

Dans les deux cas, le premier actant peut être exprimé par la dépendance subjectale adnominal (syntagme *de N*) ou par la dépendance déterminative (A-poss) tandis que le deuxième actant n'est pas réalisé syntaxiquement auprès de la lexie ayant ce régime. La seule différence se situe dans l'utilisation des variables dans la forme propositionnelle des lexies correspondantes : l'ASém1 correspond tantôt à X tantôt à Y tandis que l'ASém2 correspond tantôt à Y tantôt à Z. Le régime r.N.2.064 est celui de la lexie FACTURE(1)‡II *és: conséquence fp: ~ DE Y [QUE DOIT ASSUMER LA personne Z]*. Ici, Y et Z ont été utilisés pour rendre explicite le lien avec la lexie FACTURE(1)‡I *és: document justificatif fp: ~ faite par la personne X à la personne Z pour Y d'un montant W*. Donc, Y correspond dans les deux cas à ce qui fait l'objet de la facture (ce qui coûte quelque chose), et Z correspond à la personne qui doit assumer la facture. Mais, il n'en demeure

pas moins que Y et Z correspondent aux ASém1 et 2 de la lexie FACTURE(1)‡II. En ce sens, ce régime se rapproche du régime r.N.2.002, le régime des lexies suivantes (parmi d'autres) : CRIME‡I.b, FAUTE‡I.1a, IMAGINATION‡1, IMAGINATION‡2, LÂCHETÉ‡2, NAISSANCE‡I.1, PÉCHÉ, VOLTE-FACE‡I.

### Mise en commun des dépendants prépositionnels semblables

Comparons les régimes r.N.2.086 (XÉNOPHOBIE) et r.N.099 (INSOLENCE‡1, IMPARTIALITÉ‡a).

<p><b>r.N.2.086</b> 1 → X 2 → Y</p> <p>1 ⇔ I</p> <p>subjectale adnominale <input type="text" value="de N"/></p> <p>déterminative <input type="text" value="A-poss"/></p> <hr/> <p>2 ⇔ II</p> <p>objectale oblique <input type="text" value="Prép-envers N"/></p>	<p><b>r.N.2.099</b> 1 → X 2 → Y</p> <p>1 ⇔ I</p> <p>subjectale adnominale <input type="text" value="de N"/></p> <p>déterminative <input type="text" value="A-poss"/></p> <hr/> <p>2 ⇔ II</p> <p>objectale oblique <input type="text" value="à l'égard de N"/> <input type="text" value="à l'endroit de N"/> <input type="text" value="envers N"/> <input type="text" value="vis-à-vis de N"/></p>
--	---

Figure 5.2 – Comparaison des Régimes r.N.2.086 et r.N.2.099

Nous remarquons que la différence entre ces deux régimes est assez ténue. En effet, les dépendances utilisées pour l'expression des actants sont les mêmes, seules les prépositions possibles pour exprimer le deuxième actant par la dépendance **objectale oblique** sont légèrement différentes. Ces prépositions sont assez proches sémantiquement, elles ont la particularité d'avoir le sens de 'envers'. *Prép-envers* est utilisé pour indiquer qu'il est possible d'utiliser n'importe quelle préposition qui fait partie de cette partie du discours spécifique, c'est-à-dire une préposition qui a le sens de 'envers'<sup>1</sup>. Dans le cas du r.N.2.099, on a listé les prépositions possibles, car on a jugé que c'était les

1. Cette façon de faire est relativement récente, elle n'est pas présente dans les DECFC. On l'a introduite dans le DiCo et le LAF et on la mentionne brièvement dans l'ILEC, mais il n'est pas possible pour l'instant de retrouver une liste complète de ces sous-groupes de la partie du discours *Prép*. Sans prétendre en donner une liste exhaustive, nous fournissons dans l'annexe II les informations que nous avons recueillies à divers endroits du DiCo et certaines modifications ou ajouts que nous proposons.

seules qui permettent d'exprimer l'ASyntPII (en excluant donc les autres prépositions du groupe *Prép-envers*).

En ce sens, il serait possible de créer une catégorie de régimes où l'ASyntPII est exprimé par la dépendance **objectale oblique** à l'aide d'une préposition faisant partie du groupe *Prép-envers*. Ce groupe de régimes comprendrait également les régimes r.2.019 (ADORATION#*a*), r.N.2.045 (CRIME#*I.a*), r.N.2.049 (DÉGOÛT), r.N.2.050 (DÉPIT), r.N.2.052 (DÉSIR#*2*), r.N.2.086, r.N.2.099, r.N.2.104 (INSOLENCE#*2*), r.N.2.108 (MÉCONTENTEMENT), r.N.2.116 (PARTIALITÉ#*a*), et potentiellement les régimes multiples r.N.2.020 (ANIMOSITÉ, AVERSION) et r.N.2.041 (HAINE, HOSTILITÉ#*I*, COUP DE FOUDRE#*1*). La liste des prépositions possibles serait alors donnée dans l'article lexicographique de L plutôt que dans son Tableau de régime.

### **Commentaires, notes d'usage, expression obligatoire d'un actant**

Dans cette optique d'ontologisation des régimes, il serait également possible d'exclure les commentaires et contraintes des tableaux pour en arriver à des régimes plus simples. Prenons en exemple un régime assez simple, le r.N.2.003 (AMI#*I.1*, AMI#*I.3*, BAGUE#*2*, ENNEMI#*I*, ENNEMI#*IV*, ENVELOPPE#*I*, ÉQUIPAGE, FACTURE(2)#*a*, PARTIE(1)#*I*, PATIENT, PETIT AMI, POULE#*II.a*, VERROU#*I*). Si l'on forme un groupe avec tous les régimes qui ont cette base (ASém1 non exprimable et ASém2 exprimable par **complétive adnominale de N** ou **déterminative A-poss**), il comprendrait également certains régimes qui incluent des contraintes particulières, par exemple le r.N.2.004 (AMI#*II*, ENNEMI#*II*) où l'expression de l'ASém2 est obligatoire et où l'ASém2 est exprimable par A-poss seulement si Y est une personne).

<p><b>r.N.2.003</b> <math>1 \rightarrow X 2 \rightarrow Y</math></p> <table border="1"> <tr> <td>1 ⇔ I</td> </tr> <tr> <td>—</td> </tr> </table> <p>2 ⇔ II</p> <p>complétive adnominal <span style="border: 1px solid black; padding: 2px;">de N</span></p> <p>déterminative <span style="border: 1px solid black; padding: 2px;">A-poss</span></p>	1 ⇔ I	—	<p><b>r.N.2.004</b> <math>1 \rightarrow X 2 \rightarrow Y</math></p> <table border="1"> <tr> <td>1 ⇔ I</td> </tr> <tr> <td>—</td> </tr> </table> <p>2 ⇔ II   oblig.</p> <p>complétive adnominal <span style="border: 1px solid black; padding: 2px;">de N</span></p> <p>déterminative <span style="border: 1px solid black; padding: 2px;">A-poss   Y est une personne</span></p>	1 ⇔ I	—
1 ⇔ I					
—					
1 ⇔ I					
—					

Figure 5.3 – Comparaison des Régimes r.N.2.003 et r.N.2.004

### Regroupement des DépSyntS

Il va sans dire que ces premières étapes visent à regrouper des régimes assez semblables en offrant un patron de régime simplifié. Si l'on souhaite véritablement créer une typologie des régimes, on pourra envisager d'aller plus loin dans les regroupements, par exemple en rassemblant tous les régimes où l'ASém2 est exprimé par la dépendance **objectale oblique** (peu importe les prépositions possibles, peu importe s'il existe d'autres possibilités d'expression pour ce même ASém, peu importe le moyen d'expression de l'ASém1, etc.).

Nous envisageons la création d'une telle typologie puisque nous croyons qu'il serait utile de développer un outil qui permette de faire des généralisations à plusieurs niveaux. Cela semble nécessaire, parce que les généralisations peuvent être assez spécifiques (par ex., les lexies qui dénotent des sentiments réciproques et leur modification de diathèse) ou plutôt générales (par ex., beaucoup de lexies quasi prédicatives – classées dans la hiérarchie sous l'étiquette sémantique *entité* – ont le premier actant de leur prédicat interne fusionné avec la lexie nominale (la lexie nominale correspond à ce premier actant) : 'femme X, qui est la mère de Y' ⇔ MÈRE de N[=Y], ce qui fait que leur premier actant n'est pas syntaxiquement réalisé comme un D de L).

Il en est ainsi parce que les régimes peuvent être perçus comme des entités « lexico-syntaxiques » qui portent en elles-mêmes des sens, parce que leur structure offre la pos-

sibilité d'exprimer une structure actancielle donnée. Nous croyons donc que d'étudier ces patrons syntaxiques en tant qu'outils d'expression de groupes de sens (qui peuvent être complexes ou assez vagues) peut être intéressant.

