

Université Paris 7 - Denis Diderot

Ecole doctorale Sciences du Langage
UFR de Linguistique

Une métagrammaire pour les noms prédicatifs du français

Développement et expérimentations pour les grammaires TAG

THÈSE

présentée et soutenue publiquement le 30 Mai 2006

pour le doctorat en linguistique

par

Sébastien BARRIER

Jury

Anne ABEILLE (directrice)
Eric LAPORTE
Jean-Marie PIERREL (rapporteur)
Owen RAMBOW (rapporteur)
Pollet SAMVELIAN
Robert VIVES

Laboratoire de Linguistique Formelle - UMR 7110

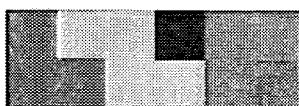


Table des matières

Introduction	i
I Etat de l'art et premiers choix	1
1 Etat de l'art	3
1.1 Point de vue lexical	5
1.1.1 Les locutions verbales	5
1.1.2 Les collocations	6
1.2 Point de vue sémantique	10
1.2.1 Sémantique du verbe support	10
1.2.2 Le lexique-grammaire des verbes supports	12
1.2.3 Sémantique du nom prédicatif	14
1.3 Point de vue syntaxique	18
1.3.1 Propriétés syntaxiques	18
1.3.2 Représentations syntaxiques	22
2 La notion de métagrammaire	37
2.1 Aperçu général	38
2.1.1 Motivations	38
2.1.2 La proposition de Vijay-Shanker et Schabes	40
2.1.3 Compilateur et hiérarchie	43
2.2 Le compilateur de Candito	53

Remerciements

Voici enfin venu le temps des remerciements. Lorsque je rédigeais ce mémoire, je pensais au moment où, la tâche de rédaction accomplie, j'adresserais mes remerciements aux personnes qui de près ou de loin m'avaient encouragé dans mes études.

J'éviterai ici les poncifs qui signalent que les moments de rédaction sont des passages parfois difficiles à vivre, pour ne retenir que les instants les plus stimulants.

Avant tout, je tiens à remercier les membres du jury, Eric Laporte, Jean-Marie Pierrel, Owen Rambow, Pollet Samvelian et Robert Vivès d'avoir accepté d'honorer de leur jugement ce travail.

Bien entendu, je remercie également Anne Abeillé, ma directrice, qui tout au long de la rédaction de mon travail m'a donné de précieux conseils.

Je souhaiterais également remercier les professeurs qui ont su me donner goût à la linguistique, lorsque jeune bachelier, j'ai été accepté à l'Université Paris 7, pour y suivre des cours de mathématiques appliquées et de sciences sociales. Je pense notamment à Michel Launay, Philippe Ségéral et Marie-Josèphe Gouesse. C'est avec cette dernière que j'ai d'ailleurs pu faire mes premières armes en tant que moniteur allocataire. Les remarques et principes qu'elle m'a inculqués me guideront toute ma vie d'enseignant.

Je remercie sincèrement Jean Lowenstamm, responsable de l'école doctorale des sciences du langage, pour son soutien et son écoute.

Merci beaucoup également à mon frère, Nicolas, qui m'a permis de réaliser l'implémentation de ma grammaire. C'est grâce à lui que la Métagrammaire de Marie-Hélène Candito a pu être utilisable sur des machines standards. Sa connaissance du compilateur m'a permis un gain de temps précieux.

Enfin, je remercie les divers participants de l'ARC RLT, notamment les membres du Loria (langue et dialogue), Bertrand Gaiffe, Azim Roussanally et Benoît Crabbé. Les discussions que j'ai pu avoir ce dernier m'ont d'ailleurs conforté dans l'idée que même si nous utilisons des compilateurs de métagrammaire différents, les problèmes que nous pouvions rencontrer étaient sensiblement les mêmes.

D'un point de vue plus personnel, je souhaiterais également remercier Michèle Lasek,

désormais jeune retraitée, pour m'avoir aidé dans les démarches administratives. Sa sympathie et sa bonne humeur m'ont toujours été agréables.

Pour finir, je souhaiterais remercier mes premiers relecteurs : ma maman tout d'abord, mais également François Monnanteuil.

Sans le soutien de ma famille et de mes amis, je n'aurais pu mener ce travail à terme. Merci à tous pour votre aide et les espoirs que vous avez placés en moi.

2.2.1	Langage de description	53
2.2.2	Calcul des arbres	54
2.2.3	Comportements déviants	55
2.2.4	Des exemples concrets	58
2.2.5	Format de sortie du compilateur	61
2.3	Problèmes de représentation	62
2.3.1	Problèmes de dominance	62
2.3.2	Problèmes d'ordre entre constituants	65
3	Choix de représentation syntaxique	73
3.1	Les constructions à verbe support	73
3.1.1	Rappel des propriétés définitoires	75
3.1.2	Les CVS : des constructions syntaxiquement régulières	79
3.1.3	La détermination du nom prédicatif dans les CVS	84
3.1.4	Les relatives de quantité	87
3.1.5	Mobilité des compléments et double analyse	88
3.2	Les groupes nominaux complexes	93
3.2.1	Les groupes nominaux en <i>par</i>	94
3.2.2	Les compléments en <i>de</i> des groupes nominaux complexes	97
3.3	Premiers choix de représentation pour les TAG	112
3.3.1	Des familles particulières et des entrées lexicales nouvelles	112
3.3.2	Représentation des CVS	117
3.3.3	Représentation des groupes nominaux complexes	122
II	Une métagrammaire pour les noms prédicatifs du français	127
4	Dimension 1 : la sous-catégorisation initiale	129
4.1	Quelles fonctions utiliser ?	130
4.1.1	Sous-catégorisation sémantique et sous-catégorisation syntaxique	130

4.1.2	Quelles fonctions syntaxiques retenir ?	131
4.1.3	Les fonctions retenues : exemples et remarques	140
4.2	Liste des familles à nom prédicatif	141
4.2.1	Conventions de nommage	141
4.2.2	Familles à arguments nominaux	143
4.2.3	Familles à argument phrastique	144
4.3	Implémentation	146
4.3.1	Limitations dûes au compilateur	146
4.3.2	Organisation hiérarchique	148
4.3.3	Vision graphique de la hiérarchie	150
4.3.4	Principales différences avec l'implémentation de Candito	152
5	Dimension 2 : les redistributions fonctionnelles	155
5.1	Généralités sur la dimension 2	156
5.1.1	La notion d'artère	156
5.1.2	Absence de syntagme verbal	157
5.2	Les différents changements de diathèses et redistributions utilisés	161
5.2.1	L'actif "de base"	161
5.2.2	Le passif	162
5.2.3	Les constructions causatives	166
5.2.4	L'impersonnel	170
5.2.5	Le passif impersonnel	171
5.2.6	Les constructions réfléchies	172
5.2.7	Les constructions moyennes	173
5.2.8	La représentation des groupes nominaux	174
5.3	Implémentation	177
5.3.1	La description des morphologies verbale et nominale	177
5.3.2	Passage des fonctions initiales aux fonctions finales	179
5.3.3	Les différentes redistributions	181

5.3.4	Remarques concernant la dimension 2	190
5.3.5	Vision graphique de la hiérarchie	190
5.3.6	Principales différences avec l'implémentation de Candito	193
6	Dimensions 3 et 4 : la réalisation et l'ordonnement des fonctions syntaxiques	195
6.1	Les réalisations canoniques	196
6.1.1	Le sujet	196
6.1.2	L'objet direct	199
6.1.3	Les a_objets, de_objets et prep_objets	199
6.1.4	Les compléments locatifs	201
6.1.5	L'attribut	202
6.2	Les réalisations non canoniques	205
6.2.1	Le sujet	205
6.2.2	Les clitiques	210
6.2.3	Les extractions	214
6.2.4	Accord du participe passé	221
6.3	Quelques remarques et exemples d'implémentation	223
6.3.1	La représentation des extractions	223
6.3.2	La représentation des clitiques	226
6.3.3	"Descente" du nom prédicatif	230
6.3.4	La représentation des syntagmes prépositionnels	231
6.3.5	Récapitulation générale : vision hiérarchisée	231
6.4	La dimension 4 : l'ordre des arguments réalisés	242
III	Expériences en corpus	245
7	Expériences sur corpus et codage lexical	247
7.1	Le corpus de Paris 7	248
7.2	Préférences lexicales	250

7.2.1	Recherches sur le verbe <i>faire</i>	250
7.2.2	Recherches sur le verbe <i>avoir</i>	253
7.2.3	Recherches sur le verbe <i>commettre</i>	255
7.2.4	Recherches sur le verbe <i>procéder</i>	257
7.2.5	Recherches sur le verbe <i>courir</i>	258
7.3	Extraction de noms prédicatifs	260
7.3.1	Les noms à 2 compléments	260
7.3.2	Autres noms prédicatifs	263
7.4	Expérimentations avec le parseur d'A. Sarkar	265
7.4.1	Analyses sur des phrases test	265
7.4.2	Analyses sur phrases de corpus	273
7.4.3	Raisons des rejets ou des suranalyses	279
7.4.4	Récapitulatif	280
Conclusion		283
Références		287
Annexes		303
A Hiérarchie exemple des familles à nom prédicatif		305
B Implémentation des problèmes d'ordre		323
B.1	La solution de K. Gerdes	323
B.2	Notre proposition	326
B.3	Les arbres générés	329
B.3.1	Famille n0Vn1	329
B.3.2	Famille n0Vn1(pn2)	329
B.3.3	Famille n0Vn1(loc2)	330
C Extrait des lexiques syntaxiques		331

D	Catégories et traits utilisés par la grammaire TAG du français	337
D.1	Catégories utilisées	337
D.2	Traits utilisés	338

Introduction

Le but général de cette thèse est l'étude et la représentation à l'aide d'une métagrammaire des combinaisons "verbe (+ préposition) + nom" ou "nom + verbe" en français, connues sous le nom de *constructions à verbe support* (dorénavant CVS).

Ces constructions ont souvent été considérées par les grammaires traditionnelles normatives comme des "maladies stylistiques", mais dans la langue parlée usuelle, "de telles expressions sont bien souvent plus naturelles que le verbe lui-même, et offrent des possibilités d'expression parfois exclues avec le verbe simple" [Giry-Schneider, 1987]. D'un point de vue plus scientifique, les CVS ont suscité davantage d'intérêt et nombreux sont les linguistes les ayant considérées dans le cadre de diverses traditions.

Dans ce travail, nous nous intéresserons aux combinaisons évoquées, comprenant tant des verbes dits "passe-partout" (comme *faire* ou *avoir*), que des verbes considérés plus appropriés ou idiomatiques.

Le phénomène des CVS remet en cause les affirmations selon lesquelles le verbe représente toujours le prédicat sémantique de la phrase. En effet, on peut dire que le verbe support est "vide" de tout signifié lexical, et que c'est le nom qui porte le poids sémantique.

Lexicalement, c'est le nom prédicatif qui sélectionne le verbe qui lui sert de support. Cette sélection est effectuée de façon phraséologique.

Enfin, syntaxiquement la particularité principale des CVS se situe dans le fait que les arguments sémantiques du nom et le nom lui-même sont réalisés comme des arguments du verbe support. Pour résumer, le rôle de ce dernier, consiste donc à servir de support syntaxique au nom qui exprime le prédicat sémantique.

Pour illustrer la productivité des CVS, nous pouvons d'ores et déjà présenter une liste d'exemples représentatifs, qui pourrait facilement être complétée :

- *Faire un commentaire, faire une promenade, faire un voyage*
- *Avoir peur, avoir envie, avoir besoin*
- *Donner un avertissement, donner un renseignement*
- *Procéder à une enquête, procéder à un interrogatoire*
- *Etre en retard, être en colère, être à la recherche*
- *Accorder une entrevue, lancer une attaque, mener une enquête*

Pour entreprendre cette étude, nous nous plaçons dans le cadre des grammaires d'arbres adjoints [Joshi et al., 1975]. Formalisme mathématique au départ, les TAGs ont trouvé une application intéressante avec la représentation de connaissances linguistiques. Basées sur des combinaisons de structures arborescentes originales (la substitution, mais surtout l'adjonction), elles redéfinissent la notion de localité et permettent de considérer les phénomènes d'interaction entre lexique, sémantique et syntaxe. Elles s'inscrivent dans le cadre des grammaires d'unification en ce sens qu'elles s'appuient sur des représentations linguistiques déclaratives et monotones, utilisent des structures de traits combinées par unification et décrivent directement des phénomènes de surface.

Mais comme le souligne [Crabbé, 2005], les grammaires fortement lexicalisées, et donc les TAGs constituent un ensemble conséquent d'unités disparates dans lesquelles aucune généralisation n'est facilement identifiable. Ainsi, la mise en œuvre ou la maintenance de telles grammaires devient rapidement une tâche difficile et démotivante : toute modification aussi minime soit elle, peut nécessiter la correction de nombreuses règles, et il n'est jamais certain que la cohérence de la grammaire décrite soit totale.

C'est pourquoi, la description des règles que nous développerons ne sera pas uniquement réalisée en TAG, mais également en utilisant un langage déclaratif de plus haut niveau : la Métagrammaire. Initialement mise au point par [Candito, 1996, Candito, 1999a] sur une proposition de [Vijay-Shanker, 1992], la métagrammaire se révèle être un formalisme de description syntaxique à vocation multilingue¹. Elle permet une représentation relativement élégante et non redondante, dans la mesure où elle utilise un partage d'information par héritage et s'appuie sur des domaines d'information linguistiques bien définis.

Les 3 domaines d'information retenus initialement par [Candito, 1996, Candito, 1999a] sont appelés Dimension. A chaque dimension est assignée une tâche bien précise :

- la dimension 1 renseigne les sous-catégorisations initiales exprimées sous forme de fonctions et de catégories
- la dimension 2 indique les redistributions fonctionnelles reliant la sous-catégorisation initiale aux différentes sous-catégorisations finales, avec tous les changements associés tant en morphologie que pour les fonctions
- la dimension 3 réalise en surface les fonctions finales (avec toutes les variations d'ordre des mots)

Après la description des verbes par [Candito, 1996, Candito, 1999a], la description des adjectifs par [Barrier, 2002, Barrier, en préparation], c'est donc maintenant au tour des verbes supports et des noms prédicatifs de venir enrichir la métagrammaire.

Ce travail s'articulera autour de sept chapitres, se répartissant en 3 parties.

¹Différentes langues ont ainsi pu être représentées. On peut citer le français, et l'italien, mais également l'anglais et le coréen (en cours de description par Sinwon Yoon).

Dans le premier chapitre, nous établirons un rapide état de l'art de la question concernant les CVS. Nous envisagerons les CVS tant d'un point de vue lexical, que sémantique ou syntaxique, dans différentes approches théoriques.

Dans le chapitre 2, nous introduirons le mode de représentation que constitue la Métagrammaire et aborderons la question de l'enrichissement de ce formalisme.

Le chapitre 3 reviendra alors sur les différentes propriétés des CVS pour définir plus précisément ce que nous entendons nous-même par CVS. Une présentation des premiers choix linguistiques pour les TAGs sera entreprise de manière à permettre au lecteur d'avoir une idée plus précise de la grammaire de bas niveau à générer. Cette présentation achèvera la première partie.

La seconde partie constituera en fait la représentation effective de la métagrammaire que nous implémenterons, chaque chapitre correspondant à la description d'une dimension.

Nous terminerons notre exposé par le chapitre 7, qui présentera notamment l'importance que peuvent revêtir les CVS en corpus journalistique. Ce chapitre fera également état des diverses expérimentations que nous avons entreprises afin de disposer d'un lexique de noms prédicatifs et d'évaluer la grammaire que nous aurons décrite.

Première partie

Etat de l'art et premiers choix

Chapitre 1

Etat de l'art

Sommaire

1.1	Point de vue lexical	5
1.1.1	Les locutions verbales	5
1.1.2	Les collocations	6
1.2	Point de vue sémantique	10
1.2.1	Sémantique du verbe support	10
1.2.2	Le lexique-grammaire des verbes supports	12
1.2.3	Sémantique du nom prédicatif	14
1.3	Point de vue syntaxique	18
1.3.1	Propriétés syntaxiques	18
1.3.2	Représentations syntaxiques	22

Nous allons, dans ce premier chapitre, dresser un rapide panorama des études consacrées aux combinaisons "verbe (+ préposition) + nom prédicatif" ou "nom prédicatif + verbe", connues sous le nom de constructions à verbe support (par la suite, CVS). Il s'agit de constructions comme *faire une promenade, donner une gifle, avoir peur, procéder à une arrestation...*, dans lesquelles la fonction principale du verbe consiste à servir de support syntaxique à un nom, appelé nom prédicatif, qui exprime un prédicat sémantique.

Les CVS ne sont pas spécifiques au français. Elles peuvent se rencontrer dans diverses langues, comme par exemple :

- l'anglais : *John has influence over Mary* (Jean a de l'influence sur Marie) [Danlos, 1992]
- l'allemand : *Hans hat Einfluss auf Maria* (Jean a de l'influence sur Marie) [Danlos, 1992]
- l'italien : *Gianni fa l'analisi della situazione* (Jean fait l'analyse de la situation) [Namer et Schmidt, 1997]
- l'espagnol : *Juan a alergia al pollen* (Jean fait de l'allergie au pollen) [Alonso Ramos, 1998]
- le danois : *John har indflydelse pa Maria* (Jean a de l'influence sur Marie) [Danlos, 1992]

ou également dans des langues typologiquement et géographiquement éloignées, comme :

- le basque : *Ni beldur izan naiz* (J'ai eu peur) [Alonso Ramos, 2001]
- le japonais : *Eigono benkyooo siteiru* (Il est en train de faire des études d'anglais) [Alonso Ramos, 2001]
- le persan : *taškile jalaserâ dâdand* (Ils ont formé l'assemblée ; litt. ils ont donné la formation de l'assemblée) [Alonso Ramos, 2001]
- le coréen : *Minuga Inaui uigyoene daehayeo chanseongeul handa* (Minu donne son approbation sur l'opinion de Ina) [Bae, 2002]
- le malayalam : *raaġaawā mantRikḳ pookuwaan anuwaadam koduttu* (Le roi a donné la permission au ministre de sortir) [Jayaseelan, 1988]
- mais aussi le chinois [Chen, 1990], le russe [Zholkovsky et Mel'čuk, 1967, Reuther, 1996], l'arabe [Ibrahim, 2002a]...

Les CVS ne sont pas un phénomène récent : on les trouve, par exemple, déjà en latin et en ancien français [Chaurand, 1991, Marchello-Nizia, 1996], et dans divers états de l'anglais [Brinton et Akimoto, 1999].

Elles ont suscité l'intérêt d'un grand nombre de chercheurs et de plusieurs écoles théoriques. En France, c'est principalement suite aux travaux effectués au sein du Laboratoire d'Automatique Documentaire et Linguistique (LADL) que les CVS ont été reconnues comme un phénomène particulier qui affecte aussi bien le lexique que la grammaire¹. De nombreuses études descriptives ont ainsi été consacrées à des verbes comme *faire* [Giry-Schneider, 1978, Giry-Schneider, 1987], *avoir* [Labelle, 1974, Vivès, 1983], *être* [Danlos, 1980], *prendre*, *perdre* [Vivès, 1983], *donner* [Gross, 1989]²...

Outre Atlantique, au Canada, les travaux élaborés dans le modèle lexicographique du *Dictionnaire explicatif et combinatoire* [Mel'čuk, 1984] proposent également une prise en compte des CVS : le verbe support est considéré comme une information nécessaire à renseigner dans l'article lexicographique du nom.

Dans la tradition anglo-saxonne, différentes études des CVS ont été proposées. [Jespersen, 1942] s'est intéressé à des constructions comme *have a walk* ou *give a kiss* et a désigné le verbe de telles constructions comme un *light verb*, un verbe léger. [Cattell, 1984], quant à lui, appelle de telles expressions des prédicats complexes. Ces études ont principalement une perspective syntaxique, mais des travaux comme ceux de [Wierzbicka, 1982, Dixon, 1991] ou [Stein, 1991] ont un point de vue plus sémantique.

Du côté germanique, [Polenz, 1963] introduit le concept de construction à verbe fonctionnel (*Funktionsverbgefüge*) qui recouvre celui de verbe support, sans pour autant être identique puisqu'il inclut également les constructions à sens causatif.

¹ A ce titre, la première attestation du terme de verbe support est attribuée à [Daladier, 1978], même si les constructions à verbes supports avaient déjà fait l'objet d'études sans pour autant recevoir une dénomination particulière [Gross, 1976, Gross, 1977, Ibrahim, 2002a]

² On trouve également dans le même cadre théorique, des études sur d'autres langues comme l'italien [Angelis, 1989], le portugais [Ranchhod, 1989], le coréen [Han et Rambow, 2000]...

Nous allons maintenant envisager les CVS selon trois points de vue : lexical, sémantique et syntaxique.

1.1 Point de vue lexical

Tout d'abord, nous allons exposer ci-après le point de vue lexical : nous verrons que les CVS ont été régulièrement assimilées à des locutions verbales, puis nous nous intéresserons à leur traitement comme collocations.

1.1.1 Les locutions verbales

1.1.1.1 Caractère idiomatique

Dans la littérature, les CVS sont souvent considérées comme des locutions verbales. C'est par exemple, le point de vue adopté par [Gougenheim, 1971, Curat, 1982, Grevisse et Goose, 1993, Riegel et al., 1996]³. Les locutions regroupent, en fait, tout ce qui dans une certaine mesure a une nature figée ou phraséologique. Elles comprennent à la fois des expressions idiomatiques comme *casser sa pipe*, *perdre les pédales*, ou des expressions semi idiomatiques comme *faire attention*, *avoir le sentiment*...

La principale raison de considérer les CVS comme des locutions semble être la synonymie qu'elles peuvent avoir avec les verbes pleins correspondants :

<i>Construction à verbe support</i>	<i>Verbe plein équivalent</i>
donner la permission	permettre
donner une gifle	gifler
faire une erreur	se tromper
faire une remarque	remarquer
faire une objection	objecter
prendre une décision	décider
lancer une accusation	accuser

Ainsi, les CVS sont traitées comme des locutions car elles ne sont pas composées d'un seul mot et expriment une idée unique. Pourtant, on peut nettement percevoir une différence entre deux expressions comme *casser sa pipe* et *faire une objection*. Dans le premier cas, l'expression est sémantiquement opaque et n'est pas le résultat de l'addition compositionnelle des sens de chacun des lexèmes la composant, alors que dans le second, elle est relativement transparente⁴.

³[Gross, 1996a] exclut les CVS des locutions verbales. Pour la distinction entre locution verbale et CVS en français en allemand, voir également [Grewé, 1993].

⁴Dans le même ordre d'idée, [Alonso Ramos, 1998] signale qu'un individu apprenant une langue étrangère comme l'espagnol par exemple, et voulant employer une expression idiomatique comme *tomar el pelo* (=taquiner)

L'idiomaticité de *casser sa pipe* se manifeste au niveau de l'expression entière, alors que pour *faire une objection*, elle se manifeste dans le sens où l'on ne peut pas prédire quel sera le verbe avec lequel le nom se combine.

1.1.1.2 Les locutions verbales : une classe hétérogène

Dans cette optique, [Alonso Ramos, 1998] estime que les locutions verbales ne constituent pas une classe homogène. Elle établit 3 distinctions parmi les séquences verbe-nom⁵ :

- Les expressions comme *casser sa pipe*, *perdre les pédales*... sont indivisibles sémantiquement, car leur sens ne correspond pas à l'addition des sens de leurs constituants. Ces expressions sont véritablement des locutions verbales.
- D'autres séquences comme *faire peur*, *donner des complexes*... sont analysables : "aussi bien le nom que le verbe ont un sens". Par exemple, pour *faire peur*, le verbe ajoute un sens de causation au sens global : "cette expression pourrait être paraphrasée par 'causer que quelqu'un éprouve la peur'". Ces expressions ne sont pas libres : "le choix du verbe est restreint (...); les noms sélectionnent certains verbes pour exprimer un sens donné".
- Enfin, il y a des séquences comme *donner un conseil*, *avoir peur*, *prendre une décision*..., dans lesquelles le verbe répète une partie du sens du nom. "Si l'on essaie d'analyser [une séquence comme *avoir peur*], on voit qu'il est impossible de définir le verbe sans répéter une composante déjà incluse dans le sens du nom". C'est ce groupe qui constitue les CVS.

1.1.2 Les collocations

1.1.2.1 Imprévisibilité du verbe support

On l'a entrevu, le choix du verbe support est directement fonction du nom prédicatif. Il y a bien sûr des régularités qui peuvent être formulées en terme de classes sémantiques, mais ces généralisations sont parfois sujettes à des exceptions [Guenthner et Blanco, 2004, Alonso Ramos, 1998] et des noms à sens proche peuvent prendre des verbes différents. Ainsi, on *donne un conseil*, mais on *fait une suggestion* ; on *joue un mauvais tour*, mais on *fait une entourloupe* ; on *apporte un complément*, mais on *fait un ajout* ; on *donne un baiser*, mais on *fait une caresse* ou on *prend une décision*, alors qu'on *fait un choix*.

Une autre preuve de l'imprévisibilité de ces verbes peut aussi être rencontrée au niveau de la traduction des CVS d'une langue à l'autre. En français, on *fait un pas*, alors qu'on le prend en

devra l'apprendre en bloc, alors que pour employer une expression comme *poner atenci ón* (=prêter attention), il devra d'abord connaître l'unité lexicale *atenci ón*. C'est "seulement quand il voudra utiliser ce nom dans un contexte phrasal qu'il aura besoin d'un verbe spécial" qu'il devra apprendre et "qui lui permettra d'exprimer les actants de *atención* sans ajouter un nouveau signifié lexical".

⁵Ses exemples sont proposés pour l'espagnol, mais les classes dégagées sont aussi valables pour le français. Nous adoptons ici les exemples pour le français.

anglais (*take a step*), ou qu'on le donne en espagnol (*dar un paso*). On *fait un rêve*, alors qu'on l'a en anglais (*have a dream*) ; on *donne un cours*, alors qu'on l'enseigne en anglais (*teach a course*), ou qu'on le lit en russe (*čitat' kurs*) [Mel'čuk, 1997] ⁶.

Ces différences d'usage sont connues sous le terme de collocations. Ce sont des associations de mots conventionnelles qui doivent être apprises comme telles et que l'on ne peut prévoir à partir des mots qui les composent. [Hausmann, 1979] note une dissymétrie au niveau des composants de la collocation : "dans la collocation le statut des deux combinés n'est pas égal". Pour Hausmann, la base constitue un élément autonome, elle garde son sens habituel. En revanche, le collocatif dépend de cet élément. La collocation est alors une combinaison restreinte et orientée.

Pour les CVS, c'est le nom prédicatif qui joue le rôle de la base, alors que c'est le verbe support qui joue celui de collocatif.

1.1.2.2 Les pratiques dictionnairiques

D'après Hausmann, les dictionnaires de définitions ne sont guère utiles pour codifier les collocations. En fait, si l'on compare des dictionnaires de langue comme *Le Petit Larousse 2005* et *Le Petit Robert 2005*, on se rend compte que les traitements ne sont pas homogènes. *Le Petit Larousse* est vague et confus : sur 7 noms prédicatifs que nous avons choisis au hasard (*sieste, somme, erreur, crime, meurtre, promesse* et *rêve*), seule l'entrée de *somme* mentionne une CVS *faire un somme*. En revanche, *Le Petit Robert* est déjà beaucoup plus régulier, puisque ces 7 noms comprennent tous un exemple de CVS : *faire une sieste, faire un somme, faire une erreur, commettre un crime, commettre un meurtre, faire des promesses, et faire des rêves*.

Le dictionnaire combinatoire du français de [Zinglé et Brobeck-Zinglé, 2003] permet déjà un meilleur traitement. En tant que dictionnaire d'expressions, il est à même de fournir des variantes de support comme par exemple *faire une erreur* et *commettre une erreur*. Malheureusement, toutes les variantes ne sont pas renseignées et l'on n'a pas par exemple *perpétrer un crime* à côté de *commettre un crime*. Or si l'on n'inclut pas cette combinaison, l'utilisateur peut penser que la seule façon d'exécuter un manquement grave à la morale ou la loi, est de *commettre un crime*. La préface de 2 pages du dictionnaire fournit quelques choix linguistiques, mais ne les explicite visiblement pas tous. Il semble que le critère de la fréquence ait été utilisé, mais comme le remarque Hausmann, ce critère n'est pas un critère d'identification des collocations :

⁶ Ainsi, les CVS suscitent l'intérêt des chercheurs en traduction automatique. [Danlos, 1994] note que la traduction du nom prédicatif est compositionnelle, alors que celle du verbe support ne l'est pas. Elle illustre son propos dans un système à transfert :

- Dans la phase d'analyse, l'expression à verbe support est repérée, le verbe ne doit pas être traduit par une règle bilingue.
- Dans la phase de transfert, on fait appel à une règle de transfert lexical pour calculer la traduction du nom prédicatif
- Dans la phase de génération, on cherche le verbe support du nom prédicatif de la langue cible, et on construit une phrase avec ce verbe en respectant les informations temporelles et aspectuelles du verbe source.

une combinaison comme *planter un arbre* est sans doute probable et même fréquente, mais ne constitue pas pour autant une collocation.

famille briographe conventionnels

Seul le *Dictionnaire explicatif et combinatoire du français* de [Mel'čuk, 1984], fournit véritablement des informations spécifiques aux CVS, grâce à l'utilisation de fonctions lexicales particulières.

"La vocation des fonctions lexicales est de fournir au locuteur la totalité des moyens lexicaux nécessaires à l'expression la plus riche, la plus variée et la plus complète de la pensée et, en même temps, de garantir le choix le plus précis de la formulation appropriée" [Mel'čuk et al., 1995].

Ainsi, les fonctions lexicales alimentent-elles un système de paraphrasage.

Trois fonctions lexicales formalisent la notion de verbe support. Ce sont **Oper_i**, **Func_i** et **Labor_{ij}**. La vocation de ces fonctions lexicales est syntaxique. Elles se distinguent entre elles par le rôle joué auprès d'elles par le mot-clé et par le rôle des actants syntaxiques du mot-clé⁷ :

- **Oper_i** : son mot-clé est son complément d'objet direct. L'indice *i* indique que le sujet grammatical de cet **Oper** est l'actant syntaxique profond *i* du mot-clé.
- **Func_i** : son mot-clé est son sujet grammatical. L'indice actanciel renvoie au complément d'objet direct de cet **Func**.
- **Labor_{ij}** : son mot-clé est son complément d'objet indirect. L'indice actanciel *i* en première position renvoie au sujet grammatical du **Labor**, et l'indice *j* en deuxième position renvoie toujours au complément d'objet direct.

Obj - indirect ?

Nous fournissons ci-dessous quelques exemples d'utilisation des fonctions lexicales⁸ :

Fonctions lexicales	Exemple
Oper₁ (remarque) = faire [ART ~]	Max fait une remarque.
Oper₁ (méfait) = perpétrer [ART ~]	Max a perpétré un méfait
Oper₁ (conseil) = donner [ART ~ à N]	Max donne un conseil à Luc
Oper₁ (ordre) = donner [ART ~]	Le général donne un ordre de retraite
Oper₁ (baiser) = donner [ART ~ à N]	Max donne un baiser à Léa
Oper₂ (baiser) = recevoir [ART ~]	Léa reçoit un baiser de Max
Oper₂ (applaudissement) = recueillir [ART ~]	Max recueille les applaudissements
Oper₃ (conseil) = recevoir [ART ~]	Luc reçoit un conseil de Max
Func₁ (responsabilité) = incomber [à N]	Cette responsabilité incombe à Luc
Func₁ (ordre) = venir [de N]	L'ordre de retraite vient du général
Labor₁₂ (interrogatoire) = soumettre [N à ART ~]	Max soumet Luc à un interrogatoire

⁷[Mel'čuk, 1984] utilisent une terminologie parallèle à celle de [Hausmann, 1979] : le mot-clé correspond à la base et le collocatif à un élément de la valeur de la fonction lexicale.

⁸Pour la détermination des indices, voir [Mel'čuk et al., 1995], page 139.

Comme on peut le remarquer, ces fonctions lexicales entretiennent des relations de conversions entre elles. Ainsi par exemple, **Oper**₁ est le conversif de **Oper**₂ (comparer *Max donne un baiser à Léa* et *Léa reçoit un baiser de Max*)⁹ ; **Func**₁ est le conversif de **Oper**₁ (comparer *Le général donne un ordre de retraite* et *L'ordre de retraite vient du général*)... C'est la fonction lexicale **Oper**_i qui est la plus productive.

Ces fonctions lexicales peuvent, par ailleurs, entrer en combinaison avec d'autres fonctions lexicales (**Incep**, **Cont** et **Fin**) qui expriment les trois phases différentes d'un état ou d'un événement : le début (**Incep**), la continuation (**Cont**) et la fin (**Fin**) :

- La fonction lexicale **Incep** combinée avec une autre fonction lexicale, produit un verbe qui signifie "commencer le fait désigné par le mot-clé".
- La fonction lexicale **Cont** combinée avec une autre fonction lexicale, produit un verbe qui signifie "continuer le fait désigné par le mot-clé".
- La fonction lexicale **Fin** combinée avec une autre fonction lexicale, produit un verbe qui signifie "cesser le fait désigné par le mot-clé".

Les valeurs résultant de cette combinaison sont des variantes phasiques des verbes supports :

Fonctions lexicales combinées	Exemple
IncepOper ₁ (voyage) = entreprendre [ART ~]	Max entreprend un voyage
IncepOper ₁ (renommée) = acquérir [ART ~]	Max acquiert une certaine renommée
IncepOper ₁ (plan) = concevoir [ART ~]	Max conçoit un plan ingénieux
ContOper ₁ (respect) = garder [ART ~]	Max a gardé un profond respect envers Luc
ContOper ₂ (respect) = conserver [le ~]	Luc a conservé le respect de Max
ContFunc ₁ (souvenir) = perdurer	Le souvenir perdure
FinOper ₁ (respect) = perdre [ART ~ / tout ~]	Max a perdu tout respect pour Luc
FinOper ₂ (respect) = perdre [le ~]	Luc a perdu le respect de Max

L'élaboration du DEC semble une entreprise colossale¹⁰. A la différence des dictionnaires de langues traditionnels ou des dictionnaires d'expression, le DEC s'inscrit dans un cadre bien formalisé et offre ainsi une systématisation et une consistance au niveau de ses descriptions, grâce à l'utilisation d'un métalangage plus abstrait. La lecture des entrées du dictionnaire n'est pas toujours très aisée, et il est facile de se perdre parmi les différentes fonctions lexicales. Néanmoins, l'expérience aidant, le DEC se révèle être une source d'informations précieuse.

⁹Les CVS décrites par **Oper**₂ correspondent ainsi à ce que [Gross, 1989] nomme constructions converses.

¹⁰Seuls 4 volumes ont pour l'instant été publiés. A l'origine, les auteurs souhaitaient continuer la publication au rythme d'un volume tous les 2 ans. Le premier volume compte 50 vocables, le second 107, le troisième 171, et le quatrième 180. Les objectifs de publication ont donc été tenus. Le choix des vocables décrits est "expérimental" : la sélection repose sur le principe qu'il faut traiter des lexies aussi différentes que possible (i.e. des lexies appartenant à diverses parties du discours, des lexies monolexémiques et multilexémiques, etc.) pour ainsi tester la méthodologie et illustrer de façon claire la démarche linguistique des auteurs.

Ainsi dégagées du fourre-tout des locutions verbales, les expressions à verbe support peuvent donc être sereinement envisagées comme des collocations.

1.2 Point de vue sémantique

Passons maintenant à l'aspect sémantique des CVS : nous nous concentrerons tout d'abord sur le verbe, en étudiant le caractère "vide" que celui-ci peut avoir, puis sur le nom, en envisageant notamment les classifications dans lesquelles il peut rentrer.

1.2.1 Sémantique du verbe support

De nombreuses discussions ont été entreprises concernant la nature sémantique des verbes supports, qui seraient relativement vides de signifié lexical propre. Plusieurs théories se sont ainsi dégagées.

[Blinkenberg, 1960] parle du sens peu spécifique du verbe dans les CVS. D'autres comme [Gougenheim, 1971] leur donnent un sens général. [Moignet, 1974], quant à lui, constate que ces verbes ont une certaine tendance à l'auxiliarité. Pour [Curme, 1935] (cité par [Cattell, 1984] et [Matsumoto, 1999]), des verbes comme *have*, *get*, *do*, *give* ou *make*, qu'il appelle *transitive copulas*, sont comparables à des copules comme *be*, mais correspondent à des états de développement différents.

Such a verb performs merely the function of announcing the predicate. It does not itself predicate, it merely links to the subject [Curme, 1935]

Enfin [Jespersen, 1942], dans son volume consacré à la morphologie, au cours d'un développement sur les conversions, a recours au terme de *light verb* pour désigner un verbe insignifiant : le verbe support. Le terme est, depuis, largement réutilisé dans la littérature anglo-saxonne.

Substantive from verbs without change

The most usual meaning of sbs derived from and identical in form with a vb is the action or an isolated instance of the action. This is particularly frequent in such everyday combinations as those illustrated in the following paragraphs after have and similar 'light' verbs [*have a rest*, *take a walk*...]. They are in accordance with the general tendency of Mod[ern] E[n]glish to place an insignificant verb, to which marks of person and tense are attached, before the really important idea... [Jespersen, 1942]

En revanche, d'autres linguistes rejettent le caractère général attribué au verbe support. C'est notamment le cas de [Gaatone, 1981, Curat, 1982] ou [Fontenelle, 1997]. Pour Gaatone,

il est plus facile d'accorder [à des verbes comme *avoir*, *faire*, *prendre* ou *donner*] un statut d'auxiliaire ou de semi-auxiliaire parce qu'il paraît plus aisé d'en décrire le sens en termes très généraux, tels que "situation" "état" pour *avoir*, "causatif ou factitif" pour *faire* et quelquefois *donner*, "inchoatif" pour *prendre*. Mais ils n'ont pas nécessairement ce sens, comme le montrent [des exemples comme] *avoir coutume*, *droit*, *recours*... [Gaatone, 1981]

Pour Curat, le caractère général d'un verbe comme *faire* tient au fait que ce dernier est très fréquent. Cela n'implique pas une sémantique moindre, mais plutôt "un sémantème plus compatible que ceux des autres verbes avec le mécanisme des locutions verbales".

Fontenelle, quant à lui, revient également sur l'appellation "sémantiquement vide" :

In fact, they [les verbes supports] cannot really be characterized as empty because they do have some (abstract) meaning such as 'to feel' or 'to make'. In fact, one could say that the verb *make* is used in conjunction with the noun *mistake* to mean what it means, which excludes the verbs *to give* or *do*. [Fontenelle, 1997]

Dans sa thèse [Alonso Ramos, 1998] convient d'une nécessité de préciser le terme "sémantiquement vide" :

- vide du point de vue paradigmatique (appelé vide 1) : le terme vide réfère au sens très général que l'on peut attribuer à certains verbes supports.
- vide du point de vue syntagmatique (appelé vide 2) : le verbe ne fait que répéter une partie du sens du nom. "Un verbe est vide dans le contexte du nom qui fonctionne comme son complément d'objet, si la sélection du verbe par le nom est automatique et le signifié du verbe répète une partie du sens du nom".

Elle propose alors une échelle de vide (vide 2), en s'appuyant sur une proposition de [Reuther, 1996] formulée pour les verbes supports russes :

1. Il y a des verbes supports purs dont le signifié est restreint à la partie taxonomique ou générale. Ce sont les verbes les plus productifs et les plus utilisés pour lesquels il est difficile de trouver un lien avec les sens de base des verbes pleins correspondants.
2. Il y en a d'autres qui gardent certains liens avec leur contrepartie libre. On peut ainsi comparer *jouir d'une situation* (sens agréable) et *jouir du paysage*.
3. D'autres verbes supports n'ont pas de contrepartie libre et ne sont que supports. C'est le cas, par exemple, de *asséner*.
4. D'autres ont une contrepartie libre, mais elle est considérée comme un homonyme du support, comme *livrer bataille* vs. *livrer une pizza*.
5. Enfin, le verbe peut être employé comme support, mais garder son signifié de base, comme dans *Jean dit des excuses à Marie* [Giry-Schneider, 1981].

1.2.2 Le lexique-grammaire des verbes supports

1.2.2.1 Verbe support et verbe ordinaire

On trouve également dans les travaux du lexique-grammaire, des discussions concernant le caractère sémantiquement vide du verbe support. Ainsi, en comparant les phrases suivantes, [Gross, 1998] signale que le verbe plein *garantir*, au contraire du verbe support *avoir* est porteur de sens :

- (1) a. Ces idées garantissent une diffusion rapide
- b. Ces idées ont une diffusion rapide

Cependant, la distinction entre verbe support et verbe ordinaire est également basée sur la distinction entre nom prédicatif et nom concret (voir [Alonso Ramos, 1998] citant [Gross, 1989]) : les verbes supports ne se combinent qu'avec des noms prédicatifs alors que les verbes ordinaires prennent aussi des noms concrets. Pour [Gross, 1989], un nom concret ne peut pas être prédicatif : il désigne un objet et n'est pas susceptible de recevoir des indications de temps, de personne et de nombre. En considérant les exemples de la série (2) suivante (à la structure $N_0 V N_1 \text{ à } N_2$), on remarque que pour la seconde phrase (2b), *Luc* est doublement sujet : sujet syntaxique du verbe support *donner*, et sujet sémantique de *démenti*. Or, pour la première phrase (2a), *Luc*, n'est pas un sujet sémantique de *pendule*.

- (2) a. Luc a donné une pendule à Max (verbe ordinaire)
- b. Luc a donné un démenti à Max (verbe support)

Le fait qu'un possessif soit obligatoirement coréférent à N_0 permet d'établir que N_0 est le sujet de N_1 : *Luc_i a donné son_i autorisation à Max*. S'il n'y a pas de coréférence, on a alors affaire à un verbe ordinaire, et à un nom concret : *Luc lui a donné ton autorisation* [document attestant d'une autorisation].

1.2.2.2 Variantes de support et extensions aspectuelles

Le lexique-grammaire établit également une différence entre un *verbe support de base* et ses *extensions ou variantes lexicales*. Ainsi, par exemple, certains noms de maladie comme *rhume*, *grippe*, *rhumatisme* peuvent se combiner non seulement avec *avoir*, mais également avec *souffrir*. Les noms d'erreur se combinent avec *faire*, mais également avec *commettre*. Les noms qui désignent des coups peuvent se combiner avec *donner*, mais également avec *appliquer*, *allonger*, *asséner*, *coller*, *décocher*, *envoyer*, *flanquer*, *foutre*, *ficher*, *lancer*, *octroyer*, *passer*... [Gross, 1989].

Un même verbe peut être la variante de plusieurs supports de base. Comparer ainsi *Max (donne + accorde) un dédommagement à Luc* et *Max (fait + accorde) une faveur à Luc*.

Certaines de ces variantes ont des allures de support, comme dans (*effectuer + accomplir*) *un long trajet* ; d'autres ont un sens plein, comme dans *écrire une lettre* ou forment avec un nom donné une combinaison très spécifique, comme dans *disputer un match* [Giry-Schneider, 1987].

Un critère pour décider si un verbe particulier est une variante d'un verbe support est la synonymie entre la phrase construite avec verbe support de base et celle construite avec la prétendue variante. Ainsi [Giry-Schneider, 1987] indique que des phrases comme *Paul fait une erreur* et *Paul commet une erreur* sont "perçues comme synonymes".

Par ailleurs, un verbe support de base peut également être différencié de ses *variantes ou extensions aspectuelles*. Ces dernières sont des "supports qui, par rapport au support standard, traduisent un aspect inchoatif, duratif [...] terminatif" ou itératif [Gross, 1989], comme dans :

- (3) a. Max entame des négociations avec Luc (inchoatif)
- b. Max garde l'espoir de sauver Luc (duratif)
- c. Max perd le moral (terminatif)
- d. Max multiplie les agressions contre Luc (itératif)

Dans cette optique, les propriétés définitoires des verbes supports sont principalement syntaxiques. [Vivès, 1983], qui a étudié les caractéristiques de *prendre* et *perdre* comme variantes du support *avoir*, indique ainsi que ces verbes conservent les propriétés caractéristiques du verbe support *avoir* (restriction sur les déterminants, double analyse, etc.)

1.2.2.3 Opérateurs causatifs

Enfin, les verbes supports sont à distinguer des opérateurs causatifs. La distinction n'est pas immédiatement évidente, car phrases à opérateur causatif et phrases à verbe support peuvent paraphraser des phrases à verbe ordinaire. [Gross, 1989] compare ainsi les emplois de *donner* dans les deux séries suivantes :

- (4) a. Luc a donné des complexes à Max
- b. Luc a complexé Max
- (5) a. Luc a donné à Max l'autorisation de jouer
- b. Luc a autorisé Max à jouer

Pourtant, ces deux emplois de *donner* sont bien différents. Pour *complexe*, le nombre des arguments est en fait augmenté d'une unité (ici *Luc*) par application de l'opérateur causatif sur le support *avoir* (*complexe* est un substantif qui a un seul argument : *Max*) :

Max a des complexes >> Luc a donné des complexes à Max

Alors que pour *autorisation*, le nombre des arguments n'est pas augmenté (*autorisation* a trois arguments : *Luc*, *Max* et *jouer*) :

Max a l'autorisation de jouer de Luc >> Luc a donné à Max l'autorisation de jouer

1.2.3 Sémantique du nom prédicatif

Jusqu'ici, nous nous sommes intéressés à la sémantique du verbe support. Nous nous tournons maintenant vers celle des noms, autres composants des CVS. Ces noms sont appelés prédicatifs, mais on trouve sous cette dénomination diverses acceptations. Ainsi, pour certains auteurs, prédicatif correspond à un nom déverbal, alors que pour d'autres, le terme renvoie à un nom abstrait. Il est également sous-entendu qu'un nom prédicatif est un nom avec structure argumentale.

accep-
tion

1.2.3.1 Noms dérivés morphologiquement

Il est souvent fait mention lors des études concernant les CVS, du caractère déverbal du nom. Le nom prédicatif serait alors un nom dérivé morphologiquement comme dans *Max gifle Luc* - *Max donne une gifle à Luc* où ces deux phrases sont associées morphologiquement et sémantiquement par une relation de synonymie.

Pourtant, il existe de nombreuses paires verbe/nom liées morphologiquement, mais pas sémantiquement :

faire un récit	≠	réciter
donner des instructions	≠	instruire
donner un avertissement	≠	avertir
faire une fleur	≠	fleurir

Inversement, il existe également de nombreuses paires verbe/nom liées sémantiquement, mais pas morphologiquement :

faire un somme	=	dormir
donner sa parole	=	promettre
faire un cadeau	=	offrir
faire un repas	=	manger

De plus, de nombreux noms combinés avec un verbe support sont isolés ou autonomes d'un point de vue morphologique :

faire un bilan	≠	*bilaniser
faire une allusion	≠	*alluder
commettre un crime	≠	*crim(in)er
donner une beigne	≠	*beigner

De nombreux auteurs s'opposent donc à l'assimilation nom prédicatif = nom déverbal. [Gross, 1989] signale dans cette optique que les études qui ont été entreprises ont montré que les substantifs non dérivés sont plus nombreux que les noms déverbaux et précise donc que :

cette relation morphologique avec un verbe n'est pas le critère de prédictivité des substantifs[...] Un prédicat nominal est défini par le fait qu'il a des arguments et par la nature de ces arguments, et non par un lien morphologique avec un prédicat verbal [Gross, 1989]

1.2.3.2 Noms concrets et noms abstraits vs. noms de résultat et noms de procès

Dans la littérature concernant les CVS, le nom prédicatif est souvent identifié à un nom abstrait, ou à un nom de procès [Giry-Schneider, 1987, Gross, 1989]. Ainsi, les noms concrets¹¹ et les noms de résultat ne pourraient pas entrer dans des CVS. Des chercheurs remettent cependant en cause cette constatation. Ainsi par exemple [Alonso Ramos, 1998] indique que des noms de résultat (qui suivent les critères de [Grimshaw, 1990]¹²) peuvent également entrer dans des CVS :

- (6) a. Max a fait des accusations très dures à Luc
- b. Max a fait beaucoup de promesses à Luc
- c. Max a fait une démonstration du théorème particulièrement claire

Pour Ramos, "étant donné que les verbes supports sont des prédicats vides, des noms comme [accusations, promesses ...] doivent disposer d'une structure argumentielle qui permette d'établir la correspondance entre les actants sémantiques et les actants syntaxiques dans une phrase à verbe support."

[Alonso Ramos, 1998] indique également que le caractère concret ou abstrait d'un nom ne l'empêche pas d'entrer dans une CVS. Elle prend ainsi l'exemple du mot concret *lettre*, comme dans *adresser une lettre à quelqu'un*. Selon elle, ce nom forme avec *adresser* une CVS, de

¹¹Au lieu de l'opposition concret/abstrait, d'autres auteurs parlent d'aliénabilité/inaliénabilité. [Cattell, 1984] signale ainsi que les noms qui forment des CVS ont le trait inaliénable (P est inaliénable pour X, si P de X ne peut pas devenir P de Y).

¹²Seuls les noms de résultat admettent le pluriel. Seuls les noms de résultat admettent l'indéfini ou le démonstratif.

même que le nom abstrait *critique* combiné avec le même verbe *adresser*, forme également une CVS. Elle compare ainsi les phrases :

- (7) a. Max adresse une critique à Luc
b. Max adresse une lettre à Luc

Dans ces deux phrases, les deux noms ont le même comportement et le sens prédicatif est exprimé par le nom. Le verbe support apporte au nom les informations de temps, de nombre, de personne... Le sujet grammatical de la phrase est aussi le sujet sémantique du nom. La formation d'un groupe nominal est par ailleurs possible :

- (8) a. La critique de Max à Luc
b. La lettre de Max à Luc

Il n'y a donc aucune raison de n'accorder qu'aux noms abstraits la possibilité de former des CVS.

1.2.3.3 Classification sémantique des noms supportés

La classification d'Emorine. Dans le but de prédire la forme du verbe support à utiliser pour un nom prédicatif donné, [Emorine, 1992] propose un classement des prédicats suivant différents critères. Elle effectue tout d'abord, une distinction entre noms statiques et noms dynamiques, puis différentes subdivisions, dont l'organisation est la suivante :

Statiques	Qualités		<i>intelligence</i>
	Etats	Emotionnels	<i>amabilité</i>
		Intellectuels	<i>certitude</i>
		Physiques	<i>faim</i>
Dynamiques	Activités		<i>dessin</i>
	Résultatifs		<i>classification</i>
	Performatifs		<i>conseil</i>
	Autres		<i>promenade</i>

La distinction statique/dynamique est acceptée par tous les chercheurs, même si les critères pour classer un état de chose comme statique ou dynamique ne sont pas toujours les mêmes. Pour Emorine, "un élément est considéré comme statique si pendant l'intervalle de temps pendant lequel il existe, il ne subit aucune transformation", alors qu'un élément est considéré comme dynamique "si pendant l'intervalle de temps lors duquel il se réalise, il signifie une transformation, cette transformation pouvant être ou non un changement d'état".

Les classes proposées par Emorine ne sont pas forcément définies sémantiquement. Ainsi par exemple, la tripartition des noms d'état se fait selon des critères sémantiques pour les états intellectuels et physiques, mais pas pour les états émotionnels : les états intellectuels admettent les classifieurs (de type *un sentiment de, une attitude de*) ; les états physiques sont compatibles avec des prédicats comme *est une sensation physique*, ou *est une maladie*. Mais si les états émotionnels sont distingués des états intellectuels c'est parce qu'à la différence de ces derniers, ils permettent certaines transformations adverbiales :

Etats Emotionnels	Max a l'amabilité de m'aider Max m'aide avec amabilité
Etats Intellectuels	Max a la certitude d'être malade *Max est malade avec certitude

Sans remettre en cause l'intérêt de cette classification, qui permet "souvent de savoir, pour un substantif donné, le support adéquat et d'en exclure d'autres", [Gross, 1996c] souligne cependant le caractère très général de cette classification, qui n'est pour lui qu'une approximation, dans la mesure où "elle ne permet pas de prédire à coup sûr la forme recherchée ni les paraphrases synonymiques".

Les classes d'objets Ainsi, même s'il est vrai que *faire* est souvent un support général pour les actions, et que *avoir* ou *être* le sont pour les états, une généralisation ne peut être effectuée. On préférera aux exemples de la série (9), les exemples faisant intervenir des variantes plus spécifiques, appelées "verbes supports appropriés" (10) :

- (9) a. ?Le chirurgien a fait une exérèse
b. *L'enfant a fait un cri
c. ? ?Max a fait un crime

- (10) a. Le chirurgien a pratiqué une exérèse
b. L'enfant a poussé un cri
c. Max a commis + perpétré un crime

Ces remarques conduisent naturellement à l'élaboration de sous-classes plus spécifiques, construites à partir de la forme des verbes supports appropriés. Il est ainsi possible par exemple de subdiviser la classe des noms d'actions¹³ en diverses classes :

¹³Voir [Gross et Kiefer, 1995] pour un traitement similaire avec les noms d'états et d'événements.

<i>Sous-classe</i>	<i>Verbe support</i>
Opérations chirurgicales	pratiquer
Bruits vocaux	pousser, émettre
Délits	commettre, perpétrer
Opérations techniques	effectuer, procéder à

Les classifications d'Emorine et Gross ont donc toutes deux comme but de neutraliser la cooccurrence lexicale restreinte. Cette neutralisation se fait à condition de trouver la classe sémantique commune aux noms qui prennent tel ou tel verbe support. Il reste cependant difficile de savoir si ces classifications peuvent vraiment avoir un poids, dans la mesure où l'on ne dispose pas encore d'une étude sémantique en profondeur (c'est-à-dire d'un dictionnaire) de tous les noms qui entrent dans les CVS. Ne va-t-on pas voir se dessiner une multitude de classes dont la complexité et la spécificité ne permettraient aucun usage efficace pour un apprenant ou pour un traitement automatique ? C'est tout au moins le point de vue pessimiste de [Alonso Ramos, 1998], qui considère "qu'une typologie sémantique n'est possible que si l'on adopte une stratégie inductive".