### 5.3 Limites de la recherche

Tout d'abord, il est évident que le travail effectué ici ne permet de mener à bien qu'une partie des objectifs fixés en début de recherche. Une certaine incertitude subsiste donc quant au caractère exhaustif de notre travail. Est-ce que toutes les DépSyntS régies par des noms du français ont été répertoriées ? Est-ce que la plupart des régimes nominaux du français ont été décrits ? La supposition la plus probable est que la presque totalité des dépendances nécessaires sont décrites, mais qu'un nombre considérable de régimes nominaux reste à décrire. Il semble évident que plus la partie du lexique couverte par ce projet s'élargira, plus de régimes seront à décrire. Toutefois, les outils techniques et la méthode utilisée devraient être efficaces pour la description de régimes d'autres unités lexicales.

De plus, nous ne prétendons pas à la complétude de la liste de dépendances que nous avons établie. Nous l'avons donnée ici dans son état actuel. Nous n'écartons pas la possibilité d'ajouter des dépendances ou d'en modifier certains critères si notre travail d'annotation actuel ou si de futurs travaux sur les dépendances le démontrent nécessaire. Par exemple, la dépendance que nous avons nommée **complétive adnominale** comprend peut-être en fait plus d'une dépendance. Nous y avons inclus des constructions qui expriment des sens différents (les relations contenant~contenu, partie~tout, de possession, de quantification, etc.), sans établir définitivement si elles partagent véritablement l'ensemble de leurs propriétés syntaxiques. Dans certains cas, le dépendant en « de N » peut être paraphrasé par A-poss ou par « à N » "fam." :

– *la bague#1 de ma grand-mère ~ sa bague#1 ~ "fam." la bague#1 à ma grand-mère*

Mais ce n'est pas toujours le cas :

– *la miette du pain ~ \*?sa miette, ~ "fam." \*?la miette au pain*

Il est donc légitime de se demander s'il ne serait pas souhaitable d'établir au moins une séparation permettant de distinguer les **complétives adnominales** qui permettent ces paraphrases de celles qui ne le permettent pas. Toutefois, nous ne l'avons pas fait, parce qu'il semble que cette possibilité soit liée à la sémantique de la lexie G et à la nature sémantique des D qu'elle contrôle (par ex., ces paraphrases sont surtout possibles si le N dans le sous-groupe de D est un individu) plutôt qu'aux propriétés syntaxiques de la dépendance. D'ailleurs, la présence d'un déterminant possessif est possible autant pour les constructions régies (*son cimetière* 'le cimetière où il est enterré') que pour les constructions libres (*son cimetière* 'le cimetière où il travaille' / 'le cimetière dont il a conçu les plans' / 'le cimetière qu'il traverse à la marche chaque matin' / etc.). Les paraphrases par A-poss sont également possibles pour les DépSyntS **subjectale adnominale** (par ex., *l'abolement* du chien ~ *son abolement*) et **objectale adnominale** (par ex., *l'abandon [par Louis] de ses enfants* ~ *leur abandon*).

Tout ceci nous porte à croire que la possibilité d'utiliser un A-poss comme paraphrase d'un dépendant en « *de N* » n'est pas une propriété définitoire (ou discriminatoire) d'une dépendance en particulier, mais que ce n'est qu'une des propriétés qui permette de reconnaître les dépendances (on a vu que dans le cas de la dépendance **attributive adnominale**, cette paraphrase n'est jamais possible). Toutefois, certains considèrent que les groupes possessifs en « *de N* » (ou les constructions possessives de l'anglais avec 's) forment un seul groupe de constructions, puisque le sens exprimé est semblable dans tous les cas. Nous croyons donc que notre choix se discute.

Certains autres problèmes persistent également. Par exemple, certaines constructions ayant d'abord été identifiées comme étant **attributive adnominale** ont plutôt été classées sous la dépendance **complétive adnominale** (par ex., *bateau de militants écologistes*, *assiette de sandwiches*, *guirlande de ballons*, *majorité d'étudiants*). Il en avait été ainsi probablement à cause de l'absence de déterminant devant N2, qui est l'une des façons de repérer la dépendance **attributive adnominale**. Toutefois,

il semble que dans ces cas, l'absence de déterminant soit plutôt due à un autre phénomène d'ordre morpho-syntaxique (règle d'haplologie ou de « cacophonie »<sup>2</sup> : DE-prép. + DES-dét.ind.plur = *de*). Ainsi, dans l'exemple *le bateau de militants écologistes* le nom *militants* n'est pas indéterminé, mais déterminé avec un article zéro qui est la variante morphologique de l'indéfini pluriel.

1. N2 est défini : *le bateau \*de les militants écologistes = le bateau des militants écologistes*
2. N2 est indéfini : *le bateau \*de des militants écologistes = le bateau de ∅ militants écologistes*

On aura longtemps cru qu'il s'agissait de la même dépendance parce que la non-référentialité des N2 (indéfinis) de ces constructions est difficile à tester de façon concluante. Il semble toutefois que le problème soit encore plus complexe que cela.

En effet, la détermination en français est quelque peu épineuse. Il semble plutôt difficile que deux noms soient déterminés avec un article indéfini au sein du même groupe nominal. Comparons les exemples suivants :

1. N1 défini, N2 défini : *le morceau de la météorite, la miette du pain, la plume de l'autruche*
2. N1 défini, N2 indéfini : *le morceau d'une météorite, la miette d'un pain, la plume d'une autruche*
3. N1 indéfini, N2 défini : *un morceau de la météorite, une miette du pain, une plume de l'autruche*
4. N1 indéfini, N2 indéfini : *?\*un morceau d'une météorite, ?\*une miette d'un pain, ?\*une plume d'une autruche*

Tout porte à croire que les sens des exemples donnés en 4 seraient plutôt exprimés par les constructions *un morceau de météorite, une miette de pain, une plume d'autruche*. Il est donc possible que l'absence de déterminant devant N2 soit également attribuable à un autre phénomène que nous n'avons pas très bien étudié jusqu'ici. Nous croyons donc que la dépendance **attributive adnominale** soit beaucoup moins présente en français que nous l'avons d'abord cru, surtout à l'intérieur des régimes. La dépendance

2. En ce qui concerne cette règle, voir : Benetti (2008, Chapitre II, point 322), Riegel et al. (1994, 166).

**attributive adnominale** serait en fait présente dans les cas où cette absence de déterminant sert exclusivement à la construction d'une structure particulière qui exprime un sens particulier et où l'opération de référence est particulière (« N1 Prép N2 » est référentiel alors que N2 pris isolément ne l'est pas).

Notons au passage que certaines constructions syntaxiques sont toujours en cours de caractérisation et que la création d'une instance de DépSyntS correspondante sera nécessaire. Cette incertitude sur la complétude de notre liste de DépSyntS nominales régies du français découle de notre corpus qui est somme toute assez petit en comparaison avec l'ensemble du lexique français. Par exemple, dans des constructions comme *ce chien de Gustave, ce gauchiste de journaliste, cet amour d'enfant*, etc., il existe une dépendance qui se rapproche des appositives (N1 et N2 ont le même référent), mais où une préposition est présente entre N1 et N2<sup>3</sup>. Nous n'avons pas terminé la caractérisation de cette dépendance particulière, parce que ce n'est probablement pas une dépendance typiquement régie et qu'elle n'est donc pas présente dans les régimes des lexies que nous avons étudiées<sup>4</sup>.

Comme dépendance qui manque potentiellement à notre inventaire, nous pouvons également noter la dépendance présente dans les cas où un D correspond à ce que M. Grevisse nomme un *pseudo-complément*, c'est-à-dire lorsque « la subordination grammaticale ne correspond pas à la subordination logique » (Grevisse et Goose 2007, 430), ou, en d'autres termes, lorsque le G syntaxique ne semble pas être ce qui est sémantiquement le plus important. Nous croyons qu'il serait peut-être utile de créer cette instance, parce que pour ces groupes nominaux, l'accord peut être déclenché par le G ou le D : *Une bande d'étudiants est venue manifester. ~ Une bande d'étudiants*

3. Pour d'autres exemples, voir : Flaux et Van de Velde (2000); Gaatone (1988); Larrivée (1994); Tellier et Valois (2003). Polguère (À paraître) propose le nom de « complétive évaluative », il faudra vraisemblablement étudier cette possibilité et ajouter la caractérisation de cette dépendance dans la hiérarchie GTN.

4. Ou du moins, elle ne l'est pas suffisamment pour que nous puissions cerner adéquatement les propriétés syntaxiques qui servent à son identification/caractérisation. Toutefois, nous croyons que cette dépendance « appositive indirecte » ou « complétive évaluative » est potentiellement présente dans le régime de BIJOU#II [*L'opéra royal est un bijou de petit théâtre.*]. Cette construction particulière est également présente dans l'article lexicographique de CHIEN#II [*ce chien de Gustave, ce chien de journaliste*], mais cela est actuellement consigné dans la zone de combinatoire lexicale plutôt que dans son régime.

*sont venus manifester*<sup>5</sup>.

Finalement, l'approche lexicographique d'un tel phénomène ne permet évidemment qu'un regard sur un aspect précis de la langue. Il sera intéressant, lorsque l'ensemble des DépSyntS (régies, non régies et mixtes, nominales, adjectivales, verbales, adverbiales, etc.) du français seront décrites de tester notre inventaire sur des corpus. Il serait éventuellement possible de mettre en branle un projet d'annotation de corpus qui permettrait de voir si les dépendances décrites permettent d'annoter les corpus dans leur ensemble ou si l'ajout d'autres dépendances serait nécessaire. Il serait également possible de vérifier dans le cadre d'un tel projet si les critères d'identification fixés sont nécessaires et suffisants pour discriminer l'ensemble des dépendances examinées effectivement dans la langue courante.

#### 5.4 Bilan

Comme nous l'avons vu dans le premier chapitre, lorsque l'on traite des unités lexicales et de leurs divers dépendants, ceux-ci sont souvent classifiés selon leur partie du discours (proposition, syntagme nominal, syntagme prépositionnel avec préposition sélectionnée – *de, envers, à cause de* ou autre –, etc.). Diverses théories linguistiques reconnaissent d'ailleurs l'utilité de cette technique dans la description du lexique (voir par exemple la c-sélection de la grammaire générative). On peut observer que cette méthode est parfois appliquée en lexicologie/lexicographie. Toutefois, en observant les grands dictionnaires, on se rend compte que lorsque cela est fait, on précise surtout certaines caractéristiques sémantiques des dépendants (*qqn, qqch.*, etc.). Une autre pratique courante est de traiter du comportement syntaxique de la lexie vedette. Pour les verbes, on utilise la plupart du temps des étiquettes de type « transitif, intransitif, pronominal, etc. », ce qui permet de savoir par exemple qu'un verbe donné s'utilise avec ou sans objet direct. L'usage de telles étiquettes, bien que souvent présentes dans les descriptions de

---

5. Grevisse considère aussi comme pseudo-compléments les D dans *l'immensité de la mer* et *le fait de n'avoir rien répondu*. Il serait peut-être bon de comparer ce dernier exemple avec les lexies CRIME#II et HONTE#II.

lexies verbales, n'est pas pratique courante dans la description du comportement syntaxique des unités lexicales des autres parties du discours. En ce qui concerne les noms, la précision quant à la nature de leurs compléments est souvent uniquement déductible des exemples donnés dans les articles.

Évidemment, toutes les techniques mentionnées ci-haut permettent de présenter de précieuses informations au sujet des dépendants syntaxiques des lexies décrites. Cependant, lorsqu'elles sont présentes dans le dictionnaire, de telles informations servent surtout à distinguer différents sens ou différents usages et ne sont pas fournies automatiquement. De plus, l'ensemble de ces pratiques ne semble pas permettre de rendre compte efficacement de la relation sémantique qui existe entre le nom gouverneur et ses dépendants. En lexicographie générale, on ne mentionne que de manière sporadique si un dépendant donné est un actant ou un circonstant du gouverneur, si un complément est facultatif ou obligatoire, etc. Il est rare que l'on précise systématiquement, surtout dans le cas des noms, quels sont les actants sémantiques de la lexie vedette et la façon dont chacun d'eux s'exprime dans une phrase vis-à-vis de celle-ci.

De plus, les descriptions en termes de simples étiquettes morphologiques (« *de N* » par exemple) ne permettent pas de rendre compte des différences entre des constructions superficiellement identiques qui possèdent pourtant des propriétés syntaxiques bien différentes. Pour illustrer ce propos, voici divers exemples où un dépendant en « *de N* » est gouverné par un nom (en gras) :

1. [*le*] **chat** *de Marie*            'le chat qui appartient à Marie'
2. [*la*] **dispute** *de Jo et Lilly*    'la dispute entre Jo et Lilly'
3. [*une*] **ceinture** *de pantalon*    'une ceinture qui se porte généralement avec un pantalon'
4. [*un*] **échange** *de joueurs*        'un échange qui implique des joueurs'
5. [*cet*] **imbécile** *de Gustave*      'Gustave, qui est un imbécile'
6. [*la*] **ville** *de Paris*            'la ville dont le nom est Paris'
7. [*l'*] **anniversaire** *de Gil*        'l'anniversaire du jour où Gil est né'

Évidemment, les complémentations en « *de N* » ci-haut, bien que superficiellement identiques, ne sont pas de même nature. Elles impliquent des relations syntaxiques différentes puisque les constructions ont des propriétés « cachées » différentes. Par exemple,

les « *de N* » des exemples (1), (2), (4) et (7) peuvent être remplacés par des déterminants possessifs (*son chat, leur dispute, [on a procédé à] leur échange, son anniversaire*), ce qui n'est pas le cas des exemples en (3), (5) et (6).

Conséquemment, nous avons voulu réfléchir sur un moyen rendant la désambiguïsation des constructions de complémentation possible. Dans notre optique d'encodage de valence, le moyen que nous avons envisagé ici est d'inscrire le nom de la dépendance syntaxique précise qui permette l'expression des actants. Dans le cadre de notre projet, divers tests syntaxiques (comme la possibilité pour le syntagme du dépendant d'être remplacé par un déterminant possessif) ont servi à l'identification des dépendances syntaxiques régies par des noms qui existent en français. De plus, leur caractérisation a également été faite en partie grâce aux sens qu'elles traduisent.

Nous avons d'abord examiné une méthode servant à l'identification des DépSyntS qui a permis l'établissement d'une liste des DépSyntS nominales du français. Nous avons ensuite réfléchi à la modélisation lexicographique des dépendances syntaxiques, suite à quoi nous avons amorcé la création de la base de données sur les régimes lexicaux du français CARNAVAL en encodant les régimes des lexies nominales du DiCo. Une telle base de données (ou sa potentielle version ontologique) pourrait éventuellement avoir diverses applications en lexicographie (édition d'articles lexicographiques, création d'articles de dictionnaire généralisés), dans l'étude de l'interface sémantique-syntaxe (notamment en ce qui a trait au problème d'*argument linking*), en TAL (parce que aide à la désambiguïsation de constructions superficiellement identiques) ou en didactique des langues (parce que permet de mettre en lumière certaines généralités de la langue).

## BIBLIOGRAPHIE

*Antidote Prisme*, 2004. [Logiciel], Druide Informatique, Montréal.

Laurence Benetti. 2008. *L'article zéro en français contemporain : Aspects syntaxiques et sémantiques*. Peter Lang, Berlin.

Morton Benson, Evelyn Benson et Robert Ilson. 1986. *The BBI Combinatory Dictionary of English : A Guide to Word Combination*. John Benjamins, Amsterdam/Philadelphia.

Claire Blanche-Benveniste. 1987. *Pronom et syntaxe : l'approche pronominale et son application au français*. Peeters Publishers, Paris-SELAF.

K. Bogacki, G. Bogustawska, A. Kleisberg, H. Lewicka, M. Tozinska, M. Thieme, L. Wojcik et T. Zielinska. 1983. *Dictionnaire sémantique et syntaxique des verbes français*. Panstwowe Wydawnictwo Naukowe, Varsovie.

Winfried Busse et Jean-Pierre Dubost. 1983. *Französisches Verblexikon. Die Konstruktion der Verben im Französischen*. Ernst Klett, Stuttgart.

Josette Caput et Jean-Pol Caput. 1969. *Dictionnaire des verbes français*. Larousse, Paris.

Noam Chomsky. 1986. *Knowledge of Language : Its Nature, Origin, and Use*. Praeger, New York.

Robert Coyne et Owen Rambow. 2009. Meaning-Text Theory and Lexical Frames. Dans David Beck, Kim Gerdes, Jasmina Milićević et Alain Polguère, éditeurs, *Proceedings of Fourth International Conference on Meaning-Text Theory (MTT'09) and International Workshop on Terminology and Lexical Semantics (TLS'09)*, pages 119–128, Montréal, 16-19 Juin 2009. Observatoire de linguistique Sens-Texte, Université de Montréal.