1.3 Point de vue syntaxique

1.3.1 Propriétés syntaxiques

Les différentes propriétés syntaxiques des CVS ont été relativement bien étudiées au sein du LADL. Nous mentionnerons tout d'abord les phénomènes liés à la double analyse, puis nous nous intéresserons au déterminant des CVS, puis à la descente de l'adverbe.

1.3.1.1 Double analyse

Les phénomènes liés à la double analyse sont certainement ceux qui mettent le plus en valeur les CVS¹⁴. Sans les nommer explicitement ou les mettre en rapport avec les verbes supports, [Gross, 1976] les a pourtant décrits le premier. Ainsi, comme le rappelle [Ibrahim, 2002b], c'est en comparant deux énoncés de même structure, qui ne diffèrent que par leur verbe conjugué (*Max a rapporté une agression contre Luc* vs. *Max a commis une agression contre Luc*), que Gross remarque que seul l'un des deux est passible d'une double analyse. Cette différence est corroborée par une série de propriétés dont notamment la transformation passive, la relativation, l'extraction (en *c'est ... que*), ou la permutation de la longueur :

¹⁴La double analyse touche spécialement les verbes supports, mais également quelques autres verbes comme *craindre* ou *financer*. Voir par exemple [Giry-Schneider, 1987] ou [Emorine, 1992].

Phrase à verbe ordinaire	Max a rapporté une agression contre Luc
Transformation passive	Une agression contre Luc a été rapportée par Max *Une agression a été rapportée par Max contre Luc
Relativisation	L'agression contre Luc que Max a rapportée est lâche *L'agression que Max a rapportée contre Luc est lâche
Extraction	C'est une agression contre Luc que Max a rapportée *C'est une agression que Max a rapportée contre Luc
Permutation de longueur	?Max a rapporté une agression des plus odieuses contre Luc *Max a rapporté contre Luc une agression des plus odieuses
Phrase à verbe support	Max a commis une agression contre Luc
Transformation passive	Une agression contre Luc a été commise par Max Une agression a été commise par Max contre Luc
Relativisation	L'agression contre Luc que Max a commise est lâche L'agression que Max a commise contre Luc est lâche
Extraction	C'est une agression contre Luc que Max a commise C'est une agression que Max a commise contre Luc
Permutation de longueur	?Max a commis une agression des plus odieuses contre Luc Max a commis contre Luc une agression des plus odieuses

Gross conclut ainsi que :

la séquence *une agression contre Luc* a, en compagnie de *commettre*, un comportement double : celui de groupe nominal objet direct et celui d'une séquence composée de l'objet direct *une agression* et d'un autre complément du verbe *contre Luc*. Habituellement ces comportements sont considérés comme exclusifs l'un de l'autre [Gross, 1976].

[Gross, 1989] note par ailleurs que la double analyse est un phénomène très régulier avec les supports *faire* ou *avoir*, mais que la situation est plus complexe avec d'autres supports comme *donner*. Dans ce cas, "la double analyse ne donne pas souvent des phrases d'une acceptabilité indiscutable" :

- (11) Max donne une gifle à Marie
 *C'est une gifle à Marie que Max donne
 C'est une gifle que Max donne à Marie

1.3.1.2 Le déterminant du nom

Les déterminants qui précèdent les noms prédicatifs ont été assez systématiquement étudiés dans le cadre des travaux du LADL. D'assez fortes contraintes ont été observées. Certains substantifs ont une détermination assez libre, d'autres moins. Ainsi, le nom *gifle* associé au support *donner* autorise de nombreux déterminants, mais pas tous, puisque les déterminants démonstratif et défini sont interdits (12b).

- (12) a. Max donne {une + des + quelques + plusieurs + deux} gifle(s) à Luc.
 b. *Max donne {cette + la} gifle à Luc.

Le nom *peur*, quant à lui, combiné avec *avoir* rejette les déterminant défini, partitif, possessif et démonstratif (13c).

- (13) a. Max a {ε + une certaine} peur
 b. Max a eu la peur de sa vie
 c. *Max a {la + de la + sa + cette} peur

Comme le souligne [Gross, 1996a], la situation est très complexe : le choix du déterminant ne dépend pas seulement de la classe sémantique du nom, mais aussi de la nature du verbe support. Par exemple, le nom *sottise* associé au support *faire* a une détermination assez libre : *Max fait {une + des + quelques + cette + une énorme} sottise(s)*, mais associé à la variante *multiplier*, le nom autorisera seulement le déterminant défini :

- (14) a. Max multiplie les sottises
 b. *Max multiplie {une + la + cette} sottise(s)

L'absence de déterminant joue, par ailleurs, un rôle crucial. Pour les chercheurs dans la mouvance de [Gougenheim, 1971], elle est perçue comme marquant un degré de cohésion syntaxique plus étroit entre les différents constituants d'une CVS. En revanche, pour d'autres, comme [Gross, 1996a], l'absence de déterminant n'est pas un critère premier. Dans le même ordre d'idée, [Alonso Ramos, 1998], à la suite de [Giry-Schneider, 1991], considère que l'absence de déterminant ne constitue "qu'une anecdote".

Comme remarqué par [Giry-Schneider, 1991], le choix du déterminant a des conséquences syntaxiques : si le nom prédicatif admet l'article indéfini, il admettra les opérations syntaxiques propres au nom comme la relativisation et la formation d'un groupe nominal :

- (15) a. Max prête {ε + *une + une grande} assistance aux pauvres.
 b. L'assistance que Max prête aux pauvres (est formidable).
 c. L'assistance de Max à ces pauvres (est formidable).

Inversement, un nom qui ne pourra pas avoir de déterminant indéfini, ne se prêtera ni à la relativisation, ni à la formation d'un groupe nominal :

- (16) a. Max donne {le + *un} fouet à Luc
 b. *Le fouet que Max donne à Luc (est ignoble)
 c. *Le fouet de Max à Luc est ignoble

Ainsi, comme le résume [Alonso Ramos, 1998] :

A partir du moment où il y a possibilité d'employer un article indéfini, le nom supporté se comporte comme un syntagme nominal régulier [...] Un nom sans article confère un plus grand degré de cohésion à la CVS, de telle façon que le lien entre le verbe et le nom tend à être perçu comme plus étroit.

Un autre aspect important de la détermination dans les CVS, concerne la coréférence obligatoire entre le déterminant possessif et le sujet dans les CVS. Lorsque cette coréférence n'est pas matérialisée, la construction est alors agrammaticale, ou cesse alors d'être une CVS :

- construction à verbe support : Max a donné son autorisation
- verbe plein : Max a donné ton autorisation = Max a remis un document particulier

1.3.1.3 "Descente" de l'adverbe

Une autre propriété remarquable des CVS a été mise en avant par [Giry-Schneider, 1987] : un adverbe et un adjectif peuvent avoir un emploi synonymique dans une phrase à verbe support. Giry-Schneider prend ainsi l'exemple de la paire *stupide-stupidement*. Elle observe qu'avec un nom comme *faux-pas*, le rapport synonymique existe, mais pas avec un nom comme *maison*¹⁵ :

- (17) a. Max fait stupidement un faux-pas
 = Max fait un faux-pas stupide
 b. Max fait stupidement une maison
 ≠ Max fait une maison stupide

Cette équivalence apparaît, on l'a vu, avec des adjectifs de sens psychologique, mais également avec des adjectifs à sens aspectuel comme *continuel* ou *fréquent* [Giry-Schneider, 1987] :

- (18) a. Max fait fréquemment des faux-pas
 = Max fait des faux-pas fréquents
 b. Max fait fréquemment des gâteaux
 ≠ Max fait des gâteaux fréquents

¹⁵Nous reprenons les exemples fournis par [Giry-Schneider, 1987] p. 31, en corrigeant les coquilles.

Elle n'est cependant pas uniquement réservée aux constructions à verbe support, comme le signale [Harris, 1968] :

- (19) Max boit rapidement un café
= Max boit un café rapide

1.3.2 Représentations syntaxiques

Nous nous tournons désormais vers différentes analyses qui ont été proposées afin de prendre en compte les CVS. Nous commencerons par étudier les propositions de [Jayaseelan, 1988], puis celles de [Cattell, 1984], et enfin [Grimshaw et Mester, 1988]. Toutes trois s'inscrivent dans le cadre théorique du Gouvernement et du liage. Dans un second temps, nous étudierons les propositions avancées dans le cadre des grammaires d'unification comme LFG avec [Matsumoto, 1996], HPSG avec [Erbach et Krenn, 1994] et TAG avec [Abeillé, 1988, Abeillé, 1991, Abeillé, 2002].

1.3.2.1 L'analyse de Jayaseelan (1998)

Nous commençons par l'analyse de [Jayaseelan, 1988] qui résume les problèmes que posent les CVS pour la théorie thématique. Ces problèmes sont le point de départ de son article :

- Les arguments du nom déverbal ne sont pas toujours réalisés à l'intérieur du SN dont le nom est tête. Ils peuvent en effet être réalisés en tant que sujet du verbe ou en tant que compléments (à l'intérieur d'un SV). Cela va donc à l'encontre de la *condition de localité* qui veut que la tête d'un SN ait des liaisons sémantiques avec ses arguments seulement à l'intérieur de son propre syntagme.
- Le deuxième problème consiste à savoir comment le verbe et le nom entretiennent des relations sémantiques avec les différents arguments. Lorsque les arguments sont promus, le verbe ne peut pas recevoir tous les arguments car il n'a plus de rôles sémantiques à assigner. Jayaseelan montre alors qu'en anglais¹⁶, les SP en italique des phrases suivantes sont des compléments du verbe. Pourtant ce dernier ne peut leur fournir un rôle sémantique :

- (20) a. John gave permission to Mary *to leave*
b. John made an offer of money *to Mary*
c. John felt hatred *toward Mary*

Dans la première phrase, *give* a seulement trois compléments à offrir : la Source (*John*), le thème (*permission*) et le But (*to Mary*). Il n'a pas une quatrième position argumentale

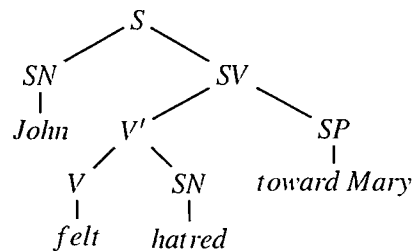
¹⁶Jayaseelan fournit également des exemples pour le Malayalam. Nous ne reproduisons ici que les exemples anglais.

et ne peut donc pas θ -marquer l'argument promu (*to leave*). Dans la seconde phrase, *make* sous-catégorise seulement deux compléments : l'Agent et le Thème. Ces rôles sémantiques sont respectivement assignés à *John* et *an offer of money*. Le verbe n'a donc plus de rôle à offrir pour *to Mary*. Enfin, dans la troisième phrase, *feel* a seulement deux arguments : un Expérienceur (*John*) et un Thème (*hatred*) et ne peut donc pas non plus θ -marquer *toward Mary*.

- Enfin, le troisième problème est l'inverse du second. Les arguments du nom ne sont pas toujours ses compléments. Les arguments ont été promus et apparaissent comme des compléments du verbe, qui ne peut pourtant pas leur assigner un rôle sémantique.

Jayaseelan propose alors que la distribution des différents rôles sémantiques soit toujours compositionnelle. L'ensemble de rôles sémantiques que peut assigner un nœud sémantique est déterminé de manière compositionnelle par ses constituants. Lorsque ces constituants ont des arguments, l'ensemble final de rôles est alors l'union des deux ensembles d'arguments.

Il illustre alors son propos, en reprenant l'exemple de *John felt hatred toward Mary*, dont la structure est la suivante :



feel peut assigner les θ -roles Expérienceur, Thème. Il marque le SN adjacent interne à V' (*hatred*) comme un Thème et promeut Expérienceur au V'. *Hatred* peut assigner les deux θ -roles Expérienceur, But, mais pas à ce niveau, donc il les monte à V'. Le nœud V' assigne alors le rôle But au SP (*toward Mary*) et promeut le rôle Expérienceur au SV. Ce dernier assigne alors Expérienceur au SN sujet (*John*).

De cette manière, Jayaseelan solutionne les problèmes mentionnés :

- *hatred* qui semblait θ -marquer *John* et *toward Mary*, tous deux en dehors de la projection maximale dont il était la tête, θ -marque en fait indirectement *John* et *toward Mary*. Les rôles sémantiques du nom sont transférés aux nœuds syntaxiques supérieurs : V' assigne le rôle But au SP et le SV assigne le rôle Expérienceur au SN sujet.
- *John* qui semblait doublement θ -marqué par le verbe et le nom, est en fait seulement θ -marqué par SV. Il reçoit le rôle d'Expérienceur.
- Enfin, le problème posé par le fait que le verbe n'avait pas de rôle sémantique pour le SP disparaît. Le SP n'est pas un complément de V, mais de V' et c'est V' qui le θ -marque.

Une conséquence de cette analyse est qu'il n'est plus nécessaire de parler d'une classe de prédicats complexes, puisque ce système d'assignation compositionnelle de rôles sémantiques est général. Les CVS ne font donc appel à aucune règle particulière.

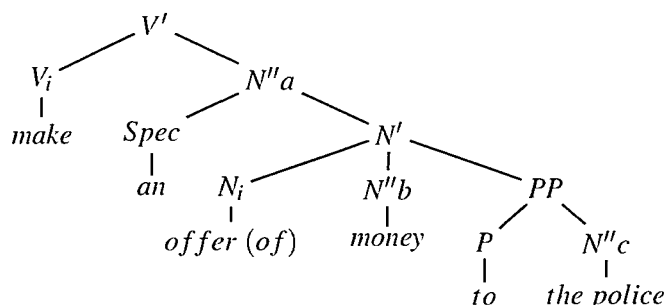
1.3.2.2 L'analyse de Cattell (1984)

Quelques années plus tôt, [Cattell, 1984] avait également proposé une analyse en vue de prendre en compte les CVS (qu'il appelait des prédicats complexes). L'idée de Cattell était de dire qu'un prédicat complexe avait deux nœuds prédicatifs. Ainsi, en s'inspirant de [Chomsky, 1981] qui suggère que *advantage*, dans *to take advantage* a le rôle de "quasi-argument", Cattell affirme que le nom *offer*, dans *to make an offer* n'est pas un argument. Il énonce le principe suivant :

"No predication node P, or any projection of it, can constitute an argument of a predicate of which P is a component." [Cattell, 1984, p.51]

Ainsi, l'article de l'élément *offer* (qui représente un nom ou un verbe) inclut les différents rôles sémantiques et les positions syntaxiques de ses arguments.

Cattell illustre plus précisément son propos, avec l'exemple *make an offer of money to the police*, dont la représentation est la suivante :



D'après le principe énoncé par Cattell, $N''a$ n'est pas un argument du prédicat. Il fait, en effet, lui-même partie d'un nœud prédicatif. Les nœuds qui dominent immédiatement *make* et *offer* sont renseignés avec un indice *i* de manière à indiquer qu'ils font partie du même prédicat complexe.

Les différents rôles sémantiques associés au nom sont assignés conjointement dans la zone dominée conjointement par V' et N' . Le prédicat complexe assigne un rôle sémantique à tout argument qui occupe une position à l'intérieur de cette zone. Cependant, le verbe assigne un rôle sémantique seulement si ce rôle est lui-même prévu dans l'article du nom¹⁷.

Cattell propose une règle de prédicat complexe, conçue comme une règle de redondance lexicale :

¹⁷Cela explique pourquoi par exemple, le verbe *give* (dans la phrase **John gave a shake of his head to Priscilla*) ne peut pas assigner le rôle But, lorsqu'il se combine avec le nom *shake* dont l'entrée renseigne seulement le rôle Thème.

Given that :

- i K and L are two lexical items such that V_K and N_L are particular categorial manifestations of K and L respectively ;
- ii L is lexically marked $V_K/PRED$ ¹⁸ ;
- iii the noun phrase of which N_L is the head in a particular structure S in the right-most noun phrase immediately dominated by the V' of which V_K is head ;

Then

$V_K...N_L$ constitutes a complex predicate in the structure S , assigning the same θ -nodes as L assign in a simple phrase. The cople predicate can assign the roles to any of the grammatical functions provided for them, in the lexical entry of K or that of L (Cattell, 1984, p.61)

Les arguments d'un prédicat complexe sont donc pris à partir de l'entrée L . Le verbe, quant à lui, peut seulement assigner les rôles prévus par l'article L . Par contre, même si les rôles sont déterminés par L , les fonctions grammaticales du verbe sont disponibles pour ces rôles. Cattell ne dit jamais que c'est le nom qui assigne les différents rôles sémantiques¹⁹.

1.3.2.3 L'analyse de Grimshaw et Mester (1988)

Si les deux premières analyses présentées s'intéressaient principalement à l'anglais, celle de [Grimshaw et Mester, 1988] se concentre essentiellement sur le verbe japonais *suru*, considéré comme un *light verb*²⁰ En tant que tel, *suru* n'apporte donc pas de signification particulière à la phrase, et sert de support à la flexion verbale.

Pour Grimshaw et Mester, la structure d'arguments de *suru* est vide, et le verbe ne peut pas assigner de rôles sémantiques : *suru* a seulement le squelette de la structure d'arguments. *suru* se combine avec le SN objet sans lui assigner de rôle. C'est la structure argumentale du nom qui justifie les arguments qui se combinent avec *suru*, même à l'extérieur du SN. *suru* permet au nom d'assigner les rôles sémantiques dans le contexte phrasal.

C'est le *transfert d'arguments* (du nom au verbe) qui fait que *suru* devient un assignateur de rôles. Le nom peut retenir certains arguments à l'intérieur du SN, ces derniers entretenant alors une relation sémantique avec la tête nominale. Comme *suru* se retrouve avoir absorbé la structure argumentale du nom, les arguments qui apparaissent sous le nœud P ont des relations sémantiques avec ce verbe. Les auteurs fournissent quelques exemples de transfert :

¹⁸Cette marque est utilisée pour rendre compte de l'imprévisibilité du support. Elle ne semble pas forcément essentielle pour Cattell : le nom *offer* est ainsi marqué comme *MAKE/PRED*, mais n'est plus marqué par la suite (p. 119). Il suffirait de dire que le nom a un trait "inaliénable" pour connaître le verbe support avec lequel il doit se combiner.

¹⁹Ainsi, dans une phrase comme *Harry gave Sue a hug*, ce n'est pas le nom *hug* qui fournit un rôle sémantique à *Sue*, mais le verbe *to hug* qui lui est associé (cf. Cattell, 1984, p.286).

²⁰Pour une approche où l'on considère que *suru* est un verbe plein, voir par exemple [Uchida et Nakayama, 1993].

Transfert partiel :

- a keikoku (Agent, But, Thème)
 suru() <acc>
 keikoku (Thème) + suru (Agent, But) <acc>
- b John-wa murabito-ni [[ookami-ga kuru-to]-no KEIKOKU]-o shita
 John-Top villager-to wolf-Nom come-Comp-Gen warn-Acc suru
 Trad. angl : John warned the villagers that the wolf was coming
 Trad. fr : John a averti les villageois que le loup venait
 (litt. John a fait un avertissement aux villageois que le loup venait)

Ici, l'argument Thème *ookami – ga kuru* (le loup venait) est resté à l'intérieur du SN dont la tête est *keikoku* (avertissement). Ceci est indiqué par le génitif *–no*.

Transfert complet :

- a keikoku (Agent, But, Thème)
 suru() <acc>
 keikoku () + suru (Agent, But, Thème) <acc>
- b John-wa murabito-ni [ookami-ga kuru-to] KEIKOKU-o shita
 John-Top villager-to wolf-Nom come-Comp warn-Acc suru

Ici, *ookami – gara kuru* forme un constituant différent de celui de *keikoku*.

Que ce soit pour le transfert partiel, ou le transfert complet, l'agent *John* et le But *murabito* (villageois) sont transférés au verbe.

On obtient comme résultat de transfert deux prédicats, dont les propriétés d'assigner des rôles sémantiques ont été modifiées, sauf lorsque tous les arguments du nom ont été transférés. Ainsi, la formation d'une CVS respecte la condition de localité, puisque l'assignation des rôles est réalisée dans deux domaines (*SN* et *P*), et par des structures argumentales différentes.

Cette analyse qui vaut pour le japonais, est également valable, selon les auteurs, pour une langue comme l'anglais. Cependant, l'ensemble est moins évident. Les verbes supports anglais ont une structure argumentale incomplète, plus spécifiée que celle totalement vide de *suru*. Il est plus difficile de vérifier le transfert d'arguments car il n'y a pas de marque de cas pour déterminer la distribution des arguments dans les CVS anglaises. Le verbe support anglais fournit la structure de cas. Il a une structure argumentale incomplète. Celle du nom est en revanche complète. Lorsque les arguments du nom sont transférés au verbe, et si le verbe support a déjà l'argument correspondant, les deux rôles sémantiques sont fusionnés.

Toutes les analyses que nous avons étudiées jusqu'à présent, essayaient d'indiquer que les différents constituants des CVS manifestaient une cohésion syntaxique plus forte qu'un syntagme normal²¹. Toutes trois considéraient également que le verbe support était un assigna-

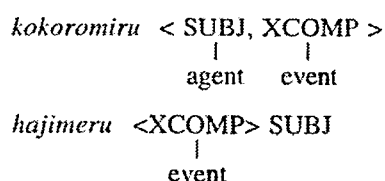
²¹ Par co-indexation de deux constituants ou par assignation de rôles sémantiques en commun.

teur de rôles sémantiques. Nous envisageons maintenant les CVS sous l'angle des grammaires d'unification et allons nous intéresser à trois nouvelles analyses : celle de [Matsumoto, 1996], celle de [Erbach et Krenn, 1994] et [Abeillé, 1988, Abeillé, 1991].

1.3.2.4 L'analyse de Matsumoto (1996)

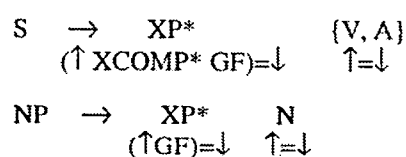
L'analyse de [Matsumoto, 1996] constitue en fait une réponse à l'analyse proposée par Grimshaw et Mester. Ces derniers déclaraient que seuls *suru* et sa forme causative *saseru* pouvaient être considérés comme des verbes légers. Matsumoto note au contraire que beaucoup d'autres verbes peuvent manifester le même comportement que *suru*. Ce sont des verbes comme *hajimeru* (commencer), *kokoromiru* (tenter), *kurikaesu* (répéter), *tsuzukeru* (continuer)... Tous ces verbes sont soit des verbes à montée, soit des verbes à contrôle : ils requièrent tous que le sujet de leur complément prédicatif soit contrôlé ou lié par un de leurs arguments.

L'analyse de Matsumoto s'inscrit dans le cadre des grammaires lexicales fonctionnelles LFG. Matsumoto décrit les entrées lexicales de verbes comme *kokoromiru* (tenter) ou *hajimeru* (commencer) de la façon suivante :



kokoromiru et les autres verbes à contrôle imposent des restrictions sémantiques à leur sujet, alors que *hajimeru* et les autres verbes à montée n'imposent pas cette restriction.

Un ensemble de règles permet aisément de rendre compte des constructions à verbe léger²². La première règle génère directement une structure dans laquelle les arguments et les ajouts du nom sont directement dominés par S. La seconde règle permet que les arguments et les ajouts du nom apparaissent à l'intérieur du NP :

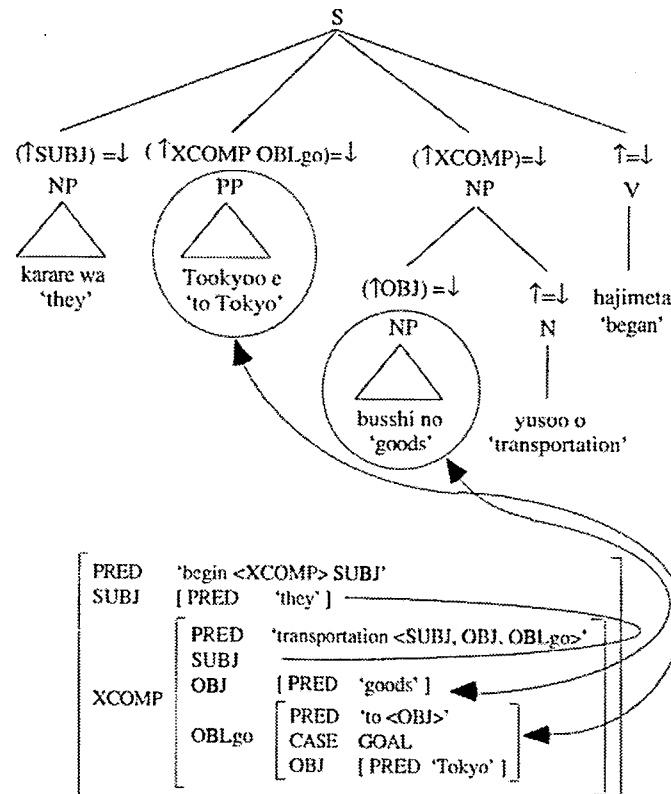


L'analyse d'une construction à verbe léger comme :

²²L'expression régulière XCOMP^* (dans laquelle * est l'étoile de Kleene) correspond à une disjonction infinie de séquences contenant de 0 à n occurrences du symbole XCOMP. GF est une abréviation qui représente une disjonction de toutes les fonctions grammaticales possibles, par exemple SUBJ, OBJ, OBJ_θ , COMP, XCOMP, ADJ...

Karera wa Tookyoo e busshi no yusoo o hajimeta.
 Ils Top Tokyo Goal marchandises Gen transport Acc commencer
 Ils commencèrent à transporter les marchandises à Tokyo

produit alors les structures suivantes :



L'analyse de Grimshaw et Mester suggérait que le verbe léger et son nom avaient des structures argumentales qui n'étaient pas complètement indépendantes. Le complexe Nom + Verbe était en un sens considéré comme un mot. L'analyse de Matsumoto suggère plutôt que le nom et son verbe léger sont deux mots indépendants à tous les niveaux (c-, f- et a-structure). Autrement dit, les constructions à verbe léger ne forment pas pour Matsumoto des prédicats complexes.

1.3.2.5 L'analyse de Erbach et Krenn (1994)

[Erbach et Krenn, 1994] soulignent le fait que la sémantique des CVS est basiquement apportée par le nom prédicatif et que le verbe support apporte simplement des informations concernant le temps, la causativité et l'Aktionsart. Le verbe remplit une de ses positions argumentales avec le nom prédicatif, et ses autres positions avec les arguments du nom prédicatif. Tout argument restant du nom prédicatif est ajouté à la liste de sous-catégorisation du verbe support.

Les auteurs illustrent leur propos avec les CVS suivantes :

- *Angst haben* (avoir peur)
- *in der Gewalt haben* (avoir le contrôle)

La sémantique des noms prédicatifs ressemble à celle des verbes. Les entrées lexicales des noms prédicatifs biargumentaux *Angst* (peur) et *Gewalt* sont les suivantes :

$$\left[\begin{array}{l} PHON \textit{Angst} \\ SYNSEM|LOC \end{array} \left[\begin{array}{l} CAT \left[\begin{array}{l} HEAD \textit{noun} \\ SUBCAT \langle (NP[gen]_{\boxed{1}}), (PP[vor]_{\boxed{2}}) \rangle \end{array} \right] \\ CONT \left[\begin{array}{l} REL \textit{angst} \\ EXPERIENCER \boxed{1} \\ PHENOMENON \boxed{2} \end{array} \right] \end{array} \right] \right]$$

Angst (peur) - Nom prédicatif

$$\left[\begin{array}{l} PHON \textit{Gewalt} \\ SYNSEM|LOC \end{array} \left[\begin{array}{l} CAT \left[\begin{array}{l} HEAD \textit{noun} \\ SUBCAT \langle (NP[gen]_{\boxed{1}}), (PP[\ddot{u}ber]_{\boxed{2}}) \rangle \end{array} \right] \\ CONT \left[\begin{array}{l} REL \textit{control} \\ CONTROLLER \boxed{1} \\ CONTROLLED \boxed{2} \end{array} \right] \end{array} \right] \right]$$

Gewalt (contrôle) - Nom prédicatif

Alors que le verbe plein *haben* a la structure de traits (21), la structure du verbe support est celle représentée par (22).

$$(21) \left[\begin{array}{l} PHON \textit{haben} \\ SYNSEM|LOC \end{array} \left[\begin{array}{l} CAT \left[\begin{array}{l} HEAD \textit{verb} \\ SUBCAT \langle NP[nom]_{\boxed{1}}, NP[acc]_{\boxed{2}} \rangle \end{array} \right] \\ CONT \left[\begin{array}{l} RELN \textit{have} \\ POSSESSOR \boxed{1} \\ POSSESSED \boxed{2} \end{array} \right] \end{array} \right] \right]$$

Haben - Verbe plein

$$(22) \quad \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{SC} \left\langle \text{NP[nom]} \text{[1]}, \text{NP[acc, Angst v ...Vertrauen]} \left[\begin{array}{l} \text{CAT|SC [] [1] + [2]} \\ \text{CONT [3]} \end{array} \right] \right\rangle \& [2] \end{array} \right] \\ \text{CONT [3]} [\text{AKTIONSPORT neutral}] \end{array} \right]$$

Valeur SYNSEM|LOCAL de *haben* - Verbe support

Si le support se combine avec un SN prédicatif, ce SN est en position d'objet. Le premier argument du nom prédicatif est coindexé avec le sujet du verbe support. Tous les arguments additionnels qui ne peuvent pas être intégrés (i.e. "mappés") à un argument du verbe support sont transmis à la liste de sous-catégorisation du verbe support. Le verbe support hérite de la relation sémantique introduite par le nom prédicatif. L'Aktionsart de la CVS est spécifiée par le verbe support.

L'instanciation de la position d'objet direct du verbe support *haben* avec le signe lexical du nom prédicatif *Angst* produit alors la structure de traits suivante :

$$\left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{NP[nom]} \text{[1]}, \\ \text{PHON Angst} \\ \text{SC} \left\langle \begin{array}{l} \text{SYNSEM|LOCAL} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{SC} \left\langle \begin{array}{l} \text{NP[gen]} \text{[1]}, \\ \text{PP[vor]} \text{[2]} \end{array} \right\rangle \right] \\ \text{CONT [3]} \left[\begin{array}{l} \text{REL angst} \\ \text{EXPERIENCER [1]} \\ \text{PHENOMENON [2]} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{CONT [3]} [\text{AKTIONSPORT neutral}] \end{array} \right]$$

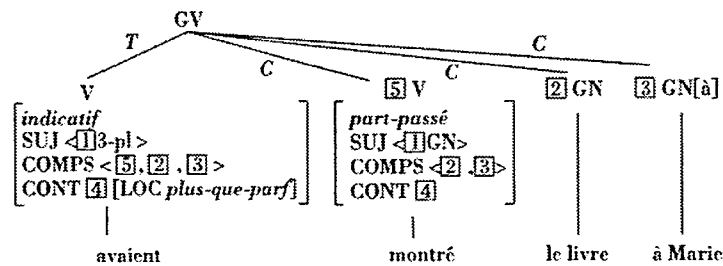
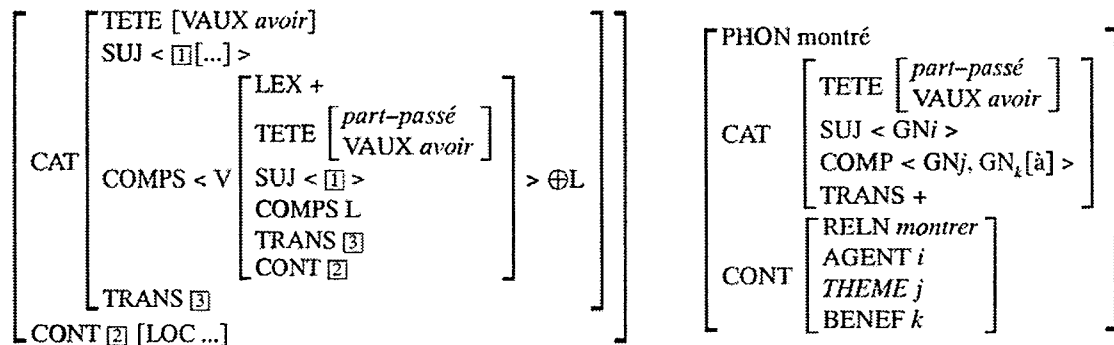
Valeur SYNSEM|LOCAL de (*Angst*) *haben*

Pour les CVS pour lesquelles le nom prédicatif apparaît dans un SP, l'objet direct de *haben* est sémantiquement co-indexé avec le second argument de la liste de sous-catégorisation du nom prédicatif :

$$\left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HEAD verb} \\ \text{NP[nom]} \text{[1]}, \\ \text{NP[acc]} \text{[2]}, \\ \text{SC} \left\langle \begin{array}{l} \text{SYNSEM|LOCAL|HEAD prep [PFORM in]} \\ \text{DTRS|COMP-DTRS} \left\langle \text{NP[dat, Gewalt]} \left[\begin{array}{l} \text{SC} \left\langle \text{[] [1]}, \text{[] [2]} \right\rangle \right] \\ \text{CONT [3]} \end{array} \right] \end{array} \right] \end{array} \right] \\ \text{CONT [3]} [\text{AKTIONSPORT neutral}] \end{array} \right]$$

Valeur SYNSEM LOCAL de (*in der Gewalt*) *haben*

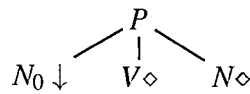
L'analyse de Erbach et Krenn, rapproche le comportement des verbes supports de celui des auxiliaires. Dans leur article, [Abeillé et Godard, 1996] ont en effet proposé d'analyser les auxiliaires comme des prédicats à valence et interprétation variables. Ils prennent un participe lexical pour complément et partagent avec lui les arguments qu'il attend (sujet et compléments éventuels).



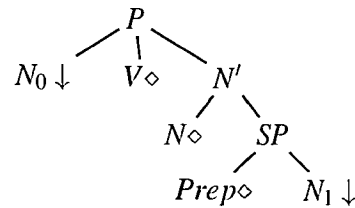
1.3.2.6 Les analyses de Abeillé (1991, 2002)

Nous terminerons notre panorama général, par la présentation des analyses de [Abeillé, 1991, Abeillé, 2002] réalisées pour la grammaire TAG du français - ces analyses s'appuyant en partie sur des propositions fournies quelques années auparavant par [Abeillé, 1988].

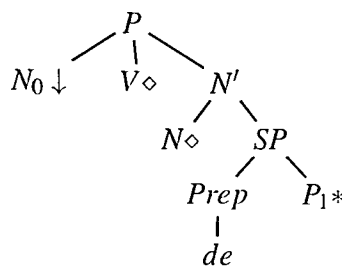
[Abeillé, 1991] considère donc que les familles à verbe support sont représentées par des arbres élémentaires à têtes multiples. Elle place sur un pied d'égalité, le verbe support, le nom prédicatif, et éventuellement le déterminant du nom et les prépositions, en les considérant comme des co-têtes. Chaque combinaison *Verbe support* - *Nom prédicatif* est notée comme une entrée.



Un arbre élémentaire valable pour
Max fait une gaffe

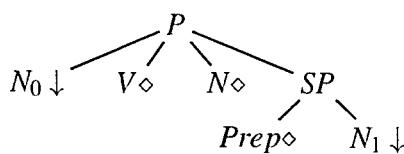


Un arbre élémentaire valable pour
Max commet une agression contre Luc

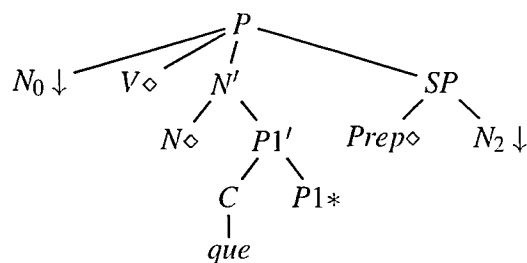


Un arbre auxiliaire valable pour
Max a envie de partir

Sont notés comme complément du verbe support, les syntagmes qui ne peuvent en aucun cas être considérés comme des compléments du nom prédicatif (l'extraction avec le nom prédicatif étant donc interdite). Les SP introduits par la préposition *à* sont donc considérés comme complément exclusif du verbe support, ce qui n'empêche pas l'incorporation régulière du SP dans le SN nominal complexe :



Un arbre élémentaire valable pour
Max fait la cour à Léa

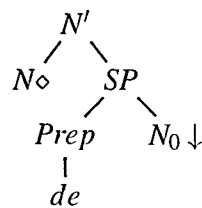


Un arbre élémentaire valable pour
Max fait la réponse que les enfants partiront à 9h à Léa

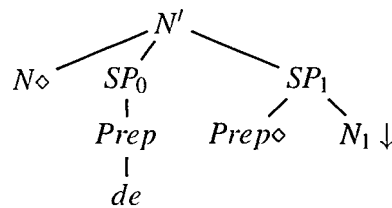
Les différentes familles proposées²³ suivent en général les mêmes règles que les familles à arguments libres, et l'on retrouve des arbres élémentaires (correspondants à chaque transforma-

²³Elles sont au nombre de 12 : 7 pour les familles à arguments nominaux ; 5 pour les familles à arguments phrastiques.

tion ou à la combinaison de deux transformations) assez semblables (cf. Abeillé, p182). Mais à la différence des familles libres, les familles à verbe support doivent également contenir des arbres élémentaires de racine N, pour permettre de représenter le SN complexe. Il s'agit là d'une différence essentielle avec les familles figées.



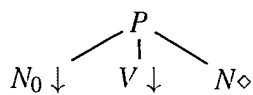
Un arbre élémentaire valable pour
La gaffe de Max



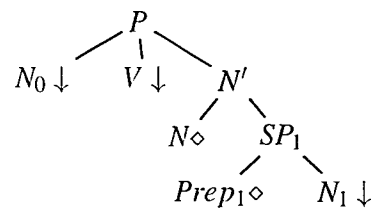
Un arbre élémentaire valable pour
L'agression de Max contre Luc

L'analyse proposée par Abeillé a l'avantage de ne pas supposer de double analyse²⁴, et de ne pas reposer non plus sur deux niveaux de représentation linguistique. La différence entre expression semi-figée et expression figée n'est cependant pas immédiatement perceptible. En outre, l'ordre entre le nom prédicatif et son complément prépositionnel ou phrastique est fixe.

Dans une version plus récente, [Abeillé, 2002] modifie cette analyse en proposant que les phrases à verbe support soient représentées par des arbres élémentaires à tête nominale. Le verbe support, perd son statut de co-tête pour être simplement substitué. Les ancres lexicales sont alors uniquement constituées du nom prédicatif et les mots fonctionnels²⁵.



Un arbre élémentaire valable pour
Max fait une gaffe



Un arbre élémentaire valable pour
Max commet une agression contre Luc

²⁴Le problème consiste en fait, à savoir si le SP est complément du verbe support ou du nom prédicatif.

Le problème peut être résolu si l'on considère la propriété "complément de" non seulement en terme de point d'attachement mais aussi comme signifiant que l'élément appartient à la même structure de base que ce dont il est complément. Ainsi, (...) il devient possible de dire [qu'un SP] est à la fois complément [d'un verbe support] et [d'un nom prédicatif], sans pour autant supposer deux arbres ou une réanalyse. [Abeillé, 1991]

²⁵16 familles sont définies : 10 à arguments nominaux et 6 à arguments phrastiques.

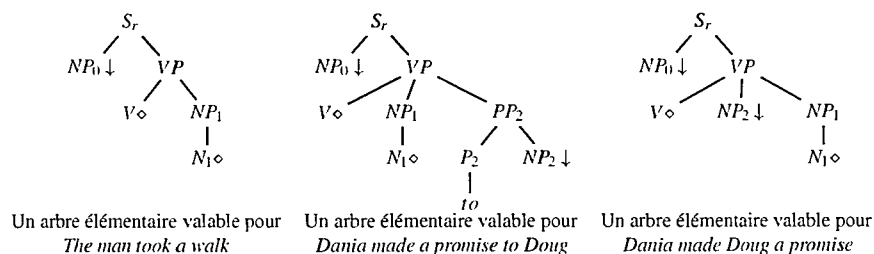
Cette proposition ne fait cependant pas l'objet d'une implémentation.²⁶

Au terme de ce chapitre, nous pouvons déjà tirer quelques conclusions :

- D'un point de vue lexical, on a vu que les CVS ne devaient pas être considérées comme des "locutions verbales". Même si le statut des CVS comme collocations n'a pas toujours été clairement indiqué, il semble pertinent d'y avoir recours. En particulier, il semble plus intéressant de déterminer la forme du verbe support de manière lexicale, plutôt que sémantique. Etant donné son caractère collocatif, un verbe support n'est donc pas une unité lexicale de plein droit.
- D'un point de vue sémantique, les auteurs ne s'accordent pas toujours sur le fameux caractère "vague" ou "vide" des verbes supports. Bien souvent, on ne sait pas toujours non plus ce qui distingue un verbe ordinaire, d'un verbe support. Néanmoins, on trouve chez certains auteurs, comme Alonso-Ramos, des tentatives de définition. En particulier, le caractère vide 2 de tout verbe support, c'est-à-dire "sa caractéristique de ne pas ajouter de signifié lexical au nom avec lequel il se combine et de ne pas être choisi d'après son sens, mais plutôt en fonction du nom qui le sélectionne", semble pertinent. En revanche, un verbe support n'est pas nécessairement vide 1 : il peut avoir un signifié lexical, mais il est forcément vide 2 dans un contexte collocationnel. Son signifié lexical ne fait alors que répéter une partie du sens du nom avec lequel il se combine.

En ce qui concerne le nom, la capacité de ce dernier à avoir des arguments a parfois été remise en question. Toujours est-il que le terme de nom prédicatif a bien souvent une interprétation trop étroite. Si l'on se limite à des noms dérivés morphologiquement d'un verbe ou à des

²⁶ A titre de comparaison, on notera que la grammaire TAG de l'anglais [XTAG, 2001] décrit des constructions à verbes légers. La description est proche de celle de [Abeillé, 1991] : nom prédicatif et verbe support forment donc des co-têtes.



A la différence du français, les familles à verbes légers se distinguent davantage des familles à verbes pleins. Ainsi, les possibilités d'extraction sont elles beaucoup plus limitées. En particulier, il n'est pas possible d'extraire le nom prédicatif. La passivation n'est également pas régulière.

Deux familles sont implémentées pour l'anglais : la famille Tnx01VN1 qui renseigne les constructions à verbe léger sans complément prépositionnel, et la famille Tnx01VN1Pnx2 qui renseigne les verbes légers bi-transitifs. Lexicalement, c'est la première famille qui est la plus représentée, puisque 259 paires verbe/nom sont décrites. La seconde famille est beaucoup plus marginale puisque seules 18 paires verbe/nom la sélectionnent.

noms ayant une structure d'arguments au sens de Grimshaw, la quantité de noms qui entrerait dans une CVS serait relativement faible. La distinction Nom prédicatif / Nom abstrait a par ailleurs été remise en question.

- Enfin, d'un point de vue syntaxique, il semble que le degré de cohésion entre le verbe et le nom ne soit qu'une trace d'un caractère phraséologique. Les analyses qui se placent dans le cadre du gouvernement et du liage indiquent que les différents constituants des CVS manifestent une cohésion syntaxique plus forte qu'un syntagme normal et considèrent également que le verbe support est un assignateur de rôles sémantiques. Si au début des années 90, très peu d'analyses avaient été envisagées en vue d'un traitement automatique, dans des cadres bien formalisés et définis, le retard semble aujourd'hui se combler.

Le chapitre suivant présentera brièvement le cadre théorique dans lequel nous nous plaçons pour notre étude.

Chapitre 2

La notion de métagrammaire

Sommaire

2.1	Aperçu général	38
2.1.1	Motivations	38
2.1.2	La proposition de Vijay-Shanker et Schabes	40
2.1.3	Compilateur et hiérarchie	43
2.2	Le compilateur de Candito	53
2.2.1	Langage de description	53
2.2.2	Calcul des arbres	54
2.2.3	Comportements déviants	55
2.2.4	Des exemples concrets	58
2.2.5	Format de sortie du compilateur	61
2.3	Problèmes de représentation	62
2.3.1	Problèmes de dominance	62
2.3.2	Problèmes d'ordre entre constituants	65

L'écriture d'une grammaire électronique à large couverture n'est pas une chose facile. Tout linguiste confronté au problème d'une implémentation effective se retrouve tôt ou tard face à des difficultés qu'une réflexion théorique ne peut imaginer : dès lors qu'une grammaire n'est plus amenée à être une grammaire jouet, toute description, toute modification peut avoir des conséquences problématiques.

Plusieurs voies ont été explorées pour décrire de telles grammaires : [Abeillé, 1991] a ainsi réalisé une grammaire *à la main*. D'autres comme [Xia, 2001] ont extrait automatiquement une grammaire à l'aide de corpus d'apprentissage. La Métagrammaire constitue en quelque sorte, une troisième voie, permettant de prendre en compte les avantages des 2 méthodes précédentes (description linguistiquement motivée, automatisation de tâches répétitives), tout en évitant leurs inconvénients (durée de développement, dépendance vis à vis des données).

Ce chapitre présente le concept général de la hiérarchisation de l'information pour les TAG, proposé par [Vijay-Shanker et Schabes, 1992], puis s'attache à la description de différents projets de métagrammaire qui en découlent : celui élaboré par [Candito, 1999a], celui élaboré par [Gaiffe et al., 2002], puis celui élaboré par [Crabbé, 2005]. Il revient ensuite sur le compilateur de métagrammaire de Candito, qui est utilisé pour la description de nos familles à nom prédicatif en explorant certains problèmes qui se posent au descripteur et qui n'ont pas forcément été mis en avant par [Candito, 1999a].

Quelques problèmes représentationnels sont ensuite abordés et le concept de dimension 4, dimension permettant de rendre compte de l'ordre entre constituants argumentaux, est alors introduit.

2.1 Aperçu général

2.1.1 Motivations

Les grammaires TAG offrent une relative souplesse qui leur permet de disposer de domaines d'informations bien différenciés. Ainsi, elles manient :

- une base d'arbres élémentaires non instanciés, classés par familles ;
- une base de lemmes ; chaque lemme étant associé à une ou plusieurs familles ;
- et une base morphologique dans laquelle chaque forme fléchie est associée à un lemme.

Malheureusement, les arbres regroupés au sein des familles ne disposent d'aucune organisation interne ; et comme le remarque [Candito, 1999a], "représenter l'ensemble des schèmes¹ sans aucune organisation revient à nier à peu près toute régularité syntaxique à l'intérieur de la phrase simple".

De plus, comme nous l'avons mentionné, quiconque est amené à travailler à l'implémentation d'une grammaire électronique se trouve face à des problèmes de plus en plus conséquents au fur et à mesure que sa grammaire s'étend ; et les grammaires TAG n'échappent malheureusement pas à cette constatation, car il s'agit de :

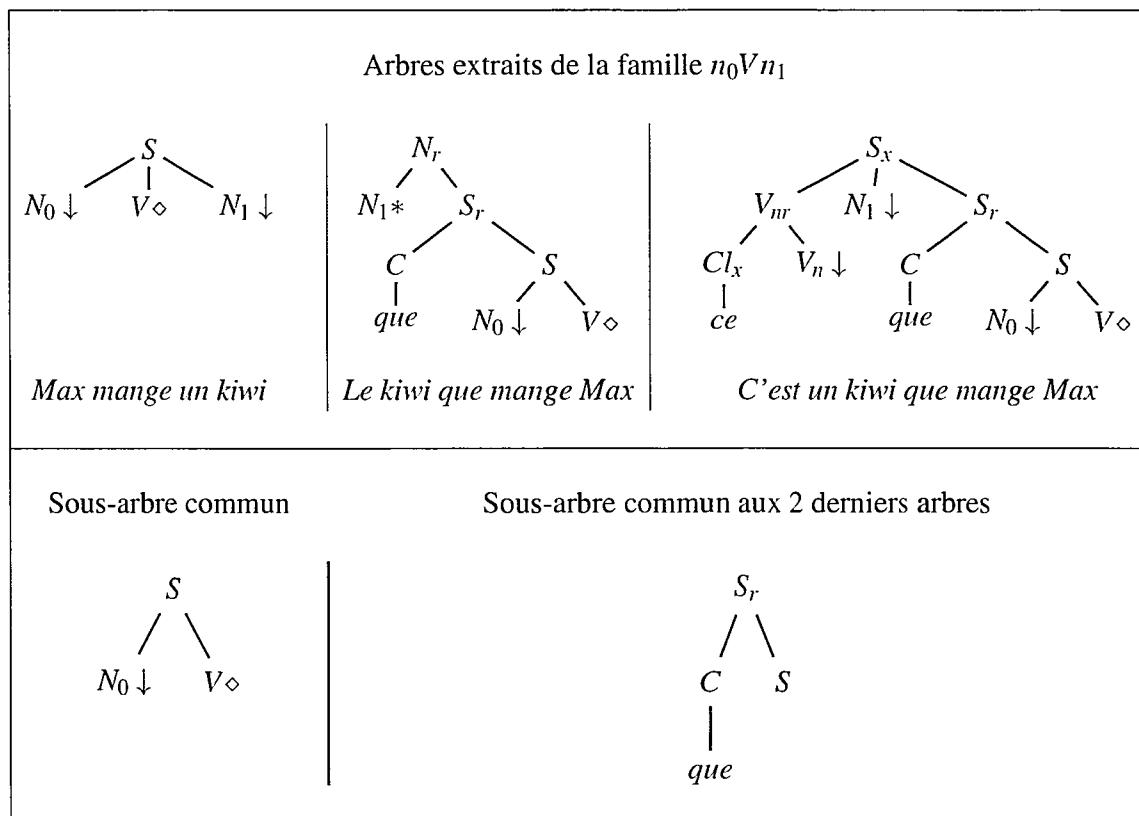
- rester cohérent au niveau des choix linguistiques opérés
- énumérer les nombreuses structures syntaxiques existantes, sans en oublier
- effectuer des maintenances, des modifications qui s'intégreront aisément et sans heurt à la grammaire existante

En particulier, il faut veiller à ce qu'un changement ou un ajout ne viennent pas rendre inutilisable une grammaire, tout en ayant à l'esprit l'ensemble de la grammaire d'origine. Tout changement aussi minime soit-il pourrait en effet nécessiter d'effectuer manuellement des mo-

¹Un schème est un arbre élémentaire dont l'ancre lexicale n'est pas instanciée. On parle également de schéma d'arbre, ou plus généralement d'arbre. Dans la littérature anglophone, le terme correspondant est "tree sketch".

difications dans des centaines d'arbres.

De façon analogue, et cela rejoint partiellement les constatations précédentes, il faut également pouvoir éliminer une grande partie de la redondance des informations fournies au niveau de la grammaire, car bien souvent ce sont les mêmes structures qui sont utilisées. Si l'on considère par exemple les 3 arbres ci-dessous, on se rend facilement compte de la répétition au sein de la description. Ici, le sous-arbre renseigné par la règle $S \rightarrow N_0 \downarrow V_{\diamond}$ est utilisée à chaque fois, alors que le sous-arbre renseigné par les règles $S_r \rightarrow C S$ et $C \rightarrow \text{que}$ est utilisé pour les deux derniers arbres.



Enfin, outre ces faits qui concernent davantage le linguiste que la machine, on peut également signaler que plus une grammaire s'étend, plus elle nécessite un espace mémoire important pour la stocker. Toutes ces constatations nous amènent à chercher un moyen de synthétiser l'information et de la rendre plus fiable.

Nous allons donc maintenant explorer une partie des solutions qui ont été proposées pour palier ces problèmes majeurs.

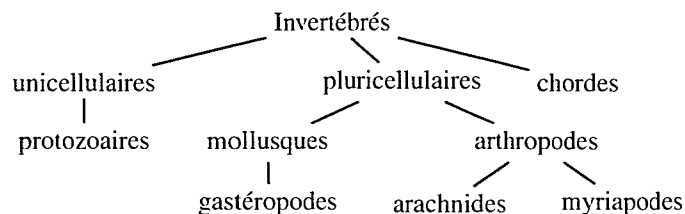
2.1.2 La proposition de Vijay-Shanker et Schabes

Si l'on observe les arbres que nous venons de fournir, on constate qu'il peuvent être ancrés par plusieurs "mots" (par exemple, les verbes transitifs comme *manger*; *aimer*; *construire*...). Il s'agit alors de se demander comment représenter le lien qui unit un mot et un arbre. [Vijay-Shanker et Schabes, 1992] proposent d'organiser les différentes entrées lexicales de manière hiérarchique en utilisant le concept d'héritage multiple²

Le principe est le suivant : chaque élément du lexique est associé à une classe, classe elle-même définie par héritage multiple d'autres classes renseignant des arbres élémentaires partiellement définis. Pendant l'héritage, l'information n'est jamais perdue ou modifiée.