- Karel Van den Eynde et Piet Mertens. 2006. *Le dictionnaire de valence DICOVALENCE : manuel d'utilisation*, 2006. URL [http://bach.arts.kuleuven.be/dicovalence/manuel\\_061117.pdf](http://bach.arts.kuleuven.be/dicovalence/manuel_061117.pdf).
- Jacques Dendien et Jean-Marie Pierrel. 2003. Le Trésor de la Langue Française informatisé : un exemple d'informatisation d'un dictionnaire de langue de référence. *TAL*, 44(2):11–37.
- DiCoInfo. URL <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi>.
- DiCoPop. URL <http://olst.ling.umontreal.ca/dicopop/index.php>.
- Christiane Fellbaum, éditeur. 1998. *WordNet : An Electronic Lexical Database*. The MIT Press, Cambridge.
- Charles J. Fillmore. 1977. Scenes-and-frames semantics. Dans Antonio Zampolli, éditeur, *Linguistic Structures Processing*, numéro 59 dans *Fundamental Studies in Computer Science*, pages 55–88. North-Holland Publishing Company, Amsterdam/New York/Oxford.
- Nelly Flaux et Danièle Van de Velde. 2000. *Les noms en français : esquisse de classification*. Ophrys, Paris.
- Ovidiu Frînculescu. 1978. *Dictionarul verbelor franceze cu constructiile lor specifice [Dictionnaire des verbes français avec leurs constructions spécifiques]*. Albatros, Bucarest, Roumanie.
- David Gaatone. 1988. Cette coquise de construction : remarques sur les trois structures affectives du français. *Travaux de linguistique*, 17:159–176.
- Maurice Grevisse. 1995. *Précis de grammaire française*. De Boeck/ Duculot, Bruxelles, 30<sup>e</sup> édition.
- Maurice Grevisse et André Goose. 2007. *Le bon usage : grammaire française*. De Boeck Université, Bruxelles, 14<sup>e</sup> édition.

- Maurice Gross. 1975. *Méthodes en syntaxe*. Hermann, Paris.
- Maurice Gross. 1979. On the Failure of Generative Grammar. *Language*, 55(4):859–885.
- Maurice Gross. 1996. *Les expressions figées en français*. Ophrys, Paris.
- Lidijia Iordanskaja et Igor Mel'čuk. 2009. Establishing an Inventory of Surface-Syntactic Relations : Valence-controlled Surface-Syntactic Dependents of the Verb in French. Dans Igor Mel'čuk et Alain Polguère, éditeurs, *Dependency in Linguistic Description*, volume 111 de *Studies in Language Companion Series*, pages 151–234. John Benjamins, Amsterdam/Philadelphia.
- Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge.
- Évelyne Jacquy et Anne-Céline Naels. 2005. Vers un lexique syntaxique du français : extraction d'informations de sous-catégorisation à partir du TLFi. Dans *Interface lexique-grammaire et lexiques syntaxiques et sémantiques*. ATALA, 12 Mars 2005. URL [http://www.atala.org/doc/JE\\_050312/Lexsynt-Jacquy.pdf](http://www.atala.org/doc/JE_050312/Lexsynt-Jacquy.pdf).
- Anne-Laure Jousse. 2003. Normalisation des fonctions lexicales. Mémoire de DEA, UFR de linguistique, Paris 7.
- Anne-Laure Jousse et Alain Polguère. 2005. *Le DiCo et sa version DiCouèbe. Document descriptif et manuel d'utilisation*. OLST, Département de linguistique et traduction, Université de Montréal. URL <http://olst.ling.umontreal.ca/dicouebe/>.
- Sylvain Kahane. 2001. Grammaires de dépendance formelles et Théorie Sens-Texte. Dans *Actes de TALN 2001*, volume 2, pages 17–76, Tours, 2-5 juillet 2001.
- Sylvain Kahane. 2003. The Meaning-Text Theory. Dans V. Agel, L. Eichinger, H.-W. Eroms, P. Hellwig, H.J. Heringer et H. Lobin, éditeurs, *Dependency and Valency*.

- An International Hand-book of Contemporary Research*, volume 1, pages 546–570, Berlin-New York, 2003. W. de Gruyter.
- Marie Laurence Knittel. 2009. Le statut des compléments du nom en [de NP]. *Canadian Journal of Linguistics/ Revue canadienne de linguistique*, 54(2):255–290.
- Pierre Larrivée. 1994. Quelques hypothèses sur les structures syntaxique et sémantique de *Ce fripon de valet*. *Revue québécoise de linguistique*, 23(2):101–113.
- Veronika Lux-Pogodalla et Alain Polguère. 2011. Construction of a French Lexical Network : Methodological Issues. Dans *Proceedings of the First International Workshop on Lexical Resources (WoLeR 2001)*, pages 55–62, Ljubljana, Slovenia, 1-5 août 2011. ESSLLI.
- Sébastien Marengo. 2009. Les dépendants syntaxiques de l’adjectif en français : vers un inventaire des relations syntaxiques de surface. Dans David Beck, Kim Gerdes, Jasmina Milićević et Alain Polguère, éditeurs, *Proceedings of Fourth International Conference on Meaning-Text Theory (MTT’09) and International Workshop on Terminology and Lexical Semantics (TLS’09)*, pages 263–273, Montréal, 16-19 Juin 2009. Observatoire de linguistique Sens-Texte, Université de Montréal.
- Sébastien Marengo. 2011. *Les adjectifs jamais attributs*. Série Champs linguistiques. De Boeck/ Duculot, Bruxelles.
- Igor Mel’čuk. 1988. *Dependency syntax : theory and practice*. SUNY series in linguistics. State University of New York Press, Albany.
- Igor Mel’čuk. 1997. *Vers une linguistique Sens-Texte*. Leçon inaugurale (faite le Vendredi 10 janvier 1997), Collège de France, Chaire Internationale.
- Igor Mel’čuk. 2001. Semantics and the lexicon in modern linguistics. URL [http://olst.ling.umontreal.ca/?page\\_id=395](http://olst.ling.umontreal.ca/?page_id=395).
- Igor Mel’čuk. 2004a. Actants in semantics and syntax I : Actants in semantics. *Linguistics*, 42/1:1–66.

- Igor Mel'čuk. 2004b. Actants in semantics and syntax II : Actants in syntax. *Linguistics*, 42/2:247–291.
- Igor Mel'čuk. 2009. Dependency in Natural Language. Dans Igor Mel'čuk et Alain Polguère, éditeur, *Dependency in Linguistic Description*, pages 1–110. John Benjamins, Amsterdam/Philadelphia.
- Igor Mel'čuk. En préparation. *Semantics : From Meaning to Text*, chapitre Government Pattern. Benjamins, Amsterdam/Philadelphia.
- Igor Mel'čuk, André Clas et Alain Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Duculot, Paris/Louvain-la-Neuve.
- Igor Mel'čuk et coll. 1984-1999. *Dictionnaire explicatif et combinatoire du français contemporain : recherches lexico-sémantiques I-IV*. Les Presses de l'Université de Montréal, Montréal.
- Igor Mel'čuk et Jasmina Milićević. 2011. *Introduction à la linguistique*, volume 1. Hermann, Paris.
- Igor Mel'čuk et Nikolaj Pertsov. 1987. *Surface Syntax of English. A Formal Model within the Meaning-Text Framework*. John Benjamins, Amsterdam/Philadelphia.
- Igor Mel'čuk et Alain Polguère. 2007. *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. De Boeck & Larcier, Bruxelles.
- Igor Mel'čuk et Alain Polguère. 2008. Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. *Revue de linguistique et de didactique des langues*, 37:99–114.
- Igor Mel'čuk et Leo Wanner. 1996. Lexical Functions and Lexical Inheritance for Emotion Lexemes in German. Dans Leo Wanner, éditeur, *Lexical Functions in Lexicography and Natural Language Processing*, pages 209–278. Benjamins, Amsterdam/Philadelphia.

- Jasmina Milićević. 1997. Étiquettes sémantiques dans un dictionnaire formalisé de type « Dictionnaire explicatif et combinatoire ». Mémoire de maîtrise, Université de Montréal.
- Jasmina Milićević. 2006. A Short Guide to the Meaning-Text Linguistic Theory. *Journal of Koralex*, 8:187–233.
- Jasmina Milićević. 2009. Schéma de régime — le pont entre le lexique et la grammaire. *Langages*, 1(1):94–116.
- Jonathan Owens. 1988. *The Foundations of Grammar : An Introduction to Mediaeval Arabic Grammatical Theory*. Benjamins, Amsterdam.
- Paul Bogaards, éditeur. 2003. *International Journal of Lexicography*, Special Issue 1, 16(3). Oxford University Press.
- Jacqueline Picoche. 1993. *Didactique du vocabulaire français*. Nathan, Paris.
- Jacqueline Picoche et Jean-Claude Rolland. 2002. *Dictitonnaire du français usuel : 15 000 mots utiles en 442 articles*. De Boeck/ Duculot, Bruxelles.
- Alain Polguère. 1998. La Théorie Sens-Texte. *Dialangue*, 8-9:9–30.
- Alain Polguère. 2002. Modélisation des liens lexicaux au moyen des fonctions lexicales. Dans *TALN 2002*, pages 37–60, Nancy, 24-27 juin 2002.
- Alain Polguère. 2007. Lexical Function Standardness. Dans Leo Wanner, éditeur, *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, volume 84 de *Language Companion Series*, pages 43–95. John Benjamins, Amsterdam/Philadelphie.
- Alain Polguère. À paraître. Les dépendances nominales régies. Dans *Actes de la conférence ComplémentationS*, GRAMM-R. Peter Lang, Berlin.
- Josette Rey-Debove et Alain Rey. *Le nouveau Petit Robert : dictionnaire alphabétique et analogique de la langue française*. [CD-ROM], 2008.

- Martin Riegel, Jean-Christophe Pellat et René Rioul. 1994. *Grammaire méthodique du français*. PUF, Paris.
- Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson et Jan Scheffczyk. 2010. *Framenet II : Extended Theory and Practice*. ICSI. URL <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>.
- Ivan Sag et Carl Pollard. 1989. Subcategorization and Head-Driven Phrase Structure. Dans M. Baltin et A. Kroch, éditeurs, *Alternative Conceptions of Phrase Structure*, pages 139–181. University of Chicago Press, Chicago.
- John Sinclair, éditeur. 1987. *Collins COBUILD English Language Dictionary for Advanced Learners*. Collins, London/ Glasgow.
- Della Summers, éditeur. 2005. *Longman Dictionary of Contemporary English*. Pearson Education Limited, England.
- Christine Tellier et Daniel Valois. 2003. *Constructions méconnues du français*. Les Presses de l'Université de Montréal, Montréal.
- Lucien Tesnière. 1959. *Éléments de syntaxe structurale*. Klincksieck, Paris.
- Ophélie Tremblay. 2009. *Une ontologie des savoirs lexicographiques pour l'élaboration d'un module de cours en didactique du lexique*. Thèse de doctorat, Université de Montréal, Département de didactique de la Faculté des sciences de l'éducation.
- Karel Van den Eynde et Claire Blanche-Benveniste. 1978. Syntaxe et mécanismes descriptifs : présentation de l'approche pronominale. *Cahiers de Lexicologie*, 32:3–27.
- Karel Van den Eynde et Piet Mertens. 2003. La valence : l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies*, 13:63–104.
- Piet van Sterkenburg, éditeur. 2003. *A Practical Guide to Lexicography*. John Benjamins, Amsterdam /Philadelphia.

Aleksandr Žolkovskij et Igor Mel'čuk. 1967. O semantičeskom sinteze [sur la synthèse sémantique]. *Problemy kibernetiki*, 19:177–238. [Il existe une traduction en français : T.A.Informations, 1970, 2, 1-85.].

Ladislav Zgusta. 2006. Methods and trends in lexicography. Dans Frederic S. F. Dolezal et Thomas B. I. Creamer, éditeurs, *Lexicography Then and Now : Selected essays*, Lexicographica Series Maior, pages 87–140. Max Niemeyer Verlag, Tübingen.

## Annexe I

### Liste des lexies étudiées et leur numéro de régime

ABAISSEMENT#I.1	r.N.1.002	ACCUSATION#I.1	r.N.3.003
ABAISSEMENT#I.2	r.N.2.007	ACCUSATION#I.2	r.N.4.002
ABAISSEMENT#II	r.N.1.002	ACCUSATION#II	r.N.2.001
ABANDON#I.1	r.N.2.008	ADMIRATION	r.N.3.004
ABANDON#I.2a	r.N.2.009	ADORATION#a	r.N.2.019
ABANDON#I.2b	r.N.1.003	ADORATION#b	r.N.2.014
ABANDON#I.3a	r.N.2.009	AGNEAU#a	r.N.1.004
ABANDON#I.3b	r.N.1.003	AGNEAU#b	r.N.1.001
ABANDON#I.4	r.N.2.010	AGNEAU#c	r.N.1.001
ABANDON#I.5	r.N.3.001	ALLUMETTE	r.N.1.001
ABANDON#II	r.N.2.011	AMI#I.1	r.N.2.003
ABAT-JOUR	r.N.1.004	AMI#I.2	r.N.3.005
ABATTEMENT(1)	r.N.1.002	AMI#I.3	r.N.2.003
ABATTEMENT(2)	r.N.4.001	AMI#II	r.N.2.004
ABCÈS#I	r.N.2.012	ANCRE	r.N.1.004
ABCÈS#II	r.N.2.006	ANGOISSE	r.N.2.016
ABDICATION#I	r.N.2.010	ANIMOSITÉ	r.N.2.020
ABDICATION#II	r.N.2.013	ANNIVERSAIRE#1	r.N.4.003
ABEILLE	—	ANNIVERSAIRE#2	r.N.2.005
ABÎME#I	r.N.1.001	APPLAUDISSEMENTS	r.N.3.006
ABÎME#II.1	r.N.1.001	APPRÉHENSION	r.N.2.016
ABÎME#II.2	r.N.2.015	ARMOIRE	r.N.2.017
ABÎME#II.3	r.N.2.006	ASSASSINAT	r.N.2.018
ABOIEMENT#I	r.N.1.002	ASSENTIMENT	r.N.3.007
ABOIEMENT#II	r.N.3.002	ASSIETTE#a	r.N.2.017
ACCLAMATION	r.N.2.010	ASSIETTE#b	r.N.1.005

AUTOPSIE	r.N.3.008	BILLET#I.2	r.N.4.005
AVERSION	r.N.2.020	BILLET#I.3	r.N.2.028
AVERTISSEMENT#I.1	r.N.3.009	BILLET#II	r.N.2.029
AVERTISSEMENT#I.2a	r.N.4.004	BILLET#III.1	r.N.2.030
AVERTISSEMENT#I.2b	r.N.3.010	BILLET#III.2	r.N.3.019
AVERTISSEMENT#II	r.N.3.011	BOÎTE AUX LETTRES	r.N.3.017
AVOCAT(1)#I	r.N.3.012	BON SENS	r.N.1.002
AVOCAT(1)#II	r.N.3.013	BONNE HUMEUR	r.N.2.001
AVOCAT(2)	—	BOUILLIE#I.1	r.N.1.008
BAGAGE#I.1	r.N.2.021	BOUILLIE#I.2	r.N.1.008
BAGAGE#I.2	r.N.2.021	BOUILLIE#II	r.N.1.001
BAGAGE#II	r.N.3.014	BOUILLIE#III	r.N.1.009
BAGARRE#a	r.N.3.015	BOURGEON	r.N.1.004
BAGARRE#b	r.N.3.016	BUVEUR#a	r.N.2.031
BAGUE#1	r.N.2.021	BUVEUR#b	r.N.2.032
BAGUE#2	r.N.2.003	CAMÉLÉON#I	—
BALAI	r.N.3.017	CAMÉLÉON#II.a	r.N.1.001
BANC(1)#I	r.N.1.004	CAMÉLÉON#II.b	—
BANC(1)#II	r.N.1.006	CANARD#I.a	—
BANC(2)	r.N.1.007	CANARD#I.b	r.N.1.001
BARBE	r.N.1.004	CANARD#II	r.N.2.033
BATAILLE#I.1	r.N.2.023	CANARD#III	r.N.2.021
BATAILLE#I.2	r.N.3.015	CATASTROPHE#I	r.N.1.001
BATAILLE#II.1	r.N.3.018	CATASTROPHE#II.1	r.N.1.001
BATAILLE#II.2	r.N.2.024	CATASTROPHE#II.2	r.N.1.001
BATEAU	r.N.2.025	CEINTURE#I.1	r.N.2.054
BÉTAIL	r.N.2.026	CEINTURE#I.2	r.N.2.021
BIJOU#I	r.N.1.004	CEINTURE#II	r.N.2.035
BIJOU#II	r.N.2.027	CÉLIBAT	r.N.1.002
BILLET#I.1	r.N.2.028	CÉLIBATAIRE	r.N.1.001

CERISE	—	COMBAT#I.3b	r.N.2.023
CHAT	r.N.1.004	COMBAT#II.1	r.N.2.023
CHIEN#I	r.N.1.004	COMBAT#II.2	r.N.2.037
CHIEN#II	r.N.2.001	COMMÉRAGE#a	r.N.4.010
CHIEN#III	r.N.1.004	COMMÉRAGE#b	r.N.3.025
CHIFFON#a	r.N.2.017	COMPLIMENT#a	r.N.3.026
CHIFFON#b	r.N.1.001	COMPLIMENT#b	r.N.3.027
CIGARETTE#1	r.N.1.004	COMPRIMÉ	r.N.3.028
CIGARETTE#2	r.N.1.001	COMPTOIR#I.1a	r.N.4.011
CIMENT#I	r.N.1.001	COMPTOIR#I.1b	r.N.3.029
CIMENT#II	r.N.2.036	COMPTOIR#I.2	r.N.3.030
CIRE	r.N.1.001	COMPTOIR#II	r.N.2.038
CLÉ#I.1	r.N.2.054	CONVERSATION#a	r.N.3.031
CLÉ#I.2	r.N.2.021	CONVERSATION#b	r.N.1.023
CLÉ#I.3	r.N.1.004	CORPS À CORPS	r.N.2.023
CLÉ#II.1	r.N.3.020	COUP DE BALAI#I	r.N.2.039
CLÉ#II.2	—	COUP DE BALAI#II	r.N.2.040
CLÉ#III	r.N.1.010	COUP DE FOUDRE#1	r.N.2.041
CLINIQUE#1	r.N.3.021	COUP DE FOUDRE#2	r.N.2.042
CLINIQUE#2	r.N.4.007	COUP DE SOLEIL	r.N.2.038
CLOCHE#I	r.N.4.008	COUTEAU	r.N.2.021
CLOCHE#II	—	COUVERCLE	r.N.1.004
CLOU	r.N.3.022	COQ	—
COLLE#I	r.N.3.022	CRAINTE	r.N.2.043
COLLE#II	r.N.3.023	CRI#I.1	r.N.2.044
CHÈQUE	r.N.4.006	CRI#I.3	r.N.1.002
COLLE#III	r.N.4.009	CRI#I.2	r.N.3.032
COMBAT#I.1	r.N.2.023	CRI#II	r.N.3.033
COMBAT#I.2	r.N.2.023	CRIME#I.a	r.N.2.045
COMBAT#I.3a	r.N.3.024	CRIME#I.b	r.N.2.002