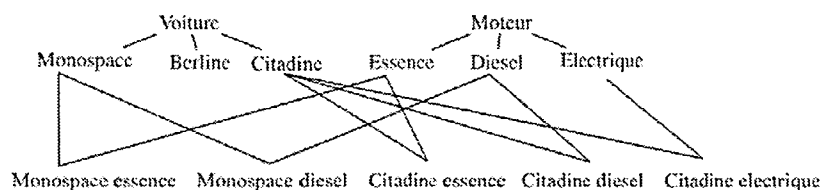
Nous illustrerons notre propos par la hiérarchie, très simple, suivante. Tous les arbres élémentaires ancrés par un verbe comprennent une branche $S \triangleleft *V$ (classe VERBE). Parmi

²Un héritage peut être simple ou multiple. Si l'héritage est simple, une classe (qui désigne un ensemble de propriétés) peut être dérivée d'au plus une autre classe et le graphe d'héritage est alors un arbre. Avec ce principe, deux types A et B représentés par des classes sont forcément apparentés de l'une des 3 façons suivantes : soit A et B n'ont pas d'éléments en commun, soit A contient tous les éléments de B ou soit B contient tous les éléments de A. La hiérarchie suivante représentant un fragment de la classification des invertébrés, représente un héritage simple :



Extrait de la classification des invertébrés (héritage simple)

Si l'héritage est multiple, deux classes A et B peuvent avoir des éléments en commun sans que A et B soient apparentées par héritage. Il suffit qu'il existe une classe C, qui hérite, même indirectement, à la fois de A et de B. Dans ce cas, les éléments de B sont communs à A et B. La hiérarchie suivante illustre ces points. Une voiture est considérée selon sa forme et sa motorisation :

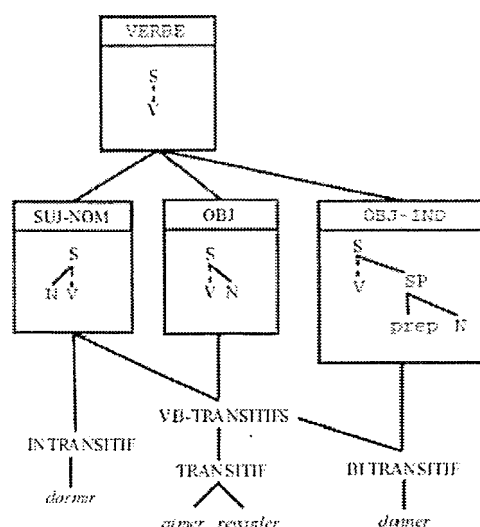


Extrait de la classification des véhicules automobiles (héritage multiple)

Dans tous les cas, la décomposition en classes et sous-classes permet de factoriser des caractéristiques communes : il est inutile de répéter dans une sous-classe une information qui a déjà été renseignée dans une classe mère.

ces verbes, on distingue ceux décrits par des arbres élémentaires possédant une branche $S \triangleleft N$ à gauche de $S \triangleleft *V$ ³ (c'est-à-dire les verbes possédant comme argument un sujet nominal : il s'agit de la classe SUJ-NOM), des arbres élémentaires ayant une branche $S \triangleleft N$ à droite de $S \triangleleft *V$ (c'est-à-dire les verbes ayant un objet nominal comme argument : il s'agit de la classe TRANSITIF), des arbres élémentaires ayant une $S \triangleleft SP$ à droite de $S \triangleleft *V$ (c'est-à-dire les verbes ayant un objet indirect comme argument).

En terme de description arborescente, les classes (INTRANSITIF), (TRANSITIF) et (BITRANSITIF) ne contiennent rien de plus que leurs parentes. Ainsi par exemple, la classe (TRANSITIF) hérite directement des informations contenues par la classe (VB-TRANSITIFS), classe héritant elle-même, des classes (SUJ-NOM), représentant la branche du sujet, et (OBJ-NOM), représentant la branche de l'objet direct.

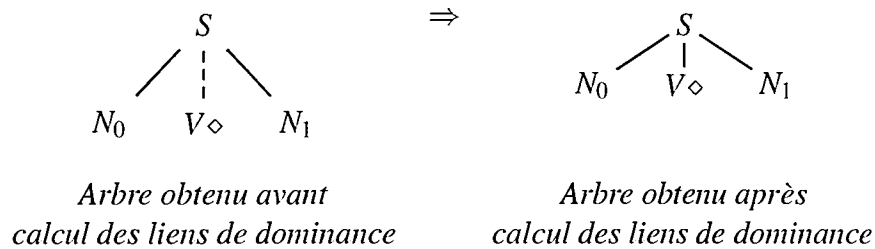


Exemple d'organisation d'un lexique verbal

Après héritage, la classe (TRANSITIF) contient donc la description arborescente suivante (les liens de dominance large sont alors totalement spécifiés⁴) :

³Les symboles $\triangleleft$ et $\triangleleft *$ désignent respectivement la dominance immédiate et la dominance "large".

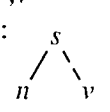
⁴La spécification des liens de dominance sera expliquée au cours des sections suivantes.



Au sein d'une hiérarchie, chaque classe est entre autres définie par les attributs suivants :

- super-classes : qui représente l'ensemble des classes immédiatement en amont dont la classe courante hérite ;
- nœuds (*nodes*) : qui représente la liste des nœuds associée à la classe ;
- description : qui fournit la description partielle de l'arbre associé à la classe. Elle renseigne la dominance immédiate, la dominance et la précédence linéaire ;
- équations contraintes (*constraint equations*) : qui fournit les équations représentant les contraintes d'unification sur les traits portés par chaque nœud ;
- précédence linéaire (*linear precedence*) : qui permet de fixer un ordre entre différents groupes ;
- ancre (*anchor*) : qui indique quel nœud sera le nœud ancre dans l'arbre décrit par la classe.

Ainsi, la classe (SUI-NOM) est-elle décrite plus formellement de la façon suivante ⁵

SUI-NOM
Super-classe : VERBE
Nœuds : s, n, v
Description : 
Equations : $n. < cat > = N,$ $n. < accord > = v. < accord >$...
Nœud argument : n
Précédence : $n < v$
ancre = v

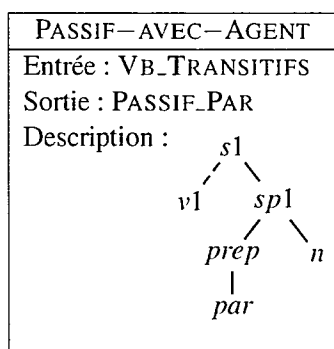
Description de la classe (SUI-NOM)

⁵Notons au passage que s, n et v désignent des variables de nœuds et non des catégories syntaxiques.

Pour récupérer l'arbre élémentaire associé à l'élément lexical, il suffit alors de :

1. sélectionner la classe associée à l'élément lexical
2. récupérer les descriptions partielles d'arbres des super-classes définies
3. identifier les nœuds de type "ancree" dans chaque description
4. déterminer l'arbre correspondant à la conjonction des descriptions

Reste que cette hiérarchisation n'est valable que pour des cas "standards" ou non marqués. Pour effectuer certaines transformations, comme par exemple un changement de diathèse, mais aussi une extraction, une inversion, on a recours à des règles lexicales, qui associent une entrée lexicale à une sortie lexicale. La règle suivante permet ainsi à partir de la classe (TRANSITIF) de construire l'arbre élémentaire du passif avec agent (avec la description fournie) et génère la classe (PASSIF-PAR).



Règle lexicale pour la construction du passif avec agent

En tant que telles, ces structures syntaxiques sont donc extérieures à la hiérarchie. Elles semblent cependant assez complexes dans la mesure où elles rendent compte à la fois des changements de diathèse et des "positions marquées". L'ordre dans lequel il faut appliquer ces règles n'est d'ailleurs pas anodin. Néanmoins, la représentation de l'information permet une certaine systématisation. Systématisation bien appréciable si l'on considère les problèmes posés au départ.

2.1.3 Compilateur et hiérarchie

2.1.3.1 La métagrammaire de Candito

[Candito, 1996, Candito, 1999a] s'appuie sur cette proposition pour réaliser ce qu'elle appelle une Métagrammaire :

”il ne s’agit pas (...) d’un ensemble de règles pour décrire des énoncés de la langue, mais d’un ensemble de règles pour décrire des règles (...), elles-mêmes utilisables par des règles combinatoires (...) pour décrire les énoncés de la langue.” [Candito, 1999a] p.33.

Tout en gardant à l’esprit les enjeux importants décrits dans les sections précédentes (simplicité d’écriture, facilité de mise à jour, cohérence), Candito définit 3 principes de base pour sa Métagrammaire. Cette dernière doit :

- reposer sur des principes linguistiques de hiérarchisation de structures syntaxiques
- représenter de manière modulaire et paramétrable les schèmes d’une TAG
- permettre de générer l’ensemble de ces schèmes

Pour cela, Candito définit une hiérarchie en 3 dimensions, dont les caractéristiques sont les suivantes :

- La dimension 1 (représentant la sous-catégorisation initiale) permet de renseigner le cadre de sous-catégorisation initiale pour une famille donnée.
- La dimension 2 (représentant les redistributions des fonctions syntaxiques) permet de rendre compte des différents changements de diathèse.
- La Dimension 3 (représentant les réalisations de surface d’une fonction syntaxique) permet d’exprimer les différentes réalisations possibles pour une fonction syntaxique particulière. Pour cette dimension, ”l’idée est de représenter le plus séparément possible la réalisation d’une fonction de celle d’éventuelles autres fonctions de la sous-catégorisation finale”.

Les règles lexicales, telles que décrites par [Vijay-Shanker et Schabes, 1992] disparaissent et sont prises en charge par les dimensions 2 et 3. En effet, une fonction syntaxique est décrite au sein de la MG par toutes ses réalisations possibles. Cela permet ainsi de rendre compte sans règle particulière, d’une extraction, ou d’une cliticisation. En outre, les différents changements de diathèse, sont traités en dimension 2. Les classes de cette dimension viennent donc modifier la sous-catégorisation déclarée en dimension 1 et les nouvelles fonctions assignées sont alors renseignées par la dimension 3.

Toutes les classes ainsi déclarées sont définies par le même type d’attributs que ceux utilisés par [Vijay-Shanker et Schabes, 1992], à ceci près, qu’il est impossible chez Candito d’avoir recours à la précedence linéaire seule⁶.

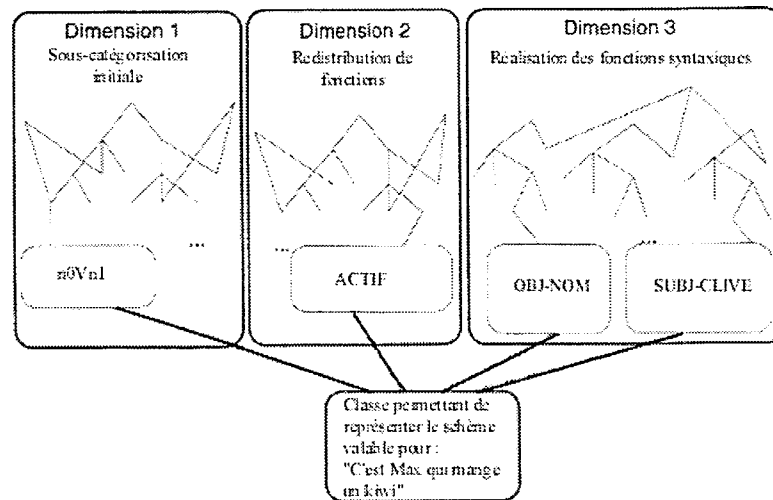
Le calcul de l’ensemble des schèmes s’effectue en deux phases. Lors de la compilation, la hiérarchie est augmentée de classes croisées, générées automatiquement par croisements systématiques de phénomènes linguistiques :

- Une classe croisée hérite d’une classe terminale de dimension 1
- Puis, d’une classe terminale de dimension 2
- Puis d’autant de classes de dimension 3, représentant les réalisations de chaque fonction de la sous-catégorisation finale

⁶Cette dernière s’exprime obligatoirement en rapport avec la dominance.

Par la suite, chaque classe croisée est traduite en un ou plusieurs schèmes (satisfaisant toutes les contraintes d'héritage)

Bien sûr, il reste possible d'intervenir au niveau des croisements. Il est ainsi possible de signifier que certaines classes sont incompatibles⁷ ou au contraire imposent la présence d'une classe particulière⁸. Il est aussi possible de spécifier que certaines sous-classes d'une classe sont mutuellement exclusives⁹.



Exemple de génération d'une classe croisée

En outre, le compilateur a été amélioré de façon à fournir également une structure de traits de type hypertags (voir [Kinyon, 2000, Barrier et Barrier, 2004]). Les hypertags résument les phénomènes linguistiques encodés dans un arbre particulier. Avec un parseur adéquat, une telle structure peut servir de base à un filtrage lors du processus d'ancrage, étant donné qu'elle permet d'atteindre des ensembles d'arbres plus fins que les familles TAG classiques.

L'exemple suivant illustre pour chaque arbre la structure de traits qui lui est associée¹⁰ pour la famille n0Vn1 (actif et passif) :

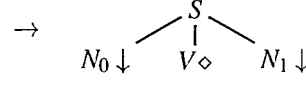
⁷Par exemple, pour une morphologie nominale, on ne réalisera pas de clitique.

⁸Par exemple, un sujet nominal inversé n'est possible que dans le cas d'une extraction d'un complément

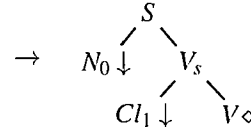
⁹Par exemple, en français, une seule extraction (interrogative, relative, clivée) est possible pour un même prédicat.

¹⁰Suivant les configurations arborescentes, un même hypertag peut correspondre à plusieurs arbres.

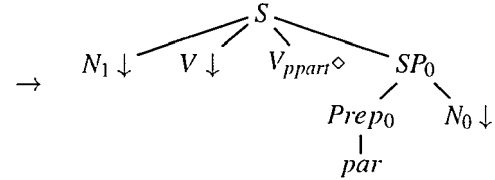
FAMILY :	n0Vn1
TREENAME :	Actif-n0Vn1-sujet_nom_can-obj_can
ANCHORS :	V
ARG.0 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Sujet} \\ \text{FINAL_FUNCTION : Sujet} \\ \text{CAT : N} \end{array} \right]$
ARG.1 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Objet} \\ \text{FINAL_FUNCTION : Objet} \\ \text{CAT : N} \end{array} \right]$
REDISTRIBUTION	Actif
SPAN :	N0:subst V:anchor N1:subst



FAMILY :	n0Vn1
TREENAME :	Actif-n0Vn1-suj_nom_can-obj_cll
ANCHORS :	V
ARG.0 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Sujet} \\ \text{FINAL_FUNCTION : Sujet} \\ \text{CAT : N} \end{array} \right]$
ARG.1 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Objet} \\ \text{FINAL_FUNCTION : Objet} \\ \text{CAT : Cl} \end{array} \right]$
REDISTRIBUTION	Actif
SPAN :	N0:subst Cl:subst V:anchor



FAMILY :	n0Vn1
TREENAME :	Passif-n0Vn1-suj_nom_can-par_obj_can
ANCHORS :	Vppart par
ARG.0 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Sujet} \\ \text{FINAL_FUNCTION : Par-objet} \\ \text{CAT : N} \end{array} \right]$
ARG.1 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Objet} \\ \text{FINAL_FUNCTION : Sujet} \\ \text{CAT : N} \end{array} \right]$
REDISTRIBUTION	Passif
SPAN :	N0:subst V:subst Vppart:anchor par N1:subst



Pour sélectionner certains arbres, il suffit alors de fournir une structure à unifier avec les hyper-tags générés. Ainsi, la première sous-structure sélectionnera tous les arbres, alors que la seconde ne sélectionnera que l'arbre du passif.

ARG.0 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Sujet} \end{array} \right]$	ARG.0 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Sujet} \\ \text{FINAL_FUNCTION : Par-objet} \end{array} \right]$
ARG.1 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Objet} \end{array} \right]$	ARG.1 :	$\left[\begin{array}{l} \text{INIT_FUNCTION : Objet} \\ \text{FINAL_FUNCTION : Sujet} \end{array} \right]$

Les hypertags présentent par ailleurs d'autres intérêts : ils permettent notamment une meilleure gestion d'une grammaire. Il est ainsi possible d'extraire des "sous-grammaires", de réaliser des interfaces de description du lexique (voir [Kinyon, 2000]).

2.1.3.2 Le point de vue de Gaiffe

Dans les années qui suivirent la création du compilateur de Candito, apparut un projet parallèle de Métagrammaire, développé par [Gaiffe et al., 2002]. L'idée était d'explorer d'autres approches ou de compléter certains points laissés de côté par [Candito, 1999a]. Il s'agissait de :

- récolter des structures de traits pour faciliter l'ancrage lexical
- généraliser le concept de Métagrammaire à plus de 3 dimensions
- pouvoir expérimenter une Métagrammaire avec d'autres notions que celles de fonctions syntaxiques
- pouvoir paramétrer le type des langages de description

Le système actuel possède de nombreux points en commun avec celui de Candito, mais tâche d'être plus générique. La dimension 3 qui n'était chez Candito qu'une collection de réalisations (du sujet, de l'objet...) est désormais scindée en autant de dimensions que de fonctions nécessaires. C'est principalement la notion de croisement qui est revue en partant de la constatation que :

- certaines classes doivent être croisées avec d'autres car elles sont incomplètes
- certaines classes ne doivent pas être croisées avec d'autres car elles sont incompatibles

En fait, une classe incomplète et qui doit donc être croisée avec une autre classe est une classe qui apporte ou qui a besoin de quelque chose. Ces besoins sont hérités tout au long de l'héritage (s'ils ne sont pas rendus nuls par une ressource), de même que les ressources (si elles ne sont pas rendues nulles par un besoin).

Pour effectuer un croisement : on cherche tous les croisements équilibrés en besoins et ressources des classes finales. Un besoin (respectivement une ressource) annulé ne peut réapparaître. Les arbres élémentaires ne sont générés que pour les classes finales équilibrées.

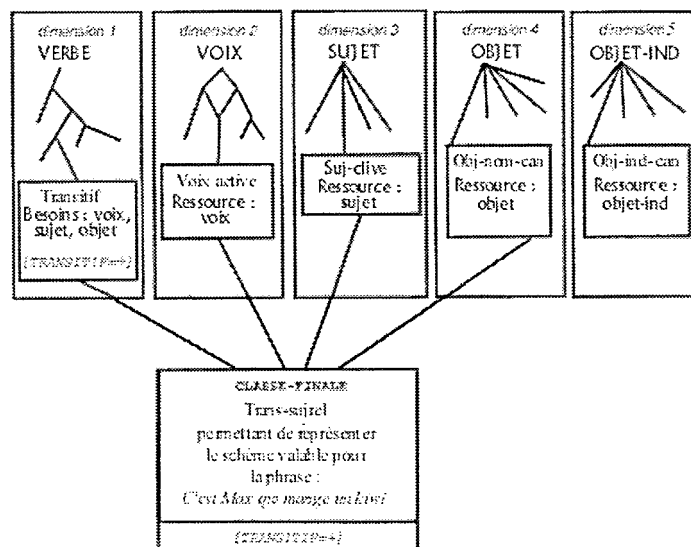
Il reste cependant à contraindre certains croisements. Ceci est réalisé par l'emploi de structures de traits hypertag, qui sont unifiées lors du croisement de deux classes finales. Si les différentes valeurs de traits sont compatibles, les classes seront croisées, sinon le croisement sera interdit ¹¹

¹¹Cela permet par exemple d'interdire les passifs pour les verbes intransitifs ou les verbes de mesure. On crée un trait [TRANSITIF -/+] et un trait [PASSIF -/+].

Les classes de la hiérarchie sont par ailleurs définies par relation d'héritage multiple. Elles reprennent le type d'attribut déjà énuméré pour [Vijay-Shanker et Schabes, 1992] et [Candito, 1999a], et ajoutent les attributs suivants :

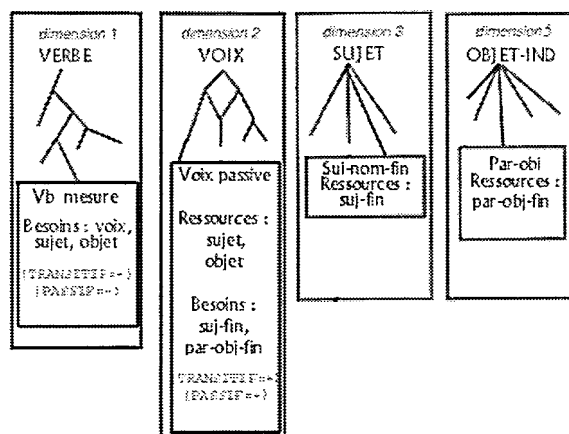
- un ensemble de besoins
- un ensemble de ressources
- une structure de traits hypertag

L'exemple ci-après illustre un croisement réussi pour représenter un schème pour un verbe transitif, dont les besoins sont un sujet, un objet et également une voix. Le sujet se réalise sous la forme d'un sujet clivé, l'objet se réalise de manière canonique, et la voix indique un actif.



Exemple de génération d'une classe croisée

Ci-dessous, un croisement non autorisé pour un passif (les valeurs du trait PASSIF n'étant pas les mêmes pour les classes de dimension 1 et 2). Noter au passage que le passif annule les besoins en sujet et objet de la classe (vb_mesure) et crée de nouveaux besoins (sujet final et complément d'agent).



Exemple de croisement non autorisé

2.1.3.3 Le point de vue de Crabbé

Si Benoit Crabbé faisait partie intégrante du projet précédent, il propose néanmoins dans sa thèse une nouvelle prise en compte des Métagrammaires (voir [Crabbé, 2005]) en s'éloignant du modèle présenté ci-avant. Les hypothèses de départ restent cependant les mêmes : il faut capturer deux types de généralisations. D'une part, les généralisations de structure, qui doivent permettre une factorisation et une réutilisation de descriptions grammaticales partielles ; d'autre part les généralisations d'alternatives dont l'objectif est de "capturer les réalisations alternatives d'une même notion".

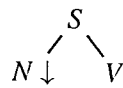
Le langage de description utilisé par Crabbé comporte deux mécanismes : un langage de contrôle, construit autour de la notion de classe, et une famille de langages de description de structures arborescentes qui permet de décrire des structures grammaticales partielles.

La notion de dimension n'étant pas explicitement présente, il faut en plus de la conjonction, qui va permettre l'assemblage de fragments arborescents, utiliser la notion de disjonction, qui permettra d'exprimer les différentes variantes alternatives de réalisation d'une même notion.

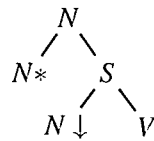
Plus précisément, le langage de description utilise les opérations suivantes :

- Nommage : le langage permet de nommer une description. En effet, associer un nom à une description permet de la réutiliser par la suite. Ainsi, par exemple, on associe à un nom (à gauche), une description correspondante (à droite).

Sujet canonique →



Sujet relativisé →

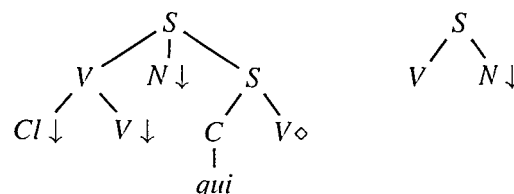


Forme verbale →

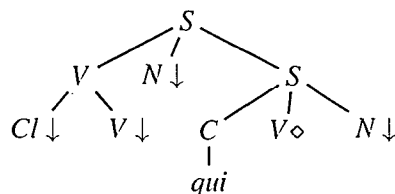


- Réutilisation : il est possible dans une description donnée de réutiliser une description déjà associée à un nom par ailleurs. Ainsi, par exemple, on réutilise la "brique" *Forme verbale* dans la description de *Verbe actif* :
Verbe actif → Forme verbale
- Alternatives : elles sont exprimées en utilisant l'opérateur logique *ou* ∨. La notion de sujet recouvre ainsi la notion de *Sujet canonique* et de *Sujet relativisé* :
Sujet → Sujet canonique ∨ Sujet relativisé
- Composition : elle permet de combiner plusieurs descriptions et utilise l'opérateur logique *et* ∧. Ainsi, par exemple, une famille intransitive est construite à partir d'un *Sujet* et d'un *Verbe actif* :
Verbe intransitif → Sujet ∧ Verbe actif

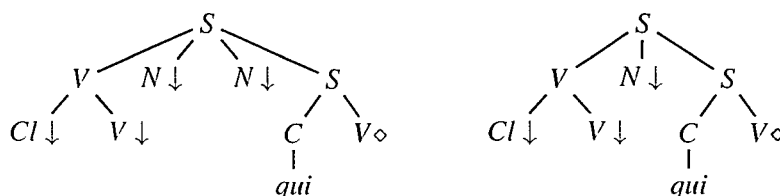
La principale nouveauté des propositions de [Crabbé, 2005] réside dans l'utilisation d'un principe de coloration qui a pour ambition de faciliter en pratique l'écriture d'une grammaire "en évitant l'obtention de modèles indésirables, souvent contre-intuitifs ou inattendus". En effet, bien souvent si les descriptions ne sont pas assez spécifiées, on obtient en plus du résultat attendu, d'autres descriptions non souhaitées. C'est le cas par exemple, de la conjonction des deux descriptions suivantes, qui va surgénérer :



On obtiendra bien l'arbre pour un sujet clivé avec un objet canonique réalisé (comme dans *C'est Max qui mange un kiwi*) :



mais également les deux arbres suivants, qui n'ont, bien entendu, aucun sens linguistique : le premier arbre inscrit en position clivée le sujet et l'objet ; alors que le second identifie sujet et objet.



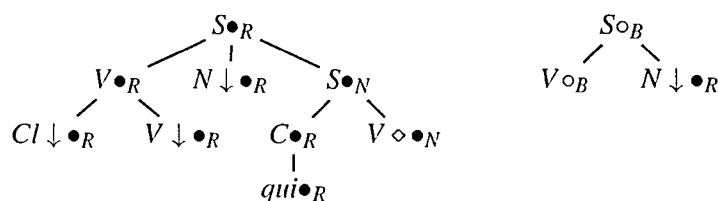
Pour remédier à ce problème, Crabbé utilise donc un langage de coloration des arbres qui comporte deux aspects :

- On rajoute au langage de base un nombre fini de propriétés (les couleurs *blanc*, *noir* et *rouge*). Chaque nœud de la description est alors associé à une couleur.
- On ajoute une contrainte de bonne formation des arbres engendrés : tout nœud d'un modèle est coloré soit en *rouge*, soit en *noir*

Lorsque l'on combine deux descriptions, celles-ci sont superposées. Les couleurs associées aux différents nœuds sont agencées selon la table de composition suivante¹² :

	$\bullet_N$	$\bullet_R$	$\circ_B$
$\bullet_N$	$\perp$	$\perp$	$\bullet_N$
$\bullet_R$	$\perp$	$\perp$	$\perp$
$\circ_B$	$\bullet_N$	$\perp$	$\circ_B$

Avec ce système, il est donc possible de résoudre le problème initial. Pour l'exemple des constructions clivées, il suffit de rajouter les informations de couleur suivantes :



¹²Le noir correspond à $\bullet_N$; le rouge à $\bullet_R$; et le blanc à $\circ_B$.

On obtiendra alors l'arbre souhaité, et l'on empêchera la génération des deux autres¹³.

Par rapport aux propositions originelles de Candito, les adaptations de Gaiffe et Crabbé essaient de simplifier l'écriture d'une métagrammaire. Mais en même temps, elles suppriment une partie des généralités linguistiques proposées. La vision hiérarchique tri-dimensionnelle n'est ainsi pas toujours conservée. Bien souvent, les auteurs souhaitent ne pas limiter le nombre de dimensions, a priori. Pourtant les choix de l'implémentation d'une grammaire de taille raisonnable sont toujours réalisés de manière à conserver une organisation tri dimensionnelle¹⁴.

L'utilisation de fonctions syntaxiques a par ailleurs été remise, indirectement, en cause, par [Gerdes, 2002] qui préfère avoir recours à la notion de champs topologiques pour sa grammaire de l'allemand.

Le désir de monotonie de Gaiffe et Crabbé semble légitime, d'un point de vue pratique. En effet, l'application successive de plusieurs redistributions n'est pas possible chez Candito, du fait même des redistributions. Néanmoins, présenter notre grammaire en terme de changement de diathèse nous paraît linguistiquement intéressant.

La communauté utilisant le compilateur de Gaiffe est plus importante que celle utilisant le compilateur de Candito. Cela tient, en partie, au choix du langage de programmation de Candito. Cette dernière avait en effet choisi d'implémenter son compilateur en LISP, sous Lucid Common Lisp, langage commercial au prix d'achat très élevé. A l'heure actuelle, le compilateur a été remanié et réimplémenté dans un langage LISP facilement disponible (cf. [Barrier, 2002]), mais il reste austère : il ne comporte pas d'éditeur graphique et manie constamment du LISP. Il faut cependant remarquer qu'il permet d'avoir un recul beaucoup plus important que ne l'offre le compilateur de Gaiffe¹⁵

A notre connaissance, le compilateur dessiné par Crabbé n'a été utilisé que pour le français. Si nous avons donc choisi d'utiliser le compilateur de Candito, c'est principalement pour des raisons pratiques qui ont fait que le compilateur était immédiatement disponible et éprouvé. Le fait que le compilateur repose sur des bases linguistiques nous semblait également important.

Reste que de nouveaux points de vue concernant les Métagrammaires ont vu le jour. [Villemonte de la Clergerie et Thomasset, 2005] avec la notion de nœuds gardés, tentent ainsi de synthétiser encore davantage l'information obtenue.

¹³Sans que cela soit expliqué, une partie de ce système semble exister chez [Candito, 1999a]. Il est en effet possible de déclarer qu'une variable ne doit pas s'unifier avec une autre. Dans ce cas, on retrouve donc l'utilisation de la couleur rouge.

¹⁴Voir [Gerdes, 2002] sur ce point, qui souhaiterait pouvoir "disposer d'outils plus généraux et plus puissants que les trois dimensions", et qui pourtant, présente sa grammaire DTAG, en terme de familles, diathèses et réalisations. [Crabbé, 2005] présente également sa description en termes tri dimensionnels.

¹⁵La première implémentation du compilateur de Candito date de 1996, alors que celle de Gaiffe date de 2001.

2.2 Le compilateur de Candito

2.2.1 Langage de description

Le compilateur de [Candito, 1999a] s'appuie plus ou moins directement sur les propositions de [Vijay-Shanker et Schabes, 1992] et [Rogers et Vijay-Shanker, 1994], mais s'en écarte de certains points. Ces derniers utilisent un langage utilisant les relations entre nœuds suivantes :

- $\triangleleft *$ (dominance large)
- $\triangleleft$ (dominance stricte)
- $<$ (précédence linéaire)
- $=$ (égalité)

[Candito, 1999a], quant à elle, constate que seule la relation de dominance est nécessaire : "grâce à la relation de dominance, on peut laisser sous-spécifié le chemin entre deux nœuds à un niveau général du graphe d'héritage, chemin qui sera éventuellement spécifié [par la suite] (...). On peut également laisser sous-spécifié l'ordre relatif de deux nœuds (...). On dispose ainsi de l'appareil formel pour dissocier dominance et précédence.". La précédence s'exprime alors chez Candito, uniquement en rapport avec la dominance, tout autre moyen étant exclu.

Les descriptions partielles d'arbres sont alors renseignées par des quasi-arbres¹⁶.

Les arbres sont écrits sous forme de liste, suivant 2 syntaxes, pour rendre compte respectivement de la dominance stricte et de la dominance large :

Dominance stricte :

$\langle \text{quasi} - \text{arbre} - \text{strict} \rangle := (\langle \text{racine} \rangle \langle \text{quasi} - \text{arbre} - \text{strict} \rangle *)$
 où $\langle \text{racine} \rangle := \text{symbol}$

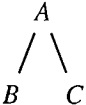
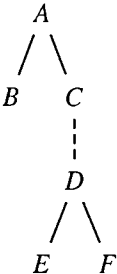
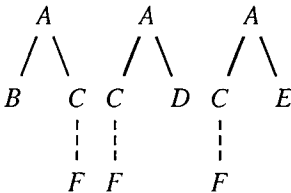
Dominance large :

$\langle \text{quasi} - \text{arbre} - \text{large} \rangle := (\langle \text{racine} \rangle \langle \text{quasi} - \text{arbre} - \text{large} \rangle *)$
 où $\langle \text{racine} \rangle := (\text{symbol} \text{booleen})$
 et $\langle \text{booleen} \rangle := \text{true} | \text{false}$
true indique un lien de dominance large
false indique un lien de dominance stricte

La figure suivante illustre quelques exemples de correspondances : le premier exemple renseigne une dominance stricte, le second une dominance large. Le dernier exemple est plus complexe et nécessite la description de 3 quasi-arbres. Il fait intervenir à la fois dominance (stricte et large) et précédence. Le choix de description des 3 quasi-arbres n'est pas unique¹⁷.

¹⁶Candito s'appuie sur la définition des quasi-arbres de [Rogers et Vijay-Shanker, 1994]. Un quasi-arbre a pratiquement tout d'un arbre, mais certaines paires de variables, au lieu d'être reliées par la précédence ou la parenté sont reliées uniquement par la dominance.

¹⁷On peut par exemple ne pas avoir recours à la branche $C \triangleleft *F$ pour le premier arbre.

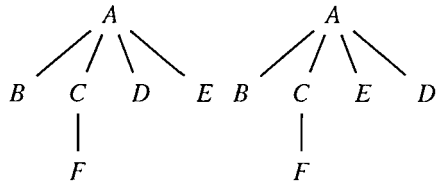
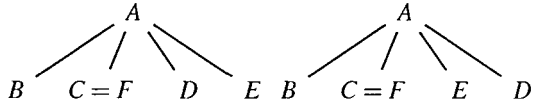
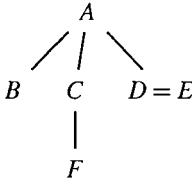
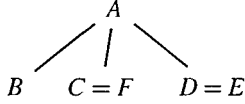
Vision arborescente	Langage de description de [Rogers et Vijay-Shanker, 1994]	Langage de description de [Candito, 1999a]
	$A \triangleleft B$ $A \triangleleft C$ $B < C$	$(A(B)(C))$
	$A \triangleleft B$ $A \triangleleft C$ $B < C$ $C \triangleleft *D$ $D \triangleleft E$ $D \triangleleft F$ $E < F$	$((Afalse)((Bfalse))$ $((Ctrue)((Dfalse)$ $((Efalse))((Ffalse))))$
	$A \triangleleft B$ $A \triangleleft C$ $A \triangleleft D$ $A \triangleleft E$ $B < C$ $C \triangleleft *F$ $F < D$ $F < E$	$((Afalse)((Bfalse))$ $((Ctrue)((Ffalse))))$ $((Afalse)((Ctrue)((Ffalse)))$ $((Dfalse)))$ $((Afalse)((Ctrue)((Ffalse)))$ $((Efalse)))$

2.2.2 Calcul des arbres

Pour passer de ces descriptions partielles à un arbre TAG bien formé, il faut renseigner totalement les liens de dominance ou de précédence laissés sous-spécifiés. Ces sous-spécifications peuvent être ou non modifiées par les descriptions partielles d'autres classes. Si elles ne le sont pas, il faut alors pour chaque couple de nœuds dans une relation de dominance large, créer le chemin le plus court possible : si les variables représentant les nœuds sont compatibles¹⁸ la longueur du chemin est ramenée à 0, à la longueur nécessaire sinon. L'ordre linéaire laissé sous-spécifié entre des nœuds, créera autant d'arbres que nécessaire.

¹⁸Le cas se présente rarement, car on fait généralement en sorte que les variables soient instanciées, à des valeurs différentes.

Pour reprendre notre troisième exemple, conjonction de 3 quasi-arbres, on obtiendra différents résultats suivants que les variables soient ou non compatibles :

Si toutes les variables sont incompatibles 
Si toutes les variables sont incompatibles, sauf celles représentant C et F 
Si toutes les variables sont incompatibles, sauf celles représentant D et E 
Si C et F sont compatibles et D et E compatibles 

2.2.3 Comportements déviants

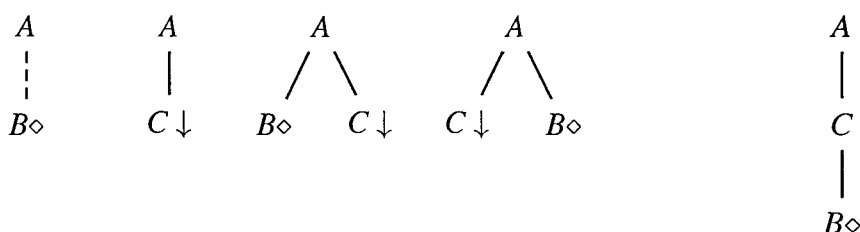
Nous abordons maintenant dans cette section, les différents comportements -pas toujours renseignés- qui obligent le descripteur à remanier sa hiérarchie pour obtenir le résultat attendu.

Le premier problème a été signalé par [Gerdes, 2002] : un nœud feuille n'ayant pas de valeur renseignée pour un attribut *qual* est par défaut considéré comme à substituer. Ce choix n'est pas justifié par Candito et semble totalement *ad hoc*, d'autant qu'on prend toujours soin de bien renseigner les nœuds feuilles¹⁹.

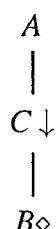
¹⁹Le fait de ne pas attribuer de valeur à l'attribut *qual* d'un nœud feuille, permet cependant d'éviter une révision

Toujours concernant les nœuds feuilles, un autre comportement gênant est à signaler : par définition ces types de nœuds ne peuvent avoir de fils. Or le compilateur passe outre ces conventions. Si un nœud est feuille - c'est-à-dire si c'est un nœud pied, une ancre ou un nœud à substituer - il peut recevoir un fils. Le compilateur supprime alors les attributs *qual* des anciennes feuilles, sans forcément pour autant les fournir aux nouvelles. En tous cas, on comprend mal pourquoi le compilateur réalise ces choix : si l'auteur d'une hiérarchie renseigne les différents attributs d'un nœud, ce n'est pas pour que ces informations soient supprimées, même si cela a pour but de rendre le nouvel arbre cohérent avec les conventions en vigueur en TAG. Si ce traitement n'était pas effectué, l'arbre généré ne répondrait même pas aux conventions d'écriture et serait donc normalement supprimé. L'exemple suivant illustre le problème : on croise la classe A avec la classe B (on suppose ici que les liens de dominance large ne sont pas ramenés à 0) :

Classe A Classe B Arbres souhaités et obtenus Arbre non souhaité mais obtenu



Si le dernier arbre devait être généré selon les spécifications de l'auteur de la hiérarchie, sans traitement particulier du compilateur, on obtiendrait :



Dans ce cas, il paraît encore plus clair que cet arbre ne devrait tout bonnement pas être généré par le compilateur.

D'autres comportements déviants du compilateur sont à signaler. Ainsi, certains croisements ne sont pas réalisés, alors qu'ils sont normalement licites. Prenons l'exemple simple suivant dans lequel on essaie de croiser la classe C1 dans laquelle les variables A et C sont instanciées (à leurs catégories respectives), mais pas la variable B, avec la classe C2 dans laquelle les variables A et C sont elles aussi toutes deux instanciées à leurs catégories respectives :

éventuelle de l'attribut. En pratique, on n'assiste qu'exceptionnellement à ce type de situation pour les familles des verbes pleins et il est possible de contourner le problème.

Classe C1	Classe C2	Croisement attendu	Résultat
$ \begin{array}{c} A \\ \vdots \\ B \\ \\ C \end{array} $	$ \begin{array}{c} A \\ \\ C \end{array} $	$ \begin{array}{c} A = B \\ \\ C \end{array} $	Aucun croisement obtenu

Plus étrange encore, le croisement de la classe C1, avec la classe C3 (version modifiée de C2) fournit deux arbres²⁰ :

Classe C1	Classe C3	Croisements obtenus	
$ \begin{array}{c} A \\ \vdots \\ B \\ \\ C \end{array} $	$ \begin{array}{c} A \\ \vdots \\ C \end{array} $	$ \begin{array}{c} A \\ \\ C \end{array} $	$ \begin{array}{c} A \\ \\ Nil \\ \\ C \end{array} $

Certains croisements laissent le compilateur totalement indifférent : aucun message d'erreur n'est fourni, ni même un résultat (il n'y a pas non plus de sortie prématurée du programme). C'est le cas par exemple avec les classes C4 et C5 (les variables B, C et D sontinstanciées à leurs catégories respectives - A ne l'est pas forcément) :

Classe C3	Classe C4	Résultat
$ \begin{array}{cc} & A \\ & \diagdown \quad \diagup \\ B & \quad C \end{array} $	$ \begin{array}{cc} & D \\ & \diagdown \quad \diagup \\ B & \quad C \end{array} $	Aucune manifestation du compilateur

Bien entendu, ces défauts de comportement ne suppriment en rien les qualités générales du compilateur. Il aurait pu cependant être souhaitable de trouver chez [Candito, 1999a] une discussion informatique plus développée qui permette de comprendre à quoi sont exactement dûs ces comportements douteux²¹.

²⁰Pour peu qu'on n'y prenne garde, le compilateur a tôt fait de créer de nœuds à Nil.

²¹On trouve très peu d'informations concernant le choix du langage de programmation. Le choix du langage de description est à peine plus renseigné, et la syntaxe d'une commande est rarement mise en valeur, de sorte que l'on ne sait que rarement quelle peut en être l'utilité effective. Seul un algorithme général de croisement est renseigné, sans pour autant décrire ou expliciter les opérations qui le composent. Candito a davantage misé sur le côté théorique de son compilateur.

2.2.4 Des exemples concrets

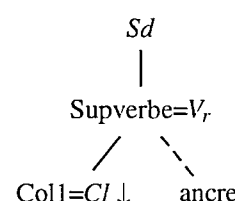
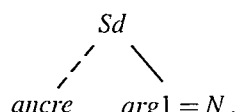
Nous fournissons maintenant un exemple réaliste et complet d'un fragment d'une hiérarchie pour les verbes pleins du français. On souhaite écrire un arbre pour la famille $n0Vn1$. La branche du sujet, réalise un clitique, et celle de l'objet un nom. Cet arbre permettra ainsi d'analyser une phrase comme *Il cultive des tomates*. Les classes (Dimension 1), (Dimension 2) et (Dimension 3) regroupent le comportement de chaque dimension²². Les classes (POSITION-SUJET) et (POSITION-OBJET) regroupent les différentes réalisations des différentes fonctions à représenter²³. Enfin, le clitique sujet se trouve à l'intérieur d'un nucleus verbal, ce qui explique que l'ancre verbale ne soit pas directement rattachée à la racine²⁴. On remarque que les variables désignant le même "objet" mais ayant un nom différent, sont reliées par leur fonction syntaxique. Aucune variable n'est donc inconnue pour le compilateur.

Dimension 1		
Les classes (SUJET-INITIAL) et (OBJET-INITIAL) héritent de (Dimension 1)		
La classe n0Vn1 hérite des classes (SUJET-INITIAL) et (OBJET-INITIAL)		
Contenu de (SUJET-INITIAL) Constante $arg0 = N_0, Cl_0, S_0$ fonction sujet	Contenu de (OBJET-INITIAL) Constante $arg1 = N_1, Cl_1, S_1$ fonction objet	(Contenu de n0Vn1) Constante $arg0 = N, Cl$ Constante $arg1 = N, Cl$
Après héritage la classe (n0Vn1) contient alors Constante $arg0 = N_0, Cl_0$ - fonction sujet Constante $arg1 = N_1, Cl_1$ - fonction objet		
Dimension 2		
La classe (MORPHO-VERBALE) hérite de (Dimension 2)		
La classe (ACTIF) hérite de la classe (MORPHO-VERBALE)		
Contenu de (MORPHO-VERBALE) Constantes $Sd = S$, ancre = $V \diamond$	Contenu de (ACTIF) Sd ⋮ ancre	
Après héritage la classe (ACTIF) contient alors Constantes $Sd = S$, ancre = $V \diamond$ $Sd = S$ ⋮ ancre = $V \diamond$		

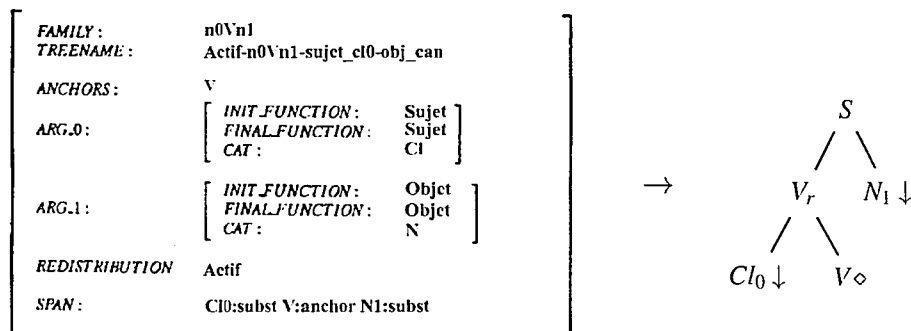
²²Leur contenu est vide.

²³Leur contenu est également vide.

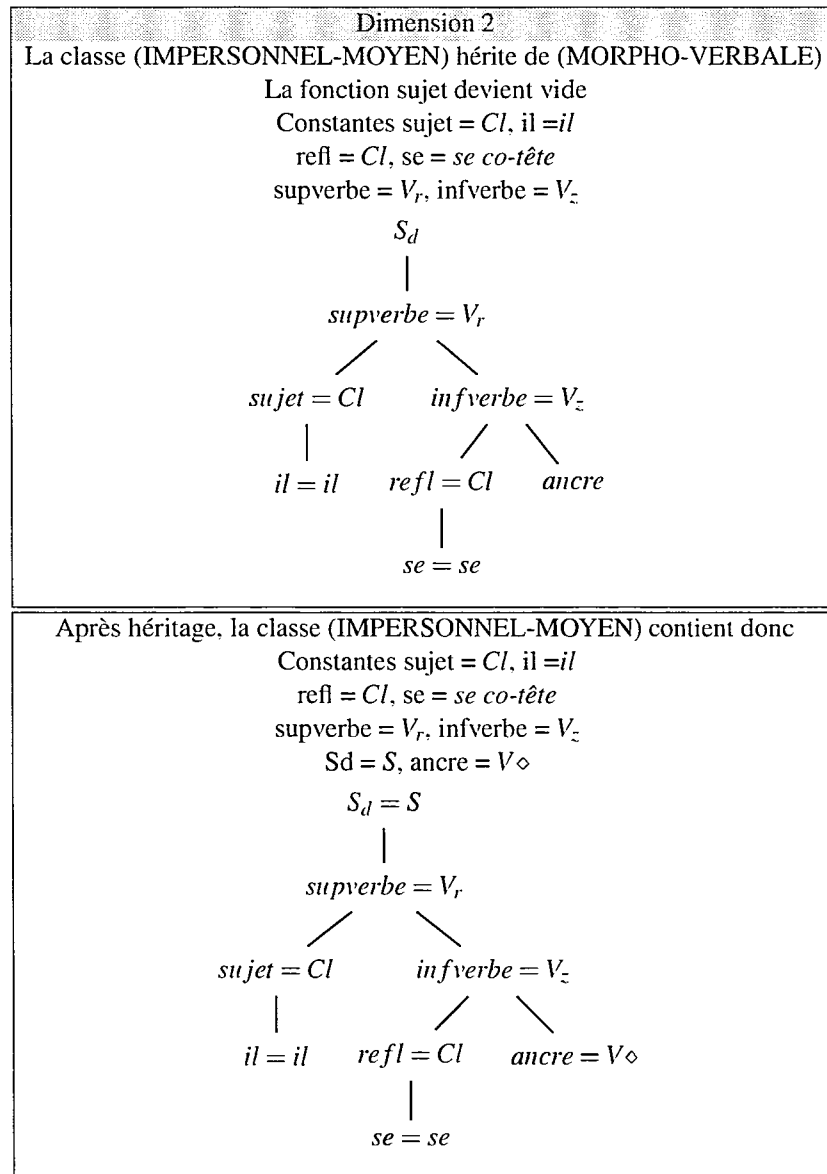
²⁴Par ailleurs, le clitique sujet est séparé des autres clitiques, rattachés quant à eux à un niveau inférieur.

Dimension 3	
La classe (POSITION-SUJET) hérite de (Dimension 3) La classe (SUJ-CLT) hérite de la classe (POSITION-SUJET)	
La classe (POSITION-SUJET) ne contient pas d'information nouvelle	Contenu de (SUJ-CLT) Constantes $\text{Supverbe} = V_r$, $\text{Col1} = Cl \downarrow$ - fonction sujet 
La classe (SUJ-CLT) reste inchangée après héritage	
La classe (POSITION-OBJET) hérite de (DIMENSION3) La classe (OBJ-NOM) hérite de (POSITION-OBJET)	
La classe (POSITION-OBJET) ne contient pas d'information nouvelle	Contenu de (OBJ-NOM) Constante $\text{arg1} = N \downarrow$ 
La classe (OBJ-NOM) reste inchangée après héritage	

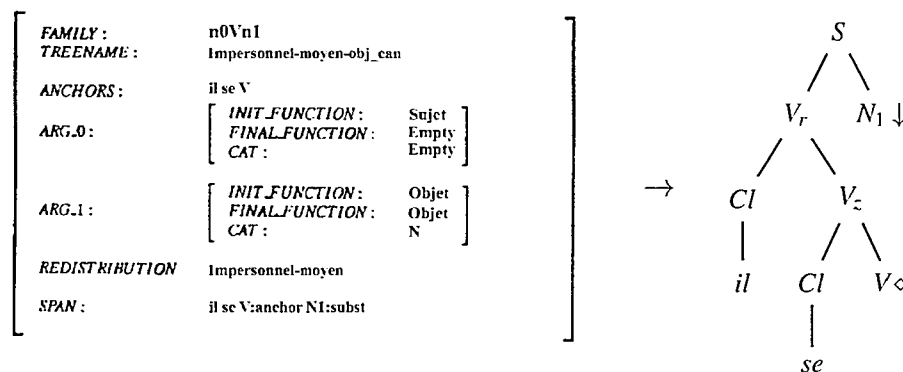
Le compilateur croise alors automatiquement les classes (n0Vn1), (ACTIF), (SUJ-CLT) et (OBJ-NOM), spécifie la relation laissée sous-spécifiée entre l'ancre verbale et la racine, et génère alors l'arbre suivant, auquel il associe également l'hypertag correspondant :



Faisons maintenant intervenir une redistribution pour représenter un arbre valable pour un impersonnel moyen. Ici, on a considéré que l'impersonnel moyen n'est pas dérivé du moyen, mais est un moyen où la promotion de l'objet n'a pas lieu (comme pour le passif impersonnel). La fonction sujet (initial) disparaît donc, alors que la fonction objet reste inchangée. Il nous faut définir une classe (IMPERSONNEL-MOYEN) capable de renseigner ce nouveau cas.



Le compilateur croise alors les classes (n0Vn1), (IMPERSONNEL) et (OBJ-NOM) et génère l'arbre suivant, permettant, par exemple de reconnaître la phrase *Il se cultive des tomates (en toutes saisons)*' :



On comprend vite avec ces exemples la nécessité d'une rigueur d'écriture. En particulier, il faut absolument veiller à ce que chaque variable soit instanciée. En outre, le développement de nouvelles classes doit se faire en respect des anciennes, ce qui n'est pas forcément toujours aisé.

2.2.5 Format de sortie du compilateur

Grâce aux divers modules de génération d'arbres, il est possible de choisir le format de sortie qui sera utilisé par le parseur. A l'heure actuelle, deux formats ont été prévus. Le premier, et le plus ancien, correspond au format LISP XTAG. Peu lisible, nous l'avons cependant utilisé au début de notre implémentation. A l'heure actuelle, le compilateur génère les arbres dans un second format, qui correspond au standard XML. Nous fournissons ci-après la sortie réalisée dans les deux formats pour notre premier exemple (Famille n0Vn1 - Actif - Sujet clitique et objet canonique réalisés). Les équations sur les traits ne sont ici pas mentionnées :

```
( "^BActif-n0Vn1-sujet_cl0-obj_can" :COMMENT-DISPLAY? NIL
  :FEATURE-DISPLAY? NIL
  :EQUATION-DISPLAY? NIL
  :UNIFICATION-EQUATIONS ""
  :COMMENTS "T_Traits T_Canonique"
  :SHAPE :NONE
  :BORDER-WIDTH 1
  :CONSTRAINT-STYLE (:SWISS :ITALIC :NORMAL)
  :CONNECTOR :LINE
  :DEFAULT-STYLE (:SWISS :BOLD :NORMAL)
  :SUBSCRIPT-STYLE (:SWISS :ROMAN :SMALL) :WHITE-SPACE 3
  :MINIMUM-NODE-SEPARATION NIL
  :LEVEL-SEPARATION NIL)
  (((("S" . "d"))) (((("V" . "r"))) (((("Cl" . "0")) :SUBSTP T))
    (((("V" . " ") :HEADP T)) ) (((("N" . "1")) :SUBST T)) )
```

Sortie au format XTAG

```

<tree name="Actif-n0Vn1-sujet_cl0-obj_can" type="init" >
<node id="S_d" cat="S" type="std">
  <node id="V_r" cat="V" type="std">
    <node id="Cl_0" cat="Cl" type="subst">
      </node>
    <node id="V" cat="V" type="anchor">
      </node>
    </node>
    <node id="N_1" cat="N" type="subst">
      </node>
  </node>
</tree>

```

Sortie au format XML

2.3 Problèmes de représentation

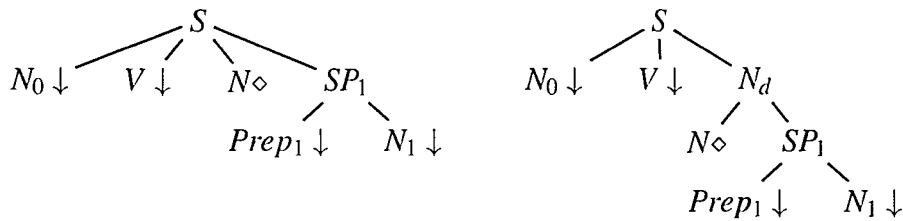
Examinons maintenant quelques difficultés que l’auteur d’une hiérarchie peut-être amené à rencontrer lors de son implémentation. Suivant les solutions qu’il retiendra, le nombre d’arbres obtenus différera. Dans tous les cas, la différence de nombre entre les arbres attendus et générés doit être le plus petit possible (idéalement 0).