CRIME#II	r.N.1.011	ÉLOQUENCE	r.N.1.002
CRIMINEL	r.N.2.001	ÉMOI	r.N.2.056
CUIR	r.N.2.046	ENCOURAGEMENT#1	r.N.2.057
DÉ À COUDRE#II.b	r.N.1.010	ENCOURAGEMENT#2	r.N.3.040
DÉ À COUDRE#I	r.N.2.021	ENGUEULADE#1	r.N.3.041
DÉ À COUDRE#II.a	r.N.2.001	ENGUEULADE#2	r.N.3.034
DÉ#I.a	r.N.2.021	ENNEMI#I	r.N.2.003
DÉ#I.b	r.N.2.001	ENNEMI#II	r.N.2.004
DÉ#II.1	r.N.2.047	ENNEMI#III.a	r.N.2.001
DÉ#II.2	r.N.1.010	ENNEMI#III.b	r.N.1.001
DÉBAT#1	r.N.3.034	ENNEMI#IV	r.N.2.003
DÉBAT#2	r.N.3.034	ENTREVUE	r.N.3.034
DÉCEPTION	r.N.2.048	ENVELOPPE#I	r.N.2.003
DÉFAITE	r.N.3.035	ENVELOPPE#II	r.N.3.042
DÉGOÛT	r.N.2.049	ENVELOPPE#III	r.N.4.013
DÉPIT	r.N.2.050	ÉPÉE#a	r.N.3.017
DÉPOSITION	r.N.3.036	ÉPÉE#b	r.N.1.001
DÉSIR#1	r.N.2.051	ÉPINGLE	r.N.3.022
DÉSIR#2	r.N.2.052	ÉQUIPAGE	r.N.2.003
DIALOGUE#1a	r.N.2.053	ÉTAGE#1	r.N.2.058
DIALOGUE#1b	r.N.2.054	ÉTAGE#2	r.N.1.004
DIALOGUE#2	r.N.3.034	ÉVIER	r.N.2.001
DIRECTIVE	r.N.3.037	EXAGÉRATION#1a	r.N.2.059
DRAP	r.N.1.004	EXAGÉRATION#1b	r.N.1.002
DRAPEAU	r.N.3.038	EXAGÉRATION#2	r.N.2.060
EFFROI	r.N.2.055	EXCLAMATION	r.N.3.043
ÉLÉPHANT#I	—	EXCUSE#I	r.N.3.044
ÉLÉPHANT#II	r.N.1.001	EXCUSE#II	r.N.3.045
ÉLOGE#1	r.N.4.012	EXTASE#1	r.N.1.002
ÉLOGE#2	r.N.3.039	EXTASE#2	r.N.2.061

FACTEUR(1)	r.N.3.047	FIERTÉ#II	r.N.2.070
FACTEUR(2)	r.N.2.015	FLÈCHE#I	r.N.2.021
FACTEUR(3)#I	r.N.2.062	FLÈCHE#II	r.N.3.051
FACTEUR(3)#II.1	r.N.2.063	FLÈCHE#III	r.N.2.001
FACTEUR(3)#II.2	r.N.1.012	FLEUVE#I	r.N.3.022
FACTURE(1)#I	r.N.4.014	FLEUVE#II	r.N.1.014
FACTURE(1)#II	r.N.2.064	FLEUVE#III	—
FACTURE(2)#a	r.N.2.003	FORÊT#I	r.N.1.015
FACTURE(2)#b	r.N.1.003	FORÊT#II	r.N.1.007
FAILLITE#I	r.N.1.013	FORTUNE(1)#1a	r.N.2.071
FAILLITE#II	r.N.1.002	FORTUNE(1)#1b	r.N.3.052
FANTAISIE#1	r.N.1.002	FORTUNE(1)#2	—
FANTAISIE#2	r.N.2.065	FORTUNE(2)#I	r.N.1.016
FARDEAU#I	r.N.1.004	FORTUNE(2)#II.1	r.N.2.072
FARDEAU#II	r.N.2.066	FORTUNE(2)#II.2	r.N.1.002
FARINE	r.N.1.010	FOUR#a	r.N.2.017
FAUTE#I.1a	r.N.2.002	FOUR#b	r.N.2.001
FAUTE#I.1b	r.N.3.048	FOURREAU	r.N.1.004
FAUTE#I.2	r.N.2.039	FRÉMISSEMENT#I.1	r.N.1.002
FAUTE#I.3	r.N.2.067	FRÉMISSEMENT#I.2	r.N.2.070
FAUTE#II	r.N.2.154	FRÉMISSEMENT#II	r.N.1.017
FAUTEUIL#I	r.N.1.004	FRÉMISSEMENT#III	r.N.1.002
FAUTEUIL#II	r.N.2.022	FRIGIDAIRE	r.N.2.021
FAVEUR#1a	r.N.2.002	FRIGO	r.N.2.021
FAVEUR#1b	r.N.2.002	FRISSON#1	r.N.3.053
FAVEUR#2	r.N.3.049	FRISSON#2	r.N.1.018
FÉLICITATIONS#a	r.N.3.026	FRONTIÈRE#I.1	r.N.4.015
FÉLICITATIONS#b	r.N.3.050	FRONTIÈRE#II.1	r.N.2.073
FIERTÉ#I.1	r.N.2.068	FRONTIÈRE#I.2	r.N.1.019
FIERTÉ#I.2	r.N.2.069	FRONTIÈRE#II.2	r.N.1.019

FUNÉRAILLES	r.N.2.072	GUICHET#1	r.N.2.054
FUSÉE#1	r.N.1.001	GUICHET#2	r.N.2.021
FUSÉE#2a	r.N.1.001	GUIRLANDE	r.N.2.080
FUSÉE#2b	r.N.1.001	HACHE	r.N.2.021
FUSIL(1)#a	r.N.3.017	HAINÉ	r.N.2.041
FUSIL(1)#b	r.N.1.001	HÉCATOMBE#I	r.N.2.081
FUSIL(2)	r.N.2.021	HÉCATOMBE#II	r.N.2.081
GALOP	r.N.1.002	HÉSITATION#1	r.N.2.082
GARE#a	r.N.2.003	HÉSITATION#2	r.N.1.002
GARE#b	r.N.1.020	HILARITÉ	r.N.2.083
GÂTEAU	r.N.1.021	HOMICIDE-n	r.N.2.084
GÉMISSEMENT#I	r.N.2.044	HONTE#I.1	r.N.3.058
GÉMISSEMENT#II	r.N.2.074	HONTE#I.2	r.N.2.085
GÉMISSEMENT#III	r.N.1.002	HONTE#II	r.N.1.011
GIFLE#I	r.N.3.054	HORLOGE	r.N.2.001
GIFLE#II	r.N.2.075	HOSTILITÉ#I	r.N.2.041
GOÛT#I.1	r.N.3.055	HOSTILITÉ#II	r.N.2.091
GOÛT#I.2	r.N.1.002	HOSTILITÉ#III	r.N.2.091
GOÛT#II.1	r.N.2.076	HUMEUR#1a	r.N.2.096
GOÛT#II.2	r.N.2.015	HUMEUR#1b	r.N.2.097
GOÛT#III	r.N.2.077	HUMEUR#2	r.N.2.098
GOÛT#IV	r.N.2.078	HURLEMENT#I.1	r.N.2.070
GRATITUDE	r.N.3.056	HURLEMENT#I.2	r.N.1.002
GRENOUILLE	—	HURLEMENT#II	r.N.1.002
GRIPPE	r.N.1.002	HURLEMENT#III	r.N.1.022
GROGNEMENT#I	r.N.1.002	IMAGINATION#1	r.N.2.002
GROGNEMENT#II.1	r.N.2.044	IMAGINATION#2	r.N.2.002
GROGNEMENT#II.2	r.N.2.079	IMPARTIALITÉ#a	r.N.2.099
GUÉRISON#a	r.N.2.067	IMPARTIALITÉ#b	r.N.1.023
GUÉRISON#b	r.N.3.057	INDICATION#I.a	r.N.3.059

INDICATION#I.b	r.N.3.060	LUTTE#II	r.N.2.107
INDICATION#II	r.N.3.061	MAISON#I.1	r.N.1.004
INDICATION#III	r.N.2.100	MAISON#I.2	r.N.1.004
INDICE(1)	r.N.3.062	MAISON#II.1	r.N.1.025
INDICE(2)#I	r.N.2.101	MAISON#II.2	r.N.4.016
INDICE(2)#II	r.N.2.102	MAISON#III	r.N.3.068
INDIGNATION	r.N.2.103	MAJORITÉ#I.1	r.N.2.162
INTERROGATOIRE	r.N.3.063	MAJORITÉ#I.2	r.N.2.001
INSOLENCE#1	r.N.2.099	MAJORITÉ#I.3	r.N.1.001
INSOLENCE#2	r.N.2.104	MAJORITÉ#II	r.N.1.002
IRRITATION#I	r.N.2.105	MARTEAU#1	r.N.2.021
IRRITATION#II.1	r.N.3.064	MARTEAU#2	r.N.2.021
IRRITATION#II.2	r.N.3.065	MÉCONTENTEMENT	r.N.2.108
JAMBON	r.N.1.001	MÉDICAMENT	r.N.3.069
JUPE	r.N.1.004	MEURTRE	r.N.2.109
JURON#1	r.N.1.002	MICROSCOPE	r.N.2.021
JURON#2	r.N.2.106	MIETTE#I	r.N.1.024
KLAXON#a	r.N.2.054	MIETTE#II.1	r.N.1.035
KLAXON#b	r.N.1.003	MIETTE#II.2	r.N.2.110
LABEUR	r.N.1.002	MINISTÈRE#I.1	r.N.4.017
LÂCHETÉ#2	r.N.2.002	MINISTÈRE#I.2	r.N.2.111
LAIT#a	r.N.2.102	MINISTÈRE#I.3	r.N.1.027
LÂCHETÉ#1	r.N.1.002	MINISTÈRE#II	r.N.1.002
LAIT#b	r.N.1.010	MINISTRE	r.N.3.070
LEVÉE DE BOUCLERS	r.N.3.066	MINISTRE DU CULTE	r.N.2.112
LION#I	—	MONTRE	r.N.2.021
LION#II	r.N.1.001	MORCEAU#I.1a	r.N.1.026
LOUP	—	MORCEAU#I.1b	r.N.1.010
LUTTE#I.1	r.N.3.067	MORCEAU#I.2	r.N.1.014
LUTTE#I.2	r.N.2.001	MORCEAU#I.3	r.N.1.022

MORCEAU#II	r.N.1.001	PARACHUTE	r.N.1.004
MORCEAU#III.1	r.N.2.113	PARAPLUIE	r.N.1.004
MORCEAU#III.2	r.N.2.021	PARTIALITÉ#a	r.N.2.116
MORT(1)#I.1	r.N.1.002	PARTIALITÉ#b	r.N.1.023
MORT(1)#I.2	r.N.1.028	PARTIE(1)#I	r.N.2.003
MORT(1)#II	r.N.1.023	PARTIE(1)#II	r.N.2.117
MORT(2)	—	PARTIE(1)#III.1	r.N.2.118
MOUCHE#I	—	PARTIE(1)#III.2	r.N.2.119
MOUCHE#II	r.N.2.021	PARTIE(2)	r.N.3.024
MOUSTIQUE	—	PARTIES GÉNITALES	r.N.1.004
MOUTON#I.a	r.N.1.004	PATIENT	r.N.2.003
MOUTON#I.b	r.N.1.001	PAUPIÈRE	r.N.1.004
MOUTON#I.c	r.N.1.001	PAUVRETÉ#I.1	r.N.1.002
MOUTON#II	r.N.1.001	PAUVRETÉ#I.2	r.N.1.002
MOUTON#III	—	PAUVRETÉ#II.1	r.N.2.120
MUSÉE	r.N.2.001	PAUVRETÉ#II.2	r.N.1.003
MUTISME#a	r.N.1.002	PÉCHÉ	r.N.2.002
MUTISME#b	r.N.2.114	PÉCHEUR	r.N.2.001
NAISSANCE#I.1	r.N.2.002	PÉCULE	r.N.1.004
NAISSANCE#I.2	r.N.1.002	PELLE(3)	r.N.1.002
NAISSANCE#II	r.N.1.002	PEINE#I.1	r.N.2.121
NAVIRE	r.N.2.025	PEINE#I.2a	r.N.3.048
OBSÈQUES	r.N.2.072	PEINE#I.2b	r.N.2.122
ONGLE	r.N.1.004	PEINE#II	r.N.4.019
ORDONNANCE#I	r.N.3.071	PELLE(1)	r.N.2.021
ORDONNANCE#II	r.N.4.018	PELLE(2)	r.N.2.001
OUTIL#I	r.N.2.115	PETIT AMI	r.N.2.003
OUTIL#II	r.N.3.072	PILULE#a	r.N.2.123
PAIN#I	r.N.1.010	PILULE#b	r.N.1.002
PAIN#II	r.N.1.010	PIPE#I	r.N.2.021

PIPE#II	r.N.2.001	PLOMBAGE#a	r.N.3.076
PIRATE#1	r.N.3.073	PLOMBAGE#b	r.N.3.077
PIRATE#2	—	PLUMAGE#1	r.N.1.004
PISTE#I.1	r.N.2.021	PLUMAGE#2	r.N.2.095
PISTE#I.2	r.N.3.074	PLUME#I	r.N.2.134
PISTE#I.3a	r.N.2.046	PLUME#II	r.N.1.004
PISTE#I.3b	r.N.2.046	PLUME#III.a	r.N.2.002
PISTE#I.4	r.N.1.001	PLUME#III.b	—
PISTE#II	r.N.2.021	PNEU	r.N.1.004
PLAIE#I	r.N.2.124	POIGNARD	r.N.3.017
PLAIE#II.1	r.N.2.002	POIGNÉE DE MAIN	r.N.2.053
PLAIE#II.2	r.N.2.125	POIGNÉE#I	r.N.1.004
PLAINE	—	POIGNÉE#II.a	r.N.1.010
PLAISANTERIE	r.N.3.075	POIGNÉE#II.b	r.N.1.010
PLAN(1)#I.1	—	POISSON#I.a	—
PLAN(1)#I.2	r.N.1.001	POISSON#I.b	—
PLAN(1)#II.1	r.N.2.126	POISSON#II	r.N.1.001
PLAN(1)#II.2	r.N.3.042	POMMADE	r.N.3.078
PLAN(2)#I	r.N.4.020	POULE#I	r.N.2.021
PLAN(2)#II	r.N.2.127	POULE#II.a	r.N.2.003
PLAN(2)#III.1	r.N.2.028	POULE#II.b	r.N.2.001
PLAN(2)#III.2	r.N.2.065	POULE#III	r.N.2.130
PLAQUE DENTAIRE	r.N.1.001	POULS#I	r.N.1.002
PLAQUE#I.1	r.N.1.014	POULS#II	r.N.1.023
PLAQUE#I.2	r.N.2.132	PRISE DE SANG	r.N.2.136
PLAQUE#I.3	—	QUITTANCE	r.N.3.079
PLAQUE#II.1	r.N.4.021	QUOLIBET	r.N.2.137
PLAQUE#II.2	r.N.2.021	RANCUNE	r.N.3.056
PLAQUE#III	r.N.1.001	RASOIR	r.N.3.017
PLEURS	r.N.2.133	RÉFRIGÉRATEUR	r.N.2.021

REGRET#I.1	r.N.2.138	RUBAN#II	r.N.1.010
REGRET#I.2	r.N.2.139	RUGISSEMENT#I	r.N.1.002
REGRET#II.1	r.N.2.140	RUGISSEMENT#II	r.N.2.070
REGRET#II.2	r.N.2.141	RUGISSEMENT#III	r.N.1.002
REMÈDE#I	r.N.3.069	RUISSEAU	—
REMÈDE#II	r.N.3.080	RUMEUR#I.1	r.N.3.084
REMONTRANCE	r.N.3.026	RUMEUR#I.2	r.N.2.148
REMORDS	r.N.2.142	RUMEUR#II	r.N.1.002
REPENTIR	r.N.2.143	SANDWICH	r.N.3.085
RÉPUGNANCE#1	r.N.2.144	SEMELLE#1	r.N.1.004
RÉPUGNANCE#2	r.N.2.145	SEMELLE#2	r.N.1.004
RÉPULSION#I	r.N.2.086	SENSIBILITÉ#I.1	r.N.1.002
RÉPULSION#II	r.N.2.146	SENSIBILITÉ#I.2	r.N.2.149
RESSENTIMENT	r.N.3.056	SENSIBILITÉ#II	r.N.2.150
RESTAURANT	r.N.3.081	SENSIBILITÉ#III	r.N.2.151
RETRAITE#I.1	r.N.3.082	SERPENT	—
RETRAITE#I.2	r.N.2.039	SERRURE	r.N.2.021
RETRAITE#I.3	r.N.1.004	SÉVÉRITÉ#I.1a	r.N.2.152
RETRAITE#II.1	r.N.1.002	SÉVÉRITÉ#I.1b	r.N.1.023
RETRAITE#II.2	r.N.3.083	SÉVÉRITÉ#I.2	r.N.1.023
REZ-DE-CHAUSSÉE	r.N.2.021	SÉVÉRITÉ#II	r.N.1.002
RICHESSÉ#I.1	r.N.1.002	SIFFLET#1	r.N.1.004
RICHESSÉ#I.2	r.N.1.004	SIFFLET#2	r.N.1.002
RICHESSÉ#II	r.N.1.004	SIFFLET#3	r.N.2.002
RICHESSÉ#III.1	r.N.2.120	SIGNE#I.1	r.N.3.086
RICHESSÉ#III.2	r.N.1.003	SIGNE#I.2	r.N.1.001
RIVIÈRE#I	—	SIGNE#II	r.N.3.087
RIVIÈRE#II	r.N.1.014	SIGNE#III	r.N.1.002
ROBINET	r.N.2.046	SILENCE#I.1	r.N.1.002
RUBAN#I	r.N.2.147	SILENCE#I.2a	r.N.1.002