2.3.1 Problèmes de dominance

Les liens de dominance large peuvent parfois se relever très problématiques. En effet, dès lors que plusieurs groupes ayant même père sont représentés, un constituant voisin peut devenir un lieu d’accueil pour l’élément dont la relation est laissée sous-spécifiée avec son père. Les représentations de dimension 3 sont directement concernées²⁵

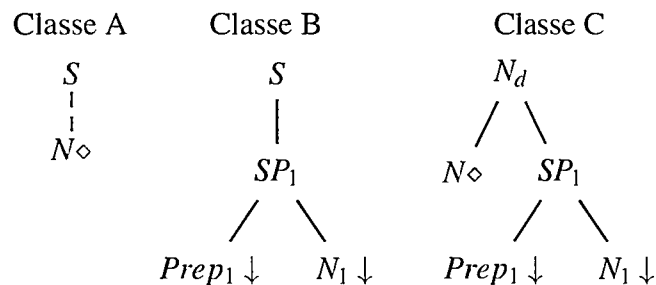
Pour illustrer notre propos, imaginons que l’on veuille rendre compte de la double analyse possible de la phrase *Max commet un crime contre Luc*. Il nous faut donc obtenir deux arbres différents :

²⁵Même si un lien est laissé sous-spécifié en dimension 2, il ne concerne qu’une tête syntaxique. Les représentations de dimension 3 mentionnent normalement le rapport à la tête syntaxique. Ce qui explique que seule la dimension 3 soit problématique.



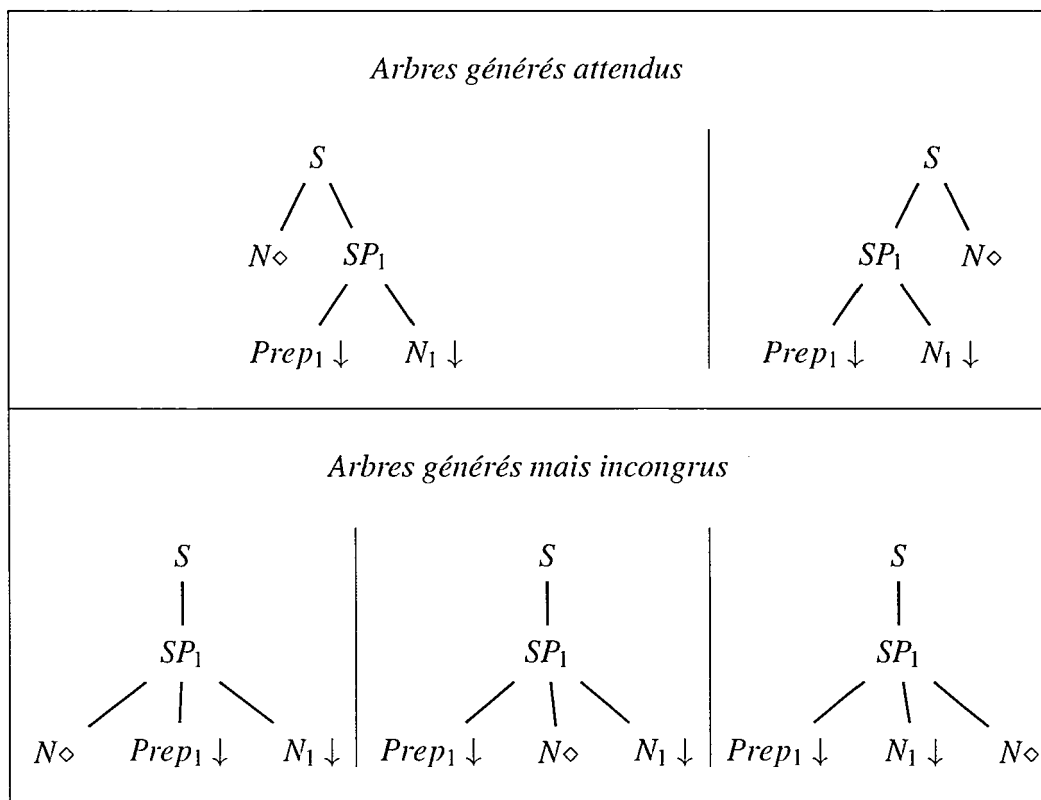
Deux arbres élémentaires pour Max commet un crime contre Luc

Nous nous intéresserons ici uniquement à la réalisation des compléments post-verbaux. L'observation de ces 2 arbres, nous permet de constater que le nom prédicatif ne peut être relié directement à S, puisqu'il peut se placer à l'intérieur du groupe nominal. Intuitivement (et peut être un peu naïvement), on est donc porté à créer 3 classes pour permettre la création des arbres souhaités :



Le croisement de la classe A avec la classe C, nous permettra d'obtenir la partie de l'arbre 2 souhaitée. Malheureusement, le croisement de la classe A avec la classe B, va engendrer des difficultés, car il va permettre la génération des arbres suivants²⁶ :

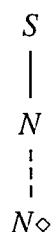
²⁶Deux autres arbres sont en outre créés. Ils correspondent à un comportement anormal du compilateur (voir sous-section 1.2.3).



Cette solution, bien qu'intuitive, n'est donc pas totalement appropriée, puisqu'elle permet l'obtention d'arbres non désirés. L'élément $N_{\diamond}$ de la classe A est amené à porter une fonction syntaxique, ce qui n'est pas toujours juste : s'il est notamment inclus dans un groupe nominal comprenant en outre un syntagme prépositionnel, c'est bien le groupe nominal qui doit porter une fonction et non sa tête.

Une deuxième solution est d'avoir recours à des méthodes écrites par Candito, mais non documentées. Il est en effet possible à l'aide de la commande *revise-qual* de modifier le qualificatif d'une variable. Ainsi, en supposant que la variable qui porte $N_{\diamond}$ et N_d soit unique, on peut modifier le contenu de la classe A en reliant directement S et $N_{\diamond}$. La classe C modifie alors la variable affectée à N_d en supprimant alors le $\diamond$ de la classe A. Les arbres générés sont alors les arbres attendus sans sur-génération. Même si elle est appréciable, cette solution n'est cependant guère séduisante, car elle n'est pas rigoureuse.

Une troisième solution est de créer au sein de la classe A, un nœud intermédiaire (portant la fonction syntaxique nécessaire) entre S et $N_{\diamond}$. Et ainsi de décrire la classe A suivante :



Dans ce cas, lors d'un croisement entre les classes A et B, le lien sous-spécifié est réduit à 0 et les variables identifiant N et $N_{\diamond}$ sont déclarées identiques, car compatibles. Dans le cas d'un croisement entre A et C, la variable identifiant N est compatible avec N_d . Au final, on obtient donc les deux arbres souhaités permettant de rendre compte de la double analyse. Cette solution semble donc la plus séduisante, même si elle exige une importante documentation pour être facilement compréhensible²⁷. En outre, le fait que la classe A comporte deux groupes N n'est pas forcément des plus intuitif.

Chacune des solutions exposées est envisageable mais présente des inconvénients. La première solution, si elle est retenue nécessitera des croisements contraints, croisements contraints dont les autres solutions n'auront pas besoin.

2.3.2 Problèmes d'ordre entre constituants

2.3.2.1 Description du problème et solutions apportées jusqu'à présent

Outre le problème de la typologie des arbres, que nous venons d'évoquer, et qui n'est d'ailleurs pas le plus fréquent, il demeure un autre souci essentiel : celui de l'ordre entre constituants. En français l'ordre des constituants est normalement libre, mais dans d'autres langues le problème se pose véritablement. Pour illustrer notre propos, nous appuierons sur l'exemple des verbes bi-transitifs anglais, cité par [Gerdes, 2002], exemple lui-même repris de [Candito, 1999a].

Pour de tels verbes, l'objet direct précède l'objet indirect avec la préposition *to*. On a ainsi :

- *Mary gives a book to Peter*
- mais pas **Mary gives to Peter a book*

Le problème qui se pose est que lorsque "la classe des verbes bitransitifs exige la réalisation d'un objet direct et d'un objet indirect, ces deux classes sont appelées indépendamment l'une de l'autre. (...) La classe de réalisation de l'objet indirect ne possède pas d'information sur la classe de l'objet direct" et inversement. K. Gerdes propose alors de faire porter à la classe de l'objet indirect, une information quant à la place de l'objet direct, en postulant que l'objet direct se trouve avant l'objet indirect. Lorsque les deux classes sont réalisées, tous les éléments sont donc décrits correctement et on évite ainsi **Mary gives to Peter a book*.

Seulement, "pour l'ergonomie de la représentation", K. Gerdes note qu'"il est (...) souhaitable de réutiliser ces classes dans un autre contexte" ; par exemple pour représenter la phrase *Mary will see to the car*²⁸. Dans ce cas, il n'y a pas d'objet direct, et l'élément le représentant n'apparaîtra pas. En théorie donc, la solution est pertinente, puisqu'elle permet la génération des arbres corrects, mais en pratique, qu'en est-il exactement ? Ce que l'on peut dire, et K. Gerdes le

²⁷Dès que l'on crée un lien de dominance large, on a tendance à penser que le lien ne sera pas nul. On pense plus facilement en terme de dominance propre que de dominance large.

²⁸Nous n'avons pas repris l'exemple initial *Mary dances to the music*, fourni par Gerdes. Il n'est en effet pas certain que le groupe prépositionnel soit dans cet exemple un complément.

remarque justement, c'est que "seule la classe d'un argument sait comment se placer par rapport à d'autres éléments, mais elle ne sait pas si ces éléments sont présents". Pourquoi alors ne pas faire porter ces informations uniquement lorsque les classes concernées sont toutes présentes ? Est-il si pertinent de vouloir parler d'un constituant dont on ne connaît même pas la présence et dont on n'est même pas sûr qu'il apparaîtra ? Pire encore, y a-t-il un sens à parler d'un objet direct, pour un verbe qui n'en attend pas ? K. Gerdes semble ne pas s'en soucier, étant bien plus occupé par le fait d'exprimer des généralités. En pratique, et pour une question de maintenance efficace, il est clair que cette solution, bien que séduisante, ne peut être retenue totalement²⁹.

En extrapolant, on peut tout de même s'y intéresser encore un peu dans le cas de l'ordre entre tête syntaxique et argument. Dans ce cas, l'ordre sera exprimé dans la classe de l'argument, et ce en rapport avec sa tête. Lorsque l'argument sera réalisé, le placement sera effectué correctement, et lorsque l'argument n'apparaîtra, le problème ne se posera même pas³⁰.

Résumons donc ce que nous avons jusqu'à présent énoncé :

- Bien qu'intéressante, la proposition de K. Gerdes concernant le placement des constituants entre arguments mérite d'être précisée, voire restreinte dans certains cas. On ne peut en effet pas porter d'hypothèse sur la réalisation d'un autre argument, même si un argument est réalisé.
- On ne peut pas forcément parler facilement d'ordre dans un cadre général de réalisation des arguments en dimension 3.

2.3.2.2 Ajout d'une quatrième dimension

Nous ne sommes pas les seuls à souligner ces faits. [Candito, 1999b] va aussi dans ce sens lorsqu'elle fait remarquer que "l'idée (de la dimension 3) est de représenter le plus séparément

²⁹[Gerdes, 2002] laisse penser que la solution qu'il propose n'est pas utilisable avec la Métagrammaire de Marie-Hélène Candito. Or, il n'en est rien. La MG dispose de l'appareillage théorique nécessaire pour obtenir les mêmes arbres bien formés qu'il propose pour son exemple (le "métalangage" d'écriture est certes différent. En particulier, la précédence ne peut pas s'exprimer chez M.H. Candito indépendamment des phénomènes de dominance). Ce qui semble gêner K. Gerdes, c'est que dans une classe croisée, après que l'unification de toutes les descriptions héritées ait été réalisée, tout nœud dont l'attribut : *qual* n'est pas spécifié sera considéré comme un nœud à substitution. Or ce comportement peut être facilement évité. Il suffit de déclarer que la variable n'a pas d'attribut : *qual*. Si par la suite, il fallait que cette variable retrouve un attribut, il suffirait d'utiliser la fonction *revise – qual*.

Pour plus de clarté reprenons un exemple, et supposons que l'on veut énoncer le fait qu'un *N*, objet direct, se trouve obligatoirement devant un *SP* (peu importe la fonction : objet indirect ou locatif). On décide arbitrairement de faire porter l'information d'ordre dans la classe de l'objet direct : on définit une variable pour le *SP*, sans l'instancier et on indique que cette variable n'a pas d'attribut : *qual*. Après croisement, si le *SP* est omis, la variable est supprimée, car elle n'a pas été nommée, si le *SP* n'est pas un nœud à substituer, le placement est celui décrit dans la classe du *N* objet direct.

Reste maintenant le cas où le *SP* représente un nœud à substituer. Dans ce cas, il y a collision entre une valeur *no – qual* et une valeur *subst*. Il faut donc utiliser la fonction *revise – qual* qui fait en sorte de modifier l'attribut précédemment défini : *no – qual* devient *subst* et l'arbre correct est ainsi généré.

³⁰Notons au passage que le choix de l'ordre dans une classe n'est pas anodin. Il vaut ainsi mieux faire porter l'information de placement dans une classe qui n'apparaît pas forcément.

possible la réalisation d'une fonction de celle d'éventuelles autres fonctions de la sous-catégorisation finale"; même si elle ajoute que "lorsque les réalisations de surface de plusieurs fonctions sont interdépendantes"³¹ [et le mot est ici *crucial*], une même classe terminale peut fixer la réalisation de plusieurs fonctions". Pourtant cela n'est pas techniquement possible en dimension 3³².

Nous proposons avec N. Barrier, une approche mixte plus restreinte qui reprend les points pertinents abordés, tout en évitant certains écueils (cf. par exemple [Barrier et Barrier, 2003]). Puisque la dimension 3 n'est pas forcément le lieu où il convient de parler d'ordre, nous créons une nouvelle dimension, adaptée à cet usage. En soi, la démarche semble naturelle : la dimension 3 concerne la réalisation des arguments et non le placement des arguments entre eux.

Principe L'idée est donc de créer une dimension spécialement destinée au placement ou à l'ordre des constituants. Comme l'algorithme de M-H. Candito ne considère que 3 dimensions, il faut trouver un moyen de relier cette nouvelle dimension avec celles existantes. En fait, sans en avoir explicitement fait mention, [Candito, 1999a] prépare la voie à cette quatrième dimension, avec ces **constrained-crossing**, qui rappelons-le, permettent de croiser une classe avec x autres, lorsque ce croisement n'est pas réalisé automatiquement.

Concrètement, qu'est-ce que cela permet ? Au lieu de signifier un ordre général dans une classe, on indique simplement un ordre uniquement lorsque toutes les classes comportant des éléments à ordonner sont concernées. L'avantage est de définir l'ordre en dehors de toute classe de dimension 3, et donc de gagner en lisibilité, l'effort d'écriture restant le même.

Pourquoi une autre dimension ? Nous avons vu jusqu'à présent des raisons théoriques qui nous amènent à postuler l'existence d'une quatrième dimension. Nous avons néanmoins très peu envisagé le côté pratique de l'écriture d'une Métagrammaire. La MG, telle qu'elle est décrite et considérée par le linguiste est un réseau d'héritage. Or, pour parler d'ordre entre constituants, la notion essentielle à utiliser n'est plus celle d'héritage, mais celle de croisement.

Il est souvent facile de confondre héritage et croisement lors de l'écriture d'une Métagrammaire. En effet, dans un cadre général, le concepteur s'emploie uniquement à définir un réseau d'héritage. Ce n'est qu'après compilation, qu'un croisement est réalisé automatiquement. Tout semble donc clair pour ne pas "mélanger" ces deux notions. Pourtant, le réseau ainsi défini s'avère en partie inefficace dans la mesure où il permet de générer des arbres sans aucun intérêt linguistique. Et c'est pour cela qu'il est possible d'interagir avec le programme pour lui indiquer quels croisements sont ou non licites.

A partir de ce point, il est facile pour un non spécialiste de confondre les deux notions³³. Il faut donc être particulièrement soigneux pour éviter cette confusion qui peut être fatale.

³¹ Elle cite ici l'exemple des clitiques

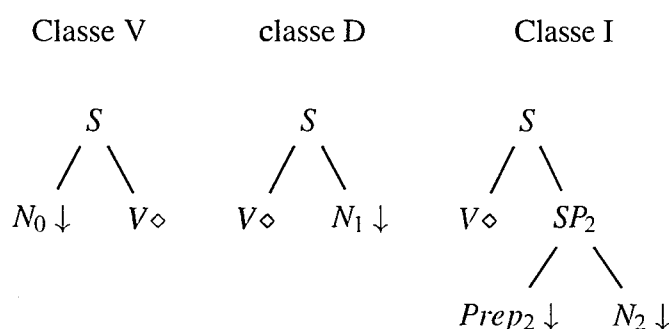
³² Par définition, une classe de dimension 3 réalise une seule fonction syntaxique et non plusieurs.

³³ Indiquer qu'une classe hérite des classes A et B ne signifie pas qu'elle va croiser avec A et B

Cette différence suffit pour nous à en faire une classe à part entière. Les phénomènes d'ordre ne concernent clairement ni la dimension 1 ni la dimension 2. La dimension 3 quant à elle n'est pas toujours adaptée.

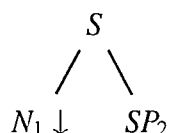
Si l'on veut donc être juste et strict, l'utilisation de la dimension 4 s'impose³⁴.

Pour finir, nous reprenons l'exemple fourni par [Gerdes, 2002]. Nous conservons donc les classes de dimension 3 qu'il a pris soin de définir, en y ôtant les notions d'ordre entre arguments.³⁵



Il nous faut ensuite définir une classe de dimension 4, indiquant l'ordre des éléments entre l'objet direct et l'objet indirect :

Classe (N<SP)

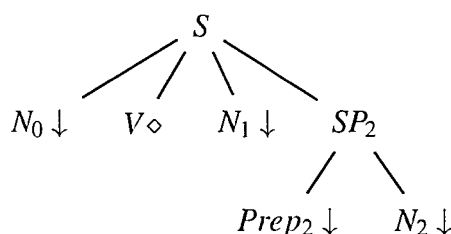


Les classes (V), (D) et (I) sont toujours croisées automatiquement. il reste cependant à croiser la (N<SP). Ceci sera réalisé lorsqu'un croisement entre les classes (D) et (I) sera réalisé.

³⁴En théorie, on pourrait spécifier toutes les notions d'ordre dans cette classe (sauf le cas du Sujet Inversé, car la Métagrammaire ne gère pas la disjonction), mais en pratique on n'utilise la dimension 4 que pour le placement des arguments entre eux.

³⁵Les classes qu'il définit ont juste valeur d'exemple. Il est clair qu'il vaut mieux parler de tête syntaxique ou d'ancre, plutôt que directement de verbe, pour une plus grande réutilisabilité.

Au final, on croise donc toutes les classes ainsi définies et on n'a fait intervenir l'ordre que lorsque toutes les classes sont concernées. Les ordres définis en dimension 4 ne viennent ainsi pas polluer inutilement les réalisations des arguments. Le seul arbre possible est donc le suivant :



Le seul arbre élémentaire pour Mary gives a book to Peter

Une implémentation effective est proposée Annexe B. La proposition de [Gerdes, 2002] est implémentée dans le formalisme de [Candito, 1999a]. Vient ensuite notre proposition.

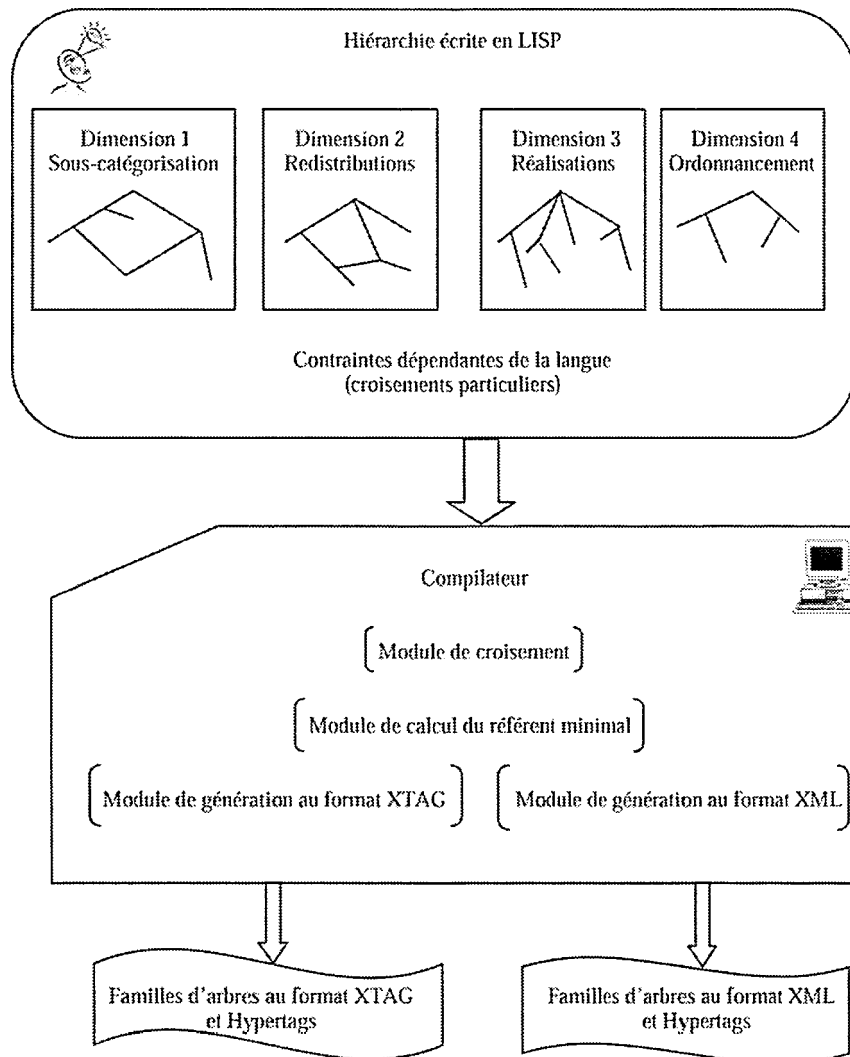
Pour le français, nous nous servirons principalement de la dimension 4 pour représenter les problèmes d'ordre entre différents clitiques.

Nous avons donc exploré au cours de ce chapitre diverses questions à la fois théoriques et pratiques concernant les métagrammaires. On a souvent eu tendance à privilégier l'aspect hiérarchique d'une description, et faire de ce privilège le point saillant d'une métagrammaire³⁶. Or, autant que le concept de hiérarchisation, c'est aussi l'aspect systématique des croisements qui est intéressant dans ce type de démarche.

Nous avons également introduit le concept de dimension 4, dimension permettant de rendre compte de l'ordre entre constituants argumentaux, et avons montré que les solutions proposées jusqu'à présent n'étaient guère souhaitables. Au final, une hiérarchie est donc définie par 3 dimensions obligatoires et une facultative :

- Dimension 1 : Sous-catégorisation initiale
- Dimension 2 : Redistribution de fonction et réagencement
- Dimension 3 : Réalisation d'un argument par rapport à la tête
- Dimension 4 : Ordre des arguments entre eux

³⁶Les nombreux articles présentant les métagrammaires font malheureusement toujours état d'une hiérarchie, sans pour autant indiquer le contenu des classes qui la composent, faute de place.



*Génération automatique d'une grammaire dans un formalisme particulier
à partir d'une hiérarchie écrite à la main*

Le compilateur de Candito souffre de problèmes résiduels, qu'il conviendra de résoudre dans un avenir proche. En particulier, toute sa documentation est encore à faire. L'approche relativement modulaire du compilateur a d'ores et déjà permis des améliorations substantielles. Néanmoins, il conviendra encore de concentrer les efforts au niveau de plusieurs aspects : une unification au niveau des structures de traits paraît absolument nécessaire. De plus, la possibilité d'avoir recours à des variables locales paraît également primordiale pour permettre une plus grande réutilisabilité. Resteront ensuite à relier éléments du lexique et arbres, qui sont pour l'instant deux entités distinctes.

Pour finir, nous voudrions insister sur le fait que même si le compilateur ne génère pour l'instant que des grammaires TAG, il pourrait être intéressant d'utiliser les hiérarchies déjà écrites pour d'autres formalismes. En effet, comme le remarque [Candito, 1999a], les différents domaines d'information retenus semblent relativement indépendants d'une théorie linguistique particulière et de nombreux modèles linguistiques utilisent des notions comparables : "l'utilisation de la notion de fonction ajoutée à la notion de constituant est un schéma que l'on retrouve dans plusieurs théories linguistiques importantes...".

A notre connaissance, peu d'expériences ont été menées dans ce domaine. [Clément et Kinyon, 2003b] se sont penchés sur cette question, sans pour autant avoir apporté de solution complètement convaincante. En effet, la hiérarchie qu'ils ont proposée est une hiérarchie qui répond à la fois aux contraintes du formalisme TAG et du formalisme LFG : le contenu typologique des classes initialement décrites pour une grammaire TAG a ainsi été enrichi en décorant chaque nœud de diverses équations fonctionnelles à la LFG³⁷. L'idée pour nous, serait plutôt de réutiliser la hiérarchie existante, et de ne modifier que le module de compilation³⁸. L'architecture du compilateur se prêterait à ce type d'expérience.

³⁷Dans un premier temps, le compilateur génère des arbres qui ne sont ni des arbres TAG, ni un ensemble de règles de réécriture à la LFG, mais plutôt des arbres hybrides, décorés à la fois au style TAG (nœuds à substitution, nœuds pieds, ancrés...) et LFG (équations fonctionnelles...). Dans un second temps, les différentes informations contenues par ces arbres sont réparties pour générer à la fois une grammaire TAG et une grammaireLfg.

³⁸Bien entendu, il conviendrait également de revoir le langage de description initial, pour le rendre moins dépendant d'un formalisme particulier.

Chapitre 3

Choix de représentation syntaxique

Sommaire

3.1	Les constructions à verbe support	73
3.1.1	Rappel des propriétés définitives	75
3.1.2	Les CVS : des constructions syntaxiquement régulières	79
3.1.3	La détermination du nom prédicatif dans les CVS	84
3.1.4	Les relatives de quantité	87
3.1.5	Mobilité des compléments et double analyse	88
3.2	Les groupes nominaux complexes	93
3.2.1	Les groupes nominaux en <i>par</i>	94
3.2.2	Les compléments en <i>de</i> des groupes nominaux complexes	97
3.3	Premiers choix de représentation pour les TAG	112
3.3.1	Des familles particulières et des entrées lexicales nouvelles	112
3.3.2	Représentation des CVS	117
3.3.3	Représentation des groupes nominaux complexes	122

Nous allons maintenant présenter les premiers choix qui nous guideront pour la description des familles à nom prédicatif de la grammaire TAG du français. Mais avant cela, nous aimerions, d'une part, revenir un peu plus précisément sur les constructions à verbe support pour définir plus exactement ce que nous entendons par CVS, et, d'autre part, nous attacher également à décrire le comportement des groupes nominaux complexes.

3.1 Les constructions à verbe support

On a vu au cours des chapitres précédents que les constructions à verbe support avaient suscité l'intérêt de nombreux chercheurs. Toutes les analyses proposées ont ainsi permis de prendre

en compte un nombre important de données. Malheureusement, la définition d'une CVS peut varier d'une étude à l'autre, et il n'est pas rare que des données prises en compte par une étude, soient rejetées par une autre. Dans cette section, nous effectuons un rapide résumé de ce qui fait pour nous, une CVS.

On sait depuis longtemps que les noms prédicatifs peuvent apparaître dans diverses positions. La plus étudiée a sans doute été celle d'objet direct (voir entre autres [Giry-Schneider, 1978, Vivès, 1983, Giry-Schneider, 1987, Gross, 1989]). D'autres positions peuvent cependant accueillir un nom prédicatif, comme l'ont par exemple indiqué [Negroni-Peyre, 1978, Danlos, 1980, Danlos, 1988, Gross, 1986, Gross, 1989]. Ainsi, ce dernier peut figurer :

- En position sujet :
 Un orage s'est produit hier soir sur la région
 Le silence règne sur cette affaire
 L'angoisse étreint Max
 L'envie prend Max de partir
 La rumeur court que le baril de pétrole va dépasser les 100\$
- En position d'objet indirect :
 Max procède à une enquête sur les indiens
 Max brûle d'amour pour Léa
 Max jouit d'une bonne santé
 Max souffre d'allergies
- En position d'attribut :
 Max est instituteur
 Max est le directeur de l'usine
 Max est ami avec Luc
- En position d'attribut
 prépositionnel :
 Max est en colère contre Luc
 Max est en conflit avec Luc
 Max est en extase devant ces résultats
 Max est dans l'impossibilité d'avouer son amour
 Max est à la recherche de la meilleure solution

Au moins deux questions se posent alors. D'une part, on peut se demander si tous ces exemples constituent des constructions à verbe support ; d'autre part, si tel est le cas, on peut s'interroger sur les propriétés définitives des CVS : est-il possible d'appliquer à cet ensemble d'exemples les propriétés déjà dégagées pour le nom prédicatif en position d'objet ? Autrement dit, les propriétés définitives d'une CVS sont-elles identiques suivant les différentes positions syntaxiques que le nom prédicatif peut remplir ou au contraire, leur sont-elles spécifiques ? Ces questions sont à notre sens, intimement liées

3.1.1 Rappel des propriétés définitoires

Dans un article récent, [Langer, 2005] s'est appliqué à résumer les différents tests permettant d'identifier une CVS. Ces tests sont d'une grande variété, mais souffrent pour la plupart d'exceptions.

Sans les citer tous¹, nous allons maintenant présenter ceux qui ont été régulièrement avancés pour le français. Certains comme l'impossibilité de nominalisation du verbe support, la "descente" de l'adverbe ou bien la double analyse du complément prépositionnel ou phrastique permettent d'identifier globalement une CVS ; d'autres comme la formation du groupe nominal complexe et la relation de N à N_0 sont absolument nécessaires.

L'impossibilité de nominalisation du verbe support s'explique par l'absence de sémantique forte de ce dernier. Elle est ainsi observable avec certains verbes supports (23), mais n'est pas généralisable à leur ensemble (24) :

- (23) a. Max donne un conseil à Luc
 *Le don d'un conseil de Max à Luc (serait déplacé)
- b. Un orage s'est produit hier soir
 *La production d'un orage (était prévisible)
- c. Max souffre d'allergies
 *La souffrance de Max d'allergies (est atroce)
- (24) a. Max prend une décision
 La prise de décision de Max (a été rapide)
- b. Les jeunes commettent des actes de vandalisme
 La commission de ces actes de vandalisme (est en régression) [JT France3 Lille]
- c. Max apporte une aide précieuse à Luc
 L'apport d'une aide précieuse à Luc par Max (est systématique)

¹Certains comme l'impossibilité de nominalisation de la CVS par un composé synthétique ne sont par ailleurs pas applicables au français :

- Kritik an etwas üben (émettre une critique sur quelque chose)
- *Die Kritikübung (* L'émission-d'une-critique)

En effet, en français, le composé synthétique n'est également pas observable dans d'autres constructions, alors qu'il l'est dans une langue comme l'allemand :

- Etwas in Frage stellen (remettre quelque chose en question)
- Die Infragestellung (La-remise-en-question)

La "descente" de l'adverbe décrite par [Giry-Schneider, 1987], qui rend compte de l'emploi synonymique que peuvent avoir un adverbe et un adjectif, concerne certaines CVS (25), mais ne leur est pas uniquement réservée [Harris, 1968] (26) :

- (25) a. Max fait rapidement la connaissance de Luc
Max fait la rapide connaissance de Luc
- b. Max procède systématiquement à une inspection
Max procède à une inspection systématique
- c. Max est extrêmement en colère contre Luc
Max est dans une colère extrême contre Luc
- d. Ces bruits courent régulièrement
Ces bruits réguliers courent

- (26) a. Max boit rapidement un thé
Max boit un thé rapide

Même les propriétés de **double analyse**, qui sont citées depuis [Gross, 1976] et ont été régulièrement reprises ou complétées entre autre par [Gross, 1977, Giry-Schneider, 1978, Vivès, 1983, Giry-Schneider, 1987, Gross, 1989] dans le cadre du lexique-grammaire ou [Abeillé et Godard, 2000] dans le cadre d'un traitement du *être* attributif en HPSG, et qui illustrent le fait que les compléments prépositionnels ou phrastiques des CVS peuvent tout aussi bien être considérés comme des compléments de verbe, que comme des compléments du nom (27), souffrent elles-aussi d'exceptions (28) :

- (27) a. Max commet un crime contre Luc
Le crime contre Luc que Max a commis est ignoble
Le crime que Max a commis contre Luc est ignoble
- b. Max a besoin d'être aimé
Le besoin qu'a Max d'être aimé n'est pas comblé
Le besoin d'être aimé qu'a Max n'est pas comblé
- c. Max est en colère contre Luc
Aussi en colère que Max soit contre Luc, il lui restera fidèle
Aussi en colère contre Luc que Max soit, il lui restera fidèle
- (28) Le chirurgien procède à l'ablation de la vésicule
*L'ablation à laquelle le chirurgien procède de la vésicule (est délicate)
L'ablation de la vésicule à laquelle le chirurgien procède (est délicate)

En revanche, une propriété qui apparaît comme primordiale est celle de **la formation**

d'un groupe nominal complexe contenant le nom prédicatif et ses compléments. Cette formation est systématique :

- (29) a. Max fait le portrait de Léa
Le portrait de Max de Léa (est fidèle)
- b. Max a de l'intelligence
L'intelligence de Max (est évidente)
- c. Max a une passion sans borne pour le sport
La passion de Max pour le sport (est sans limite)
- d. Max procède à la construction de la bibliothèque
La construction de Max de la bibliothèque (est achevée)
- e. Max brûle d'amour pour Léa
L'amour de Max pour Léa (est sincère)
- f. Le silence règne sur cette affaire
Le silence sur cette affaire (n'est pas anodin)
- g. Une envie soudaine de partir en voyage prend Max
L'envie soudaine de partir en voyage de Max (a surpris tout le monde)
- h. Max est en colère contre Luc
La colère de Max contre Luc (est justifiée)

Elle rejette définitivement les constructions qui placent le nom prédicatif en position d'attribut :

- (30) a. Max est conducteur de trains
*Le conducteur de trains de Max

Malheureusement, cette propriété se révèle insuffisante. En effet, de nombreux noms peuvent autoriser la formation d'un groupe nominal et apparaître avec des verbes comme *avoir* ou *faire* sans que ces derniers soient clairement identifiés comme des verbes supports. Les cas les plus emblématiques sont ceux faisant intervenir :

- des noms concrets : Max fait une maison
 La maison de Max

 Max a une voiture
 La voiture de Max
- des noms relationnels : Max a un frère
 Le frère de Max

 Max a un ami
 L'ami de Max
- des noms de maladie : Max a la grippe aviaire
 La grippe aviaire de Max

Il faut donc lui ajouter la propriété que [Giry-Schneider, 1987] nomme **la relation de N à N_0** et qui stipule qu'un nom prédicatif doit également avoir un sujet, pour permettre une identification plus précise des CVS². Cette propriété permet ainsi de rejeter³ les noms concrets comme intervenant avec une CVS, puisque la formation des phrases suivantes est tout à fait légitime :

- (31) a. Max fait la maison de Luc
 b. Max a la voiture de Luc

En revanche, les acceptabilités concernant les deux autres classes de nom sont moins nettes et varient selon les locuteurs. Nous considérons pour notre part que les noms de maladie⁴ et les noms relationnels interviennent avec des verbes supports, puisque la formation de phrases suivantes est exclue dans un cadre général (32), tout comme elle l'est d'ailleurs avec des verbes apparaissant clairement comme support⁵ (33) :

- (32) a. *Max a le frère de Luc
 b. *Max a la grippe aviaire de Luc

²[Gross, 1989] utilise la même notion lorsqu'il déclare qu'une CVS spécifie qu'il y a une relation de sujet à prédicat entre N_0 et le nom prédicatif.

³Des exemples comme *Le vent souffle* ou *Une malédiction plane sur cette assemblée* sont également exclus.

⁴Le cas des noms de maladie est difficile à cerner. [Labelle, 1984a] considère qu'ils interviennent avec des verbes supports. Certains locuteurs acceptent cependant sans difficulté des phrases comme *Max a la grippe de Luc*. Pour notre part, nous considérons que de telles phrases sont interprétables comme signifiant *Max a la même grippe que Luc*. Pour appuyer notre choix, nous signalerons le fait qu'il reste possible de trouver des phrases clairement rejetées par la majorité des locuteurs et faisant intervenir un nom de maladie, dès lors qu'un complément prépositionnel entre en jeu : **Max a le cancer aux intestins de Luc*.

⁵Les phrases fournies en exemple peuvent être totalement rejetées, ou avoir un sens particulier. Par exemple, la première phrase signifie que Max a commis le même crime que Luc a commis contre Léa

- (33) a. *Max a commis le crime de Luc contre Léa
 b. *Max fait les spéculations de Luc sur les ovnis
 c. *Max a fait l'erreur de Luc
 d. *Max procède à l'étude de Luc sur la population
 e. *L'envie de Léa prend Max de partir

Au final, on peut considérer que les propriétés permettant d'identifier clairement une CVS sont peu nombreuses. Mises à part la possibilité de formation d'un groupe nominal et la relation de N à N_0 , les propriétés décrites ci-avant ne sont alors simplement ni nécessaires ni forcément suffisantes. Nous reviendrons sur la double analyse au cours de la section (3.1.5)

Reste à savoir si les CVS manifestent des comportements syntaxiques plus particuliers. Pour Langer, la notion de verbe support est "une idéalisation (nécessaire) d'un champ très complexe situé entre les expressions compositionnelles et les idiomes". La question est alors de déterminer si les CVS tiennent plus des constructions figées que des constructions à verbes pleins.

3.1.2 Les CVS : des constructions syntaxiquement régulières

On sait qu'une construction est d'autant plus figée qu'elle a moins de propriétés transformationnelles. Bien souvent ont été relevées des impossibilités, notamment au niveau de l'interrogation, de la pronominalisation et de la coordination du nom prédicatif. Mais ces dernières sont-elles systématiques ? Nous allons envisager successivement ces possibilités et également nous intéresser à la détermination, puisque celle-ci a régulièrement été considérée comme un critère décisif pour identifier des expressions figées.

3.1.2.1 L'interrogation du nom prédicatif

Les impossibilités d'interrogation (34) en *que* (objet), *prep + quoi* (objet indirect), *comment* (attribut) ou *qu'est-ce qui* (sujet) portant sur le nom prédicatif ont souvent été mentionnées, notamment chez des linguistes comme [Giry-Schneider, 1987, Gross, 1989, Abeillé, 1991]⁶. Le verbe support n'ayant pas de sens prédicatif, ce comportement trouve naturellement une explication⁷

⁶[Giry-Schneider, 1987] note que les grammairiens allemands ont fait la même remarque à propos des constructions à Funktionsverben pour l'allemand. [Ogata, 1982] signale également le même phénomène pour le japonais, pour lequel le questionnement par *nani* est uniquement autorisé pour la construction à verbe indépendant

⁷Le verbe support étant sémantiquement vide, il n'est pas possible d'interroger le prédicat. Cela serait équivalent à interroger le participe d'un temps composé :

- Max a bronzé
- *Qu'a Max ? bronzé

et semble donc rapprocher les CVS des expressions figées, pour lesquelles l'interrogation est également inenvisageable (35) :

- (34) a. Max a fait une remarque à Léa
 *Qu'a fait Max à Léa ? Une remarque
 b. Max est en colère contre Luc
 ?Comment est Max contre Luc ? En colère
- (35) a. Max a perdu la boussole
 *Qu'a perdu Max ? La boussole
 b. Max a pris la mouche
 *Qu'a pris Max ? La mouche

Il reste cependant possible de trouver des emplois de support pour lesquels l'interrogation est autorisée. On sait ainsi que l'on commet un acte répréhensible et que l'on procède à une opération technique ou complexe (36) :

- (36) a. Qu'a commis Max ?
 Un assassinat horrible
 b. A quoi a procédé Max ?
 A une analyse fine

Autrement dit, il apparaît que dès que le type de l'objet est globalement déterminé par le verbe support, et qu'il reste simplement à préciser son identité, la question est autorisée. L'interrogation portant sur le groupe prédicatif n'est donc pas impossible. Elle est certes limitée, mais cela est dû en grande partie au fait que la majorité des verbes supports présentent également un emploi de verbes pleins. Or dans ce cas, l'interprétation à verbe plein prévaut sur celle à verbe support dans la question.

3.1.2.2 La pronominalisation

Dans certains cas, le nom ne peut pas être pronominalisé (37) :

- (37) a. *Léa a tiré avantage_i de la situation. Il_i s'est avéré primordial pour sa candidature.

Anne Abeillé (communication personnelle) nous fait remarquer que cette possibilité existait en moyen français, avec le sens *Qu'a fait Max ?*

- Bien entendu, comme le signale [Gross, 1989] il reste possible d'interroger le déterminant du nom prédicatif :
- Quel bruit court-il ?
 - Quelle analyse Max a-t-il faite des données ?
 - Quel démenti Max a-t-il donné à Luc ?

- b. *Léa a rendu hommage_i aux victimes de l'attentat. Il_i a même fait la une de la plupart des journaux.

Cependant, pour le français, c'est l'absence de détermination qui explique cette impossibilité (38). [Simatos, 1997] a ainsi montré que le déterminant seul permet la construction d'une référence actuelle (et non virtuelle)⁸. C'est donc la présence de ce dernier qui permet la pronominalisation du groupe prédicatif, comme l'illustrent les exemples suivants :

- (38) a. Léa a tiré [un avantage certain]_i de la situation. Il_i s'est avéré primordial pour sa candidature.
 b. Léa a rendu [un vibrant hommage]_i aux victimes de l'attentat. Il_i a même fait la une de la plupart des journaux.

Les CVS ne présentent donc pas de comportement particulier à l'égard de la pronominalisation.

3.1.2.3 La coordination

Les impossibilités de coordination relevées par [Rohrer, 1967] (39) ont été expliquées par [Gaetone, 1981] : ce qui rend ces exemples agrammaticaux, ce n'est pas l'impossibilité de coordination, mais plutôt l'inexistence de CVS comme *donner instructions* ou *avoir joie*. Pour que la coordination soit possible, il faut et il suffit que les noms coordonnés fassent tous deux partie d'une CVS à verbe support identique. Ainsi, si l'on coordonne deux noms qui sélectionnent le même verbe support, la phrase est tout à fait correcte (40) :

- (39) a. *Le général a donné instruction et ordre définitif
 b. *Max a grand plaisir et joie immense d'annoncer la naissance de Léa
- (40) a. Max a faim et soif
 b. Max perd patience et courage très rapidement
 c. Max fait un voyage et une excursion
 d. Max procède au drainage et à l'assainissement de la rivière
 e. Le bruit et la rumeur se répandent que le pétrole augmente

La coordination de verbes supports est en revanche plus difficile à observer, dans la mesure où c'est le nom prédicatif qui sélectionne son verbe support. Le sens des CVS coordonnées

⁸Dans d'autres langues, la situation est différente. [Alonso Ramos, 1998] cite ainsi un cas pour l'espagnol, dans lequel la pronominalisation est possible, même si le nom n'a pas de déterminant :

– Confesô que hacia campaña porque todos la hacem

Comme on l'a dit, ce type d'exemple est non recevable en français où la phrase correspondante ? *Il avoue faire campagne parce que tous la font* est exclue. Avec la variante *mener*, l'exemple est encore moins bon : **Il avoue mener campagne parce que tous la mène*.

doit donc être globalement identique, ce qui implique que les verbes supports doivent être des variantes d'un même support de base (41). Si ce n'est pas le cas, la coordination est impossible (42) :

- (41) a. Max caresse et nourrit l'espoir de réussir
 b. Max prend et perd rapidement l'avantage sur Luc
 c. Max a commis et perpétré de nombreux crimes
 d. Max est tombé et est resté définitivement en émerveillement devant cette fresque
 e. Le bruit court et se répand que le baril de pétrole dépassera les 100\$
- (42) a. *Max a donné et fait un coup remarquable

3.1.2.4 L'absence de détermination du nom prédicatif

Les problèmes de détermination et particulièrement l'absence de déterminant devant le nom, ont souvent été perçus comme la marque d'une cohésion plus étroite entre les différents constituants d'une CVS.

Pourtant, l'absence obligatoire de déterminant n'est guère observable dans les CVS. Les exemples fournis par [Gross, 1989] comme *Max donne congé à Luc* ou *Max a carte blanche* ne sont pas des CVS et sont d'ailleurs considérés depuis [Gross et Valli, 1991] comme des constructions à verbe figé. Dans le même ordre d'idée, les exemples fournis par [Giry-Schneider, 1987] comme *Max a fait mouche* ou *Cet exemple fera date*, sont considérés par [Giry-Schneider, 1991] comme des expressions à éléments figés (voir [Gross, 1982]).

Il est par contre possible d'observer un déterminant optionnel, comme l'attestent les exemples suivants :

- (43) a. Max a fait (un) signe à Léa
 b. Max a fait (une) allusion à ce problème
 c. Max a fait (le) vœu de ne plus dormir tard
 d. Max a donné (un) rendez-vous à Léa
 e. Max a donné (l') ordre de partir

Or dans ce cas, la construction est très régulière puisqu'elle admet la formation d'un groupe nominal :

- (44) a. Le signe de Max à Léa
 b. L'allusion de Max à ce problème

- c. Le vœu de Max de ne plus dormir tard
- d. Le rendez-vous de Max à Léa
- e. L'ordre de Max de partir

En fait, comme le rappelle [Giry-Schneider, 1987], à la suite de [Kuroda, 1968], il suffit que le nom supporté admette l'article indéfini pour qu'il puisse y avoir relativisation et formation d'un groupe nominal complexe.

C'est donc ce qui explique que nous ne considérons pas la plupart des constructions dans lesquelles le déterminant est figé comme des CVS. En effet, soit le déterminant figé n'est pas un indéfini et alors la formation du groupe nominal complexe est impossible, soit il l'est et alors deux situations sont observables :

1. La construction envisagée n'est pas une CVS :

- (45) Ce film fait un malheur
*Le malheur de ce film

2. La construction envisagée est une CVS⁹ :

- (46) Max a une dent contre Luc
La dent de Max contre Luc

Différentes thèses s'affrontent afin de déterminer si l'absence de détermination a des conséquences sur l'interprétation du nom. Pour [Anscombe, 1986, Anscombe, 1991], cette absence traduit un caractère plus événementiel ou processif. [Alonso Ramos, 1998] remet cependant en cause cette constatation : à son sens, une CVS comme *faire usage* n'est pas perçue comme plus proche du verbe *user/utiliser*, qu'une CVS comme *faire un usage efficace*.

Ce qui est certain, c'est que l'absence de détermination donne parfois à l'énoncé un caractère officiel ou administratif :

- (47) a. L'article premier donne autorisation de percevoir les impôts existants
b. Le premier ministre a donné ordre de planifier l'évacuation de 17 colonies de la bande de Gaza

Mais cela n'est pas toujours le cas :

- (48) a. Max a rendez-vous avec Léa
b. Max a fait allusion à ce problème dans son exposé
c. Max mène campagne contre le bruit

⁹L'exemple est emprunté à [Giry-Schneider, 1991].

Pour Ramos, "s'il faut trouver une différence de sens, l'on pourrait dire que [le déterminant] rend plus concret l'événement désigné par le nom ; d'une certaine façon, [le déterminant] le rend plus ponctuel". Et, bien que l'on "puisse trouver certaines différences sémantiques entre l'absence ou la présence du déterminant dans les CVS", il ne lui semble "pas nécessaire de considérer que le sens du nom soit plus prédicatif (i.e. plus verbal) sans déterminant".

En conséquence, on peut conclure dans la lignée de [Alonso Ramos, 1998], mais également [Giry-Schneider, 1991] que malgré l'importance accordée à l'absence de déterminant du nom dans les CVS, elle ne constitue qu'une "anecdote".

Nous continuerons à nous intéresser à la détermination dans la sous-section suivante, mais au terme de cette partie, nous pouvons d'ores et déjà affirmer que les CVS manifestent un comportement syntaxique relativement uniforme et régulier. Dans la plupart des cas, ce comportement est le même que celui d'un syntagme argumental libre. Dans d'autres, il est dû à la phraséologisation inhérente de toutes les collocations.

3.1.3 La détermination du nom prédicatif dans les CVS

3.1.3.1 Observations générales

[Gross, 1996a] signale que les contraintes portant sur la détermination du nom prédicatif dans les CVS n'ont "jamais fait l'objet d'un bilan systématique qui mettrait en évidence des régularités". Certaines observations semblent cependant assez générales.

Ainsi, la distinction nom massif / nom comptable permet, comme dans le cas des substantifs argumentaux de prévoir le déterminant à utiliser. Ainsi, les noms comptables autorisent l'article indéfini, les adjectifs numéraux et les expressions quantitatives :

- (49)
- article indéfini : Max a pris une douche
 - adjectif numéral : Max a pris deux douches
 - expression quantitative : Max a pris {diverses, quelques, plusieurs, un grand nombre de} douches

Les noms massifs, quant à eux autorisent l'article partitif et les expressions quantitatives :

- (50) – article partitif : Max a du courage
 – expression quantitative : Max a {un peu, beaucoup, moins, plus} de courage

A cette liste, on peut également ajouter l'absence de détermination (51) et l'obligation de modifier le nom prédicatif si un déterminant indéfini est utilisé (52) :

- (51) Max a peur
- (52) Max a peur du loup
 *Max a une peur du loup
 Max a une peur terrible du loup

Il ne faut cependant pas donner à ces distinctions une valeur de vérité absolue. Les chercheurs du LADL ont ainsi montré que la nature du déterminant dépend non seulement du nom prédicatif, mais aussi de son verbe support. Ainsi, le nom *attention* peut sélectionner entre autres, les supports *apporter*, *porter* ou *prêter*. Seul le premier n'autorise pas l'absence de détermination ; le dernier n'autorise pas le partitif.

- (53) a. Max apporte { *attention / une attention particulière / de l'attention } aux préparatifs du voyage
 b. Max porte { attention / une attention particulière / de l'attention } aux préparatifs du voyage
 c. Max prête { attention / une attention particulière / *de l'attention } aux préparatifs du voyage

On le voit, les contraintes sur le déterminant sont complexes. Elles peuvent même être différentes à l'actif (54) et au passif (55), comme l'a indiqué [Giry-Schneider, 1976] :

- (54) a. Max a pris { part / une part importante } à l'action
 b. Max a pris { *décision / la décision } de partir tôt
- (55) a. { *Part / Une part importante } a été prise à l'action par Max
 b. { Décision / La décision } a été prise de partir tôt par Max

3.1.3.2 La possessivation

Seule la possessivation a vraiment fait l'objet d'une étude systématique dans l'article de [Buvet, 2003]. Elle avait parfois été mise de côté par certains chercheurs qui soulignaient "simplement" qu'elle fournissait souvent des exemples peu acceptables.

C'est principalement lorsque le déterminant est coréférent à N_1 que l'on peut observer des régularités. Ainsi, la possibilité d'observer le déterminant possessif, lorsque celui-ci est un représentant du groupe prépositionnel qu'il reprend, est corrélée à l'obligation de régir N_1 par *de* dans la CVS (56) :

- (56) a. Max a fait le portrait de Luc
 Luc_i, Max a fait son_i portrait
 b. Max a procédé à l'inspection de cette salle hier soir
 [Cette salle]_i, Max a procédé à son_i inspection hier soir

- (57) a. Max a fait des ennuis à Luc
 *Luc_i, Max a fait ses_i ennuis
 b. Max a dé l'avance sur Max
 *Luc_i, Max a son_i avance

*fait le com à Luc
à la présence sur Luc*

Cette condition, bien que nécessaire, n'est cependant pas suffisante :

- (58) Max a la passion des vieilles Citroën
 *[Les vieilles Citroën]_i, Max a leur_i passion

Il faut également que le nom prédicatif soit un nom processif.

Lorsque le déterminant possessif n'est pas coréférent au complément du nom prédicatif, il doit se combiner avec un complément en *à* (et non plus *de*). Il a alors une interprétation aspectuelle et marque :

- (59) – la modalité ou la nécessité : Max a donné son avancement à Luc = Max a donné à Luc l'avancement qu'il devait lui donner
 – le fréquentatif : Max a donné son avancement à Luc = Max a donné à Luc l'avancement qu'il lui donne habituellement

Lorsque le déterminant possessif est coréférent au sujet du verbe support, la situation est plus complexe, et ne dépend plus de principes syntaxiques¹⁰. Buvet distingue deux cas, suivant que le nom prédicatif soit processif ou statique.

¹⁰Buvet mentionne que le déterminant possessif coréférent à N_0 peut être figé. Malheureusement, les exemples fournis ne nous paraissent pas vraiment probants, et il est toujours possible de trouver des contextes dans lesquels d'autres déterminants sont observables :

- Max a sa fierté
- Max a une fierté démesurée

- Max a son caractère
- Max a un caractère bien trempé

Lorsqu'il se combine avec un nom prédicatif processif, le déterminant possessif est généralement le reflet d'un caractère :

- (60) – fréquentatif : Max a fait sa sieste = Max a fait la sieste qu'il fait d'habitude
 – modal : Max a fait sa sieste = Max a fait la sieste qu'il devait faire
 – déictique : Max a fait sa sieste = Max a fait la sieste qui a été mentionnée

Avec un nom statique, la fonction de la possessivation est limitée à la mise en avant de la relation entre le prédicat et son premier argument, ou à la stipulation de la spécificité de cette relation.