SILENCE#I.2b	r.N.2.114	TERRE#I.1	—
SILENCE#II	r.N.1.001	TERRE#I.2a	r.N.2.017
SONNETTE#a	r.N.1.001	TERRE#I.2b	r.N.1.030
SONNETTE#b	r.N.1.001	TERRE#II.1	—
SOUPÇON#I	r.N.3.088	TERRE#II.2	—
SOUPÇON#II	r.N.1.010	TERRE#III	—
SOUPIR#II	r.N.1.002	TERRITOIRE#1a	r.N.1.030
SOUPIR#III	r.N.1.001	TERRITOIRE#1b	r.N.1.004
SOUPIR#I	r.N.2.044	TERRITOIRE#2	r.N.2.156
STRATAGÈME	r.N.3.089	TESTAMENT#I	r.N.3.094
SUCRE#a	r.N.2.153	TESTAMENT#II	r.N.2.157
SUCRE#b	r.N.2.021	TESTICULE	r.N.1.004
SUICIDE	r.N.1.002	TÊTE-À-QUEUE	r.N.1.002
TACT	r.N.1.002	TEXTE	r.N.3.095
TACTIQUE-n#a	r.N.3.089	THÉÂTRE#I.1	r.N.2.001
TACTIQUE-n#b	r.N.3.089	THÉÂTRE#I.2	r.N.1.004
TAMBOUR#I.1	r.N.1.004	THÉÂTRE#I.3	r.N.3.017
TAMBOUR#I.2	r.N.1.001	THÉÂTRE#II	r.N.2.154
TAMBOUR#II	r.N.1.029	THÉORIE(1)#1a	r.N.3.096
TÉMOIGNAGE#I.1	r.N.3.090	THÉORIE(1)#1b	r.N.3.097
TÉMOIGNAGE#I.2	r.N.3.091	THÉORIE(1)#2	r.N.1.001
TÉMOIGNAGE#II.1	r.N.2.154	THÉORIE(2)	r.N.1.031
TÉMOIGNAGE#II.2	r.N.3.092	THERMOMÈTRE#I.1	r.N.2.001
TÉMOIGNAGE#II.3	r.N.4.022	THERMOMÈTRE#I.2	r.N.1.016
TÉMOIN#I.1	r.N.3.093	THERMOMÈTRE#II	r.N.2.154
TÉMOIN#I.2	r.N.2.155	THÈSE#1	r.N.3.097
TÉMOIN#II.1	r.N.2.046	THÈSE#2	r.N.4.023
TÉMOIN#II.2	—	TICKET#I	r.N.2.022
TÉMOIN#II.3	—	TICKET#II	r.N.2.131
TÉMOIN#III	r.N.3.042	TICKET#III	r.N.4.024

TIGRE	—	USINE	r.N.3.101
TIGRESSE#I	—	VACHE	r.N.2.021
TIGRESSE#II	r.N.1.001	VAINQUEUR	r.N.3.102
TIMBRE(1)#1	r.N.3.098	VAISSELLE#1	r.N.2.001
TIMBRE(1)#2	r.N.2.158	VAISSELLE#2	r.N.2.089
TIMBRE(2)#I	r.N.1.002	VAPEUR#a	r.N.1.001
TIMBRE(2)#II	r.N.1.001	VAPEUR#b	r.N.1.032
TIR#1a	r.N.3.099	VERROU#I	r.N.2.003
TIR#1b	r.N.2.159	VERROU#II	r.N.2.001
TIR#2a	r.N.2.095	VERVE	r.N.1.002
TIR#2b	r.N.2.095	VEXATION#I	r.N.2.129
TIR#3	r.N.2.160	VEXATION#II	r.N.2.128
TIRAGE(1)#1	r.N.3.100	VICTOIRE#I	r.N.3.103
TIRAGE(1)#2a	r.N.2.095	VICTOIRE#II	r.N.2.088
TIRAGE(1)#2b	—	VIE#I.1	r.N.1.033
TIRAGE(2)	r.N.2.094	VIE#I.2	r.N.1.002
TIRAGE(3)	r.N.1.002	VIE#I.3	—
TISANE#a	r.N.2.022	VIE#II	r.N.1.033
TISANE#b	r.N.2.001	VOLTE-FACE#I	r.N.2.002
TONNEAU#I	r.N.2.092	VOLTE-FACE#II	r.N.2.087
TONNEAU#II	r.N.2.001	WAGON	r.N.2.022
TONNEAU#III	r.N.1.018	XÉNOPHOBIE	r.N.2.086
TORCHON#I	r.N.2.092	YAOURT#a	—
TORCHON#II	r.N.1.001	YAOURT#b	—
TUYAU#I.1	r.N.3.104	YOGA	r.N.1.001
TUYAU#I.2	—	ZIGZAG	r.N.1.002
TUYAU#II	r.N.4.025	ZÈLE	r.N.2.090
ULTIMATUM	r.N.4.026	ZONE	r.N.1.034
UNANIMITÉ#1	r.N.2.093		
UNANIMITÉ#2	r.N.3.022		

## Annexe II

### Les sous-groupes de la partie du discours Prép.

#### Recensement des informations que l'on peut trouver dans le DiCo :

- Prép-envers (pour 12 lexies) = {*à cause de, à l'égard de, à l'encontre de, à l'endroit de, à propos de, au sujet de, contre, devant, envers, face à, pour, vis-à-vis de*}
- Prép-devant (pour 2 lexies) = {*à cause de, à l'égard de, à la vue de, au sujet de, devant, en face de, face à*}
- Prép-pour (pour 3 lexies) = {*à la vue de, à propos de, au sujet de, en face de, face à, pour*}
- Prép-cause (pour 1 lexie) = {*à cause de, à l'égard de, au sujet de, contre, devant, envers, face à, vis-à-vis de*}
- Prép-percep (pour 1 lexie) = aucune liste
- Prép-propos (pour 35 lexies) = aucune liste

Pour nous, le nom Prép-pour n'est pas très révélateur. Nous l'avons renommé soit Prép-propos, soit Prép-devant. Comme Prép-percep n'était présent qu'une seule fois et ce sans liste qui nous aurait permis de bien saisir quel sens ce groupe de prépositions exprime, nous l'avons remplacé par Prép-devant. Nous avons créé le groupe Prép-datif. Dans certains groupes, nous avons également ajouté quelques prépositions qui nous semblaient exprimer le même sens. Voici donc une liste temporaire non exhaustive des sous-groupes de la partie du discours Prép.

#### Liste après ajouts et modifications :

- Prép-envers = {*à cause de, à l'adresse de, à l'égard de, à l'encontre de, à l'endroit de, à propos de, au sujet de, contre, devant, envers, face à, pour, vis-à-vis de*}
- Prép-devant = {*à cause de, à l'égard de, à la vue de, au sujet de, devant, en face*

*de, face à, vis-à-vis de*

- Prép-cause = { *à cause de, à l'égard de, à la suite de, au sujet de, contre, devant, du fait de, en raison de, envers, face à, vis-à-vis de* }
- Prép-datif = { *à, à l'adresse de, à l'attention de, à l'endroit de, à l'égard de, à l'intention de, en direction de, pour* }
- Prép-propos = { *à propos de, au sujet de, pour* }

# Index

actant

- ~ sémantique, 43
- ~ syntaxique de surface, 47
- ~ syntaxique profond, 46

dépendance

- ~ sémantique, 40, 44
- ~ syntaxique de surface, 41
  - ~ régie ~ non régie ~ mixte, 75
- ~ syntaxique profonde, 41

diathèse, 54

étiquette sémantique, 49

forme propositionnelle, 45

lexie, 48

nom sémantique, 44

prédicat sémantique, 43

quasi-prédicat sémantique, 44

valence

- ~ active, 52
- ~ passive, 5, 51
- ~ sémantique, 49