La possessivation donne lieu à une simple focalisation, et ne peut pas généralement indiquer le fréquentatif¹¹ :

- (61) Cette histoire avait son épilogue = *Cette histoire avait l'épilogue qu'elle a habituellement

Ces cas sont rares et irréguliers. En revanche, le déterminant possessif est beaucoup plus fréquent et régulier lorsqu'il contribue à mettre en avant la relation entre N et N₀, et permet de stipuler la spécificité de cette relation :

- (62) Max avait ses exigences

De cette brève présentation, on tirera comme conclusion que les contraintes portant sur le déterminant sont généralement d'ordre sémantique et qu'elles dépendent donc du type sémantique du nom prédicatif et de son interprétation. Bien entendu, les faits présentés ici ne sauraient avoir un caractère de généralité. Ce sont simplement des tendances que l'on essaiera de prendre en compte le plus fidèlement possible.

3.1.4 Les relatives de quantité

On l'a vu, même les noms prédicatifs qui apparaissent sans déterminant dans une CVS, permettent la construction d'une relative dès qu'ils sont accompagnés d'un article. La relativisation d'un nom prédicatif est donc relativement semblable à celle d'un nom argument, si ce n'est qu'elle peut également être le reflet d'une double analyse, illustrant, rappelons-le, le fait qu'un groupe prépositionnel ou phrastique peut tout aussi bien être considéré comme un complément de verbe, que comme un complément de nom :

¹¹La valeur de fréquentatif est envisageable avec certains noms massifs, lorsque ces derniers se combinent avec un prédéterminant du type *accès* :

– Max avait son accès de colère = Max avait l'accès de colère qu'il a habituellement

- (63) Max mène une croisade contre le mercantilisme
 La croisade que Max mène contre le mercantilisme
 La croisade contre le mercantilisme que mène Max

Un point que nous n'avons pas mentionné jusqu'à présent, concerne l'interprétation de la relative. De manière régulière, on observe dans de nombreuses grammaires une dichotomie au sein des relatives. Lorsque la relative est nécessaire à l'identification de l'antécédent, on dit qu'elle est restrictive. Lorsqu'elle n'a pas de rôle dans cette identification, on dit qu'elle est appositive :

- (64) – Relative restrictive : Les américains qui adorent la science-fiction se réjouissent du retour de Star Wars
 – Relative appositive : Les américains, qui sont plus de 290 millions, sont gros consommateurs d'énergie

[Carlson, 1977] remet cette bi-participation en cause. Selon lui, il existe trois types de relatives : les relatives restrictives, les relatives appositives, et les relatives de quantité (*amount relatives*). Or, il est intéressant de remarquer que les relatives formées à partir de CVS, sont un exemple de relatives de quantité. Ainsi, pour certains noms prédicatifs, l'interprétation appositive n'est pas possible. *La peur que Max a eue* ne peut recevoir une interprétation appositive. Cet exemple n'est pas non plus une relative restrictive, puisque l'interprétation qui indiquerait *une peur parmi un ensemble de peurs* est impossible.

En fait, cette relative est plutôt l'indication d'un degré de peur, c'est-à-dire l'indication du degré de la peur de Max. C'est ce que Carlson appelle une relative de quantité.

Cette interprétation quantificationnelle n'est cependant pas observable avec toutes les CVS. Une relative comme *le crime que Max a commis* n'est pas le reflet d'un degré. Néanmoins, il nous semble intéressant de signaler que certaines CVS peuvent permettre la construction d'une relative de quantité.

3.1.5 Mobilité des compléments et double analyse

A présent, nous aimerions revenir sur les phénomènes de double analyse, cités régulièrement depuis [Gross, 1976]. Lorsque cette dernière est observable, elle semble aller de pair avec une mobilité entre les compléments prépositionnels et le complément prédicatif. L'ordre canonique semble être l'ordre dans lequel le groupe prédicatif est placé avant le complément prépositionnel. L'ordre complément prépositionnel/groupe prédicatif semble beaucoup moins naturel (65b) et est même totalement exclu lorsque le nom prédicatif n'admet pas de déterminant (66b, 67b) :

- (65) a. Max commet un crime contre Luc
 b. ?Max commet contre Luc un crime

- (66) a. Max a peur du loup
b. *Max a du loup peur

- (67) a. Max fait pression sur Luc
b. *Max fait sur Luc pression

C'est la même absence de détermination qui bloque les possibilités de double extraction (68), alors que les groupes prédicatifs admettant un déterminant ne subissent en général pas ce blocage (69), comme le signalent [Abeillé et Godard, 2004]. Dans ce cas, dès que le groupe comportant le nom prédicatif est alourdi, l'ordre nom prédicatif/complément semble plus libre et la double analyse s'observe généralement.

- (68) a. *C'est peur que Max a du loup
*C'est peur du loup que Max a
b. *C'est pression que Max fait sur Luc
*C'est pression sur Luc que Max fait

- (69) a. Max fait une réclamation auprès des services postaux
Max fait auprès des services postaux une réclamation justifiée
C'est une réclamation justifiée que Max fait auprès des services postaux
C'est une réclamation auprès des services postaux que Max fait
- b. Max mène une lutte contre les fourmis
Max mène contre les fourmis une lutte sans merci
C'est une lutte sans merci que Max mène contre les fourmis
C'est une lutte contre les fourmis que Max mène
- c. Max a des contacts avec les autorités
Max a avec les autorités des contacts réguliers
Ce sont des contacts réguliers que Max a avec les autorités
Ce sont des contacts avec les autorités que Max a
- d. Max fait des annotations dans la marge
Max fait dans la marge des annotations très pertinentes
Ce sont des annotations très pertinentes que Max fait dans la marge
Ce sont des annotations dans la marge que Max fait
- e. Max a peur du loup
Max a du loup une peur terrible
C'est une peur terrible que Max a du loup
C'est une peur terrible du loup que Max a
- f. Un silence règne sur cette affaire
Un silence assourdissant sur cette affaire règne
C'est un silence assourdissant qui règne sur cette affaire

C'est un silence assourdissant sur cette affaire qui règne

Pourtant, comme on l'a déjà dit, la double analyse ne concerne pas l'ensemble des CVS. Les chercheurs du LADL, et principalement [Giry-Schneider, 1987] et [Gross, 1989] ont ainsi signalé qu'avec certains verbes, et notamment *donner*, elle peut (70) ou non s'appliquer (71). Or, même dans ce cas, la mobilité est observable.

- (70) Max donne une réponse à cette interrogation
 Max donne à cette interrogation une réponse argumentée
 C'est une réponse argumentée que Max donne à cette interrogation
 C'est une réponse à cette interrogation que donne Max
- (71) a. Les soldats ont donné la charge à l'ennemi
 Les soldats ont donné à l'ennemi une charge extraordinaire
 *C'est une charge extraordinaire à l'ennemi que les soldats ont donnée
 C'est une charge extraordinaire que les soldats ont donnée à l'ennemi
- b. Max a donné un désaveu
 Max a donné à Luc un désaveu public
 *C'est un désaveu à Luc que Max a donné
 C'est un désaveu public que Max a donné à Luc

Les exemples de la série (71) sont pourtant bien des exemples de CVS, puisqu'il permettent l'incorporation du complément prépositionnel en *à* dans le groupe nominal complexe (72). Dès lors, on peut en conclure que dans ce cas, le complément en *à* doit uniquement être considéré comme un complément du verbe¹².

- (72) a. La charge extraordinaire des soldats à l'ennemi (avait été longuement planifiée)
 b. Le désaveu public de Max à Luc (a fait du bruit)

Par ailleurs, il est également possible d'observer des exemples qui n'autorisent ni mobilité, ni double analyse. [Danlos, 1980] avance ainsi des exemples pour les noms prédicatifs en position d'attribut prépositionnel (73), mais cette situation ne se limite pas qu'à eux seuls (74) :

- (73) Max est resté dans l'attente de sa venue
 *Max est resté de sa venue dans une attente fébrile
 *C'est de sa venue que Max est resté dans l'attente
 C'est dans une attente fébrile de sa venue qu'est resté Max
- (74) a. Max procède à l'analyse de ces échantillons
 *Max procède de ces échantillons à une analyse méticuleuse

¹²[Abeillé, 1991] considère tous les compléments prépositionnels en *à* comme des compléments du verbe.

*C'est à une analyse méticuleuse que Max procède de ces échantillons

C'est à une analyse méticuleuse de ces échantillons que Max procède

mais : C'est de ces échantillons que Max procède à une analyse

Le critère unique de la position n'explique cependant pas toutes ces impossibilités. Mobilité et double analyse peuvent en effet être observées avec d'autres noms prédicatifs acceptant pourtant les mêmes supports :

- (75) a. Max est resté dans un conflit perpétuel avec sa hiérarchie
 Max est resté avec sa hiérarchie dans un conflit perpétuel
 C'est avec sa hiérarchie que Max est resté dans un conflit perpétuel
 C'est dans un conflit perpétuel avec sa hiérarchie qu'est resté Max
**C'est dans un conflit perpétuel que Max est resté avec sa hiérarchie dans*
- b. Max procède à des essais sur ces échantillons
 Max procède sur ces échantillons à des essais méticuleux
 C'est à des essais méticuleux que Max procède sur ces échantillons
 C'est à des essais sur ces échantillons que Max procède

Il semble en fait que ce soit la conjonction de deux critères qui interdise à la fois la mobilité et la double analyse. Pour que ces dernières ne puissent être observées, il suffit de postuler que le groupe prédicatif doit être en position d'attribut prépositionnel ou d'objet indirect, et que son complément prépositionnel soit introduit par *de*¹³. *→ complément de l'attribut (mais peut être explicable seul quant même)*

Pour les compléments phrastiques, la mobilité est plus difficilement observable. Les acceptabilités ne sont pas aussi nettes que dans les cas précédents, et les exemples suivants de la série (76) sans être jugés excellents sont parfois jugés "lourds" ou "peu naturels" par la plupart des locuteurs. Seules les CVS qui placent le nom prédicatif en position sujet autorisent de manière naturelle le déplacement du complément phrastique (77) :

- (76) a. Max fait la constatation que tous sont partis
 ??Max fait que tous sont partis la constatation amère
- b. Max a une envie folle de partir
 ?Max avait de partir une envie folle
- c. Max a besoin d'être estimé
 ? Max avait d'être estimé un besoin terrible
- d. Max a tendance à exagérer
 ?? Max a à exagérer une tendance désagréable

¹³Le cas des noms prédicatifs objets directs prenant un complément en *de* est différent. Avec le verbe *faire*, ils autorisent bien la double analyse mais la mobilité n'est pas toujours facilement observable. En effet, si le complément en *de* est déplacé, le verbe semble perdre pour certains locuteurs sa valeur de support pour prendre la signification d'un verbe comme transformer :

- Max fait un récit de ses aventures.
- Max fait de ses aventures un récit passionnant

*non: Max a fait un récit passionnant de ses aventures
 Max a fait de ses aventures un récit passionnant*

Bien entendu, cela n'exclut pas que l'interprétation initiale à verbe support soit impossible, mais l'interprétation à attribut de l'objet indirect semble prévaloir dans ce cas.

- (77) a. Le bruit court que tous sont partis
Le bruit que tous sont partis court depuis une semaine
- b. L'envie prend Max de partir
L'envie de partir prend Max

[Labelle, 1984b] indique que les compléments phrastiques se prêtent mal à l'extraction. On sait en effet que les complétives ne peuvent pas être focalisées en utilisant la construction clivée. Le recours à la pseudo-clivée reste cependant possible :

- (78) Ce que Max fait c'est la constatation que tous sont partis
Ce dont Max ait la constatation c'est que tous sont partis

En revanche, les infinitives se prêtent sans grande difficulté à la construction clivée et la double analyse est alors observable, quelque soit la position du non prédicatif.

- (79) a. C'est une envie folle de partir qu'a Max
C'est une envie folle qu'a Max de partir
- b. C'est un terrible besoin d'être estimé qu'a Max
C'est un terrible besoin qu'a Max d'être estimé
- c. C'est une tendance désagréable à exagérer qu'a Max
C'est une tendance désagréable qu'a Max à exagérer

Les complétives apparaissent à première vue problématiques. Cependant, même si elles sont peu mobiles, leur statut de complément de verbe ne fait pas de doute :

- (80) a. lors d'un passif : Constatation a vite été faite par Max que tous étaient partis
- b. lorsqu'un complément prépositionnel supplémentaire est utilisé (voir à ce sujet [Gross, 1989]) :

Max a donné à Léa l'assurance qu'il viendrait

Max a donné l'assurance à Léa qu'il viendrait

**Max a donné qu'il viendrait l'assurance à Léa*

Pour résumer, il semble donc que le comportement des compléments prépositionnels et phrastiques au sein des CVS ne soit pas uniforme. Dans la plupart des cas, ils peuvent aussi bien être considérés comme des compléments de verbe, que comme des compléments du nom. Mais dans d'autres, ils ne peuvent pas revêtir cette double fonction ; on les considère alors soit comme des compléments de verbe, soit comme des compléments de nom, mais pas les deux à la fois.

Il est intéressant de constater qu'une situation partiellement semblable existe dans des

langues proches du français. [Alonso Ramos, 1998] a ainsi montré que tous les compléments en *de* devaient être considérés comme des compléments du nom en espagnol. Dans ce dernier cas, il n'y a pas de possibilité de double analyse :

- (81) *el recuento que Pedro hizo de las victimas
 el recuento de las victimas que Pedro hizo

La situation est, comme on l'a vu, plus complexe en français, où des exemples correspondants sont permis :

- (82) le recensement que Pierre fait des victimes
 le recensement des victimes que Pierre fait

Au terme de cette section, nous pouvons donc affirmer que la plupart des propriétés décrivant les CVS ne sont pas nécessaires. Elles souffrent pour la plupart d'exceptions. Nous avons par ailleurs montré que les CVS manifestaient un comportement syntaxique relativement uniforme et régulier et que dans la plupart des cas, ce comportement était le même que celui d'un syntagme argumental. Enfin, nous avons montré que le comportement syntaxique des compléments prépositionnels ou phrastiques n'était pas régulier. Certains peuvent à la fois être considérés comme complément de verbe et complément du nom, d'autres uniquement comme complément de verbe ou comme complément du nom. Cette constatation remet également en cause les phénomènes de double analyse qui ne sont donc pas aussi réguliers qu'on pourrait le penser avec les CVS.

3.2 Les groupes nominaux complexes

A présent, nous allons nous intéresser à la formation des groupes nominaux complexes ayant pour tête un nom prédicatif. On a vu au cours de la section précédente que la formation d'un groupe nominal était un critère obligatoire pour identifier les CVS. Le recours à ces dernières permet non seulement de déduire la structure de compléments du nom (83), mais également de décrire le comportement de substantifs qui n'ont pas d'équivalent verbal (84) :

- (83) a. Max admire {Léa / *pour Léa}
 Max a de l'admiration pour Léa
 L'admiration de Max pour Léa
 b. Max s'intéresse {à / *pour} la linguistique
 Max a de l'intérêt pour la linguistique
 L'intérêt de Max pour la linguistique

- (84) a. Max fait le portrait de Léa
Le portrait de Max de Léa
- b. Max fait une farce à Léa
La farce de Max à Léa
- c. Max fait un cadeau à Léa
Le cadeau de Max à Léa

Beaucoup de discussions ont eu lieu concernant les nominalisations (voir par exemple [Daladier, 1978, Milner, 1982, Giry-Schneider, 1987, Gross, 1989, Godard, 1992, Samvelian, 1995]), mais pas toujours dans le cadre des CVS. Pour notre part, nous allons nous pencher sur deux questions. La première consistera à déterminer le plus précisément possible quand la formation d'un groupe nominal dit "passif" (en *par*) sera possible. La seconde concernera la représentation des groupes nominaux comportant des compléments génitifs. Est-il toujours possible de faire cohabiter des compléments en *de* et peuvent-ils toujours donner naissance à un déterminant possessif ?

3.2.1 Les groupes nominaux en *par*

De la même façon qu'il est possible à partir d'une CVS à l'actif de construire un groupe nominal "actif", il est possible de former à partir d'une CVS au passif, un groupe nominal "passif" :

- (85) a. Max fait le recensement des victimes
Le recensement de Max des victimes
- b. Le recensement des victimes est fait par Max
Le recensement des victimes par Max (était nécessaire)

L'impossibilité de passivation d'une CVS expliquerait alors l'impossibilité de formation d'un groupe nominal passif :

- (86) Max a une bonne connaissance des lieux
*Une bonne connaissance des lieux est eue par Max
*La connaissance des lieux par Max (n'est pas surprenante)

Cette hypothèse n'est cependant pas suffisante. En effet, si le verbe *avoir* refuse la passivation, ses variantes ne sont pas sujettes à cette interdiction. Pourtant, même dans ce cas, la formation d'un groupe nominal en *par* n'est pas pour autant autorisée :

- (87) Les villageois {ont/ressentent} une peur terrible du loup-garou
Une peur terrible du loup-garou est {*eue / ressentie} par les villageois
*La peur du loup-garou par les villageois (est ridicule)

A l'inverse, il pourrait être tentant de croire que l'existence d'une CVS au passif suffise à autoriser la formation d'un groupe nominal en *par*. On sait cependant qu'il n'en n'est rien. Les noms prédicatifs sans complément N_1 n'admettent ainsi pas de groupes "passifs" :

- (88) Max prend une douche
 Une douche est prise par Max
 *Une douche par Max (serait salulaire)

En fait, seuls les noms prédicatifs (non psychologiques) ayant pour complément un groupe introduit par *de* [Daladier, 1978, Samvelian, 1995] permettraient la formation d'un groupe nominal "passif".

[Gross, 1989] remet cependant cette conclusion en cause. Pour lui, la possibilité de formation d'un groupe nominal en *par* est observable avec d'autres types de compléments, dès que le déterminant indéfini est utilisé :

- (89) a. Le gouvènement donne une adhésion claire à ce projet
 Une adhésion claire à ce projet est donnée par le gouvernement
 Une adhésion claire à ce projet par le gouvernement (serait souhaitable)
 *L'adhésion claire à ce projet par le gouvernement (serait souhaitable)
- b. Le premier ministre apporte une contribution significative à ce débat
 Une contribution signification à ce débat a été apprtée par le premier ministre
 Une contribution significative à ce débat par le premier ministre (serait la bienvenue)
 *La contribution significative à ce débat par le premier ministre (serait la bienvenue)

De tels groupes nominaux semblent cependant assez rares et nombreux sont les locuteurs qui les rejettent. Le seul exemple que nous ayons trouvé en corpus est le suivant :

- (90) Elles avaient pour origine des propos au Nouvel Observateur par M. Abel Farnoux

L'hypothèse qui voudrait qu'un groupe nominal "passif" puisse être formé à partir d'une CVS passive demande donc a être clarifiée. D'autant qu'il est possible de trouver des groupes nominaux sans CVS passive correspondante : *nm*

- (91) a. L'attraction du fer par l'aimant (est inévitable)
 b. L'agression de Luc par Max (était préméditée)

Nous pensons dans ce cas qu'une deuxième source peut être envisagée, et venir expliquer ces nominalisations si l'on considère les constructions converses décrites par [Gross, 1989]. Ainsi, *L'attraction du fer par l'aimant* ne peut être déduite de la CVS passive *Une attraction est exercée par l'aimant sur le fer*. En revanche, si l'on considère la construction *Le fer subit l'attraction {de/ ?par} l'aimant*, construction converse de la CVS précédente, la formation du groupe nominal semble plus légitime.

De la même manière, *l'agression de Luc par Max* ne sera pas déduite de la CVS passive *Une agression est commise par Max contre Luc*, mais par la construction converse *Luc subit une agression {de / ?par} Max*.

L'existence d'une construction converse ne saurait cependant pas constituer une condition suffisante à la formation d'un groupe "passif" :

- (92) Luc reçoit une promesse {de / ?par} Max
 *La promesse de Luc par Max (est impossible à tenir)

Ainsi, si l'on résume, deux sources peuvent être envisagées pour la formation d'un groupe nominal "passif"¹⁴ :

1. l'une à partir de la construction où le verbe support apparaît sous forme passive :
 - Une adhésion claire est donnée par le gouvernement à ce projet
 - Une adhésion claire par le gouvernement à ce projet (serait la bienvenue)
2. l'autre à partir d'une construction converse :
 - Ce projet de loi reçoit une adhésion claire {du / ? par le} gouvernement
 - Une adhésion claire de ce projet de loi { ?du / par le} gouvernement (serait la bienvenue)

Mais ces formations ne sont pas systématiques. D'une part, parce que certains supports n'admettent pas de passif ou de converse ; d'autre part parce qu'il ne suffit pas qu'un support accepte la passivation ou qu'une construction converse existe pour que cela implique forcément la formation d'un groupe nominal.

Il se pose donc la question de savoir comment contraindre l'existence d'un groupe nominal passif. Dans sa thèse, [Samvelian, 1995] a montré que toute tentative d'explication tâchant d'assigner le rôle sémantique d'agent au complément en *par* était vouée à l'échec. En effet, il n'est pas vrai que seuls les agents puissent apparaître dans cette construction. On peut par exemple y trouver des Buts (93a) ou des instruments (93b) :

¹⁴Lorsque le nom prédicatif a un complément en *de*, la CVS passive et la construction converse donnent naissance au même groupe nominal :

- Max a donné une lecture claire de cet ouvrage
- Une lecture claire de cet ouvrage a été donnée par Max
- Une lecture claire de cet ouvrage par Max (était nécessaire)
- Cet ouvrage a reçu une lecture claire {de / ?par} Max
- Une lecture claire de cet ouvrage par Max (était nécessaire)

- (93) a. Le fer subit l'attraction de l'aimant
L'attraction du fer par l'aimant
- b. La machine fait le lavage du linge
Le lavage du linge par la machine

En revanche, la prise en compte des propriétés aspectuelles du groupe nominal semble éclairante. Ainsi, pour Samvelian, il existe une corrélation entre le caractère événementiel ou processif de la nominalisation et la possibilité de former un groupe nominal en *par*, cela quel que soit le rôle du complément en *par*.

C'est ce qui explique, par exemple, les différences observées dans ces deux groupes, où des noms tels que *désert* ou *muraille* ne peuvent pas impliquer un changement d'état, car ils n'ont pas la possibilité d'un mouvement, d'une évolution, ou d'une apparition (94). Il suffit que le nom introduit par *par* ait cette autonomie pour que le groupe nominal devienne acceptable (95) :

- (94) a. L'encerclement de la maison par la police
?* L'encerclement de ce village par le désert
- b. La protection de la ville contre l'attaque de l'ennemi par les soldats
*La protection de la ville contre l'attaque de l'ennemi par les murailles qui l'entourent

- (95) L'encerclement progressif du village par la forêt vierge

? L'isolation thermique de l'immeuble par une fenêtre - est de 60%
? La protection du bateau par une sentinelle est double

L'aspect joue donc un rôle déterminant dans l'acceptabilité des groupes nominaux en *par*. Seules les nominalisations préservant le caractère événementiel ou processif d'une CVS peuvent donner lieu à une construction en *par*.

3.2.2 Les compléments en *de* des groupes nominaux complexes

A présent, nous allons nous tourner vers la structuration des groupes nominaux complexes comportant des compléments en *de*. On l'a vu, le recours aux CVS, permet non seulement de déduire la structure de compléments du nom, mais également de décrire le comportement de substantifs qui n'ont pas d'équivalent verbal. Néanmoins, la formation du groupe nominal n'est pas toujours acceptable, lorsque deux compléments en *de* sont utilisés :

- (96) a. Les ennemis procèdent au bombardement de la ville
?Le bombardement de la ville des ennemis (est continué)

- b. Marie Curie a fait la découverte du radium
?La découverte du radium de Marie Curie (lui valut le prix Nobel)
- c. Max procède à l'internement des patients les plus dangereux
?L'internement de Max des patients les plus dangereux (interviendra dans la soirée)

La situation est d'autant plus troublante qu'elle ne semble pas systématique. Aussi d'autres groupes nominaux comportant deux compléments en *de* sont-ils parfaitement acceptables :

- (97) a. Max a une parfaite maîtrise de la navigation
La parfaite maîtrise de la navigation de Max (n'est pas surprenante)
- b. Max a une bonne connaissance de ces lieux
La bonne connaissance de Max de ces lieux (est rassurante)
- c. Léonard fait le portrait de Mona-Lisa
Le portrait de Léonard de Mona-Lisa (sera fini dans une heure)
- d. Max a fait une sortie remarquée de la salle de réunion
La sortie remarquée de Max de la salle de réunion (a fait scandale)

Quoiqu'il en soit, le remplacement du complément en *de* correspondant au sujet du verbe support par le déterminant possessif est toujours autorisé, et les groupes nominaux à acceptabilité douteuse semblent alors beaucoup plus naturels :

- (98) a. Leur bombardement de la ville (est continu)
- b. Sa découverte du radium (valut le prix Nobel à Marie Curie)
- c. Son internement des patients les plus dangereux (interviendra dans la soirée)

En revanche, pour les autres compléments en *de*, cette association ne semble pas constante. A côté d'exemples où les déterminants possessifs sont observables (99), on remarque ainsi d'autres exemples où les déterminants possessifs sont rigoureusement interdits (100) :

- (99) a. Max fait une description du suspect
Sa (= du suspect) description (est concordante avec les autres témoignages)
- b. Max a une parfaite maîtrise de la navigation
Sa (= de la navigation) maîtrise (n'est pas surprenante)
- c. Les habitants ont une bonne connaissance des lieux
Leur (= des lieux) connaissance (est rassurante)

- (100) a. Max donne sa démission de son poste
 *Sa (=de ce poste) démission (en a surpris plus d'un)
- b. Max fait une sortie remarquée de la salle de réunion
 *Sa (=de la salle de réunion) sortie (fait scandale)
- c. Max a l'espoir d'une vie meilleure
 *Son (=d'une vie meilleure) espoir (est vital)

Il suffit par ailleurs que le complément correspondant au sujet du verbe support soit présent pour que l'association du complément avec le déterminant possessif soit impossible :

- (101) a. *Sa (=du suspect) description de Max (est concordante)
- b. *Sa (=de la navigation) maîtrise de Max (n'est pas surprenante)
- c. *Leur (=des lieux) connaissance des habitants (est rassurante)

De tels faits sont troublants. En effet, un groupe nominal comme *La description de Max* est ambigu entre une interprétation dans laquelle *c'est Max qui fait une description* et une interprétation dans laquelle *on fait une description de Max*. Un groupe nominal comme *sa description* préserve la même ambiguïté. On s'attendrait donc à ce qu'un groupe nominal comme *sa description de Max* ait lui aussi deux interprétations. Pourtant la seule interprétation possible est celle où *l'on fait une description de Max* (voir [Milner, 1982]).

Il convient donc de trouver des explications permettant de rendre compte de ces comportements particuliers. D'une part, il faut pouvoir expliquer pourquoi la coprésence de deux compléments en *de* n'est pas toujours possible au sein d'un groupe nominal ; d'autre part, il faut également expliquer pourquoi certains compléments en *de* ne peuvent pas toujours être associés à un déterminant possessif. Plusieurs hypothèses ont été développées hors du cadre des constructions à verbe support. La plupart d'entre elles établissent une distinction entre les différents compléments en *de*. Certains sont des compléments directs, dans lesquels *de* est une marque casuelle ; d'autres sont de véritables groupes prépositionnels. Les critères de définition de chaque complément varient souvent d'une étude à l'autre ; néanmoins, toutes s'accordent à dire que la différence de structure de ces compléments rend compte de différences interprétatives entre les groupes nominaux construits autour d'une nominalisation.

Nous commencerons par l'analyse de [Milner, 1982], qui fait figure de référence dans la littérature portant sur les nominalisations, puis nous nous tournerons vers l'analyse plus récente [Godard, 1992]. Enfin, nous exposerons les critiques que ces analyses peuvent soulever, notamment chez deux auteurs : [Samvelian, 1995] et [Kupferman, 2004].

3.2.2.1 La structure interne des groupes nominaux chez Milner (1982)

[Milner, 1982] établit une distinction entre les nominalisations proprement dites (les nominalisations processives) et les groupes nominaux ordinaires qui regroupent non seulement les noms non-dérivés, mais également les nominalisations résultatives.

Le groupe nominal construit autour d'une nominalisation processive, reçoit une interprétation globale parallèle à une phrase ayant pour élément central le verbe correspondant. Ainsi, à la différence des groupes nominaux ordinaires qui ont pour référence un "état stable" de la réalité, les nominalisations ont, tout comme les constructions phrastiques parallèles, pour référence un processus événementiel et évolutif, susceptible de valeur de vérité. Cela se marque par la possibilité de former une nominalisation négative de la forme *non-N*, dans laquelle *non* a la même valeur que la négation totale *ne-pas* dans la phrase correspondante.

Les nominalisations processives se distinguent également par une relation fonctionnelle au verbe dont elles sont issues en conservant la structure de compléments du verbe : la nominalisation a les mêmes compléments prépositionnels et directs que le verbe. Selon qu'ils correspondent au sujet ou à l'objet direct du verbe, les compléments en *de* de la nominalisation sont respectivement appelés *génitif subjectif* et *génitif objectif*.

Un grand nombre de noms dérivés peuvent présenter les deux emplois processif et résultatif. [Milner, 1982] propose donc un ensemble de critères permettant de distinguer chaque emploi d'un nom :

1. L'emploi du déterminant indéfini *un* ou d'un déterminant pluriel est difficilement observable avec une nominalisation processive.
2. Les pronoms forts sont exclus aux génitifs objectif (102) et subjectif (103) :
 - (102) a. *l'agression de moi par un sadique
 b. *la dénonciation de toi par un ennemi
 c. *la consultation de lui par un malade
 - (103) a. *la venue de moi à Paris
 b. *l'arrivée de toi à Paris
 c. *l'intervention de lui auprès de son frère

Par contre, dans les groupes nominaux ordinaires, les mêmes types de compléments peuvent être associés à un pronom fort, comme l'illustrent les groupes nominaux suivants à interprétation résultative : *La dénonciation de toi dont on m'a parlé* désigne "l'objet qui résulte de l'acte de dénoncer" ; *la description de toi dont on m'a parlé* désigne soit "la description que tu as faite", soit "la description qui a été faite de toi".

3. Les génitifs objectif (104) et subjectif (105) n'ont pas d'emploi attributif :

(104) a. *L'agression est d'une jeune fille

b. *La dénonciation était d'un innocent

c. * La consultation est d'un médecin réputé

(105) a. *l'agression de la jeune fille était d'un sadique

b. *la dénonciation de cet innocent était de son meilleur ami

c. *la consultation du médecin était d'un incurable

L'interprétation stable rend de telles structures acceptables. Autrement dit, lorsqu'une telle construction est observable, le nom a obligatoirement un emploi résultatif :

(106) a. la dénonciation de cet innocent était de son meilleur ami

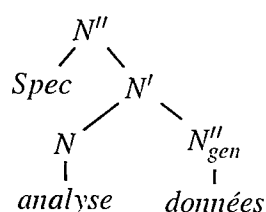
b. La description de la victime est d'un témoin oculaire

c. la description de Pierre est d'un paysage italien

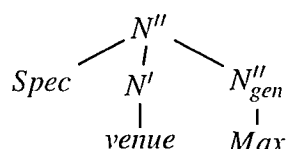
Outre ces différences interprétatives, les nominalisations proprement dites et les nominalisations résultatives présentent une différence de structure interne. Ainsi, certaines structures excluent l'interprétation processive, alors que d'autres sont incompatibles avec une interprétation résultative.

1. Dans une nominalisation processive, les compléments correspondant au sujet et à l'objet direct du verbe sont des N'' au cas génitif. *De* est alors une marque casuelle et non une préposition.

Le génitif objectif est le N'' qui est au même niveau que N. Ainsi, la représentation du groupe nominal *L'analyse des données (qui a duré 5 heures)* est la suivante :

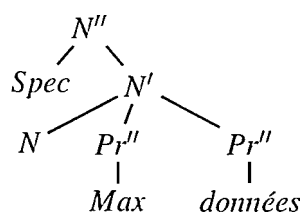


Le génitif subjectif est également un N'', mais rattaché directement au N'' racine. Ainsi, la représentation du groupe nominal *La venue de Max* est la suivante :

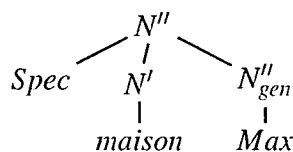


La co-présence des deux compléments directs dans le groupe nominal est interdite par le principe de non redondance fonctionnelle¹⁵. Ce qui implique que tout groupe nominal construit autour d'une nominalisation issue d'un verbe transitif direct et contenant deux compléments en *de* a alors nécessairement une interprétation résultative.

2. Dans une nominalisation résultative, les compléments en *de* sont des groupes prépositionnels. Ainsi, dans le groupe nominal *L'analyse de Max des données (qui est sur le bureau)*, *de Max* et *des données* ne sont pas des N''. *De* est une vraie préposition. Le groupe nominal a alors la structure suivante :



La seule possibilité pour qu'une nominalisation résultative accepte un complément direct N'' au cas génitif, est que ce dernier ait une interprétation possessive. La représentation est alors la suivante :



La représentation du Génitif possessif et du Génitif subjectif est donc identique, mais le N'' rattaché au N'' racine ne reçoit pas la même interprétation. Dans le premier cas c'est un possesseur, dans le second c'est un agent. La présence d'un Génitif possessif

¹⁵Une des formulations possibles de ce principe est la suivante : *une catégorie X ne peut contenir qu'une seule occurrence d'une fonction déterminée*. Une hypothèse équivalente est formulée par [Vergnaud, 1985], pour qui un groupe nominal ne peut contenir qu'un seul génitif ; le terme de génitif chez Vergnaud désignant uniquement les génitifs directs.

dans un groupe nominal exclut d'emblée toute interprétation processive pour ce groupe.

Malgré ces différences interprétatives et représentationnelles, les deux types de compléments en *de* peuvent tous deux être associés à un déterminant possessif, comme l'illustre le groupe nominal *sa description du bandit*, pouvant recevoir à la fois une interprétation processive et une interprétation résultative :

Interprétation processive	Sa description du bandit dura 5 heures
Interprétation résultative	Sa description du bandit est sur le bureau

Avec le génitif direct, le déterminant possessif est généré directement dans la position de Spec. En revanche, avec le génitif indirect, le déterminant n'est pas généré directement : le pronom, sans la préposition *de* est déplacé vers la position Spec par une transformation de mouvement ; une règle de surface efface la préposition *de* dans la forme de surface. La trace, laissée dans la position d'origine permet de fournir l'interprétation Agent au possessif.

Ainsi, les compléments prépositionnels (comme les compléments Agent et Objet des noms iconiques) peuvent donc être réalisés comme des déterminants possessifs¹⁶ :

- (107) Le portrait de Léonard de Vinci de la Joconde (est exposé au Louvre)
 Son portrait de la Joconde (est exposé au Louvre)
 *Son portrait de Leonard de Vinci (est exposé au Louvre)
 Son portrait par Léonard de Vinci (est exposé au Louvre)

3.2.2.2 L'analyse de Godard (1992)

[Godard, 1992] s'intéresse quant à elle uniquement aux conditions d'extraction des compléments en *de*. Elle adopte en grande partie les hypothèses de [Milner, 1982] :

1. Il existe à l'intérieur du groupe nominal deux niveaux d'attachement : le sujet est directement dominé par N'' ; le complément d'objet est dominé par un niveau N' intermédiaire.

¹⁶Dans le cas où Génitif Agent et Génitif Objet sont tous deux présents, seul le Génitif Agent peut donner naissance au déterminant possessif. La seule possibilité pour le Génitif Objet de donner naissance au déterminant possessif lorsqu'il apparaît avec un agent, est que ce dernier soit réalisé par un SP en *par*.

2. Les compléments en *de* sont soit des compléments directs (des génitifs), soit des compléments prépositionnels. Un groupe nominal ne peut contenir deux compléments génitifs.

Elle s'en éloigne cependant sur plusieurs points. D'une part, le terme de génitif n'a pas le même emploi : il désigne pour Milner à la fois les compléments directs et prépositionnels en *de*. D'autre part, les critères de définition des compléments en *de* sont différents : alors que pour Milner, la possibilité d'être associé à un déterminant possessif ne distingue pas les génitifs directs, des génitifs indirects, Godard stipule que seuls les compléments directs peuvent être réalisés comme déterminant possessif.

Godard prend en compte de nombreux types de noms. Elle s'intéresse ainsi aux noms ordinaire, représentationnel, processif, psychologique, relationnel ou de propriété :

Type de nom	Argument		Exemple
	Sémantique	Catégorie	
Nom ordinaire	Possesseur	NP	La maison <i>de Paul</i>
	Agent	NP ou PP	L'immeuble <i>du Corbusier</i>
Nom représentationnel	Possesseur	NP	Les portraits <i>du collectionneur Barnes</i>
	Thème	NP	Les portraits <i>de la laitière</i>
	Agent	NP ou PP	Les portraits <i>de Vermeer</i>
Nom processif déverbal	Thème	NP	La construction <i>de l'Opéra</i>
	Agent	NP	La venue <i>de Paul</i>
Nom de propriété	Thème	NP	L'intelligence <i>de Paul</i>
Nom psychologique	Expérienceur	NP	La passion du jeu <i>des aristocrates</i>
	But	PP	La passion <i>du jeu</i> des aristocrates
Nom relationnel	Theme	NP	L'ami <i>de Paul</i>

Ne peuvent être extraits par *en* ou *dont* que les compléments génitifs qui sont le premier argument du groupe nominal dans lequel ils apparaissent¹⁷. Il y a deux moyens d'être premier argument : soit l'argument est le premier argument d'origine, soit il le devient. En effet, un second argument peut être promu et alors accéder au statut de premier argument, si le premier argument d'origine est supprimé ou retrogradé ("downgraded"). Cette promotion ne s'applique cependant pas à tous les seconds arguments. Ces derniers doivent nécessairement être des compléments directs (i.e. des génitifs) pour pouvoir accéder à un tel statut. Ainsi, même si le premier argument n'est pas réalisé, les compléments prépositionnels ne peuvent pas accéder au rang de premier argument.

¹⁷Le premier argument est défini comme le premier constituant qui sature la première a-variable dans la liste des a-variables. Dans la s-structure, il apparaît plus haut que les autres arguments. Il peut être supprimé ou ne pas être réalisé.

Par exemple, un nom comme *portrait* dispose de deux arguments pouvant être extraits : l'agent et le thème ; l'agent étant le premier argument. L'extraction de ce dernier s'effectue indépendamment de la présence ou non du thème. En revanche, le thème ne peut être extrait que si l'agent n'est pas réalisé, comme l'attestent les exemples suivants :

- (108) Corot dont le portrait de cette jeune femme se trouve à la fondation Barnes en a malheureusement peint très peu

*La jeune femme dont le portrait de Corot se trouve à la Fondation Barnes est inconnue

La jeune femme dont le portrait est à la fondation Barnes est inconnue

En revanche, le complément prépositionnel d'un nom psychologique ne pourra pas être extrait même si le premier argument n'est pas réalisé :

- (109) *Le jeu dont la passion des aristocrates russes était immense est destructeur

*Le jeu dont la passion était immense chez les aristocrates est destructeur

Pour résumer, peuvent être premier argument :

1. les premiers arguments d'origine :

- l'agent d'un nom processif déverbal
- le thème d'un nom de propriété
- l'expérimenteur d'un nom psychologique
- le thème d'un nom relationnel
- le possesseur d'un nom ordinaire
- le possesseur d'un nom représentationnel
- l'agent d'un nom représentationnel
- l'agent d'un nom ordinaire

2. les premiers arguments dérivés (en l'absence du premier argument d'origine) :

- le thème d'un nom processif déverbal
- le thème d'un nom représentationnel

3.2.2.3 L'analyse de Samvelian (1995)

[Samvelian, 1995] remet partiellement en cause les hypothèses et conclusions des analyses précédentes. C'est principalement la structure des compléments en *de* qui est source d'interrogations. En effet, en élargissant les données, Samvelian remarque que des facteurs sémantiques entrent en compte.

A l'inverse de ce qui était affirmé fréquemment dans les études portant sur les génitifs, Samvelian signale ainsi qu'un groupe nominal peut contenir simultanément deux compléments directs. Les données sont particulièrement difficiles à appréhender. Bien souvent, les acceptabilités entre les locuteurs varient, et il est difficile d'accorder un jugement de valeur stable ou sûr. En fait, les jugements portés peuvent aller du "tout à fait exclu" au "tout à fait correct" en passant par toute une gamme de variation.

Toujours est-il, qu'en principe, seules les nominalisations résultatives devraient autoriser deux compléments en *de* ; les nominalisations processives, quant à elles, ne le pourraient pas. C'est en tous cas, ce qu'attestent les exemples suivants :

- Nominalisation résultative : La description de Max du suspect
 La preuve de Max de ce théorème
- Nominalisation processive : *L'agression de Max du bandit
 *La pollution des voitures de la capitale

Or il est tout à fait possible de trouver des exemples de nominalisations non résultatives contenant deux complément en *de* qui peuvent être acceptables. Certains sont lourds ou peu élégants (110), mais d'autres sont très naturels (111) :

- (110) a. ?? La conquête éclair des maquisards du versant nord de la colline
 b. ? L'acquisition récente de Luc de trois tableaux de De Staël
 c. ? L'oubli volontaire de Luc de tous les détails qui peuvent le déranger
- (111) a. La descente des Champs-Élysées du Général de Gaulle
 b. L'ascension de l'Everest de Edmund Hillary et Tenzing Norgay
 c. La traversée de la Manche de Louis Blériot

De tels exemples devraient pourtant être exclus par le principe de non redondance fonctionnelle : leurs compléments en *de* ne répondent pas aux critères de définition des compléments prépositionnels :

- (112) a. *La descente de lui des Champs-Élysées
 *La descente du Général de Gaulle est des Champs-Élysées
 *La descente des Champs-Élysées est du Général de Gaulle
- b. *L'ascension d'eux de l'Everest
 *L'ascension de l'Everest est de Edmund Hillary et Tenzing Norgay
 *L'ascension de Edmund Hillary et Tenzing Norgay est de l'Everest
- c. *La traversée de lui de la Manche
 *La traversée de la Manche est de Blériot

*La traversée de Louis Blériot est de la Manche

De plus, à la différence des nominalisations résultatives, ils sont compatibles avec les modifieurs temporels ou processifs :

- (113) a. La descente des Champs-Élysées du Général de Gaulle en Août 1944 (fut un événement majeur)
 b. L'ascension de l'Everest de Edmund Hillary et Tenzing Norgay en neuf semaines (fait partie des records établis à ce jour)
 c. La traversée de la Manche de Louis Blériot en 1909 (fut le point culminant de sa carrière de pilote)

Les nominalisations statiques posent également un problème du point de vue de la redondance fonctionnelle. [Samvelian, 1995] fournit ainsi des exemples de groupes nominaux pour lesquels les compléments introduits par *de* correspondent au sujet et au complément d'objet direct de la construction verbale :

- (114) a. La maîtrise de Max de cette technique
 b. La connaissance de Max de l'informatique
 c. L'ignorance des responsables des vrais problèmes

Ces nominalisations sont issues de verbes transitifs statiques et préservent les propriétés du verbe dont elles sont issues en désignant, tout comme le verbe, un état. Elles sont ainsi compatibles avec les modifieurs temporels apparaissant dans la construction verbale :

- (115) Les responsables ignorent les vraies dimensions du problème depuis les premiers signes du mécontentement
 L'ignorance des responsables des vraies dimensions du problème depuis les premiers signes du mécontentement

On n'est donc pas en présence de nominalisations résultatives. Ces nominalisations présentent toutes les caractéristiques des nominalisations proprement dites (présence obligatoire du complément d'objet, formation d'une nominalisation négative de la forme *non* – *N*, où *non* a la même valeur que la négation totale *ne* – *pas*).

Les compléments en *de* de ces nominalisations ne peuvent donc pas être considérés tous les deux comme des compléments prépositionnels. La seule solution pour que le principe de redondance ne soit pas violé est alors de considérer qu'un des compléments directs du verbe doit se transformer en complément prépositionnel dans le groupe nominal. Le complément correspondant au

sujet de la construction verbale se réaliserait comme un génitif subjectif dans le groupe nominal, celui correspondant au complément d'objet se réaliserait comme un complément prépositionnel. Or, aucun des critères proposés par [Milner, 1982] n'atteste cette hypothèse :

1. Le critère du pronom fort n'est pas pertinent, car les compléments non humains ne peuvent être associés à un pronom fort
2. Le complément d'objet ne peut apparaître dans une construction attributive :

(116) *L'ignorance des responsables est des vraies dimensions du problème

Ainsi, le nom *ignorance* se comporte comme une nominalisation proprement dite, puisqu'il conserve les propriétés aspectuelles et la structure de compléments du verbe dont il est issu. Les deux compléments introduits par *de* manifestent les propriétés de compléments directs, et peuvent pourtant être simultanément présents dans le groupe nominal, ce qui devrait en principe, être exclu par le principe de redondance fonctionnelle.

De tout cela, Samvelian conclut que les différences constatées ne peuvent pas recevoir une explication en termes d'une différence de structure. L'hypothèse qui voudrait que deux génitifs directs ne puissent pas être cooccurents dans un groupe nominal est donc remise en cause¹⁸.

Mais dans ce cas, comment expliquer ces phénomènes ? En fait, pour Samvelian, la prise en compte de facteurs sémantiques pourrait les expliquer plus adéquatement. En effet, elle pose l'hypothèse que :

Le degré d'acceptabilité d'un groupe nominal contenant deux compléments génitifs décroît en fonction de deux facteurs :

1. L'accroissement du degré auquel l'argument réalisé comme un génitif objectif est affecté par le procès ou l'événement décrit par le nom tête du groupe nominal.
2. L'accroissement du degré de la participation "externe" de l'argument réalisé comme un génitif subjectif au même procès ou événement.

Autrement dit, plus l'argument désigné par le génitif objectif est "affecté" et plus l'argument désigné par le génitif subjectif a une participation "indirecte" ou "externe" au procès décrit par le prédicat, et moins le groupe nominal est acceptable [Samvelian, 1995, p174]"

¹⁸Le principe non redondance fonctionnelle est également problématique avec les déterminants possessifs qui apparaissent dans des nominalisations processives. En effet, pour [Milner, 1982], un groupe nominal comme *sa description du suspect* peut recevoir deux lectures : une résultative et une événementielle. Il peut donc correspondre à deux structures différentes. L'une (correspondant à l'interprétation résultative), où les compléments en *de* sont des groupes prépositionnels, l'autre (l'interprétation processive) où les deux compléments sont des génitifs directs.

La première ne viole pas le principe de non redondance, mais la seconde est équivalente à celle d'un groupe nominal comme **La description de Luc du suspect (dura 5 minutes)*, exclue par le principe de non redondance. Il semble difficile d'expliquer pourquoi une structure est autorisée dans un cas, et exclue dans l'autre.

Ainsi, les Champs-Élysées restèrent-ils identiques après que le Général de Gaulle les eut descendus ; de la même façon que l'Everest et la Manche restèrent inchangés après le passage d'Edmund Hillary et de son sherpa Tensing Norgay, ou de Louis Blériot.

En revanche, dans un groupe nominal tel que **l'adoucissement du vent de la température*, la température subit un changement : elle devient, sous l'action du vent, plus douce qu'elle ne l'était auparavant.

En outre, ces principes sémantiques permettent également de rendre compte des nominalisations statiques qui autorisent quasi-systématiquement deux compléments introduits par *de*. En effet, les noms statiques ne désignent pas un événement ou un procès. Ils ont donc des "compléments d'objet" non affectés par excellence.

3.2.2.4 Vers une synthèse des analyses

Une grande partie des travaux que nous avons présentés jusqu'à présent, tentait d'expliquer certaines propriétés des compléments en *de* (possibilité d'être associé à un déterminant possessif, d'être extrait par *en* ou *dont*), ainsi que leurs différentes possibilités combinatoires par la structure interne du groupe nominal. Les données et conclusions apportées par Samvelian nous laissent plutôt penser qu'il est assez difficile de voir dans la distinction complément direct vs. complément prépositionnel une explication des phénomènes observés.

Ce qui est certain, c'est que les compléments en *de* sont parfois difficiles à distinguer. Les critères permettant d'affirmer que tel complément est direct ou que tel autre est prépositionnel sont loin d'être évidents et fournissent parfois des résultats contradictoires. [Kupferman, 2004] signale ainsi que "les nominaux dérivés dont Milner avait montré qu'ils interdisaient aux génitifs la position attribut ou la représentation pronominal" peuvent être améliorés et admettre ces distributions :

- (117) a. La prochaine consécration sera d'un écrivain inconnu
 La consécration idéale ne peut être que de l'Académie française
 Une telle consécration de lui en tant qu'écrivain a tourné la tête à Luc
- b. Le meilleur accueil qu'il ait eu était d'amis fidèles
 Le meilleur accueil qui ait été fait au festival a été d'un film à petit budget
 Il avait compris qu'un meilleur accueil de lui comme l'acteur de l'année exigeait d'autres moyens

Dès lors, les compléments en *de* pourraient tous être considérés comme des compléments prépositionnels. Cette conclusion serait par ailleurs étayée par le fait que *de* peut porter sur une coordination de SN [Abeillé, et al., 2003], ce qui n'est pas possible si *de* est une marque casuelle :

- (118) a. La venue de [la mère et la fille] était programmée de longue date
 b. La photo de [mon père et ma mère] assis côte à côte près de la cheminée date des années 60.

La proposition de Samvelian paraît très intéressante et englobe davantage de données que les analyses précédentes. Cependant considérer le caractère plus ou moins affecté d'un argument nous paraît difficile à mettre en œuvre dans le cadre de notre grammaire. Samvelian signale d'ailleurs qu'il s'agit là, pour elle, d'une "ébauche d'explication". Elle se dit par ailleurs "consciente du caractère vague et approximatif des termes tels que 'affecté', 'participation interne'" et signale que ces propriétés "sont difficiles, voire impossibles, à formaliser".

L'analyse de Godard concernant la structure argumentale, nous paraît en revanche tout à fait formalisable dans le cadre de notre métagrammaire. Certes, elle ne s'intéresse pas explicitement aux possibilités d'association d'un complément en *de* à un déterminant possessif, mais elle nous semble également applicable dans ce cas de figure. En effet, les similarités distributionnelles mentionnées depuis [Brunot, 1926] entre le *en* adnominal et la réalisation du déterminant comme un possessif, sont assez régulières, hormis dans le cas d'un nom tête psychologique comme l'indiquent [Samvelian, 1995], et [Kupferman, 1996] à sa suite :

- (119) a. La hantise de l'exclusion est devenue générale
 L'exclusion, la hantise en est devenue générale
 *L'exclusion, sa hantise est devenue générale
 b. La nostalgie du passé est grandissante
 Le passé, la nostalgie en est grandissante
 *Le passé, sa nostalgie est grandissante

De plus, les critiques que l'on pourrait formuler envers l'analyse de Godard concernant les possibilités d'extraction des compléments, nous semblent non pertinentes dans le cadre de la détermination possessive. En fait, les acceptabilités de Godard vis-à-vis de l'extraction pourraient être trop restrictives. C'est en tous cas ce que mettent en évidence les données initialement prises en compte par [Ruwet, 1972] :

- (120) ?J'ai lu la description de Pline du cataclysme
 ?(Ce cataclysme) j'en ai lu la description de Pline
 ?Le cataclysme dont j'ai lu la description de Pline est horrible

Mais dans le cadre de la détermination possessive, les exemples correspondants nous semblent totalement exclus :

- (121) (Pline), sa description du cataclysme est saisissante
 *(Ce cataclysme), sa description de Pline est saisissante

Pour notre part, nous nous en tiendrons donc à des considérations d'ordre général, étant admis que les acceptabilités peuvent varier d'un locuteur à l'autre. Concernant les possibilités d'apparition de deux compléments en *de*, on considérera que cela est toujours possible. Concernant la détermination du groupe nominal par un déterminant possessif, on retiendra que :

1. Le complément en *de* correspondant au sujet du verbe support peut toujours être associé à un déterminant possessif : c'est un premier argument. Les autres compléments en *de* ne le peuvent si le complément en *de* correspondant au sujet du verbe support est présent, car ils n'accèdent pas au statut de premier argument. Pour que ces autres compléments échappent à cette situation, il faut que le complément en *de* correspondant au sujet ne soit pas réalisé, ou bien apparaisse sous la forme d'un syntagme en *par* :

(122) Le ministère fait un bilan pessimiste du compostage en France

Le ministère, son bilan du compostage est pessimiste

*Le compostage, son bilan du ministère est pessimiste

Le ministère, son bilan est pessimiste

Le compostage, son bilan est pessimiste

Le compostage, son bilan par le ministère est pessimiste

2. Les nominaux psychologiques ne répondent pas à ce schéma général. Même si le complément en *de* correspondant au sujet du verbe support n'est pas réalisé, l'association avec un déterminant possessif reste impossible

(123) Les chiens ont une haine féroce des chats

Les chiens, leur haine des chats est féroce

*Les chats, leur haine des chiens est féroce

Les chiens, leur haine est féroce

*Les chats, leur haine est féroce

Bien entendu, on pourra objecter qu'il s'agit ici davantage d'un traitement en largeur qu'en profondeur. On l'a vu, le problème n'est pas trivial. Pour notre part, la question n'est pas de savoir ce que l'on doit générer, mais plutôt de savoir comment on doit le contraindre. Les conclusions dégagées permettent d'ores et déjà une prise en compte relativement consensuelle des données.

3.3 Premiers choix de représentation pour les TAG

On l'a dit, la description des CVS proposée par [Abeillé, 1988, Abeillé, 1991] pour les grammaires TAG comporte plusieurs inconvénients. Outre le fait qu'elle n'a pas été systématisée par l'utilisation d'une métagrammaire, elle place sur un pied d'égalité verbe support et nom prédicatif. Pourtant, une nette dissymétrie est lexicalement observable. De plus, la plupart des compléments prépositionnels n'avaient pas de véritable statut et étaient simplement considérés comme les feuilles d'un arbre à tête verbale. Aucune mobilité des compléments ne pouvait être observée.

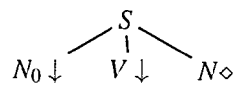
Les nouvelles familles que nous allons définir doivent non seulement rendre compte des CVS, mais également des groupes nominaux complexes qui leur sont associés. On l'a vu, dans une CVS, le nom prédicatif porte des fonctions syntaxiques variées et est considéré comme une tête sémantique, alors que le verbe support est considéré comme une tête syntaxique. Ce dernier est par ailleurs optionnel puisqu'il n'apparaît pas dans un groupe nominal complexe.

3.3.1 Des familles particulières et des entrées lexicales nouvelles

Notre grammaire doit être le reflet de ces constatations. De manière globale, une famille à nom prédicatif est donc décrite par son nom prédicatif seul¹⁹. Ce dernier sert d'ancre, alors que son verbe support, dépourvu de sens prédicatif, est substitué. C'est donc seulement pour la propagation des traits morpho-syntaxiques que le verbe support peut être dit "tête" de la phrase.

En plus des arbres phrastiques, les familles à nom prédicatif contiennent également des arbres élémentaires de racine N, permettant de rendre compte du SN complexe. La relation systématique entre la construction du SN et la CVS correspondante, est ainsi assurée :

¹⁹Un tel traitement contraste clairement avec les solutions adoptées par [Abeillé, 1988, Abeillé, 1991]. La distinction avec une expression figée est alors évidente.



INDEX : pirouette/1
 ENTRY : pirouette
 POS : N
 FAM : n0vN
 FS : V :<vsup>= faire
 EX : Max fait une pirouette

Deux arbres de la famille n0vN,
représentant une CVS, et un groupe nominal complexe

Lexique syntaxique du nom prédicatif *pirouette*
décrivant la famille n0vN

A l'image des verbes attributifs et causatifs, les verbes supports sont associés à l'arbre αV . Cet "arbre" rudimentaire ne contient pas d'équations. Ces dernières sont intégralement apportées par le lexique syntaxique. Par exemple, les verbes *faire* et *avoir* ont les descriptions suivantes :

$V\Diamond$ INDEX : faire/4
 ENTRY : faire
 POS : V
 TREES : αV
 FS : V.b :<caus>=, V.b :<cop>=, V.b :<aux>=avoir,
 V.b :<vsup>=faire
 EX : Max fait une promenade

Arbre αV

Lexique syntaxique du support *faire*

$V\Diamond$ INDEX : avoir/5
 ENTRY : avoir
 POS : V
 TREES : αV
 FS : V.b :<caus>=, V.b :<cop>=, V.b :<aux>=avoir,
 V.b :<vsup>=avoir, V.b :<passif>=
 EX : Max a une peine immense

Arbre αV

Lexique syntaxique du support *avoir*

La gestion de la forme du verbe support est réalisée par le trait $\langle vsup \rangle$. Au niveau lexical, renseigné par le nom prédicatif, ce trait indique quelles sont les formes autorisées par le nom. Porté par le verbe support, le trait indique la forme du verbe support. Ce trait $\langle vsup \rangle$ prend la valeur d'une liste de verbes supports (éventuellement réduite à un seul élément) qui indique quels sont les verbes supports autorisés d'un nom prédicatif particulier pour une sous-catégorisation particulière.

Ainsi, par exemple, alors que le nom prédicatif *gifle* attend, entre autres, les verbes supports *donner*, *flanquer* ou *passer*, le nom prédicatif *fessée* attend, quant à lui, les verbes supports *donner* ou *flanquer*. Une phrase comme **Max a passé une fessée* est naturellement rejetée, puisque le verbe support *passer* aura comme trait $\langle vsup \rangle = passer$, alors que le nom prédicatif aura comme trait $\langle vsup \rangle = donner/flanquer$ ²⁰.

On trouvera ci-après quelques exemples (simplifiés) extraits des lexiques syntaxiques, permettant au lecteur de se faire une idée plus précise des choix effectués²¹. Rappelons simplement ici que le champ INDEX indique le lemme de base (sans complémentation) et sa sémantique. Le chiffre qui suit le lemme renvoie à la sémantique particulière du lemme : tous ceux qui sont suivis par le même chiffre ont en principe même signification. Le champ ENTRY fait apparaître le lemme avec sa complémentation, renseignée par les introducteurs de complément. Le champ POS (pour Part of Speech) indique les parties du discours en rapport avec l'ENTRY. Le champ FAM indique le nom de la famille concernée. Le champ FS (pour *feature structure* en anglais) indique les traits éventuellement associés :

INDEX : capacité/1

ENTRY : capacité

POS : N

FAM : n0vN(des1)

FS : V : $\langle vsup \rangle = avoir$

EX : Max a la capacité de s'entendre avec tous

INDEX : capacité/1

ENTRY : capacité

POS : N

FAM : n0vN(as1)

FS : V : $\langle vsup \rangle = avoir$

EX : Max a la capacité à rendre sérieuse cette analyse

²⁰De manière analogue, le trait $\langle vconv \rangle$ indique quels sont les conversifs possibles d'un nom prédicatif donné.

²¹Une description plus détaillée des familles sera réalisée au cours du chapitre suivant concernant la dimension 1 de la métaграмmaire.

INDEX : électrochoc/1
 ENTRY : électrochoc
 POS : N
 FAM : n0vN(an1)
 FS : V :<vsup>=faire
 EX : Max fait un électrochoc au patient

INDEX : électrochoc/1
 ENTRY : électrochoc sur
 POS : N PREP1
 FAM : n0vNpn1
 FS : V :<vsup>=faire/pratiquer/réaliser
 EX : Max pratique un électrochoc sur ce patient

INDEX : électrochoc/1
 ENTRY : électrochoc sur
 POS : N PREP1
 FS : V :<vsup>=procéder
 EX : Max procède à un électrochoc sur ce patient

INDEX : admiration/1
 ENTRY : admiration
 POS : N
 FAM : n0vN(an1)
 FS : V :<vsup>=porter
 EX : Max porte une admiration sincère à Léa

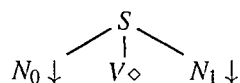
INDEX : admiration/1
 ENTRY : admiration pour
 POS : N PREP1
 FAM : n0vNpn1
 FS : V :<vsup>=avoir
 EX : Max a une admiration sans borne pour Léa

INDEX : traitement/1
 ENTRY : traitement
 POS : N
 FAM : n0vN(an1)
 FS : V :<vsup>=accorder / apporter / donner
 EX : Max accorde un traitement particulier à cette nouvelle

INDEX : traitement/2
 ENTRY : traitement sur contre
 POS : N PREP1 PREP2
 FAM : n0vaNpn1pn2
 FS : V :<vsup>=procéder
 EX : Max procède à un traitement sur le bois
 contre les vers

INDEX : traitement/2
 ENTRY : traitement contre
 POS : N PREP2
 FAM : n0vaNden1pn2
 FS : V :<vsup>=procéder
 EX : Max procède à un traitement du bois contre les vers

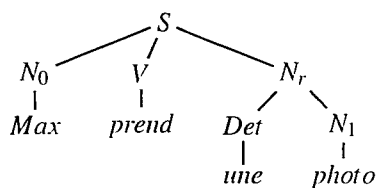
Une telle organisation permet de différencier clairement le comportement des familles à nom prédicatif, de celui des familles verbales. En effet, même si les CVS peuvent entrer en concurrence avec des phrases à verbe plein d'apparence semblable, leur analyse sera différente. L'analyse d'une phrase ambiguë comme *Max prend une photo* produira un arbre dérivé globalement semblable pour les deux types de familles. C'est au niveau de l'arbre de dérivation que se fera la différence, les unités lexicales entrant en jeu étant différentes.



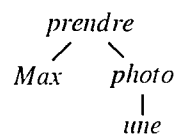
Arbre de la famille n0Vn1(an2)
permettant l'analyse de la phrase à verbe plein

INDEX : prendre/3
ENTRY : prendre
POS : V
FAM : n0Vn1(an2)
FS :
EX : Max a pris un sylo à Léa

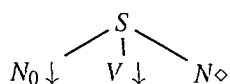
Lexique syntaxique associé à *prendre*



Arbre dérivé



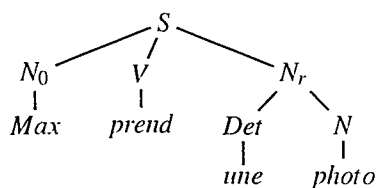
Arbre de dérivation



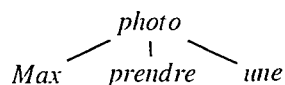
Arbre de la famille n0vN(den1)
permettant l'analyse de la phrase à verbe support

INDEX : photo/1
ENTRY : photo
POS : N
FAM : n0vN(den1)
FS : V : <vsup>=prendre
EX : Max a pris une photo de Léa

Lexique syntaxique associé à *photo*



Arbre dérivé



Arbre de dérivation

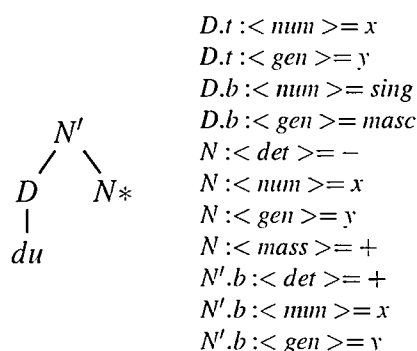
Exemples d'analyses pour la phrase *Max prend une photo*

3.3.2 Représentation des CVS

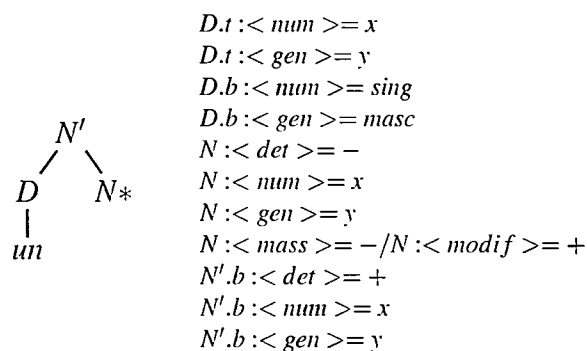
3.3.2.1 La détermination du nom dans les CVS

On l'a vu, les phénomènes touchant à la détermination du nom prédicatif sont complexes, et dépendent en général de facteurs sémantiques. La grammaire TAG du français dispose déjà d'un système qui permet de rendre compte de certains phénomènes touchant à la détermination. Dans ce système, les déterminants sont adjoints [Abeillé, 1991]. Nous le reprenons donc en partie.

Ainsi signale-t-on que le déterminant partitif n'est autorisé qu'avec un nom massique, en indiquant à son nœud pied l'équation $\langle mass \rangle = +$; les noms prédicatifs massifs portent quant à eux l'équation $N.t : \langle mass \rangle = +$, apportée par le lexique syntaxique. Dès lors, n'est autorisée l'adjonction du déterminant partitif que sur un nom massique.



Pour autoriser l'adjonction du déterminant indéfini avec un nom massique particulier, on marque au niveau du déterminant indéfini une disjonction $\langle mass \rangle = - / \langle modif \rangle = +$ à son nœud pied. Le déterminant indéfini a donc la représentation suivante :



On autorise ainsi la représentation de CVS à nom massique comme (124), tout en excluant (125) :

- (124) a. Max a du courage
b. Max a un courage remarquable

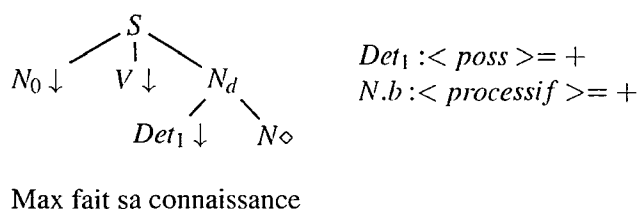
- (125) *Max a un courage

Pour prendre en compte des exemples comme (126), il faut obliger l'adjonction simultanée du déterminant et du modifieur. On indique ceci par l'équation $N.t : \langle det \rangle = N.t : \langle modif \rangle$.

- (126) a. Les enfants ont une peur bleue du loup
b. *Les enfants ont une peur du loup

En effet, au niveau des arbres, le nom prédicatif porte toujours le trait $N.b : \langle modif \rangle = -$. L'adjonction du modifieur est donc obligatoire si en top de N, on a $N.t : \langle modif \rangle = +$. Lorsque le nom n'est pas déterminé, il porte le trait $\langle det \rangle = -$. On a donc sur le nom prédicatif l'équation $N.t : \langle modif \rangle = -$. L'adjonction d'un modifieur est donc impossible. En revanche, lorsque le nom est déterminé, il porte le trait $\langle det \rangle = +$. Dès lors, l'adjonction du modifieur est donc obligatoire.

Le traitement le plus original concerne le déterminant possessif. Lorsqu'il est libre, il est traité comme les autres déterminants, c'est-à-dire qu'il est adjoint. Mais lorsqu'il est une reprise d'un complément prépositionnel en *de*, il est alors à substituer, puisqu'il reprend un argument :



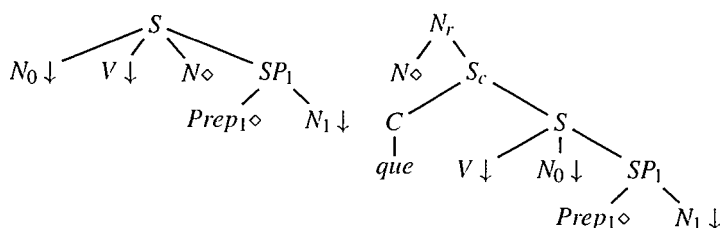
3.3.2.2 Représentation des compléments d'une CVS

Si les données menant à la double analyse de [Gross, 1976] ne peuvent être remises en cause, elles soulèvent néanmoins un problème important. En effet, même si les phrases à verbe support ne sont pas perçues comme ambiguës, elles reçoivent pourtant deux analyses. Or, comme le souligne [Abeillé, 1991], "c'est rompre la relation habituelle entre analyse syntaxique et analyse sémantique que d'assigner deux structures différentes à une phrase non ambiguë".

Nous pensons pour notre part qu'il n'est pas nécessaire d'avoir une double structure simultanée pour une phrase particulière. La grammaire peut considérer les compléments comme des compléments de verbe ou comme des compléments du nom, mais doit toujours effectuer un

choix. Autrement dit, pour un complément particulier, il peut y avoir double fonction et double structure, mais pas simultanée ; pour une phrase particulière, il n'y a ni double fonction ni double structure.

Lorsque les compléments sont mobiles et qu'ils autorisent la double analyse, ils sont considérés comme des compléments du verbe support. Ils ne sont considérés comme complément du nom, que s'ils sont extraits avec le nom prédicatif²² :

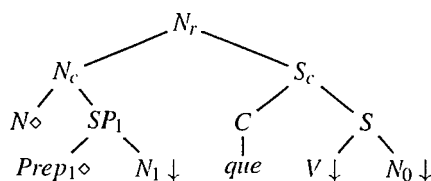


$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{prep_objet}$

Max fait des réclamations
auprès de la Poste

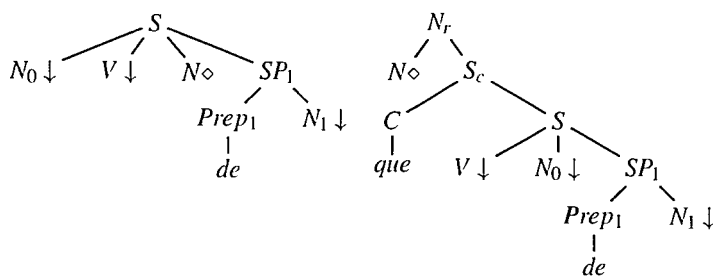
$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{prep_objet}$

Les réclamations que fait
Max auprès de la Poste



$N_c : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{cdn_prep}$

Les réclamations auprès de la Poste que fait Max



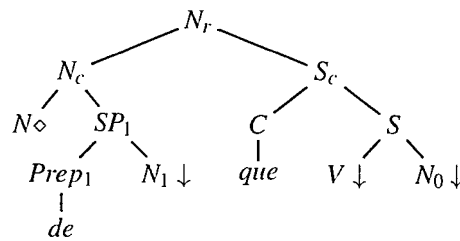
$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{de_objet}$

Max a peur du loup

$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{de_objet}$

La peur qu'a Max du loup

²²Nous introduisons dans les arbres proposés un trait $< \text{fonc} >$ explicite suivant en cela [Abeillé, 2002]. Ce trait n'apparaît pas explicitement dans les arbres de [Candito, 1999a].

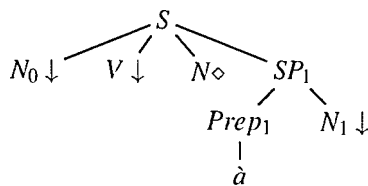


$N_c : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{cdn_de}$

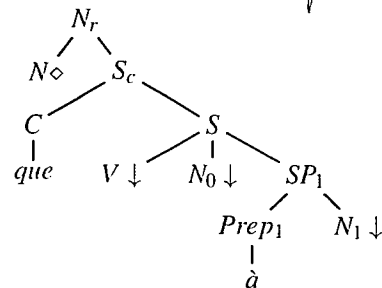
La peur du loup qu'a Max

Si l'extraction du complément avec le nom prédicatif est impossible, le complément n'est alors considéré que comme un complément de verbe. Le trait *fonc* sert alors à sélectionner les arbres adéquats, en permettant de rejeter les arbres de la famille à complément du nom. Au niveau des arbres à complément du nom, le trait *fonc* du syntagme prépositionnel porte la valeur *cdn_a*. Au niveau du lexique, on indique que le complément prépositionnel ne peut porter que la fonction *a_objet*. Ce qui permet donc la sélection des arbres adéquats.

*peu lisible
comme codage des
possibilités d'extraction*



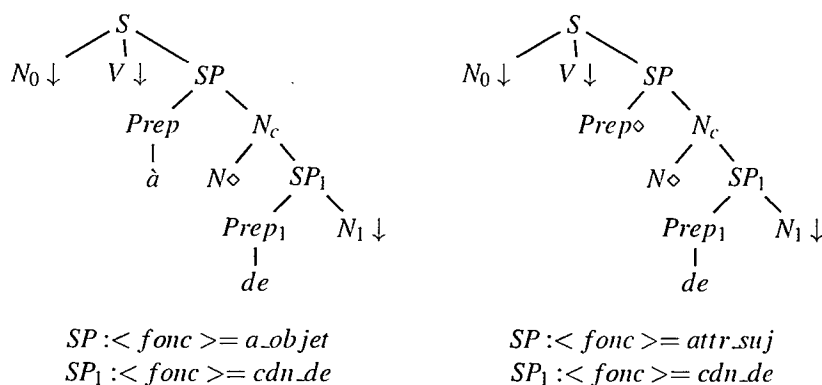
$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{a_objet}$



$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{a_objet}$

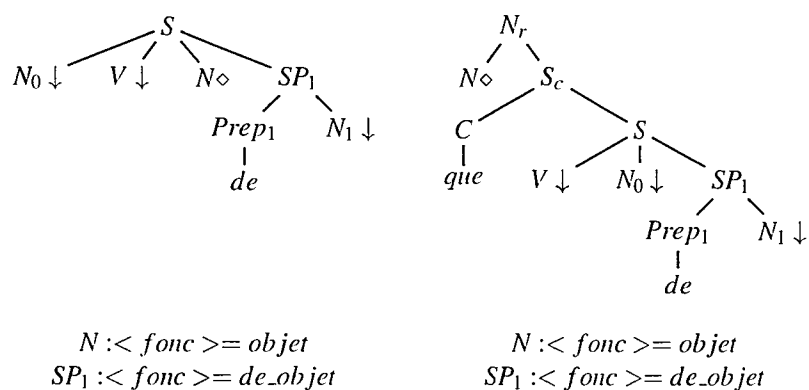
Les soldats donnent la charge à l'ennemi La charge que donnent les soldats à l'ennemi

En revanche, lorsqu'aucune mobilité ne peut être constatée et que la double analyse ne s'applique pas, les compléments sont uniquement représentés comme des compléments du nom. Cette particularité concerne les compléments en *de* des familles qui placent le nom prédicatif en position d'objet indirect ou d'attribut prépositionnel.

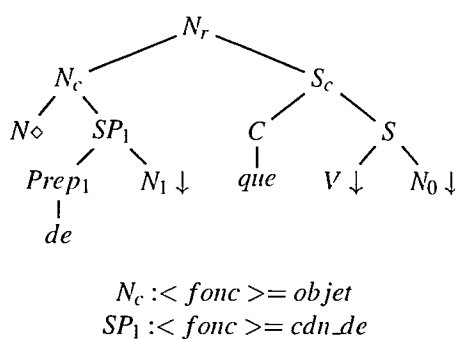


Max procède à l'analyse des données Max est dans l'attente de cet événement

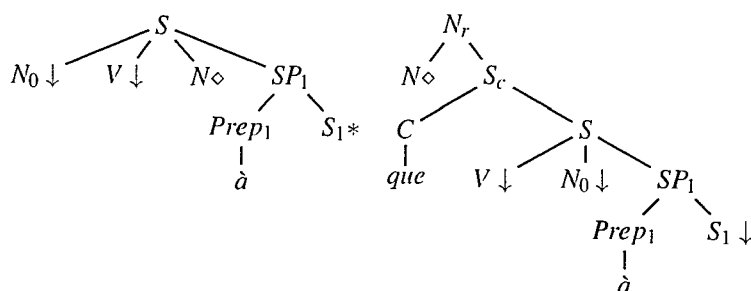
Pour les cas intermédiaires, pour lesquels la mobilité est plus difficilement observable, mais dans lesquels la double analyse est généralement constatée, on propose un traitement des compléments comme complément du verbe support. Là encore, les compléments n'apparaîtront comme complément du nom que lorsqu'ils seront extraits avec le nom prédicatif :



Max fait une étude des corps solides L'étude que fait Max des corps solides



L'étude des corps solides que fait Max

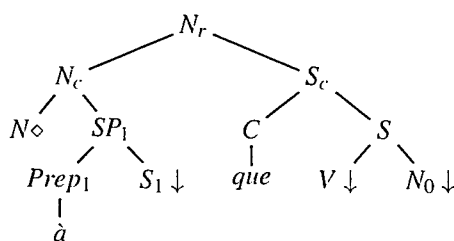


$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = a_objet$

$N : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = a_objet$

Max a tendance à exagérer

La tendance qu'a Max à exagérer



$N_c : < \text{fonc} > = \text{objet}$
 $SP_1 : < \text{fonc} > = \text{cdn_a}$

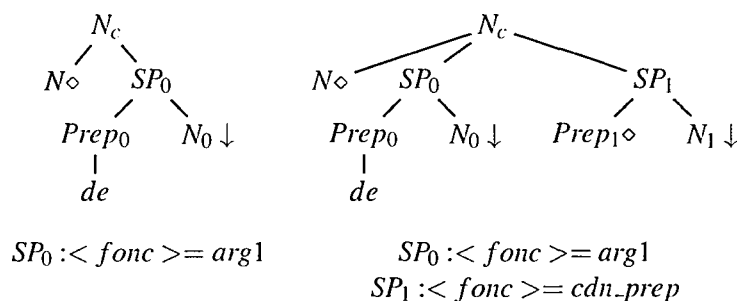
La tendance à exagérer qu'a Max

3.3.3 Représentation des groupes nominaux complexes

La représentation des groupes nominaux complexes soulève plusieurs points délicats. Il faut non seulement rendre compte de la réalisation de N_0 comme complément en *de* ou en *par*, mais également rendre compte de l'apparition du déterminant possessif.

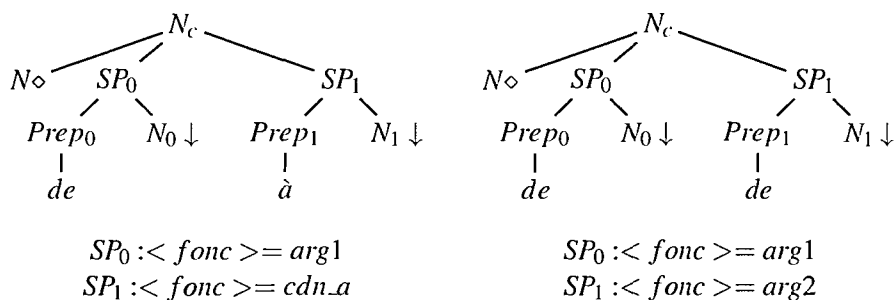
3.3.3.1 Groupes nominaux "actifs" et "passifs"

Dans les groupes nominaux "actifs", la préposition *de* qui introduit le sujet est placée directement dans l'arbre. La préposition du complément est la même que celle de la CVS :



La promenade de Max

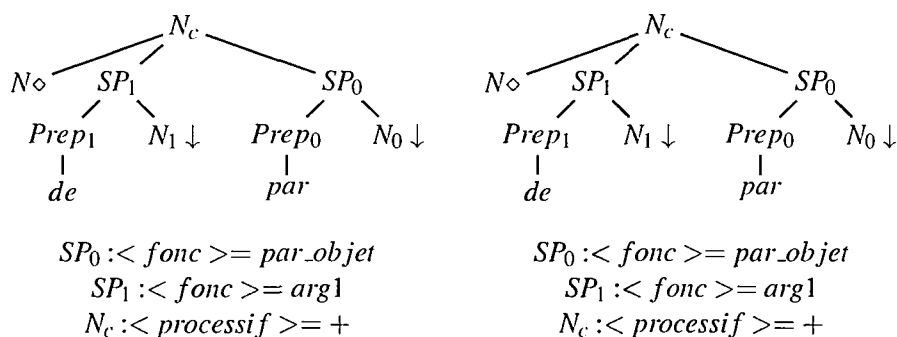
L'accusation de Max contre Luc



Le chantage de Max à Luc

Le portrait de Doisneau des Amoureux

Les groupes nominaux "passifs" ont deux sources possibles. Ils peuvent être les correspondants des arbres phrastiques passifs de la phrase à verbe support, ou bien les correspondants de la construction converse. Dans tous les cas, la préposition *par* est introduite directement dans l'arbre et le caractère processif du groupe est indiqué. On considère que le complément du nom prédicatif est obligatoire.



Le portrait des amoureux par Doisneau

L'agression de Luc par Max

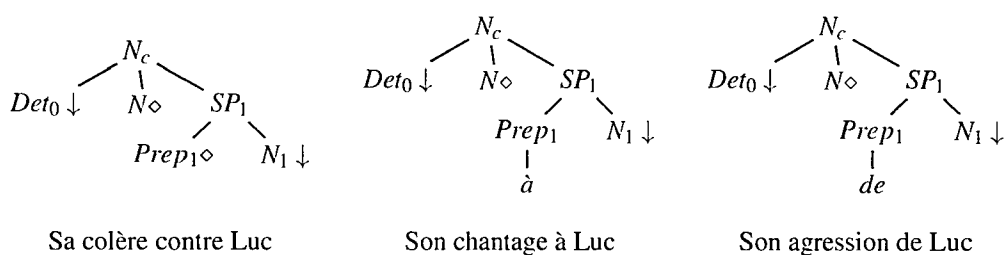
La formation des groupes nominaux "passifs" est réservée aux familles à nom prédicatif avec complément en *de*²³. Lorsque le verbe support ne peut être passivé, la formation du groupe nominal passif est également exclue. Cependant, au niveau de l'organisation de notre grammaire, il ne suffit pas que l'arbre phrastique passif soit exclu, pour que le groupe nominal "passif" soit automatiquement exclu. En effet, si les arbres phrastiques passifs sont ignorés, c'est parce qu'ils renseignent au niveau du trait bottom du verbe le trait $V.b : < passif > = -$. Cette information est apportée par le verbe support. Lors de la formation du groupe nominal passif, le verbe support n'est pas présent. L'information est donc perdue.

Seule l'information concernant le caractère processif du nom prédicatif est toujours accessible. C'est donc cette dernière qui va permettre ou non la formation du groupe nominal. Par exemple, un nom prédicatif psychologique apportera par son lexique l'information $< processif > = -$, mais comme l'arbre nominal "passif" renseignera au niveau du nom le trait $< processif > = +$, la formation sera interdite.

peu lisible

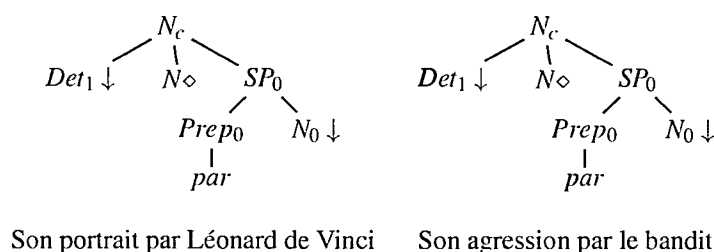
3.3.3.2 Le déterminant possessif

L'alternance des groupes en *de N* avec le déterminant possessif est représentée suivant les conclusions des sections précédentes. La réalisation comme déterminant possessif n'est ainsi possible que pour les groupes qui portent la fonction de premier argument. On note donc que le déterminant possessif est toujours possible lorsqu'il est coréférent au sujet du verbe support. Le déterminant est alors à substituer (puisqu'il représente un argument) et sa forme est contrainte par l'équation $Det_0 : < poss > = +$:

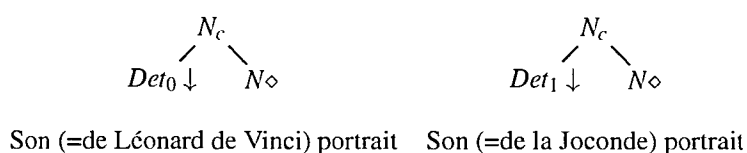


Lorsque *de N* correspond à un autre argument que le sujet, il peut être associé à un déterminant possessif, si le complément *de N* correspondant au sujet n'est pas réalisé. Là encore, la forme du déterminant est contrainte (mais par l'équation $Det_1 : < poss > = +$). Les noms psychologiques échappent cependant à cette règle, puisque même dans ce cas, l'association est tout de même interdite. On note donc le trait $N : < psycho > = -$ pour limiter les possibilités de formation du déterminant possessif. Un nom psychologique portera le trait $< psycho > = +$, ce qui interdira toute identification avec ces arbres :

²³ On rappelle que l'on n'a trouvé en corpus qu'une seule occurrence d'un groupe nominal "passif" ayant pour origine une CVS dont le complément prépositionnel est introduit par une autre préposition que *de*.



On notera qu'un groupe nominal ambigu, comme *son portrait* recevra deux représentations suivant que le déterminant corresponde au sujet du verbe support ou à son autre complément :



Au terme de cette première partie, nous pouvons donc nous faire une idée de ce que sera la prise en compte des CVS pour la grammaire TAG du français. Nous reviendrons plus précisément sur certaines caractéristiques représentationnelles plus spécifiques des CVS aux cours des chapitres suivants qui concerneront le Métagrammaire, mais nous avons d'ores et déjà dégagé les grandes lignes qui conduiront notre travail : les compléments prépositionnels ou phrastiques peuvent être considérés soit comme des compléments de verbe, soit comme des compléments du nom, soit à la fois comme des compléments de verbe ou des compléments de nom. Ces considérations guideront les possibilités de double analyse, qui ne seront donc pas régulières. Pour l'instant, nous n'avons fait qu'esquisser les choix qui sous-tendent de manière générale les représentations des CVS, mais les chapitres suivants mettront en lumière le côté plus pratique de l'implémentation de notre grammaire par le biais de la Métagrammaire.

Deuxième partie

Une métagrammaire pour les noms prédicatifs du français

Chapitre 4

Dimension 1 : la sous-catégorisation initiale

Sommaire

4.1	Quelles fonctions utiliser ?	130
4.1.1	Sous-catégorisation sémantique et sous-catégorisation syntaxique . . .	130
4.1.2	Quelles fonctions syntaxiques retenir ?	131
4.1.3	Les fonctions retenues : exemples et remarques	140
4.2	Liste des familles à nom prédicatif	141
4.2.1	Conventions de nommage	141
4.2.2	Familles à arguments nominaux	143
4.2.3	Familles à argument phrastique	144
4.3	Implémentation	146
4.3.1	Limitations dûes au compilateur	146
4.3.2	Organisation hiérarchique	148
4.3.3	Vision graphique de la hiérarchie	150
4.3.4	Principales différences avec l'implémentation de Candito	152

Le présent chapitre illustre les choix que nous avons entrepris en vue de la description des familles à verbe support. Certains s'imposent d'eux mêmes, nos travaux s'inscrivant dans la continuité de ceux de [Abeillé, 1991, Abeillé, 2002] et [Candito, 1999a], d'autres sont plus spécifiques. Avant d'entrer précisément dans le détail des différentes fonctions initiales retenues, nous effectuons quelques rappels, tant d'un point de vue pratique que théorique.

4.1 Quelles fonctions utiliser ?

4.1.1 Sous-catégorisation sémantique et sous-catégorisation syntaxique

Jusqu'à présent, lorsque nous nous sommes intéressé à l'organisation interne d'une métaگرامmaire, nous nous sommes contenté de signaler que la dimension 1 était le domaine de la sous-catégorisation initiale fonctionnelle. Nous ne nous sommes guère interrogé sur la notion de fonction, ni même sur celle de sous-catégorisation. Or, il n'est absolument pas certain que [Candito, 1999a] ait vu dans cette dimension un domaine d'information uniquement syntaxique¹.

Rien n'empêche en effet de considérer que la sous-catégorisation puisse également s'effectuer en termes sémantiques. Ce qui demeure invariant, c'est que la notion de fonction, qu'elle soit syntaxique ou sémantique, désigne toujours le rôle qu'un élément particulier joue par rapport à la tête. C'est pourquoi, la fonction se définit toujours en termes relationnels.

Le cas des verbes pleins est un exemple typique où l'on n'observe pas de divergence entre une dépendance syntaxique et une dépendance sémantique. Lorsque l'on dit *Max court*, le verbe *courir* gouverne le nom *Max* à la fois syntaxiquement et sémantiquement.

Il reste cependant possible d'examiner des cas où des divergences s'opèrent :

- Le rapport entre gouverneur syntaxique et gouverné syntaxique, peut être inversé sémantiquement. Le gouverneur syntaxique est alors gouverné sémantiquement, alors que le gouverné syntaxique est lui-même gouverneur sémantique. Ce cas peut être illustré par les adjectifs épithètes. Par exemple, dans *un linguiste remarquable*, l'adjectif *remarquable* gouverne sémantiquement le nom *linguiste*, alors que syntaxiquement, il est gouverné par lui.
- Il peut également y avoir un rapport sémantique entre deux entités, sans que ces deux entités soient reliées syntaxiquement. Là encore, les adjectifs peuvent être un cas d'illustration. Considérons cette fois-ci l'emploi attributif de l'adjectif. Dans *ce linguiste est remarquable*, *remarquable* gouverne là-encore *ce linguiste*, mais il n'y a aucune dépendance syntaxique directe entre les deux, si l'on considère que c'est le verbe *être* qui est tête syntaxique.
- Le dernier est le cas inverse du précédent : il peut y avoir dépendance syntaxique sans qu'il y ait pour autant dépendance sémantique. Ce cas est illustré par [Mel'čuk, 1988] pour le français avec l'exemple *Une plainte confuse de l'enfant*, où l'adjectif *confus* est syntaxiquement dépendant du mot *plainte*, alors qu'il est sémantiquement lié à *enfant*.

On a bien vu au cours du chapitre précédent que tête sémantique et tête syntaxique ne se confondent pas avec les CVS : c'est le verbe support qui est tête syntaxique, alors que le nom prédicatif est tête sémantique. Il n'y a que lors de la formation du groupe nominal complexe que le nom prédicatif peut revêtir les deux rôles.

¹[Candito, 1999a] fait p.285 la remarque suivante : "la MG considère que la sous-catégorisation d'un lexème est le reflet des arguments sémantiques".

Ceci pose donc la question de savoir ce que doit représenter la sous-catégorisation dans la dimension 1. La réponse n'est pas évidente. Si l'on choisit une sous-catégorisation sémantique, on ne va attribuer aucune fonction au nom prédicatif. Or, on l'a vu, les CVS sont des constructions syntaxiquement régulières, le nom prédicatif peut non seulement apparaître dans de nombreuses positions, mais également être relativisé ou clivé. Dans ce cas, ce serait la dimension 2 qui devrait prendre en compte ces différentes réalisations. La démarche ne nous semble pas impossible, mais elle nous semble très contraignante, d'autant qu'il faudra également rendre compte des changements de diathèse (passif, constructions moyennes...)².

Le plus simple nous semble de considérer le verbe support comme la tête syntaxique. Le nom prédicatif sera signalé comme prédicat sémantique, au niveau de la grammaire de bas niveau. De cette manière, on met ainsi bien en évidence que les CVS sont des constructions syntaxiquement régulières. La hiérarchisation des données sera ainsi globalement semblable à celle des verbes pleins. Seul le contenu des classes sera différent.

4.1.2 Quelles fonctions syntaxiques retenir ?

Nous savons maintenant que les fonctions que nous allons utiliser seront des fonctions syntaxiques, dont le gouverneur sera généralement le verbe support. Mais il nous faut trouver d'une part, une liste de fonctions adéquates, et d'autre part des éléments définitoires qui leur correspondent.

La notion de fonction a longtemps divisé les chercheurs en linguistique. Dans les années 60, [Chomsky, 1965] et [Gross, 1969] les avaient critiquées. Il a fallu attendre le début des années 80 et les grammaires lexicales fonctionnelles [Bresnan et Kaplan, 1982] pour que la notion soit réhabilitée comme primitive de la théorie syntaxique. On s'accorde ainsi à dire que les fonctions syntaxiques permettent de mieux représenter les universaux langagiers.

Bien entendu, il n'est pas question pour nous de faire un résumé de tout ce qui a pu être entrepris sur ce vaste sujet. Les approches peuvent par ailleurs grandement différer d'un linguiste à l'autre et varier également d'un pays à l'autre selon la tradition grammaticale en vigueur.

²[Candito, 1999a] reste relativement vague quant à la sous-catégorisation quand elle n'est pas syntaxique. On sait par exemple que les adjectifs sont considérés comme tête sémantique, mais aucune implémentation en terme de métagrammaire n'est vraiment renseignée. Le nom, tête syntaxique des constructions épithètes porte la fonction modifieur. La dimension 2 est à peine renseignée et en particulier, elle ne stipule pas si elle prendra en compte les adjectifs épithètes, les attributs du sujet et les attributs de l'objet. La mise en œuvre semble ici possible dans la mesure où la dimension 2 ne comporte pas beaucoup de classes finales pour les adjectifs. Notons ici que cette implémentation n'a pas du tout été suivie pour la réalisation des familles adjectivales pas [Barrier, 2002].

Pour le français, on sait depuis longtemps que la catégorie et la position ne suffisent pas seules à déterminer les propriétés des syntagmes. Comme le signale [Abeillé, 2002], "dans un modèle basé sur des structures de surface, on doit ajouter la notion de fonction pour distinguer les comportements syntaxiques de syntagmes de même catégorie occupant la même position". Ainsi, un syntagme nominal en position post-verbale peut occuper toujours la même position, tout en ayant des fonctions différentes, comme l'illustre la série (127) où *une jeune fille* est objet en (127a)³, attribut en (127b)⁴ ou sujet en (127c)⁵ :

- (127) a. Max regarde la jeune fille
 b. Léa sera bientôt une jeune fille
 c. Le lac où se baigne cette jeune fille (est magnifique au printemps)

Les grammaires d'unification font usage de la notion de fonction de manière plus ou moins importante. En HPSG par exemple, on dispose d'une liste courte de fonctions, toutes regroupées sous Compléments. Au contraire, en LFG, la liste de fonctions est beaucoup plus détaillée. L'approche est même hybride, puisqu'elle incorpore la catégorie au nom de la fonction (ACOMP, NCOMP...).

4.1.2.1 Les fonctions utilisées par Candito

Il nous reste donc à dresser un inventaire des fonctions grammaticales disponibles. Pour les verbes pleins, Candito a, dans sa thèse, retenu un certain nombre de fonctions syntaxiques. Même s'il n'est absolument prouvé à ce stade qu'elles seront toutes nécessaires, leur examen peut se révéler éclairant.

Les distinctions opérées permettent de prendre en compte des propriétés syntaxiques différentes. Elles peuvent également être liées aux représentations TAG :

- (128) Sujet : MAX dort
 Objet : Max mange UN KIWI
 A1-objet : Max parle À LUC

³Le groupe peut être cliticisé par *la*.

⁴Le groupe peut être cliticisé par *le*.

⁵Le groupe peut être cliticisé par *elle*. On n'a cependant pas **Le lac où se baigne elle*, mais *Le lac où elle se baigne*.

A2-objet : Max pense À LUC
 Gen-objet : Max rêve DE LÉA
 Obl-objet : Max compte SUR LÉA
 Locatif : Max va À VISAN
 Locatif de provenance : Max vient DE VISAN
 Manière : Max va BIEN
 Infinitive de but : Max court à la librairie ACHETER UN LIVRE
 Infinitive de perception : Max voit Léa ACHETER UN LIVRE
 Interrogative indirecte : Max demande à Léa SI LA TERRE EST RONDE
 Attribut du sujet : Max est HEUREUX
 Attribut de l'objet : Max trouve Léa MERVEILLEUSE

D'emblée, il apparaît que certaines fonctions ne sont pas purement syntaxiques, et pourraient correspondre à des variantes de réalisation. Les compléments phrastiques semblent ainsi suivre un traitement particulier. Mais leur justification n'est pas immédiatement évidente. En particulier, est-il nécessaire de distinguer une interrogative indirecte, alors que la fonction objet existe déjà⁶ ? Ce choix, semble contredire la remarque de Candito p.36 qui mentionne "qu'il est dommage de distinguer plusieurs fonctions dès que les catégories diffèrent".

Le recours aux fonctions infinitive de but et infinitive de perception semble également déroutant. La raison apparaît néanmoins chez [Candito, 1999b] : ces fonctions doivent être utilisées, car les infinitives de but et de perception doivent être distinguées en italien. Il faudrait donc établir la même distinction pour le français.

Hormis ces cas discutables, les autres fonctions retenues sont plus habituelles. Les objets indirects sont scindés selon la préposition qui les introduit (et son éventuelle valeur sémantique). Ceci permet de rendre compte des différentes possibilités de cliticisation du complément prépositionnel. A ce titre, on observe que les compléments en *à* sont dédoublés⁷. Cela tient là encore aux possibilités de cliticisation du complément :

- (129) a. Max parle à Luc
 Max lui parle
 *Max y parle

⁶On sait tout de même que les interrogatives indirectes et les autres objets phrastiques ont des comportements différents vis-à-vis des dépendances à longue distance. Les interrogatives indirectes forment systématiquement un îlot pour l'extraction en français, ce qui n'est pas nécessairement le cas des objets complétifs.

⁷Ces deux fonctions sont regroupées sous la fonction A-objet dans l'implémentation réalisée par Candito.

- b. Max pense à Luc
 *Max lui pense
 Max y pense

Reste qu'à notre sens toutes ces fonctions ne sont pas des fonctions initiales. L'attribut de l'objet peut être introduit en dimension 2 sans problème majeur⁸, si l'on considère que la tête lexicale est l'adjectif et non le verbe.

Si l'on examine maintenant ces fonctions, on observera que toutes ne sont pas nécessaires pour la représentation des CVS. L'emploi des fonctions infinitives est ainsi exclu, alors qu'il pourrait être potentiellement attendu :

- (130) a. Max traverse la rivière rejoindre Léa
 *Max fait la traversée de la rivière rejoindre Léa
- b. Max voit Léa rentrer dans la librairie
 *Max a la vision de Léa rentrer dans la librairie

De la même manière, les interrogatives indirectes sont exclues :

- (131) a. Max demande si la terre est vraiment ronde
 *Max fait la demande si la terre est vraiment ronde
- b. Max demande qui a fait cela
 *Max fait la demande qui a fait cela

Au total, on pourrait donc pour l'instant retenir les fonctions : sujet, objet, a-objet, gen-objet, obl-objet, locatif, locatif de provenance et attribut du sujet.

4.1.2.2 De nouvelles fonctions nécessaires

Ces fonctions se révèlent cependant insuffisantes. On a vu au chapitre précédent que certaines familles sous-catégorisaient des compléments du nom. Il conviendra donc d'ajouter également la fonction de complément du nom en *de*, permettant de rendre compte d'exemples comme :

- (132) Max procède à l'analyse de ce texte
 Max est dans l'attente de ces résultats

⁸En fait, Candito introduit cette fonction directement en dimension 1 car elle considère que les familles à attribut du sujet et à attribut de l'objet doivent être séparées. Les nouvelles familles adjectivales développées par [Barrier, 2002] regroupent non seulement au sein d'une même famille à tête adjectivale les attributs du sujet et les attributs de l'objet, mais également les épithètes.

4.1.2.3 Les fonctions syntaxiques retenues

Nous disposons donc maintenant d'un jeu de fonctions syntaxiques. Encore convient-il de les définir, car à ce stade, ce ne sont pratiquement que de simples étiquettes.

Le sujet Nous reprenons en grande partie les définitions proposées par [Riegel et al., 1996]. Rappelons ici que nous ne nous intéressons qu'aux différentes propriétés syntaxiques de ces fonctions et que nous essayons autant que faire se peut d'éliminer les définitions sémantiques.

Le sujet n'est pas effaçable et précède normalement le verbe (en phrase non marquée) :

- (133) Max mange un kiwi
 * Mange un kiwi
 * Mange un kiwi Max

Il régit l'accord du verbe en personne et en nombre (et aussi en genre lorsque le participe passé est conjugué avec l'auxiliaire être).

Il est le seul élément qui puisse être extrait par une clivée de la forme *C'est ... qui* :

- (134) Max mange un kiwi.
 C'est Max qui mange un kiwi.

Il peut revêtir la forme d'un syntagme nominal (135a), d'une complétive (135b) ou d'une infinitive (135c), mais pas d'un syntagme adjectival (136a) ou d'un syntagme prépositionnel (136b) :

- (135) a. Max mange un kiwi
 b. Que Max mange un kiwi est primordial
 c. Manger un kiwi donne des forces à Max
- (136) a. *Fort est la seule façon de taper sur ce clou
 b. *Sous la table est une mauvaise cachette

Enfin, lors d'un changement de diathèse, au passif, le sujet devient complément d'agent.

Nous considérons que le sujet est un complément comme les autres et doit donc rentrer à ce titre dans la sous-catégorisation. On peut en effet noter que les verbes imposent les mêmes types de contraintes à leur sujet et à leurs compléments (cf [Abeillé, 1996b]). De la même manière qu'un verbe contraint la catégorie nominale ou phrastique de ses compléments (137), il peut également contraindre celle de son sujet (138) :

- (137) Max comprend le problème - Max comprend qu'un problème se trame.
 Max analyse problème - *Max analyse qu'un problème se trame.
 Max accepte les critiques - Max accepte qu'on le critique.
 Max repousse les critiques - *Max repousse qu'on le critique.
- (138) Ce départ ennuie Max - Que Léa soit partie ennuie Max.
 Ce départ rebute Max - *Que Léa soit partie rebute Max.
 L'envie de partir démange Léa - ?Partir démange Léa - * Qu'elle parte maintenant démange Léa.

Et de la même façon que le verbe contraint le mode de sa complétive (infinitif, indicatif ou subjonctif) (139), le verbe peut également imposer ce type de sélection à son sujet phrastique (infinitif ou subjonctif) (140) :

- (139) Max ose leur parler sur ce ton - *Max ose qu'on parle sur ce ton.
 Max trouve qu'il a bonne mine - *Max trouve avoir bonne mine.
- (140) Que Léa soit amoureuse a transpiré dans la presse - *Aimer Max a transpiré dans la presse.
 *Que tu ailles à Venise est cher - Aller à Venise est cher.

L'objet direct L'objet direct peut prendre toutes les formes du groupe nominal et de ses équivalents pronominaux ou propositionnels. Il est généralement caractérisé par les critères suivants :

- Interrogative de la forme *Qui/que regarde-t-il ?*
- Pronominalisation en *le/la/les*.
- Variation de diathèse : tout comme le sujet, l'objet peut lui aussi être affecté par la mise au passif. Il est alors promu sujet

Malheureusement, ces propriétés ne sont pas toutes convergentes. [Gross, 1969] note qu'elles sont indépendantes, c'est-à-dire que selon les cas, certaines d'entre elles peuvent ou non s'appliquer. De plus, le passif peut s'appliquer à des compléments indirects (*Max pardonne à Luc - Luc est pardonné de Max*), et la pronominalisation en *le* peut également s'appliquer aux attributs (*Max est heureux - Max l'est*).

[Abeillé, 1996b] retient deux critères pour caractériser l'objet direct : l'emploi du *en* quantitatif qui permet de distinguer l'objet du complément indirect et du sujet (même inversé), mais pas de l'attribut nominal (141), et la négation en *pas ... de*, qui permet d'isoler un objet direct, d'un attribut nominal et d'un objet indirect, mais pas d'un sujet inversé (142) :

- (141) Objet : Max a vu trois Porsche dans la rue - Max en a vu trois
 Attribut : Max est un bon conducteur - Max en est un
 Sujet : Trois Porsche dorment dans le garage - *Trois en dorment dans le garage.
 Sujet inversé : Le jour où trois Porsche dormiront dans le garage... - *Le jour où en dormiront (trois) dans le garage...
 Objet indirect : Max parle à trois jeunes filles - *Max en a parlé à trois.
- (142) Objet : Max a vu des Porsche - Max n'a pas vu de Porsche.
 Attribut : *Max n'est pas de bon conducteur
 Sujet : *De Porsche ne dorment pas dans le garage.
 Sujet inversé : Le jour où ne dormiront pas de Porsche dans le garage...
 Objet indirect : *Max n'a pas parlé à de jeunes filles

Une réaction positive à ces deux tests permet d'isoler l'objet direct nominal, mais pas l'objet phrastique.

On remarquera par ailleurs, que ces deux derniers critères permettent également d'analyser comme objet ce que les grammaires traditionnelles appellent le sujet logique des constructions impersonnelles (143) et causatives (144) :

- (143) Il est arrivé plusieurs accidents
 Il en est arrivé plusieurs
 Il n'est pas arrivé d'accident
- (144) Max fait conduire les élèves.
 Max en fait conduire (plusieurs)
 Max n'a pas fait conduire les élèves

En définitive, les critères pour identifier l'objet sont moins nombreux et moins fiables que pour le sujet.

L'objet indirect Passons maintenant à l'objet indirect qui est encore plus problématique que les deux premiers arguments que nous venons d'évoquer, dans la mesure où il est parfois difficile de le différencier des circonstants⁹.

⁹Par exemple, [Aït-Mokhtar et al., 2003] établit une comparaison des entrées verbales dans deux dictionnaires : dans l'un, "un argument de type sur+SN était présent pour *agir* (agir sur quelque chose) ; un argument de+SN était présent pour *saigner* (saigner du nez)". Ces constructions étaient pourtant toutes absentes de l'autre.

[Wilmet, 1998] note que les grammaires "énumèrent les prépositions les plus aptes à construire" un objet indirect :

- (145) à : Max pense à Marie.
 de : Max rêve d'un kiwi bien mûr.
 en : Max part en vacances.
 dans : Cette problématique s'insère dans un cadre plus général.
 sur : Max compte sur son charme.
 contre : Max lutte contre l'insécurité.
 avec : Max danse avec Léa.
 chez : Max habite chez ses parents.
 par : Max est passé par des épreuves douloureuses.
 pour : Max vote pour le candidat idéal.

Mais plus on descend dans cette liste, plus "les doutes s'amplifient" : s'agit-il d'un complément ou d'un circonstant ? Ce que l'on peut dire, c'est que la pronominalisation des compléments indirects s'effectue selon 3 modèles :

- Les compléments introduits par *à* se pronominalisent par les formes conjointes *lui/leur/y* ou les formes disjointes *lui/elle(s)/eux* selon la nature du verbe.
- Les compléments introduits par *de* se pronominalisent par *en* s'ils sont non animés, ou par *de lui/d'elle(s)/d'eux* s'ils sont animés.
- Les compléments introduits par d'autres prépositions se pronominalisent par une forme disjointe s'ils sont animés.

[Riegel et al., 1996] notent que pour établir le caractère de complément du syntagme, "le critère décisif reste l'existence d'un double rapport de dépendance avec le verbe :"

- Rapport sémantique : le complément d'objet indirect est un actant dont le rôle sémantique est appelé par le sens du verbe.
- Rapport syntaxique : le verbe contrôle la construction du complément dont il détermine la plupart du temps la préposition introductrice.

L'objet indirect n'est en outre pas affecté par les changements de diathèse.

Le complément locatif Les compléments locatifs sont par ailleurs particularisés. Ces derniers étant caractérisés par une pronominalisation, un relatif et un pronom interrogatif spécifique :

- (146) Max va à Batz sur Mer.
 Pronominalisation : Max y va.
 Relative : Batz sur Mer où Max va (est magnifique en cette saison).
 Interrogative : Où va Max ?

Le locatif de provenance Le locatif de provenance se différencie de ce complément locatif. Il est également caractérisé par un relatif et un pronom interrogatif particulier (147), et peut par ailleurs être suivi d'un syntagme prépositionnel ou d'un adverbe locatif (148) :

- (147) Max vient de Bordeaux
 Pronominalisation : Max en vient
 Relative : Bordeaux d'où vient Max (est magnifique en cette saison)
- (148) Un bruit bizarre vient de sous le capot
 Un bruit bizarre vient de là

L'attribut du sujet Tout comme l'objet direct, l'attribut du sujet peut être pronominalisé. Mais à la différence des cas où il pronominalise les objets directs, le clitique *le* ne supporte aucune variation en terme de genre ou de nombre¹⁰ :

- (151) Léa semble heureuse.
 Léa le semble.
 *Léa la semble.

Son verbe introducteur est le verbe *être* ou un verbe d'état susceptible de lui être substitué.

De nombreux éléments peuvent porter cette fonction. Aux classiques syntagmes nominaux et adjectivaux, on peut par exemple ajouter les syntagmes prépositionnels (152a) et les infinitives (152b).

- (152) a. Max est en colère
 b. Souffler n'est pas jouer

¹⁰[Dupont, 1991] fait remarquer que "le clitique *le, la, les* semble respecter massivement les lois de l'accord du français, en genre et en nombre, du XVII^e siècle à nos jours, lorsqu'il est en relation sémantique avec un syntagme nominal défini ou assimilable" :

- (149) ... la Reine... je la suis
 ... sa maîtresse... je ne la serai jamais

L'accord peut aussi être observé avec des adjectifs :

- (150) ... sage... je la suis devenue
 ... perverse... je la deviens ; amoureuse... je la suis

L'emploi du clitique *le* quels que soient le genre et le nombre semble en tous cas avoir été plus tardif pour les syntagmes nominaux que pour les adjectifs (cf. [Dupont, 1991])

Le complément du nom Le complément du nom est souvent considéré par les grammaires traditionnelles comme un modifieur du nom. A ce titre, des grammaires comme celles de [Riegel et al., 1996] lui retiennent deux propriétés essentielles :

- il est toujours postposé au nom qu’il détermine et dont il restreint l’extension
- il est récursif, puisqu’il est à la fois constituant facultatif d’un syntagme nominal et qu’il a obligatoirement comme constituant un syntagme nominal qui peut comporter lui-même un syntagme prépositionnel

L’acception que nous donnons au terme de complément du nom est différente, en ce sens qu’il s’agit d’un complément syntaxique, qui entre dans la sous-catégorisation du nom prédicatif.

4.1.3 Les fonctions retenues : exemples et remarques

Nous retenons donc les fonctions syntaxiques initiales suivantes pour notre hiérarchie. Leurs propriétés seront complétées dans le chapitre concernant la dimension 3, mais il nous paraissait préférable de tenter de justifier les fonctions retenues. Nos fonctions sont en outre différenciées suivant le type de complément qu’elles peuvent représenter (nominal ou phrastique) et également par la préposition qui introduit leur groupe¹¹.

(153) Sujet (nominal) : MAX prend une douche.

Sujet (phrastique) : QUE MAX POSE NU fait scandale.

Objet direct : Max fait UNE SIESTE.

Complément introduit par *à* (nominal) : Max fait du chantage À LUC.

Complément introduit par *à* (phrastique) : Max a de la peine À RÉUSSIR.

Complément introduit par *de* (nominal) : Max fait la censure DE CETTE PAGE.

Complément introduit par *de* (phrastique) : Max a l’espoir DE RÉUSSIR.

Complément introduit par une autre préposition que *à* ou *de* (nominal) : Max commet un crime CONTRE LUC.

Complément introduit par une autre préposition que *à* ou *de* (phrastique) : Max fait un effort POUR RÉUSSIR.

Complément locatif : Max fait un pèlerinage À LOURDES.

¹¹ Cela tient au fait que les possibilités de cliticisation du complément prépositionnel sont différentes. Pour les attributs du sujet, on fait également une différence suivant la préposition qui introduit l’attribut, les prépositions *en* et *dans* pouvant commuter, à la différence de *à* :

- Max est en colère
- Max est dans une colère noire

Complément locatif de provenance : Max a fait une sortie remarquée DE LA SALLE DE RÉUNION.

Attribut prépositionnel du sujet (introduit par la préposition *à*) : Max est À SA RECHERCHE.

Attribut prépositionnel du sujet (introduit par la préposition *de*) : Max est DE BONNE HUMEUR.

Attribut prépositionnel du sujet (introduit par les prépositions *en* ou *dans*) : Max est EN COLÈRE.

Complément du nom introduit par *de* : Le conseil procède à l'examen DES ACTIVITÉS.

Quelques remarques s'imposent. En premier lieu, on notera que quatre fonctions sont obligatoires pour représenter les familles à nom prédicatif. La première est tout naturellement le sujet, les autres étant celles portées par le groupe prédicatif (objet, objet indirect introduit par *à*, attribut du sujet).

Le nom prédicatif porte lui aussi une fonction syntaxique, et ce même s'il est tête sémantique. Il n'est donc pas différencié à ce stade des autres compléments plus standards.

Nous manions pour l'instant un jeu de fonctions relativement réduit. Ce dernier sera complété en dimension 2 par de nouvelles fonctions : les fonctions finales.

Nous pouvons maintenant nous intéresser aux différentes familles que nous retenons pour notre implémentation.

4.2 Liste des familles à nom prédicatif

4.2.1 Conventions de nommage

Les noms de familles sont normalement transparents et suivent le schéma d'oblicité des arguments par rapport au verbe. La catégorie de l'ancre est indiquée en majuscule. Conformément aux conventions du LADL, on note le sujet avec l'indice 0, et les compléments de 1 à 2. Les syntagmes nominaux sont renseignés par *n*, les compléments phrastiques par *s*. Les syntagmes prépositionnels suivent un traitement différent selon que la préposition est ou non régie. Ainsi, on note aX^{12} un SP_à, avec *à* préposition régie, deX un SP_{de}, avec *de* préposition régie, et pX les autres SP. Certaines informations d'ordre sémantique sont aussi renseignées : c'est le cas de *loc* qui indique un complément locatif. Les arguments optionnels sont entre parenthèses.

Comme pour les familles des verbes pleins et des adjectifs, nous établissons une différence entre compléments nominaux et compléments phrastiques. Ceci ne se traduit cependant pas par

¹²X représente soit un *n*, soit un *s*

une sous-catégorisation fonctionnelle différente, mais juste par des familles différentes (c'est-à-dire que la sous-catégorisation en termes catégoriels est affectée). Même si l'on peut noter que de nombreux verbes acceptent également un complément nominal, lorsqu'ils acceptent un complément phrastique, les familles à argument phrastique n'établissent pas cette commutation¹³. L'argument phrastique admet une double réalisation : l'une comme complétive, l'autre comme infinitive.

L'ensemble de ces conventions est explicité en détail dans [Candito, 1999a]. Rappelons ici cependant qu'une famille se définit comme le nombre maximal de structures élémentaires correspondant à un cadre de sous-catégorisation donné, ce qui ne veut pas dire qu'un élément particulier entrera dans toutes ces structures à la fois (cf. [Abeillé, 1991]). Autrement dit, une famille est une classe d'équivalence avec régularité dans la variété de réalisation.

Les sous-sections suivantes vont s'attacher à décrire l'ensemble des familles retenues pour les familles à nom prédicatif du français. Toutes n'ont pas été implémentées. En particulier, le codage des familles en *être prep* reste à faire. Chaque sous-section est organisée de la manière suivante : un premier tableau établit la relation entre fonction et catégorie, puis un second tableau fournit un exemple typique pour les familles ainsi décrites.

Les abréviations sont les suivantes :

- N indique un syntagme nominal
- SP indique un syntagme prépositionnel
- Cl indique un clitique
- S indique un complément phrastique

Les fonctions renseignées concernent tous les compléments, y compris le nom prédicatif, qui rappelons-le, est dépendant syntaxique du verbe support.

Afin de faciliter la lecture, et d'identifier rapidement la tête sémantique, le groupe prédicatif a été grisé.

28 familles à nom prédicatif ont ainsi été renseignées. On a choisi de faire rentrer dans nos familles des exemples comme *Le vent souffle* ou *Un orage s'est produit sur la région*, suivant en cela [Mel'čuk, 1984] et [Vivès, 2004b].

¹³On peut voir en cela un besoin pratique, pour limiter le nombre d'arbres à générer. Les familles à arguments nominaux seraient alors des sous-ensembles des familles à arguments phrastiques. On imagine facilement le nombre d'arbres augmenter de manière artificielle. Néanmoins, il faut garder à l'esprit que ce choix n'est pas sans conséquences : le lexique doit être décrit de manière à différencier clairement arguments nominaux et phrastiques.

Famille	Exemple
n0vN	Max prend une douche
n0vN(an1)	Max fait du chantage à Luc
n0vN(den1)	Max fait la censure de cette page
n0vN(pn1)	Max commet un crime contre Luc
n0vN(loc1)	Max fait un pèlerinage à Lourdes
n0vN(deloc1)	Max fait une sortie remarquée de la salle de réunion
n0vN(den1)(an2)	Max fait le récit de ses aventures à Léa
n0vN(den1)(pn2)	Max fait la division de ce nombre par 9
n0vN(den1)(loc2)	Max fait une expédition de colis en Somalie
n0vN(pn1)(pn2)	Max fait une plaisanterie sur cette situation avec Luc
n0vAN(pn1)	Max procède à une enquête sur les indiens
n0vAN(den1)	Max procède à l'aménagement de sa chambre
n0etrePN	Max est de bonne humeur
n0etreEnN(den1)	Max est dans l'attente de ces résultats
n0etreEnN(pn1)	Max est en colère contre Luc
n0etreAN(den1)	Max est à la recherche de la meilleure solution
Nv	Le vent souffle
Nvpn1	Une malédiction pèse sur cette assemblée
Nvn0	Le chagrin brûle Max

4.2.3 Familles à argument phrastique

Famille	Fonctions initiales					Exemples
	S U J E T	O B J E T	A - O B J E T	D E - O B J E T	P R E P - O B J E T	
n0vN(as1)	N, Cl	N	S			peine
n0vN(des1)	N, Cl	N		S		espoir
n0vN(ps1)	N, Cl	N			S	effort
n0vN(an1)(des2)	N, Cl	N	SP, Cl	S		ordre
s0vN	S	N				scandale
s0vN(an1)	S	N	SP, Cl			plaisir
s0vN(den1)	S	N		SP, Cl		joie
s0vN(pn1)	S	N			SP	impression
Nvs1	N			S		bruit
Nvn0(des1)	N	N, Cl		S		envie

Famille	Exemple
n0vN(as1)	Max a de la peine à se décider
n0vN(des1)	Max garde l'espoir de réussir
n0vN(ps1)	Max fait des efforts pour réussir
n0vN(an1)(des2)	Max donne l'ordre à Luc de partir
s0vN	Poser nu fait scandale
s0vN(an1)	Que Max vienne fait plaisir à Léa
s0vN(den1)	Que Max doive partir fait la joie de Luc
Nvs1	Le bruit court que le baril de pétrole dépassera les 100\$
Nvn0(des1)	L'envie prend Max de partir

Par comparaison avec le nombre de familles à verbes pleins définies par [Candito, 1999a], nous disposons, pour les familles à noms prédicatifs, d'un nombre inférieur. Candito présentait ainsi dans sa thèse 56 familles : 24 familles à arguments nominaux et 32 à argument phrastique¹⁴. Plusieurs raisons peuvent expliquer ce nombre moindre :

- Nous n'avons pas de famille à objet direct phrastique (cette fonction est généralement portée par le nom prédicatif). Candito définit 14 familles à objet direct phrastique. Elles sont différenciées selon leur complémenteur éventuel (*de* ou *à*).
- Les familles à arguments phrastiques sont également distinguées chez Candito par le mode de réalisation de la complétive dans une phrase assertive¹⁵.
- Certaines familles sont également distinguées chez Candito suivant l'auxiliaire qu'elles utilisent¹⁶.

Notre sous-catégorisation est cependant plus riche en termes de fonctions employées. Les catégories sous-catégorisées sont également plus nombreuses¹⁷

¹⁴Le nombre de familles verbales est cependant plus important.

¹⁵Le mode de la complétive n'est pas directement renseigné au niveau des arbres, car le mode est soumis à diverses contraintes pas uniquement lexicales (la négation par exemple). Les familles qui attendent un subjonctif, ou un indicatif n'ont pas les mêmes propriétés (inversion possible du sujet de la complétive, coréférence du sujet de complétive et du verbe principal), cf [Abeillé, 2002].

¹⁶Les intransitifs qui sélectionnent des auxiliaires différents sélectionnent des arbres différents au niveau des participiales. On pourra noter cependant que [Abeillé, 2002] propose de regrouper les arbres issus de deux familles (différant par l'auxiliaire) au sein d'une seule et même famille en faisant porter l'informant concernant l'auxiliaire au verbe au sein du lexique syntaxique.

¹⁷[Candito, 1999a] ne sous-catégorise que des N ou des S. Si un élément doit se réaliser différemment, comme clitique par exemple, l'élément nominal initial n'est pas supprimé et a un fils vide. C'est cet élément vide qui porte la fonction, et non le clitique.

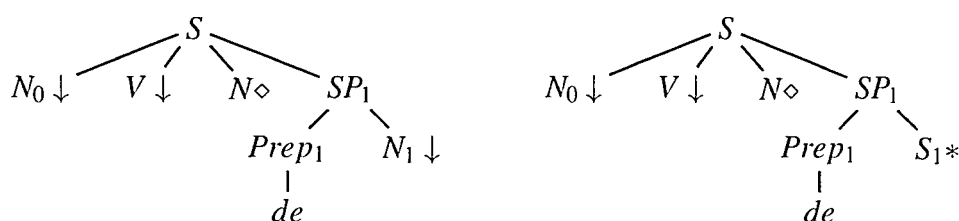
4.3 Implémentation

4.3.1 Limitations dûes au compilateur

4.3.1.1 Un problème de sous-catégorisation catégorielle

Certains des groupes auxquels nous avons assigné une fonction syntaxique, ne portent pas, dans l'implémentation de la Métagrammaire actuelle, cette fonction. En effet, c'est l'élément discriminatoire qui est amené à porter la fonction. Il serait sinon par exemple impossible de différencier un SP comportant un complément phrastique, d'un SP comportant un complément nominal. En toute rigueur, ces compléments ne devraient être considérés que comme des compléments de la préposition¹⁸.

L'exemple ci-dessous illustre la représentation du complément prépositionnel introduit par *de*. Ces 2 arbres appartiennent à 2 familles différentes : n0vN(den1) pour le premier et n0vN(des1) pour le second.



Deux arbres élémentaires permettant respectivement d'analyser les phrases
Max fait le procès de l'irresponsabilité (1) et *Max fait le rêve d'épouser Léa* (2)

Il faut en outre être cohérent au niveau du traitement de la représentation des différents groupes prépositionnels. Dans l'implémentation actuelle du compilateur, certains phénomènes de dimension 2 nécessitent en effet, de faire porter la fonction au niveau du complément de la préposition. Par exemple, pour un verbe à l'actif sous-catégorisant un sujet et un objet nominaux, l'équivalent passif fera correspondre au sujet nominal initial, un complément d'agent en *par* et à l'objet nominal initial, un nouveau sujet nominal. Malheureusement, même si les règles de redistribution vont effectivement modifier les fonctions initiales, elles ne vont pas changer les catégories initialement concernées¹⁹. L'objet initial nominal se réalisera sous forme d'un sujet final nominal, et ne sera donc pas problématique. En revanche, si l'on indique que c'est au syntagme prépositionnel en *par* de porter la fonction de complément d'agent, aucun arbre ne sera créé, car la fonction complément d'agent sera maladroitement associée au nom, idéalement complément de la préposition..

¹⁸Au niveau de la grammaire résultante, c'est cependant bien le groupe prépositionnel qui porte la fonction, car le compilateur différencie trait et méta-trait.

¹⁹Dans le même ordre d'idée, il faut également prévoir en dimension 1, la réalisation du complément du nom en *de* comme un déterminant possessif. Le sujet initial indique alors également qu'il sous-catégorise un Déterminant.

En considérant que le problème des règles de redistribution est facilement résolvable, on peut néanmoins proposer diverses solutions pour résoudre le problème de base.

Une première solution serait d'avoir recours aux possibilités du compilateur non décrites par [Candito, 1999a]. Comme l'unification sur les traits n'est pas réalisée par le compilateur, Candito avait senti le besoin de créer un attribut permettant de palier ce problème²⁰. Ainsi, pour déclarer qu'un sujet phrastique ne pouvait être à l'infinitif, elle indiquait que le sujet portait un attribut *type* dont la valeur était *ind* (pour indicatif) ou *subj* (pour subjonctif). L'idée serait ici la même, mais elle ne serait pas justifiée par une mauvaise unification sur les traits. On renseignerait simplement si le SP sous-catégorisé est de type nominal ou phrastique²¹. Ainsi le SP de la phrase *Max fait la censure de cette page* aurait un attribut *type* à valeur *nominale*, et le SP de la phrase *Max donne l'ordre de partir*, un attribut à valeur *phrastique*. La cohabitation des valeurs serait aisée, et sans problème, dans la mesure où chaque groupe serait réalisé par une classe différente. Ainsi pour *Max donne l'ordre à Luc de partir*, le premier SP a pour attribut *type nominal*, et le second *phrastique*.

Une autre solution pourrait être envisagée : il s'agirait dans ce cas d'utiliser des fonctions syntaxiques à la LFG [Bresnan, 2001]. Dans cette optique, on reprendrait donc des fonctions comme *Vcomp* pour désigner des infinitives, *Comp* pour désigner les complétives... De cette manière, les syntagmes prépositionnels seraient clairement distingués :

- Les fonctions *pObj*, *aObj*, *deObj* renverraient à des compléments nominaux
- Les fonctions *Vcomp* et *Comp* renverraient à des compléments phrastiques.

Ce type de solution est rejeté par [Candito, 1999a] qui trouve dommage de distinguer des fonctions dès que les catégories diffèrent.

Nous préférons quant à nous la première solution que nous venons de développer, même si elle n'entre pas tout à fait dans l'optique qu'envisageait Candito.

4.3.1.2 Absence de gouverneur syntaxique

Un autre problème est soulevé par le compilateur : il n'est pas possible d'indiquer quel est le gouverneur d'un syntagme qui porte une fonction. Si cette impossibilité passait relativement inaperçue dans l'implémentation des verbes pleins (le verbe restant de manière invariable la tête syntaxique de chaque groupe dont les fonctions avaient été définies), le problème se pose au niveau de l'implémentation des familles à nom prédicatif. En effet, on a vu au cours des chapitres précédents que les compléments ne sont pas tous des compléments verbaux. En particulier, certains compléments prépositionnels peuvent apparaître comme complément du nom. Mis à part ce nommage explicite, il ne sera formellement pas possible d'établir le gouverneur de ce groupe. Dans notre implémentation, les différents groupes portent donc une fonction sans

²⁰Il s'agit de l'attribut *:type*.

²¹On n'aurait d'ailleurs pas dans ce cas de classes à démultiplier inutilement, contrairement à ce qui se passe avec la différenciation du mode.

donner davantage de précision. En dimension 1, tous ces groupes dépendent majoritairement du verbe support. On pourrait indiquer sous forme de trait quel est le gouverneur syntaxique, mais cela ne serait pas d'un apport notable dans la mesure où aucun travail d'unification sur les traits n'est effectué par la métaگرامmaire. Cela aurait cependant l'avantage de faciliter la lecture dans une optique de maintenance et de débogage. Nous tâchons dans notre implémentation d'être le plus explicite possible de manière à éviter toute confusion.

4.3.2 Organisation hiérarchique

Nous allons maintenant nous attacher à la description de quelques classes de la métaگرامmaire. Ceci permettra au lecteur de se faire une idée plus précise de l'encodage de la hiérarchie pour le compilateur de Candito, et pourra éventuellement lui servir de tutoriel s'il souhaite se lancer dans la description d'une nouvelle hiérarchie.

Les classes de cette dimension, tout comme celles des autres dimensions sont typées, c'est-à-dire qu'il est possible en remontant le graphe d'héritage de déterminer quelle dimension est décrite.

Nous allons commencer notre description par la classe représentant le sujet initial. On a vu que le sujet peut prendre diverses catégories. Il s'agit du nom, du clitique, du complément phrastique, mais également du déterminant possessif.

La classe (SUJET_INITIAL) hérite de (DIMENSION1)
Elle définit un argument 0 qui peut être de catégorie N, Cl, S ou Det. Sa fonction est sujet
Syntaxe : (SUJET_INITIAL (DIMENSION1) (init-constants '(((arg0) :cat ("N" "Cl" "S" "Det") :function sujet))))

Cette classe n'est cependant pas assez précise. On a vu qu'on distinguait en TAG, les familles selon leur type de complément : phrastique ou nominal. On va donc établir cette distinction :

La classe (SUJET_PHR) hérite de (SUJET_INITIAL)
Elle indique que l'argument 0 est uniquement de catégorie S, et que son indice est 0
Syntaxe : (SUJ_PHR (SUJET_INITIAL) (init-constants '(((arg0) :cat ("S") :ind ("0"))))))
De manière analogue, la classe (SUJ_NOMINAL) indique que l'argument 0 peut porter les catégories N, Cl ou Det, mais ne précise pas son indice
Syntaxe : (SUJ_NOMINAL (SUJET_INITIAL) (init-constants '(((arg0) :cat ("N" "Cl" "Det") :ind ("0"))))))

Comme le sujet nominal peut être soit un argument, soit la tête sémantique, il faut également préciser quel est l'attribut porté par la variable argument 0.

La classe (SUJ-N-HEAD) hérite de (SUJET_NOMINAL)
Elle indique que l'argument 0 constitue l'ancre de l'arbre élémentaire et restreint encore sa réalisation sous forme unique de N. L'argument ne porte pas d'indice.
Syntaxe : (SUJ-N-HEAD (SUJET_NOMINAL) (init-constants '(((arg0) :cat ("N") :qual head))))
La classe (SUJ-N-SUBST) hérite de (SUJET_NOMINAL)
La classe (SUJ-N-SUBST) va indiquer que l'argument 0 est à substituer et qu'il porte l'indice 0. Aucune restriction sur la forme du sujet n'est à effectuer
Syntaxe : (SUJ-N-SUBST (SUJET_NOMINAL) (init-constants '(((arg0) :qual subst))))

Les autres fonctions suivent un traitement analogue, qui permet de distinguer les catégories à utiliser ainsi que l'attribut à faire porter aux variables. A partir de ce moment, on dispose donc d'un ensemble permettant de construire les cadres de sous-catégorisation adéquats. Il s'agit donc de construire les regroupements pertinents en suivant les cadres de sous-catégorisation fournis précédemment.

Par exemple, pour décrire la famille s0vN, on écrira que la classe (s0vN) hérite des classe (SUJ_PHR) et (OBJ-N-HEAD), ce qui permettra d'obtenir la sous-catégorisation pertinente²² :

La classe (s0vN) hérite des classes (SUJ_PHR) et (OBJ-N-HEAD)
La classe (s0vN) reprend ainsi les informations des classes précédentes et précise que le complément phrastique est substitué Elle indique également que le sujet ne pourra être qu'au subjonctif <i>Attention</i> : elle ne le fait pas en utilisant différentes équations sur les traits
Syntaxe : (s0vN (SUJ_PHR SUJET_INITIAL) (init-constants '(((arg0) :cat ("S") :type subj))) (init-name "s0vN" :direction :left))
Après héritage, la classe (s0vN) indiquera donc que l'on attend : - un élément sujet de catégorie "S" d'indice 0, au mode subjonctif, substitué - un élément objet, ancre, de catégorie "N"

²²La fonction `init-name` est utilisée pour indiquer le nom de l'arbre final. Elle a donc un rôle purement décoratif.

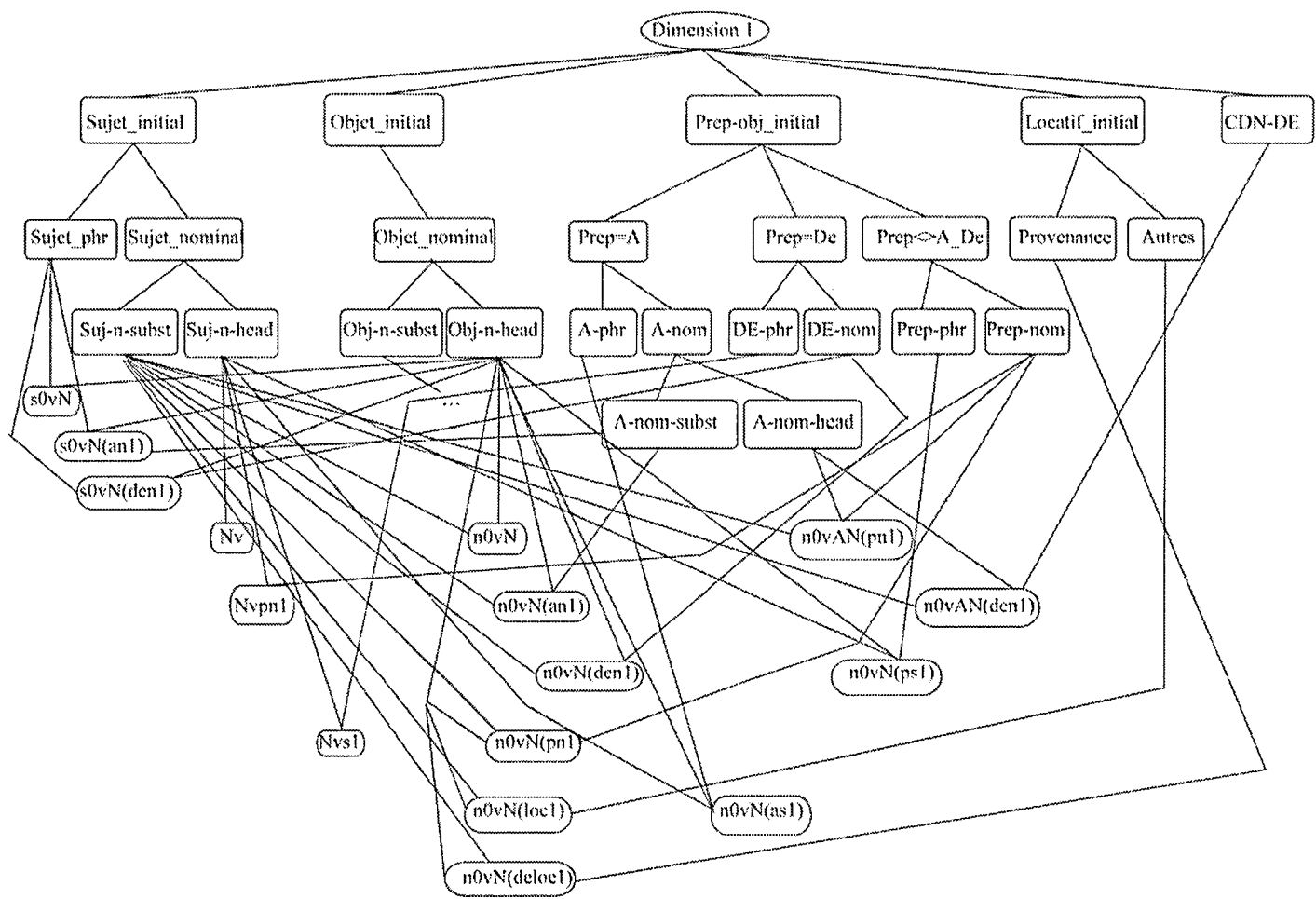
Pour finir sur un autre exemple, la classe (n0vN_pn1_) sera représentée de la façon suivante :

La classe (n0vN_pn1_) hérite des classes (OBJ-N-HEAD) (PREP-NOM) et (SUJ-N-SUBST)
Elle reprend le contenu de ces 3 classes
Syntaxe : (n0vN_pn1_ (PREP-NOM OBJ-N-HEAD SUJ-N-SUBST) (init-name "(pn1)" :direction :right)))
Après héritage, la classe (n0vN_pn1_) indiquera alors que l'on attend : - un élément sujet de catégorie "N" "CI" ou "Det" et d'indice 0 - un élément objet, ancre, de catégorie "N" - un élément objet-indirect de catégorie "N" et d'indice 1

L'ensemble des différents cadres de sous-catégorisation suit le même traitement. Les définitions de la dimension 1, en termes d'implémentation, sont donc très faciles, une fois les choix linguistiques effectués.

4.3.3 Vision graphique de la hiérarchie

Pour résumer une partie de l'organisation de la dimension 1, nous fournissons maintenant une vision graphique des classes qui la composent. Il est malheureusement impossible de fournir la vision globale de cette hiérarchie, tant les croisements sont nombreux. S'il fallait faire figurer sur ce graphique, l'ensemble des croisements effectués, ainsi que la totalité des familles, l'exemple perdrait tout caractère explicatif, tant il serait difficile de distinguer les croisements effectués.



4.3.4 Principales différences avec l'implémentation de Candito

On peut remarquer au travers de cette description que nous employons peu de classes intermédiaires : ceci pour faciliter la lecture et la maintenance future de la hiérarchie. En effet, plus le nombre de classes intermédiaires est important, plus la circulation au sein de la hiérarchie est difficile. Des noms de classe mnémoniques pour un rédacteur ne le sont pas forcément pour un autre.

4.3.4.1 Expression des généralités

La vision de la hiérarchie que nous adoptons pour la dimension 1 est assez différente de celle qu'en avait [Candito, 1999a] ; cela au moins sur 3 points...

Les classes qui pour nous sont les plus générales contiennent déjà toutes les informations. Par exemple, la classe (SUJET_INITIAL), renseigne immédiatement les éléments sujets comme des éléments de catégorie "N", "CI" ou "S". Elle peut être précisée par une autre classe dont la classe (SUJET_INITIAL) sera le parent. Candito, quant à elle, renseignait peu de choses sur le sujet, si ce n'est pour indiquer que l'élément qui allait remplir cette fonction serait d'indice 0. Ce n'est que dans les classes terminales qu'elle précisait quel élément pouvait ou non porter cette fonction. Si au final, le résultat reste le même, il montre cependant que l'on peut exprimer des généralités de deux façons :

- La classe la plus haute peut contenir tous les renseignements pertinents, et être précisée par des classes dont elle sera le parent.
- La classe la plus haute peut contenir les renseignements communs à toutes les classes dont elle sera le parent.

Nous préférons quant à nous la première solution dans la mesure où elle a l'avantage d'exprimer des généralités pour une langue donnée²³.

4.3.4.2 Déclaration des fonctions

Nous n'indiquons pas explicitement en dimension 1 que les fonctions syntaxiques que nous retenons sont des fonctions initiales²⁴, puisque par définition cette dimension ne renseigne que des sous-catégorisations initiales. On utilise donc le méta-trait général de *fonction*, réservé chez [Candito, 1999a] à la description des fonctions finales. Les fonctions sont donc distinguées suivant leur lieu d'utilisation :

²³Pour reprendre notre exemple, on précise ainsi dès le départ que certains éléments ne peuvent pas porter la fonction sujet.

²⁴C'est-à-dire que nous n'utilisons pas le méta-trait *fonction* – *i* défini par [Candito, 1999a].

- Les fonctions initiales sont décrites en dimension 1.
- Les fonctions finales sont décrites en dimension 2.

Définir explicitement des fonctions initiales en dimension 1, posait, selon nous, des problèmes de représentation. En effet, une fonction initiale non redistribuée ne peut être réalisée²⁵. Ceci amenait Candito à déclarer une sous-catégorisation en dimension 1 de manière double : une fois en utilisant des fonctions initiales, une autre fois en utilisant des fonctions finales. Ainsi, une fonction restée sans redistribution se trouvait tout de même réalisée²⁶. Cette déclaration, bien que pratique, ne nous paraît en tous cas pas justifiée théoriquement²⁷...

4.3.4.3 Un arbre en dimension 1 ?

Enfin, dernier point : contrairement à Candito, nous ne commençons à décrire aucun arbre en dimension 1, ceci étant normalement réservé aux dimensions 2 et 3. L'idée était pour Candito de représenter l'artère de la phrase, c'est à dire la séquence de nœuds en relation de dominance immédiate comprenant le verbe et ses projections jusqu'au premier nœud non verbal²⁸. Ce traitement précoce permettait selon elle, de représenter une sorte de morphologie par défaut. Mais pourquoi un traitement si avancé était-il nécessaire ? Ce type de description ne dispense en effet nullement de créer une classe en dimension 2 représentant ce cas non marqué. En adoptant cette position, Candito départicularise à la fois les dimensions 1 et 2 :

- La dimension 1 : en ne la réduisant plus à une simple déclaration de sous-catégorisation
- La dimension 2 : en tentant de lui ôter un cas de redistribution

L'autre problème qui se pose avec ce type de solution est qu'il n'est pas toujours réalisable. Le cas de la description des familles à nom prédicatif en est un exemple. Si l'on effectuait une déclaration identique de l'artère, on ne pourrait alors plus rendre compte des syntagmes nominaux à tête nominale prédictive. Ce type de traitement proposé par [Candito, 1999a] est donc approprié aux verbes pleins, mais non généralisable²⁹.

²⁵La dimension 3 réalise des fonctions finales et non initiales

²⁶Ceci est la justification théorique de la double déclaration de fonctions chez Candito. En pratique, les besoins de définition des fonctions finales en dimension 1 sont absolument nécessaires. Une révision de fonction ne peut en effet être effectuée que sur des fonctions déclarées comme finales par Candito. Les fonctions initiales chez cette dernière sont purement décoratives. Ou alors, il faut considérer qu'une fonction se définit comme un couple (fonction initiale, fonction finale). Mais même dans ce cas, il faudrait revoir le statut de la dimension 3 qui ne s'intéresse qu'à des fonctions finales

²⁷Théoriquement, cela est du ressort de la dimension 2 : la classe (SANS-REDISTRIBUTION), correspondant à la morphologie active devrait identifier fonctions initiales et fonctions finales de manière transparente.

²⁸Autrement dit, l'artère représente le chemin qui va de la racine à la tête syntaxique

²⁹Tout au plus peut-on déclarer en dimension 1, la tête lexicale. Le pronom des verbes intrinséquement pronominaux peut également être représenté et placé par rapport au verbe, puisqu'il ne disposera pas, par définition, d'une fonction associée.

Nous venons donc d'aborder au cours de ce chapitre des problèmes touchant à la sous-catégorisation : nous avons défini sommairement les fonctions syntaxiques que nous retenons et avons jeté les bases de la Métagrammaire des noms prédicatifs.

Chapitre 5

Dimension 2 : les redistributions fonctionnelles

Sommaire

5.1	Généralités sur la dimension 2	156
5.1.1	La notion d'artère	156
5.1.2	Absence de syntagme verbal	157
5.2	Les différents changements de diathèses et redistributions utilisés	161
5.2.1	L'actif "de base"	161
5.2.2	Le passif	162
5.2.3	Les constructions causatives	166
5.2.4	L'impersonnel	170
5.2.5	Le passif impersonnel	171
5.2.6	Les constructions réfléchies	172
5.2.7	Les constructions moyennes	173
5.2.8	La représentation des groupes nominaux	174
5.3	Implémentation	177
5.3.1	La description des morphologies verbale et nominale	177
5.3.2	Passage des fonctions initiales aux fonctions finales	179
5.3.3	Les différentes redistributions	181
5.3.4	Remarques concernant la dimension 2	190
5.3.5	Vision graphique de la hiérarchie	190
5.3.6	Principales différences avec l'implémentation de Candito	193

Au cours du chapitre précédent, nous avons vu qu'il n'était pas forcément souhaitable de commencer à dessiner des arbres en dimension 1. La dimension 2 que nous allons maintenant décrire est, par contre tout à fait adaptée à cette tâche. Après avoir présenté la notion d'artère, nous exhiberons les différentes redistributions qui entrent en jeu pour la représentation des familles à noms prédicatifs.

5.1 Généralités sur la dimension 2

On a souvent tendance à croire que la dimension 2 ne sert qu'à présenter les différents cas de redistributions, c'est-à-dire le passage d'une fonction initiale à une fonction finale¹. Or, cela n'est pas tout à fait le cas : si les redistributions font partie intégrante de cette dimension, il faut cependant remarquer qu'elles sont parfois inutiles si elles n'entrent pas en jeu avec un autre élément que nous allons maintenant décrire².

5.1.1 La notion d'artère

Cet élément primordial, est ce que Candito appelle *artère*. Deux définitions sont proposées par Candito : l'une représentée par [Candito, 1999b] où l'artère est définie comme représentant "l'ancre verbale et ses projections", l'autre représentée par [Candito, 1999a] où l'artère est définie comme "la séquence de nœuds en relation de dominance immédiate comprenant" la tête syntaxique "et ses projections jusqu'au premier nœud non verbal".

Cette bipartition est en fait le reflet des choix d'implémentation développés pour les TAG avec les verbes pleins. En effet, avec les verbes pleins, l'ancre verbale est souvent la tête syntaxique de la phrase minimale représentée par l'arbre élémentaire. Cela n'est cependant pas toujours le cas : les arbres du passif comportent une position pour l'auxiliaire du passif (l'ancre au participe passé étant placée à droite de l'auxiliaire), et les arbres représentant le causatif comportent une position pour le verbe causatif (l'ancre étant le verbe à l'infinitif). Dans cette optique, ancre et tête syntaxique diffèrent donc.

Choisir l'une ou l'autre de ces approches revient à déterminer si l'on se place dans un cadre sémantique, ou dans un cadre syntaxique. Il reste même possible de choisir un cadre hybride, acceptant les deux approches, et c'est finalement la solution retenue par [Candito, 1999a] lors de l'implémentation. L'artère est alors à la fois représentée par l'ancre, tête sémantique, et le verbe acceptant les marques de flexion.

Cette position n'est cependant tenable que parce qu'un verbe plein est un prédicat sémantique. Dans les familles à nom prédicatif, c'est le nom qui est prédicat sémantique. Or ce dernier est considéré comme un argument du verbe support, et se voit réalisé de diverses façons. Vouloir le représenter en dimension 2 n'est donc pas d'un grand intérêt, d'autant qu'il faut également prendre en compte le verbe support qui, quant à lui, ne se voit pas réalisé en dimension 3. Dans le cadre des familles à nom prédicatif, l'artère ne peut donc être définie que comme faisant intervenir des têtes syntaxiques.

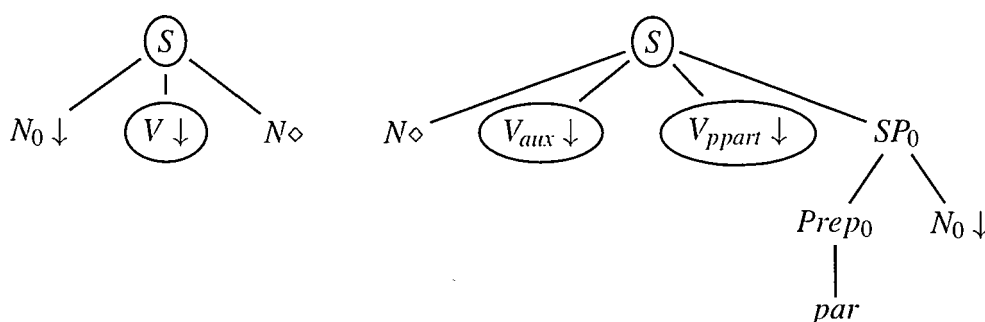
Dans le cas général, tête syntaxique et verbe support se confondent. Mais dans le cas particulier

¹ Nous restreignons le terme de redistribution à cette unique définition. Il n'est pas certain que Candito lui donne le même rôle. En particulier, elle considère un changement de diathèse comme une redistribution.

² Autrement dit, cela signifie qu'écrire une règle comme "Objet devient Sujet" est inutile si l'on n'indique rien de plus

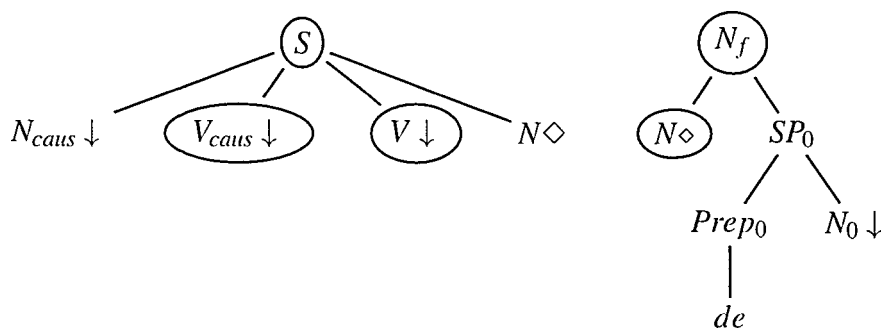
du passif et du causatif, on est amené à définir une tête syntaxique, différente du verbe support. Les seuls cas où tête syntaxique et tête sémantique pourront se confondre avec les familles à nom prédicatif seront effectués lorsque l'on représentera le groupe nominal complexe : le nom prédicatif sera alors une ancre, tout en étant également une tête syntaxique.

Plus généralement, l'artère représente donc tout ce qui n'est pas de l'ordre de la représentation des arguments. Une fois la diathèse fixée, elle représente en fait une sorte d'invariant et constitue une colonne de soutien. C'est un chemin de partage des traits.



Max prend une douche

Une douche est prise par Max



Léa fait prendre une douche aux enfants

La douche de Max

Disposant désormais de la notion d'artère, nous pouvons maintenant détailler les premiers choix qui sous-tendent la dimension 2. En particulier, nous allons rappeler pourquoi la grammaire TAG du français ne dispose pas de syntagme verbal.

5.1.2 Absence de syntagme verbal

Une des particularités de la grammaire FTAG est de ne pas avoir recours à des syntagmes verbaux pour représenter la phrase. [Abeillé, 1991] avait montré l'inadéquation de ce syntagme comme constituant intermédiaire en français. Nous présentons brièvement cet argumentaire, repris et complété par [Abeillé, 1996b, Abeillé, 2002].

5.1.2.1 Tests à orientation phonétique

La liaison obligatoire a souvent été mentionnée pour rendre compte de la bipartition des structures phrastiques. Ainsi [Bally, 1932] s'appuyait-il sur un argument de ce genre en constatant que dans la phrase *Cet enfant est intelligent*, la liaison était possible après *est*, mais pas après *enfant*. Cela lui permettait de conclure que la copule appartenait au prédicat et n'était pas un élément isolé.

Pourtant ce type de test a davantage tendance à regrouper clitiques et verbes, que verbes et compléments en français. Le cas des clitiques est clair :

- (154) Ils [z] aiment le chocolat.
Paul les [z] aime.

La liaison entre verbes et compléments est beaucoup plus hasardeuse, d'une part parce qu'elle n'est pas toujours obligatoire, et peut varier en fonction des locuteurs, d'autre part parce qu'elle n'est pas toujours possible :

- (155) Optionnalité : Il mangeait [t] une pomme.
Impossibilité : *On croirait [t] un blockaus du mur de l'Atlantique (ex. repris du corpus d'Orléans).

De plus, [Abeillé, 2002] note qu'une liaison optionnelle entre sujet et verbe n'est pas forcément exclue :

- (156) Toutes les filles [z] étaient là.
Plusieurs [z] étaient déjà arrivées.

Le test de la liaison semble donc peu enclin à éclairer ou à justifier l'existence d'un syntagme verbal.

Les pauses et la coordination sont aussi fréquemment citées pour justifier l'existence d'un syntagme verbal. Considérons les exemples suivants (# indique une pause ou une frontière de groupe) :

- (157) a. Jean adore les bananes et déteste les kiwis.
b. Jean adore # mais Marie déteste (#) les bananes que je vous ai apportées hier.
c. Jean donnera un livre # et Paul offrira un disque (#) à Marie.

(157a) est analysée comme une coordination de syntagmes verbaux sans ellipse, (157b et 157c) sont analysées comme des coordinations de phrases avec ellipse (cf. phénomènes de montée du noeud droit).

On observe de manière similaire, avec une pause certes moins forte, le même contraste avec les clitiques sujet ou complément :

- (158) Il admire (#) et il apprécie (#) ton travail.
Jean la voit et lui parle.

Il n'y a pas de pause pour la coordination de séquences compléments/verbe, mais il y en a une potentielle pour la coordination de séquences sujet/verbe. On aurait ainsi deux types de comportement pour les clitiques, selon qu'ils soient sujet ou complément, alors que les tests de liaison obligatoire tendent à les regrouper pareillement dans un sous-ensemble clitiques/verbe.

[Abeillé, 2002] note que l'"on peut citer d'autres exemples de coordinations de sous-séquences d'une structure "plate" qui ne présentent pas de pause intonative, sans pour autant former des sous-constituants" :

- (159) Coordination de 2 compléments avec mise en facteur du verbe : Jean offrira un disque à Marie et un livre à Paul.
Coordination d'un complément et d'un ajout avec mise en facteur du verbe : Jean va à la piscine le lundi et à la gym le mercredi.

En tout état de cause, il paraît clair que les coordinations ne sont pas de bons indicateurs de constituance. Les tests à orientation phonétique ne sont donc guère utiles pour justifier l'existence d'un syntagme verbal. Tout au plus, sont-ils utiles pour montrer qu'un sous-constituant clitique/verbe semble bien attesté.

5.1.2.2 Tests syntaxiques

Reprise verbale en *le faire* L'un des tests régulièrement cité pour mettre en évidence le syntagme verbal, est le test en *le faire*. *Le faire* s'emploie dans différents contextes : coordination, comparaison, subordination ; et permet l'ellipse du verbe et de ses compléments comme l'illustrent les exemples suivants :

- (160) Max lit les aventures d'Astérix et Léa le fait aussi
Max donnera un livre à sa grand-mère et Léa le fera aussi
Max étudie autant qu'on peut le faire

Le test est même valable pour distinguer compléments essentiels et modificateurs (cf. [Bonami, 1999], cité par [Abeillé, 2002]) : les compléments sont ainsi repris obligatoirement par *le faire* en contexte coordonné :

- (161) ?Max compte sur Léa et Luc le fait aussi
*Max compte sur Léa et Luc le fait sur Ida

Max étudie dans sa chambre et Léa le fait dans son bureau

Malheureusement, certains compléments obligatoires ne sont pas forcément repris par *le faire*. [Boons et al., 1976] fournissent ainsi les exemples suivants :

- (162) ?*Jean a opté pour des chaussettes rouges et Marie l'a fait aussi
Jean a opté pour des chaussettes rouges et Marie l'a fait pour des vertes.

D'autres exemples sont cités par [Abeillé, 2002] en observant les constructions comparatives :

- (163) Reprise du verbe et de ses compléments : Jean a menti à Marie comme Jacques l'avait fait avant lui
Reprise du verbe seul : Jean a menti à Marie comme Jacques l'avait fait à Paule
Reprise du verbe et de ses compléments : Jean a donné un gâteau à Marie comme Jacques l'avait fait avant lui
Reprise du verbe et d'un de ses compléments : Jean a donné un gâteau à Marie comme Jacques l'avait fait à Paule
Reprise du verbe et d'un de ses compléments : Jean a donné à Marie une religieuse comme Jacques l'avait fait d'un éclair³

Il faudrait alors une structure SV à deux étages pour rendre compte de ces phénomènes, ce qui ne paraît pas facilement justifiable... [Abeillé, 2002] donne encore d'autres exemples en considérant les verbes à deux compléments prépositionnels et arrive à la conclusion que "les reprises par *le faire* ne sont (...) pas un critère pour la mise en évidence d'un constituant SV".

L'antéposition verbale avec *le faire* semble par contre plus convaincante pour appuyer l'existence d'un constituant SV. En effet, l'antéposition du verbe sans ses compléments est peu naturelle :

- (164) ?Obéir ainsi, Robin ne le fera qu'au bon roi Richard
Obéir au bon roi Richard, Robin le fera volontiers
?* Donner de l'argent, Robin le fera volontiers aux pauvres
Donner de l'argent aux pauvres, Robin le fera volontiers
?Changer de carrosse, Robin ne le fera jamais avec le sherif
*Changer avec le sherif, Robin ne le fera jamais de carrosse
Changer de carrosse avec le sherif, Robin ne le fera jamais

³La séquence *le faire* ne peut être suivie d'un complément direct, puisque le clitique *le* est déjà présent, mais le complément direct non repris peut se réaliser comme un complément prépositionnel, généralement introduit par *de*.

[Abeillé, 2002] note cependant que "rien n'empêche d'analyser le constituant antéposé comme une phrase". Un argument fourni en faveur de cela est donné par la pronominalisation en *cela*, qui est généralement considéré comme une pro-forme phrastique. De plus, la reprise en *le faire*, tout comme en *faire cela* peut englober des circonstants :

(165) Voler de bonne heure, Robin fait cela souvent

Voler de bonne heure, Robin le fait souvent

Donner de l'argent aux pauvres (de bon matin), Robin fait cela souvent

Donner de l'argent aux pauvres (de bon matin), Robin le fait souvent

Autant dire que ces tests, même s'ils paraissent à première vue séduisants, ne sont guère éclairants quant à l'existence d'un constituant SV. Pour notre implémentation des familles à noms prédicatifs, nous suivons donc [Abeillé, 1991, Abeillé, 2002] et [Candito, 1999a] et n'avons donc pas recours au syntagme verbal. Nous verrons dans le chapitre suivant que nous ne considérons pas non plus de syntagme verbal comme complément.

5.2 Les différents changements de diathèses et redistributions utilisés

Nous commençons maintenant par rentrer dans le vif du sujet, et entamons la description des différents cas de morphologie verbale et nominale.

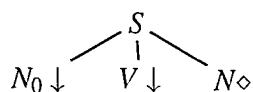
5.2.1 L'actif "de base"

L'actif "de base", autrement dit, l'actif sans redistribution fonctionnelle, est le grand absent de la thèse de Candito, qui ne le mentionne pas du tout en dimension 2. Pourtant, sa description est nécessaire : on se privera sans lui de la plupart des arbres de la voix active⁴.

Tous les arguments ont par défaut leur fonction finale comme fonction initiale.

L'arbre suivant illustre un arbre élémentaire à l'actif :

⁴Le fait que Candito décrive un schéma général en dimension 1 n'y change rien. Si aucune classe de dimension 2 n'est créée pour rendre compte des actifs, les arbres ne seront pas créés. Candito aurait dû présenter une classe qui n'aurait renfermé aucune redistribution. Par contre, la métagrammaire de démonstration des verbes de l'anglais, fournie dans sa thèse en annexe B, renferme bien une classe sans redistribution. Malheureusement, aucun commentaire n'est fourni avec cette hiérarchie. Il semble donc s'agir d'un oubli malheureux, d'autant que des articles postérieurs [Abeillé et Candito, 2000] corrigeront cela.



Max a commis une erreur

5.2.2 Le passif

A la différence des verbes pleins, on peut considérer pour les CVS deux types de passifs : l'un est syntaxique (également appelé analytique), l'autre est lexical (appelé synthétique).

5.2.2.1 Le passif syntaxique

Le passif est l'un des phénomènes qui a été le plus régulièrement étudié en français. On citera par exemple [Gatone, 1998a], [Milner, 1986] et les nombreuses grammaires qui lui consacrent un chapitre ou une large section.

Le passif se résume comme la conjonction de plusieurs phénomènes (le schéma de base étant l'actif) :

- "Démotion" du sujet : L'argument qui était sujet à l'actif devient soit complément introduit par *par* (Ia), soit complément introduit par *de* (Ib), ou est tout simplement supprimé (Ic).
- Promotion de l'objet : L'argument qui était objet à l'actif devient sujet (II).
- Morphologie passive : Le verbe prend la forme du participe passé raccordé au sujet par le verbe copule (auxiliaire du passif) *être* (III)

Ainsi, on trouvera (166), mais également (167) :

- (166) Phrase de base : Max commet un crime contre Luc
 Un crime contre Luc est commis par Max (Ia+II+III)
 Un crime est commis contre Luc par Max (Ia+II+III)
 Un crime est commis par Max contre Luc (Ia+II+III)
 Un crime contre Luc est commis (Ic+II+III)
 Un crime est commis contre Luc (Ic+II+III)

- (167) Phrase de base : Poser nu fait scandale
 Scandale est fait de poser nu (Ib+II+III)

On notera que le choix entre complément introduit par *par* et complément introduit *de* n'est en général pas libre. En français classique, la préposition *de* était très largement majoritaire, mais l'usage moderne lui substitue préférentiellement *par*.

- (168) Phrase de base : Tous ressentent cette peur
 Cette peur est ressentie de tous
 Cette peur est ressentie par tous

- (169) Phrase de base : Tous lui donnent ce conseil
 *Ce conseil lui est donné de tous
 Ce conseil lui est donné par tous

[Riegel et al., 1996] indiquent que *de* est toujours substituable par *par* ; l'inverse n'étant pas vrai. Cette constatation peut toutefois être remise en cause. On peut en effet trouver un passif en *de*, sans que le passif en *par* soit pour autant possible :

- (170) Phrase de base : Une sonate et un concerto composent ce spectacle
 Ce spectacle est composé d'une sonate et d'un concerto
 ?Ce spectacle est composé par une sonate et un concerto

Comme le note [Giry-Schneider, 1978], les passifs des CVS ont ceci de particulier que leur sujet peut apparaître sans déterminant. Cette possibilité n'apparaît pas avec les verbes pleins, mais peut en revanche également exister avec certaines expressions figées :

- (171) Phrase de base : Le commandant donne l'ordre aux troupes de partir
 Ordre de partir a été donné aux troupes par le commandant
 L'ordre de partir a été donné aux troupes par le commandant

Phrase de base : Max lui fait la promesse de partir tôt
 Promesse de partir tôt lui a été faite par Max
 La promesse de partir tôt lui a été faite par Max

- (172) Phrase de base : Max admire les grands linguistes
 *Grands linguistes sont admirés par Max
 Les grands linguistes sont admirés par Max

- (173) Phrase de base : Max a fait la lumière sur cet événement
 Lumière a été faite par Max sur cet événement
 La lumière a été faite par Max sur cet événement

Ce sont essentiellement les noms prédicatifs qui admettent un complément phrastique qui ont ces possibilités.

Enfin, comme le rappelle [Giry-Schneider, 1978], certaines CVS passives n'ont pas de formes actives correspondantes. Elles seraient alors un contre exemple à l'utilisation d'une redistribution :

(174) Actif : * Les policiers nous font mise en demeure de circuler

Passif : Mise en demeure nous a été faite par les policiers de circuler

Les noms prédicatifs qui n'ont pas d'actif correspondant sont cependant peu nombreux. Ce sont des noms comme *invitation, menace, rappel, interdiction, sommation, contrainte, obligation...*

Pour représenter les arbres passifs en TAG, il faut déterminer le rôle que joue l'auxiliaire⁵. Dans les premières versions de la grammaire FTAG, le verbe auxiliaire était adjoint au participe pour faire l'analyse du passif.

Ce n'est que plus récemment que le verbe auxiliaire a été substitué avec la description de [Candito, 1999a]. Peu d'arguments sont cependant fournis par cette dernière pour justifier un tel choix. [Abeillé, 1996a], mais aussi [Milner, 1986] apportent des arguments plus convaincants, arguments eux-mêmes repris par [Abeillé, 2002]. Il existe des différences syntaxiques entre *être* auxiliaire de temps, et *être* auxiliaire du passif : "la contribution temporelle d'*être* au passif est différente, et les surcomposés sont également différents :

(175) Max est arrivé hier soir (passé composé)

Max est soigné par Léa (présent)

(176) Quand Max a été arrivé / * a eu été arrivé...

Quand Max a été soigné / a eu été soigné ...

De plus, l'accord du participe passé est obligatoire au passif, mais pas aux temps composés :

(177) Les enveloppes seront remplies / *rempli

Les filles se sont écrit / *écrites

Enfin, [Abeillé, 2002] note que le réfléchi est impossible au passif, comme dans la construction à copule, mais qu'il reste possible avec un auxiliaire de temps :

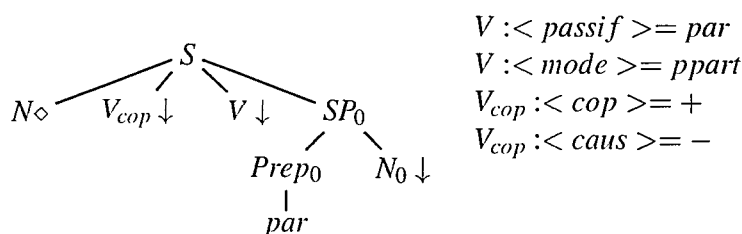
(178) Max s'est parlé à lui-même

Max (*se) sera opposé à lui-même

Max (*s') est fidèle à lui-même

Tous ces arguments sont donc en faveur d'un traitement du verbe auxiliaire *être* par substitution dans l'arbre du passif, par analogie avec la représentation des constructions attributives (voir par exemple [Boonen, 2001], [Barrier, 2002] et [Barrier, N., en préparation]). La substitution de l'auxiliaire permet en outre de prendre en compte son éventuelle variation lexicale.

⁵On a vu précédemment que l'auxiliaire du verbe était substitué, mais on n'a pas expliqué pourquoi.



Une erreur est commise par Max

Le trait *< passif >*, déjà utilisé sur le nœud V pour bloquer le passif de certains verbes comme *avoir*, est également utilisé pour indiquer que certaines CVS passives n'ont pas d'équivalent actif. Les noms prédicatifs qui acceptent la passivation renseignent un trait *< passif >* qui peut prendre la valeur *par* ou *de*, afin de spécifier quel type de complément doit être utilisé.

5.2.2.2 Le passif lexical

Les descriptions établies par [Gross, 1989, Gross, 1993a], nous permettent également de considérer les constructions converses comme des cas particuliers de redistribution. On observe en effet entre le passif verbal et les constructions converses des propriétés communes.

Ainsi, dans les deux constructions, l'inversion des arguments quand on passe de l'actif au passif est-elle de mise :

- (179) a. Max a avalisé ce projet
Ce projet a été avalisé par Max
- b. Max a donné son aval à ce projet
Ce projet a reçu l'aval de Max

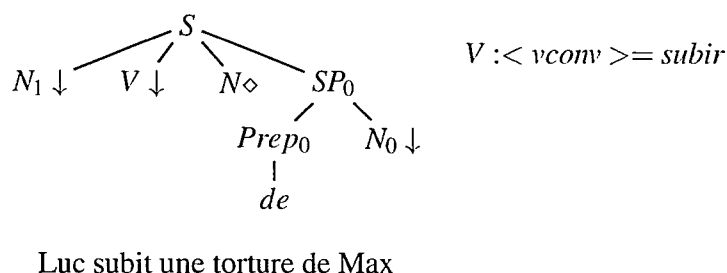
De même que l'effacement de l'agent est autorisé :

- (180) a. Max a dénoncé Luc
Luc a été dénoncé (par Max)
- b. Max a donné de bons conseils à Luc
Luc a reçu de bons conseils (de Max)

Bien entendu, des différences existent, la plus importante du point de vue linguistique, résidant ici dans le fait que l'opérateur est ici un substantif. L'autre différence majeure tient dans le fait que les redistributions entrant en jeu sont différentes : dans les constructions converses, le sujet correspond en fait au complément indirect de la construction standard :

- (181) a. Max fait une remarque désobligeante à Luc
 Luc a reçu une remarque désobligeante de Max
- b. L'aimant exerce une attraction sur le fer
 Le fer subit l'attraction de l'aimant

En TAG, les constructions converses n'ont donc pas la morphologie particulière du passif. A première vue, les arbres qui les décrivent peuvent ressembler à ceux de l'actif, sans redistribution. Néanmoins, le trait $\langle vconv \rangle$, porté par le support converse indique bien qu'il s'agit là d'un passif (lexical).



Le choix du conversif est effectué de manière analogue au support de base : on indique au niveau du lexique la forme du support conversif attendu grâce au trait $\langle vconv \rangle$. Les verbes conversifs renseignant également leur forme au niveau de leur trait, la sélection du support converse adéquat est facilement assurée.

5.2.3 Les constructions causatives

Tout comme dans le cas du passif, on peut considérer pour les CVS deux types de causatifs, l'un analytique, l'autre synthétique.

5.2.3.1 Les causatives analytiques

Les constructions causatives ont elles aussi été largement étudiées pour le français. On peut par exemple citer [Miller, 1992] ou bien encore [Abeillé, 1996b, Abeillé et al., 1998]. Dans de

telles constructions, le verbe causateur et son infinitif se comportent comme un prédicat complexe. Différents arguments peuvent être apportés pour considérer les phrases comportant des constructions causatives comme des phrases simples. D'une part, le placement du clitique se fait sur le verbe causateur et non sur l'infinitif (184), d'autre part, une seule négation est possible ; l'infinitif ne pouvant être nié (185)⁶.

(184) Max fait faire une étude du milieu à Luc

Max en fait faire une étude à Luc

*Max fait en faire une étude à Luc

(185) Max fait commettre une erreur à Luc

Max ne fait pas commettre d'erreur à Luc

*Max fait ne pas commettre d'erreur à Luc

De plus, le verbe causateur peut varier. Au traditionnel *faire*, on peut par exemple ajouter *laisser*, *voir*, *entendre*....

Dans tous les cas, un nouvel argument est introduit : il s'agit du causateur qui prend la fonction du sujet. Dans le cas général, le sujet initial recevra les fonctions :

- d'objet si l'infinitif est un verbe intransitif : *Max fait courir un bruit*⁷
- d'Aobjet ou de complément d'agent, si le verbe est transitif : *Max fait faire un voyage aux enfants*, ou bien *Max fait faire un voyage par les enfants*

Pour la représentation en TAG, le verbe causateur se combine avec l'infinitif pour former un prédicat complexe. En tant que tel, le verbe causateur est donc un verbe déficient dans la mesure où sa sous-catégorisation est déterminée par le verbe enchassé. Il ne subit donc pas un traitement en 3 dimensions.

Pour représenter le verbe causateur, on ajoute un noeud V_{caus} à substitution, qui devient alors la tête morpho-syntaxique de la structure. Ce noeud est placé avant l'infinitif [Abeillé, 2002]. On a recours à une structure plate puisque la séquence verbe infinitif/compléments ne se comporte

⁶Il existe également une construction causative concurrente, appelée causative à contrôle. A la différence des causatives à composition, ces causatives sont prises en compte par des familles particulières, et suivent une description multi-dimensionnelle. Elles ont un comportement différent vis-à-vis des propriétés énoncées :

- Les arguments cliticisés de l'infinitif se placent devant l'infinitif :

(182) Il les laisse faire la censure de cette page

Il les laisse en faire la censure

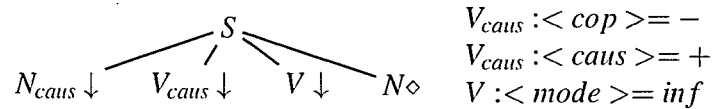
*Il les en laisse faire la censure

- La négation de l'infinitif est meilleure (sans être pour autant parfaite) :

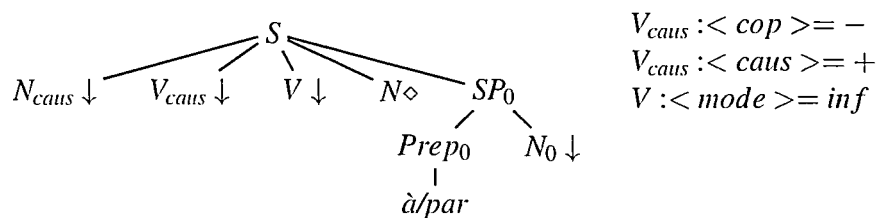
(183) La concentration les fait ne pas commettre d'erreur

⁷La pronominalisation par le *en* quantitatif, la négation en *pas ... de* prouvent bien qu'il s'agit d'un objet.

pas comme un syntagme et que le causataire et les compléments de l'infinitif peuvent librement permuter.



Max fait courir une rumeur



Max fait faire un voyage {aux / par} les enfants

5.2.3.2 Les causatifs synthétiques

A l'image des supports converses, les CVS nous permettent également d'effectuer un traitement original des opérateurs causatifs en les considérant comme des cas particuliers de redistribution. On sait en effet depuis [Gross, 1981] que les verbes qui entrent dans les constructions (186) ne sont pas de véritables supports, puisque la formation d'un groupe nominal complexe est impossible :

- (186) Les événements ont donné du poids à cet argument
 *Le poids des événements à cet argument (est indéniable)

Ces résultats ont fait de la peine à Max
 *La peine de ces résultats à Max (est immense)

Max a mis Léa au désespoir
 *Le désespoir de Max à Léa (était prévisible)

Le verbe est en fait un opérateur causatif de verbe support, et apporte un nouvel argument : le causateur. Le verbe support initial disparaît au profit de l'opérateur causatif, et le sujet ini-

tial du verbe support devient alors soit objet indirect introduit par *à* (187b et 188b⁸ ou objet direct (189b) de l'opérateur causatif. Dans ce dernier cas, la fonction du groupe prédicatif est également modifiée. L'attribut du sujet initial devient attribut de l'objet :

- (187) a. Cet argument a du poids
 b. Les événements ont donné du poids à cet argument
- (188) a. Max a de la peine
 b. Ces résultats font de la peine à Max
- (189) a. Léa est au désespoir
 b. Max a mis Léa au désespoir

On notera au passage que les CVS ne sont pas les seules concernées par ces opérateurs causatifs. Certaines expressions figées peuvent également s'y soumettre :

- (190) Léa est en cloque
 Max a mis Léa en cloque

Comme l'ont remarqué [Giry-Schneider, 1987], [Gross, 1989] et [Danlos, 1980], les opérateurs causatifs prennent en général la forme de *donner* ou *faire*, lorsque le verbe support initial est *avoir* ; ou *mettre*, lorsque le verbe support initial est *être prep.*

Les constructions à attribut de l'objet suivent un traitement très similaire⁹. C'est pourquoi, on les considère comme des variantes de ce phénomène :

⁸Notons que l'on ne fait pas rentrer des exemples comme *Le loup fait peur à Max* dans ce paradigme. En effet, il n'y a dans ce cas pas ajout d'un argument par rapport à la phrase initiale à support *avoir* (*Max a peur du loup*). Ces exemples se rapprocheraient davantage du passif conversif (l'agent n'est cependant pas introduit par *de*).

⁹Rappelons quelques propriétés de l'attribut de l'objet :

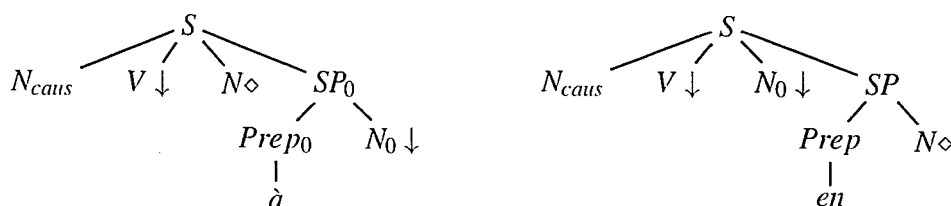
- L'attribut de l'objet, n'est pas un constituant interne du groupe nominal objet
 - En général, on peut faire une double analyse des phrases comme *Léa a trouvé les chiens en colère*, ambiguës entre les lectures épithétique et attributive du groupe prépositionnel. Le verbe *trouver* a dans ce cas deux sens distincts : dans le cas d'une lecture attributive, *trouver* a le sens de *juger* ; dans le cas d'une lecture épithétique, il a le sens de *découvrir*. Pour n'obtenir qu'une lecture attributive, il faut placer l'attribut de l'objet, avant l'objet.
- [Boonen, 2001] note que les verbes qui décident de la construction à complément attributif sont de trois types :
- Les verbes causatifs : Léa rend les chiens en colère
 - Les verbes à montée de l'objet : Léa croit les chiens en colère
 - Les verbes comme *vouloir*, *aimer*, *craindre*, *espérer*, *falloir*, *préférer* et *souhaiter* : Léa veut ses chiens en colère

(191) Léa est en colère

Opérateur causatif : Max a mis Léa en colère

Attribut de l'objet : Max a trouvé Léa en colère

La représentation en TAG suit le traitement déjà effectué pour les supports ordinaires et les converses. C'est donc le trait $V : < v_{caus} >$ qui détermine la forme de l'opérateur causatif.



Ces événements donnent du poids à cet argument

Max a mis Léa en colère

5.2.4 L'impersonnel

En français, le terme d'impersonnel est utilisé pour désigner un ensemble de constructions qui font apparaître notamment un sujet syntaxique non référentiel dont la forme est *il*¹⁰. La construction étant définie pour les verbes intransitifs (au sens large)¹¹, peu de familles à nom prédicatif sont concernées.

(192) Le bruit court que la gare sera détruite

Il court le bruit persistant que la gare sera détruite

Dans cette construction, un sujet *il* figé est introduit et le sujet nominal initial perd ses propriétés pour prendre les propriétés syntaxiques d'un objet direct (pronominalisation par *en* quantitatif (193b), partitif négatif en *pas ... de* (193c)), si ce n'est qu'il ne peut être cliticisé par *le* (193d) :

(193) a. Il court une rumeur persistante

b. Il en court une persistante

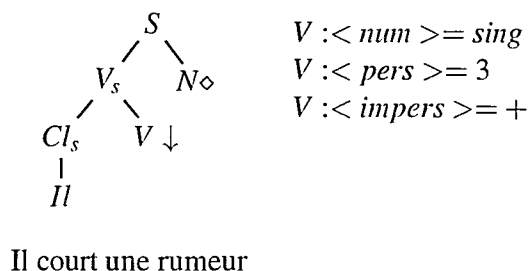
c. Il n'a pas couru de rumeur depuis longtemps

d. *Il la court

¹⁰Suivant [Candito, 1999a] et [Abeillé, 2002], nous ne parlons d'impersonnel, que lorsque des constructions personnelles peuvent alterner avec des constructions impersonnelles comme c'est le cas dans les exemples de cette sous-section. Les verbes et locutions impersonnels sont représentés par des familles spécifiques, et suivent donc un traitement multi-dimensionnel. Nous renvoyons à [Candito, 1999a] pour davantage de détails concernant ces familles particulières.

¹¹ou également certains passifs à sujet phrastique.

La représentation en TAG indique donc un clitique sujet figé, le sujet argumental étant postposé.



5.2.5 Le passif impersonnel

Le cas du passif impersonnel est intéressant dans la mesure où ce type de passif n'a pas toutes les propriétés du "passif ordinaire". En particulier, l'objet n'est pas promu sujet. Seul le sujet devient complément d'agent, la fonction de sujet étant, quant à elle, occupée par le "il impersonnel". La morphologie passive n'est pas modifiée : le verbe prend la forme du participe passé raccordé au sujet par *être*.

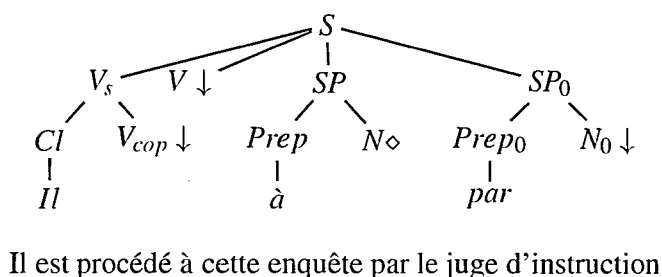
Nous considérons que le passif impersonnel n'est pas la composition des deux phénomènes pris séparément que sont l'impersonnel et le "passif ordinaire".

En effet, on peut noter qu'il n'existe pas forcément de variante "passif personnel" pour un "passif impersonnel" lorsque le nom prédicatif est en position d'objet indirect :

(194) Phrase de base : Les policiers procèdent à une enquête.

Passif : *Une enquête est procédée par les policiers

Passif impersonnel : Il est procédé à une enquête par les policiers



5.2.6 Les constructions réfléchies

Nous considérons également les constructions réfléchies comme des cas de redistribution. Dans ces constructions, le verbe porte le clitique SE, et son sens est réfléchi ou réciproque. Le SE entre alors en distribution complémentaire avec le complément prépositionnel du verbe support portant la fonction a-objet (195), ou bien avec l'objet direct (196) :

- (195) Max donne une claque à Luc
Max se donne une claque

- (196) Max met Luc en colère
Max se met en colère

La raison principale qui nous amène à considérer les réfléchis comme des redistributions vient du fait que l'on observe pour tous les réfléchis un changement d'auxiliaire de temps (être). Ce changement ne s'observe pas lorsque le clitique est la marque d'une simple alternance de réalisation.

- (197) a. Phrase de base : Max a donné une claque à Luc
b. Max s'est donné une claque
c. Max lui a donné une claque

- (198) a. Phrase de base : Max a mis Luc en colère
b. Max s'est mis en colère
c. Max l'a mis en colère

De plus, lorsque le SE peut entrer en distribution complémentaire avec un objet direct, le verbe cesse d'être transitif, comme le prouve son insertion dans les constructions causatives [Abeillé et al., 1998].

- (199) Max fait se mettre en colère Luc
*Max fait se mettre en colère à Luc

Dans ce dernier cas, SE apparaît donc comme un marqueur et ne porte plus de fonction syntaxique.

L'accord du participe passé est déclenché dans le cas du SE entrant en distribution complémentaire avec un objet, mais pas dans le cas du SE entrant en distribution complémentaire avec un a-objet.

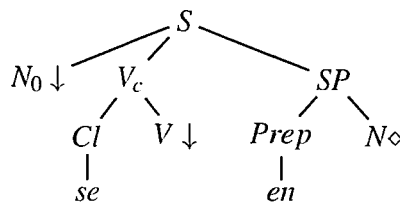
(200) Léa s'est mise dans une colère noire

*Léa s'est mis dans une colère noire

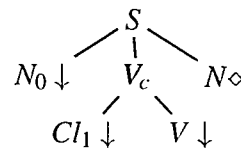
*Léa s'est donnée une claque

Léa s'est donné une claque

La représentation en TAG distingue donc les clitiques SE selon qu'ils entrent en distribution complémentaire avec un objet direct, ou un a_objet. Lorsque SE est un marqueur, il ne porte pas de fonction syntaxique et est réalisé comme une co-tête. En revanche, lorsqu'il entre en distribution complémentaire avec un a_objet, SE porte sa fonction d'a_objet, et est considéré comme un argument. Nous suivons en cela les conclusions dégagées par [Candito, 1999a].



Max s'est mis en colère



Max s'est donné une claque

Nous renvoyons à [Candito, 1999a] pour la description des traits permettant la gestion de l'accord du participe passé avec le sujet.

Notons que la réalisation du clitique SE comme marqueur, n'est possible que pour les familles en *être prep*, après application d'un opérateur causatif.

5.2.7 Les constructions moyennes

Un autre cas de redistribution est représenté par les constructions moyennes. Dans de telles constructions, le sujet initial est omis, alors que l'objet initial devient sujet. Le verbe porte un clitique dont la forme ne varie ni en nombre ni en personne :

(201) Phrase de base : Max commet un crime

Un crime se commet difficilement

Un crime, ça se commet difficilement

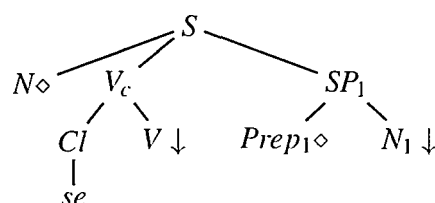
(202) Il se commet difficilement un crime

Les contraintes distributionnelles qui portent sur ce type de construction sont assez strictes, comme le rappellent [Fellbaum et Zribi-Hertz, 1989] et [Abeillé, 2002] :

- Impératif interdit
- Interprétation générique : on utilisera préférentiellement un présent (*De nos jours, un crime se commet beaucoup trop facilement*) ou un imparfait (*De mon temps, un crime se commettait très difficilement*)
- Adverbe de manière généralement obligatoire : même si les phrases ne sont pas forcément agrammaticales, elles paraissent tout de même incomplètes si aucun adverbe (ou négation ou intonation spécifique) n'est présent : ? *Un crime se commet, ? ? Une douche se prend.*

Les constructions moyennes semblent donc avoir des points communs avec les constructions passives et on leur confère généralement un "sens passif" : le sujet initial est démis, l'objet initial est promu sujet. Mais dans tous les cas, le sujet démis ne peut accéder à la fonction de complément d'agent : **Un crime, ça se commet difficilement par Max*. La comparaison s'arrête donc ici.

Les arbres des constructions moyennes sont définis pour toutes les familles à nom prédicatif objet direct. Le clitique réfléchi est directement renseigné dans l'arbre de la construction moyenne.



Un crime se commet contre les plus démunis (très facilement)

5.2.8 La représentation des groupes nominaux

On a vu au cours des chapitres précédents qu'une des caractéristiques des CVS était de présenter une double analyse des compléments prépositionnels ou phrastique, qui peuvent être considérés soit comme des compléments du verbe, soit comme des compléments du nom. Une autre caractéristique est la formation d'un groupe nominal complexe comprenant le sujet et les autres compléments.

5.2.8.1 La double analyse

Les phénomènes de double-analyse sont considérés comme des cas de redistribution. Le complément prépositionnel ou phrastique, initialement considéré comme un complément du verbe support devient alors un complément du nom prédicatif.

Pour garder les informations concernant les co-têtes introduites par le nom prédicatif soit par le lexique syntaxique, soit par le nom de la famille, on a choisi de séparer les fonctions de complément du nom, suivant les prépositions qui les introduisent. On crée donc autant de fonctions complément du nom que de fonctions indirectes initialement choisies. Ainsi, on crée une

fonction complément du nom introduit par *à*, correspondant à la fonction initiale de a-objet ; une fonction complément du nom introduit par *de*, correspondant à la fonction initiale de de-objet ; une fonction complément du nom introduit par une autre préposition, correspondant à la fonction initiale de prep-objet.

Cette redistribution soulève tout de même un problème. Les changements de fonctions sont certes effectués, mais aucune ossature n'est dessinée. A ce stade, remarquons que ce n'est pas l'ossature nominale qui sera pertinente ici. En effet, cette ossature sera apportée par la réalisation du nom prédicatif en dimension 3¹².

C'est en fait toute la morphologie verbale qui sera nécessaire, de manière à représenter les actifs, les passifs, les causatives, les impersonnels, les passifs impersonnels, les constructions réfléchies et les constructions moyennes. Il s'agira donc là d'un cas de composition de redistributions.

5.2.8.2 Le groupe nominal complexe

La représentation des groupes nominaux complexes est un cas tout différent. En effet, la sous-catégorisation concerne alors totalement le nom prédicatif. Le groupe prédicatif qui portait initialement une fonction par rapport au verbe support voit cette fonction supprimée, et le nom prédicatif présente alors les deux emplois de tête syntaxique et de tête sémantique.

On l'a vu, le groupe nominal complexe peut avoir plusieurs sources : la CVS active (203), la CVS passive (204) et la construction converse (205) :

(203) Max fait un transfert sur Luc

Le transfert de Max sur Luc (est évident)

Léonard a fait le portrait de Mona-Lisa

Le portrait de Mona-Lisa de Léonard (est formidable)

(204) Un agrandissement de cette page a été fait par Max

Un agrandissement de cette page par Max (serait nécessaire)

Le portrait de Mona-Lisa a été fait par Léonard

Le portrait de Mona-Lisa par Léonard (a inspiré de nombreux artistes)

(205) Le fer subit une forte attraction de cet aimant

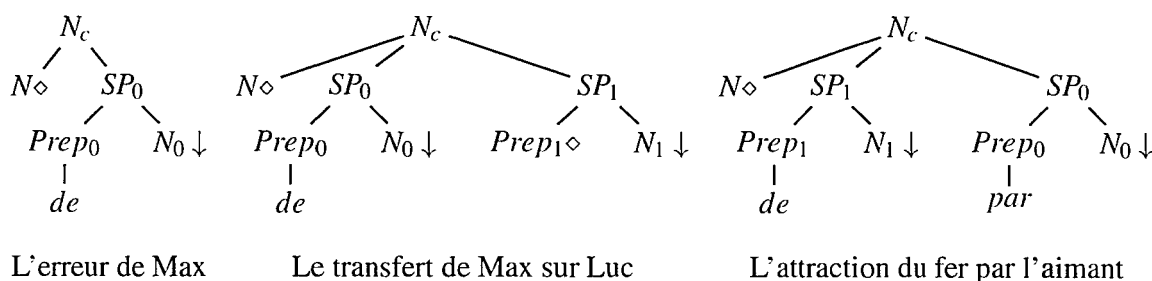
L'attraction du fer par cet aimant (est maximale)

¹²Le groupe prédicatif n'a pas perdu sa fonction. Seule la fonction des autres compléments du verbe support (hors sujet) a été modifiée.

Les redistributions pourront donc avoir lieu après les redistributions actives, passives et converses. Il s'agira là encore d'un cas de composition de redistribution.

Le sujet du verbe support (initial ou non) deviendra alors, suivant les propositions de [Godard, 1992] un premier argument. Il n'entrera pas en conflit avec l'autre complément prépositionnel indirect en *de*, issu celui-ci du complément prépositionnel du verbe support, qui quant à lui portera la fonction de second argument. Ce dernier ne pourra d'ailleurs devenir premier argument, que lorsque le premier argument d'origine sera supprimé¹³.

Ces cas de redistribution peuvent sembler complexes, mais ils permettent de représenter la relation systématique qui existe entre la formation du groupe nominal complexe et la phrase à verbe support correspondante.



Au terme de cette section, nous savons donc comment représenter les phénomènes liés aux morphologies verbale et nominale. Nous avons vu les redistributions à effectuer au niveau des fonctions définies initialement et avons introduit de nouvelles fonctions (complément d'agent, complément du nom (en *de*, à ou *prep*) ainsi que les fonctions de premier argument et second argument). Il reste à rentrer plus dans le détail de la hiérarchisation des données, et principalement à montrer le détail de l'implémentation qui éclaireront non seulement l'analyse mais également le fonctionnement du compilateur.

¹³Rappelons ici que les réalisations du premier argument et du second argument sont différentes. Le second argument ne peut pas se réaliser sous la forme d'un déterminant possessif, à la différence du premier argument.

5.3 Implémentation

5.3.1 La description des morphologies verbale et nominale

Nous allons donc maintenant nous attacher à la description plus précise des classes mises en jeu en dimension 2. Nous commencerons donc par la morphologie verbale qui nous permettra de rendre compte des phrases à verbe support. Puis, nous continuerons notre exposé avec la description de la morphologie nominale permettant de rendre compte de la formation des groupes nominaux complexes.

5.3.1.1 La morphologie verbale

On a vu en début de ce chapitre que l'artère devait différencier la tête syntaxique, du verbe support. Si ces 2 éléments se confondent généralement, cela n'est pas toujours le cas, notamment dans les cas de représentation des passifs et des constructions causatives.

Plusieurs choix s'offrent au descripteur de la métagrammaire afin de rendre compte de ces données. La première solution est de représenter dans la morphologie verbale ce qui est commun à toutes les classes. Dans ce cas, on représente le chemin qui va de la racine au verbe support. La deuxième solution consiste au contraire à représenter à la fois le verbe qui va recevoir les marques de flexion et le verbe support. C'est cette deuxième solution que nous allons explorer, car elle nous semble la moins contraignante pour la suite de la description et notamment pour les classes représentant le sujet de la dimension 3.

La classe (MORPHO-VERBALE) hérite de (DIMENSION2)
Elle définit 3 variables : - <i>S_d</i> représentant la racine de l'arbre, - <i>vsupp</i> représentant le verbe support, - <i>vmarq</i> représentant le verbe qui reçoit les marques flexionnelles. A ce stade, la longueur entre la racine et les verbes est laissée sous-spécifiée. En effet, un clitique peut-être réalisé, et alors amener un niveau intermédiaire.
Syntaxe : (MORPHO-VERBALE) <pre>(DIMENSION2) (init-constants '(((S_d) :cat ("S")) ((vsupp) :cat ("V") :qual subst) ((vmarq) :cat ("V")))) (init-description '((S_d nil) ((vmarq t)) ((vsupp t)))))</pre>
Récapitulation graphique : visuellement, le bout d'arbre suivant est créé. <pre> S_d / \ vmarq = V vsupp = V ↓ </pre>

Cette classe sera la mère de toutes les classes de la morphologie verbale. On va alors créer deux sous-classes : l'une, la classe (VSUPP_NON_MARQUE), qui représentera les cas où le verbe support est différent du verbe qui porte les marques de flexion (pour le passif et le causatif) et l'autre, la classe (VSUPP_MARQUE), où le verbe support se confondra avec le verbe qui porte les marques de flexion.

La classe (VSUPP_NON_MARQUE) hérite de la classe (MORPHO_VERBALE)
Dans cette classe, on sait désormais que c'est <i>vmarg</i> (la variable représentant la copule ou le verbe causatif), qui pourra être un hôte pour les cliti-ques. La liaison entre <i>Sd</i> et <i>vsupp</i> peut donc être ramenée à une hauteur de 1. On précise également que <i>vmarg</i> représente un verbe qui sera substitué.
Syntaxe : (VSUPP_NON_MARQUE (MORPHO_VERBALE) (init-constants '(((vmarg) : cat ("V") :qual subst))) (init-description '((S_d nil) ((vmarg t)) ((vsupp nil)))))
On obtient donc l'arbre suivant : $\begin{array}{c} S_d \\ \swarrow \quad \searrow \\ vmarg = V \downarrow \quad vsupp = V \downarrow \end{array}$

Le rôle de la classe (VSUPP_MARQUE) sera d'établir que les variables *vmarg* et *vsupp* ne font qu'une. Ceci sera réalisé ainsi :

La classe (VSUPP_MARQUE) hérite de la classe (MORPHO_VERBALE)
Elle lie les variables <i>vmarg</i> et <i>vsupp</i> déclarées dans la classe (MORPHO_VERBALE)
Syntaxe : (VSUPP_MARQUE (MORPHO_VERBALE) (init-constants '(((vmarg vsupp)))))

On dispose ainsi de tout le matériel nécessaire pour représenter les redistributions verbales.

5.3.1.2 La morphologie nominale

La morphologie nominale se révèle quant à elle beaucoup plus simple. Le nom prédicatif porte ici à la fois les rôles de tête sémantique et de tête syntaxique.

La classe (MORPHO-NOMINALE) hérite de (DIMENSION2)
Elle définit 2 variables : - <i>S_d</i> représentant la racine de l'arbre, - <i>anc</i> représentant le nom prédicatif
Syntaxe : (MORPHO-NOMINALE) (DIMENSION2) (init-constants '(((S _d) :cat("N") :ind "c") ((anc) :cat("N") :qual head))) (init-simple-desc '(S _d (anc))) (init-name "Groupe-Nominal-" :direction :left))
Récapitulation graphique : visuellement, le bout d'arbre suivant est créé. $\begin{array}{c} S_d = N_c \\ \\ anc = N_\diamond \end{array}$

Racine et nom prédicatif sont ici reliés directement puisqu'aucun élément ne peut venir s'intercaler entre eux-deux.

5.3.2 Passage des fonctions initiales aux fonctions finales

Muni des morphologies verbale et nominale, nous pouvons maintenant nous intéresser aux différentes règles lexicales qui composeront la dimension 2. Plutôt que d'établir chaque règle lexicale au sein de chaque changement de diathèse, nous allons définir de manière générale, les différentes règles lexicales. Lors des changements de diathèse, il ne restera alors plus qu'à appeler les classes ainsi décrites. Cela permettra un bon partage de l'information.

La classe (DEMOTION_DU_SUJET) regroupera les cas où le sujet descend dans la hiérarchie des fonctions. Au contraire, la classe (PROMOTION_VERS_SUJET) représentera les cas, où les fonctions remontent dans la hiérarchie des fonctions.

On va ainsi créer plusieurs classes, dont la classe (DEMOTION_DU_SUJET) regroupera les comportements suivants :

- La classe (SUJ>AGT_OBJ) où le sujet devient complément d'agent : Un crime est commis PAR MAX
- La classe (SUJ>DE_OBJ) où le sujet devient un complément introduit par *de* : Cette peur est ressentie DE TOUS LES VILLAGEOIS
- La classe (SUJ>A_OBJ) où le sujet devient complément introduit par *à* : Léa fait faire une erreur À MAX
- La classe (SUJ>vide) où le sujet disparaît : Une douche se prend en 5 minutes

De manière analogue, la classe (PROMOTION_VERS_SUJET) regroupera les comportements suivants :

- La classe (CAUS>SUJ) où un causateur sera introduit : LÉA fait faire une erreur à Max
- La classe (IMPERS>SUJ) où un sujet impersonnel sera introduit : IL court un drôle de bruit
- La classe (OBJ>SUJ) où l'objet devient sujet : UNE DOUCHE se prend en 5 minutes
- La classe (A_OBJ>SUJ) où le complément introduit par *à* devient sujet : LUC reçoit une dédicace de Max
- La classe (DE_OBJ>SUJ) où le complément introduit par *de* devient sujet : CE CORPS a reçu l'identification de Max
- La classe (PREP_OBJ>SUJ) où le complément introduit par une préposition devient sujet : LE FER subit l'attraction de l'aimant

D'autres règles lexicales peuvent également avoir lieu :

- L'attribut du sujet peut devenir attribut de l'objet : Max a mis Léa EN COLÈRE
- Les compléments prépositionnels du verbe peuvent devenir complément du nom : C'est la censure DE CETTE PAGE que Max a faite
- La fonction du groupe prédicatif peut même être supprimée, lors de la création du groupe nominal complexe

Sans détailler complètement cet ensemble de règles, nous allons maintenant observer la syntaxe des commandes à réaliser. Ici les classes (REG_LEX) et (DEMOTION_VERS_SUJET) vont être créées pour regrouper des comportements :

La classe (REG_LEX) hérite de la classe (DIMENSION2)
Elle regroupe l'ensemble des règles lexicales
Syntaxe : (REG_LEX (DIMENSION2))
La classe (DEMOTION_DU_SUJET) hérite de la classe (REG_LEX)
Elle regroupe l'ensemble des règles lexicales concernant la démotion du sujet.
Syntaxe : (DEMOTION_DU_SUJET (REG_LEX))
Les classes (SUJ>AGT_OBJ) et (SUJ>A_OBJ) héritent de la classe (DEMOTION_DU_SUJET)
Elles vont véritablement faire intervenir les changements de fonction
Syntaxe : (SUJ>AGT_OBJ (DEMOTION_DU_SUJET) (revise-fonction 'sujet 'par-objet)) (SUJ>A_OBJ (DEMOTION_DU_SUJET) (revise-fonction 'sujet 'a-objet))

La description de la classe (PASSIF-DE) est sensiblement la même que celle décrite précédemment (modulo la redistribution du sujet).

Pour les constructions converses, le traitement est sensiblement le même que pour le passif, à ceci près que c'est le conversif qui porte les marques de flexion. La classe générale du (PASSIF_CONVERSIF) hérite donc de la classe (VSUPP_MARQUE). Chaque type de passif conversif utilise alors les règles de redistribution pertinentes. On crée donc trois classes finales pour ce passif, suivant la fonction initiale portée par le complément indirect. Le tableau suivant illustre le cas pour la fonction initiale de prep_objet.

La classe (PASSIF_CONVERSIF) hérite de la classe (VSUPP_MARQUE) Elle sert à regrouper les différentes constructions converses
Syntaxe : (PASSIF_CONVERSIF (VSUPP_MARQUE))
La classe (CONVERSIF_PREP) hérite de la classe (PASSIF_CONVERSIF) Elle effectue les changements de fonction sujet → de_objet et prep_objet → sujet
Syntaxe : (CONVERSIF_PREP (PASSIF_CONVERSIF PREP_OBJ>SUJ SUJ>DE_OBJ) (init-name "Converse-" :direction :left))

5.3.3.3 Le causatif

Pour représenter les causatifs à prédicat complexe, on définit une classe (MORPHO_CAUSATIVE) qui va hériter de la classe (VSUPP_NON_MARQUE). Cette classe va ajouter les traits pertinents pour la morphologie causative.

La classe (MORPHO_CAUSATIVE) hérite de la classe (VSUPP_NON_MARQUE) Elle indique que la variable <i>vmarg</i> prendra l'indice "caus" (pour signaler un verbe causatif) On indique que le verbe initial est à l'infinitif et que le verbe causateur est de type causateur.
Syntaxe : (MORPHO_CAUSATIVE (VSUPP_NON_MARQUE) (init-constants '(((vmarg) :cat ("V") :ind "caus"))) (init-description '((S_d nil) ((vmarg t)) ((vsupp nil)))) (init-equations '(, (cr-eq vsupp () "neg" "-") , (cr-eq vsupp () "inv" "-") , (cr-eq vsupp () "mode" "inf") , (cr-eq vmarg () "caus" "+") , (cr-eq vmarg () "cop" "-"))) (init-name "Causatif-" :direction :left))
Au final, on obtient donc l'arbre suivant : $\begin{array}{c} S_d \\ \swarrow \quad \searrow \\ vmarg = V_{caus} \downarrow \quad vsupp \downarrow \end{array}$

On peut alors prendre en compte les différents causatifs et effectuer les redistributions adéquates :

- (CAUSATIF_PAR) : Léa fait faire l'étude de ces données par Max
- (CAUSATIF_A) : Léa fait prendre une douche à Max
- (CAUSATIF_I) : Max fait courir une rumeur

La classe (CAUSATIF-PAR) hérite de (MORPHO-CAUSATIVE)
La variable <i>causateur</i> est introduite : elle renseigne le sujet causateur. A ce stade, la variable est simplement renseignée : le sujet ne peut pas être directement placé dans l'arbre, puisqu'il pourra se réaliser de différentes manières.
Syntaxe : (CAUSATIF-PAR (MORPHO-CAUSATIVE SUJ>PAR_OBJ CAUS>SUJET) (init-constants '(((causateur) :cat ("N" "Cl") :function sujet :ind "caus")))))
L'arbre reste donc inchangé

Les classes (CAUSATIVE_A) et (CAUSATIVE_OBJ) suivent un traitement similaire. Seules les règles lexicales utilisées sont différentes.

Le traitement ainsi proposé des constructions causatives est cependant coûteux : plus de la moitié des arbres générés pour une famille sont des arbres causatifs. Si théoriquement cette remarque est sans importance, elle est en pratique cruciale car elle ralentit le traitement lors de l'analyse : il est plus facile de trouver un arbre adéquat parmi 100 candidats plutôt que parmi 200. D'autre part, il n'est pas non plus dit qu'un parseur puisse facilement accepter un nombre d'arbres important par famille. Il y a donc nécessairement un choix à faire entre efficacité et couverture linguistique. Un compromis semble plus difficile : pourquoi privilégier telle tournure plutôt qu'une autre ? Tous les arguments doivent-ils être obligatoirement réalisés ? Peut-on en omettre ? Lesquels omettre ? Ces questions dépassent largement le cadre de ce chapitre, mais elles montrent que tout linguiste, même s'il se veut théoricien doit à un moment ou à un autre faire des choix pratiques¹⁴.

Pour représenter les opérateurs causatifs, il faut noter que l'on utilise la classe (VSUPP_MARQUE), puisque c'est le verbe opérateur qui va porter les marques de flexion. Le sujet initial devient alors de_objet. La représentation de la classe ne pose pas de problème particulier.

La classe (OP_CAUS) hérite de la classe (VSUPP_MARQUE)
Elle effectue le changement de fonction sujet → a_objet
Syntaxe : (OP_CAUS (VSUPP_MARQUE SUJ>A_OBJ) (init-name "op-causatif" :direction :left))

¹⁴Il reste la solution idéaliste : celle qui consiste à dire qu'un jour ou l'autre on sera capable de traiter un très grand nombre d'informations très rapidement.

5.3.3.4 L'impersonnel

L'impersonnel est considéré comme une redistribution faite de la suite de deux redistributions partielles : le sujet initial devient objet, alors que la place laissée vacante par le sujet permet la réception du sujet "il" impersonnel.

Les verbes transitifs ne peuvent accéder à cette redistribution. En effet, la fonction d'objet ne peut être portée plusieurs fois.

La classe (IMPERSONNEL) hérite de la classe (VSUPP_MARQUE) et des redistributions adéquates. Elle dessine également le sous-arbre correspondant au sujet impersonnel.

La classe (IMPERSONNEL) hérite de (VSUPP_MARQUE) et des classes définissant les règles lexicales
Elle effectue les redistributions sujet → objet et introduit un nouveau sujet. Elle introduit la variable <i>il</i> qui indique le sujet impersonnel
Syntaxe : (IMPERSONNEL (VSUPP_MARQUE IMPERS>SUJ SUJ>OBJ) (init-constants '(((il) :cat ("il") :function sujet :qual no_qual))) (init-name "actif-impers" :direction :left))

A la différence de Candito, nous ne réalisons donc pas le sujet impersonnel en dimension 2 ; nous prévoyons juste sa forme de réalisation. De cette manière, nous distinguons clairement les classes de dimension 2 et les classes de dimension 3. La fonction sujet en dimension 3 se réalisera ainsi également sous la forme d'un sujet impersonnel¹⁵.

5.3.3.5 Le passif impersonnel

Nous l'avons dit, nous considérons que le passif impersonnel n'est pas une combinaison des redistributions passive et impersonnelle. En ce sens, nous n'utiliserons donc pas les classes finales que nous avons décrites pour le passif, mais plus simplement la morphologie passive (sans règle lexicale). Il reste donc à introduire les différentes règles lexicales : le sujet initial devient complément d'agent ; la place laissée vacante par le sujet est occupée par le "il impersonnel", qui prend donc la fonction sujet (non révisable).

¹⁵Chez Candito, la classe représentant le sujet impersonnel n'est pas renseignée en dimension 3. La classe impersonnel hérite d'une classe de dimension 2 représentant le sujet impersonnel.

La classe (PASSIF_IMPERSONNEL) hérite de la classe (MORPHO-PASSIVE)
On précise l'index de la variable <i>vmarq</i> La variable <i>vsupp</i> est rattachée directement à <i>S_d</i> .
Syntaxe : (PASSIF_IMPERSONNEL (MORPHO-PASSIVE IMPERS>SUJ SUJ>PAR_OBJ) (init-constants '(((il) :cat "il" :function sujet :qual no_qual) ((vmarq) :ind "cop")))) (init-simple-desc '(S _d (vsupp))) (init-name "Passif-impers" :direction :left)) Au final, on obtient donc l'arbre suivant : $\begin{array}{c} S_d \\ \swarrow \quad \searrow \\ vmarq = V_{cop} \quad vsupp \end{array}$

Là encore, le sujet impersonnel n'est pas réalisé.

5.3.3.6 Les constructions réfléchies

Pour représenter les constructions réfléchies, nous établissons une distinction entre le SE marqueur, et le SE entrant en distribution complémentaire avec l'a_objet. Dans tous les cas, l'auxiliaire du verbe est obligatoirement *être*.

La description de la classe (REFLECHI) est donc la suivante :

La classe (REFLECHI) hérite de la classe (VSUPP_MARQUE)
La valeur mode de <i>vmarq</i> est modifiée et prend désormais la valeur <i>etre</i>
Syntaxe : (REFLECHI (VSUPP_MARQUE) (init-equations-masquantes '(', (cr-eq vmarq () "aux" "etre"))))

Pour représenter la classe (SE-OBJ), le clitique *se* est directement représenté dans l'arbre, puisqu'il ne porte plus de fonction.

La classe (SE-OBJ) hérite de la classe (REFLECHI)	
Elle stipule que l'objet n'est plus réalisé et déclare les variables :	
- <i>infsv</i> qui introduit le niveau supplémentaire des clitiques compléments - <i>cl</i> qui indique la catégorie du réfléchi - <i>se</i> représentant le réfléchi	
Syntaxe : (SE-OBJ (OBJ>VIDE REFLECHI) (init-constants '(((se) :cat ("se") :qual no_qual) ((cl) :cat ("Cl") :qual subst) ((infsv) :cat ("V") :ind "c")))) (init-equations '(, (cr-eq se () "refl" "+"))) (init-description '((S_d nil) ((infsv t)))) (init-simple-desc '(infsv (cl (se)) (vmarg))))	
Le bout d'arbre suivant est donc créé : $\begin{array}{c} S_d \\ \\ infsv = V_c \\ \swarrow \quad \searrow \\ cl = Cl \quad vmarg \\ \\ se = se \end{array}$	

En revanche, la classe (SE-AOBJ) renseigne simplement que la variable portant la fonction d'a_objet est un clitique réfléchi :

La classe (SE-AOBJ) hérite de la classe (REFLECHI)	
Elle introduit la variable <i>se</i> représentant le réfléchi	
Syntaxe : (SE-AOBJ (REFLECHI) (init-constants '(((se) :cat ("se") :function a_objet :qual no_qual))))	

5.3.3.7 Les constructions moyennes

Les constructions moyennes vont être représentées par la classe (MORPHO_MOYENNE). A la différence des classes de la (MORPHO_PASSIVE), cette classe hérite de la classe (VSUPP_MARQUE), puisque les marques de temps et de mode sont portées par le verbe support. Dans cette classe, on va introduire le clitique se réfléchi et le représenter directement dans l'arbre, car il ne porte pas de fonction syntaxique¹⁶.

¹⁶La situation est donc différente des 3 autres redistributions (impersonnel, passif impersonnel et construction réfléchie) décrites ci-avant.

Tout comme pour la morphologie passive, on n'introduit pas à ce stade de règle lexicale.

La classe (MORPHO-MOYENNE) hérite de (VSUPP_MARQUE)	
Elle introduit la variable <i>se</i> qui représente le clitique réfléchi. Ce clitique est par définition toujours de type 3ème personne et réfléchi. L'auxiliaire du verbe support est désormais <i>etre</i> . On introduit également les variables <i>infsv</i> et <i>refl</i> .	
<pre> Syntaxe : (MORPHO-MOYENNE (VSUPP_MARQUE) (init-constants '(((refl) :cat ("Cl")) ((se) :cat ("se") :qual co_head) ((infsv) :cat ("V") :ind "c")))) (init-description '((S_d nil) ((infsv t)))) (init-simple-desc '(infsv (refl) (vmarg))) (init-simple-desc '(refl (se))) (init-equations '(, (cr-eq refl "b" "pers" "3")))) , (cr-eq refl () "refl" "+") , (cr-eq refl () "pers" infsv "t" "pers") , (cr-eq refl () "num" infsv "t" "num")))) (init-equations-masquantes '(, (cr-eq anc "b" "aux" "etre")))) (init-name "Moyenne-" :direction :left))</pre>	
L'arbre obtenu est donc le suivant :	<pre> S_d infsv = Vc / \ refl = Cl vmarg se = se </pre>

L'architecture générale étant maintenant posée, il ne nous reste plus qu'à définir la classe (MOYENNE) qui contiendra les redistributions adéquates :

La classe (MOYENNE) hérite de la classe (MORPHO-MOYENNE)
Elle effectue deux redistributions : le sujet initial disparaît et laisse sa place à l'objet initial
<pre> Syntaxe : (MOYENNE (MORPHO-MOYENNE OBJ>SUJ SUJ>vide))</pre>
L'arbre reste donc inchangé, seules les fonctions syntaxiques ont évolué

5.3.3.8 Les groupes nominaux

Ils nous reste maintenant à examiner les redistributions qui nous permettront de rendre compte de la double analyse et du groupe nominal complexe. On dispose d'ores et déjà de tout le matériel représentationnel nécessaire. Il "suffit simplement" de composer les classes existantes pour en créer de nouvelles.

La double analyse Pour représenter la double analyse, il suffit que le complément prépositionnel ou phrastique du verbe support devienne complément du nom. Mais il se pose un problème : le fait de faire une révision de fonction ne suffira pas à créer les arbres adéquats. Si les bons arbres sont créés lorsque les groupes prépositionnels sont compléments du verbe, c'est parce que l'organisation du compilateur n'est pas perturbée : une classe de dimension 1, croise avec une classe de dimension 2 (le passif, le causatif, le moyen...), puis avec autant de classes de dimension 3 que nécessaire. C'est la dimension 2 qui apporte l'ossature adéquate.

Or ce n'est pas le cas si l'on effectue juste une redistribution : on va croiser une classe de dimension 1, avec une classe de dimension 2 (celle qui contiendra uniquement la redistribution), puis avec autant de classe de dimension 3 que nécessaire. La classe de dimension 2 ne renfermera en son sein aucune armature ; elle aura juste créé un changement de sous-catégorisation. Il faudra donc établir différents héritages pour une seule et même règle.

En d'autres termes, il faudra, une fois la règle lexicale effectuée, la combiner avec les morphologies verbales déjà créées.

Ces descriptions sont peu coûteuses en effort de conception, mais leur intérêt se révèle limité vis-à-vis de la dimension 2. C'est pourquoi, dans l'implémentation pratique nous n'avons pas choisi de réaliser de redistribution concernant la double analyse. Les fonctions indirectes introduites sont en fait considérées comme des super-fonctions qui réalisent à la fois la description d'un complément de verbe et la description d'un complément du nom.

Les groupes nominaux complexes La description des groupes nominaux complexes échappe à cet écueil, dans la mesure où dans ce cas la morphologie nominale est forcément mise en jeu. Il faut cependant veiller à supprimer la fonction initiale portée par le groupe prédicatif. Le sujet initial du verbe support devient ensuite premier argument, alors que les autres compléments accèdent au rôle de complément du nom ou de second argument (pour le groupe indirect introduit par *de*).

Sans rentrer dans le détail de toutes les combinaisons à envisager, nous allons maintenant décrire quelques compositions :

La classe (GN_DE_DE) hérite de la morphologie nominale Le sujet initial devient premier argument Le de_objet initial devient second argument La fonction initiale portée par le groupe prédicatif (par exemple l'objet) est supprimée
Syntaxe : (GN_DE_DE <pre>(MORPHO_NOMINALE SUJ>1ARG DE_OBJ>2ARG) (revise-function 'objet 'empty) (init-name "gn-complexe" :direction :left))</pre>

La classe (GN_DE_A) hérite de la classe (MORPHO_NOMINALE) Le sujet initial devient premier argument Le de_objet initial devient second argument La fonction initiale portée par le groupe prédicatif (par exemple l'objet) est supprimée
Syntaxe : (GN_DE_A <pre>(MORPHO_NOMINALE SUJ>1ARG A_OBJ>CDN_A) (revise-function 'objet 'empty) (init-name "gn-complexe" :direction :left))</pre>

La classe (GN_DE_PREP) hérite de la classe (MORPHO_NOMINALE) Le sujet initial devient premier argument Le de_objet initial devient second argument La fonction initiale portée par le groupe prédicatif (par exemple l'a_objet) est supprimée
Syntaxe : (GN_DE_PREP <pre>(MORPHO_NOMINALE SUJ>1ARG PREP_OBJ>CDN_PREP) (revise-function 'a_objet 'empty) (init-name "gn-complexe" :direction :left))</pre>

Pour les groupes nominaux complexes pouvant également être formés à partir des CVS passives et des constructions converses, on ne va pas utiliser la morphologie verbale déjà mise en jeu, mais simplement les règles lexicales déjà décrites pour le passif et les conversifs. La description pour les groupes nominaux complexes formés à partir des règles lexicales formant le passif est par exemple la suivante :

La classe (GN_PASSIF) hérite de la classe (MORPHO_NOMINALE) Le sujet initial devenu par_objet devient alors cdn_par L'objet initial devenu par sujet est supprimé Le prep_objet initial devient cdn_prep
Syntaxe : (GN_PASSIF (MORPHO_NOMINALE OBJ>SUJ SUJ>PAR_OBJ) (revise-function 'par_objet 'cdn_par) (revise-function 'sujet 'empty) (revise-function 'de_objet 'larg) (init-name "gn-passif" :direction :left))

5.3.4 Remarques concernant la dimension 2

On l'a vu, la description de la dimension 2 est relativement complexe. Les principaux problèmes sont de plusieurs types :

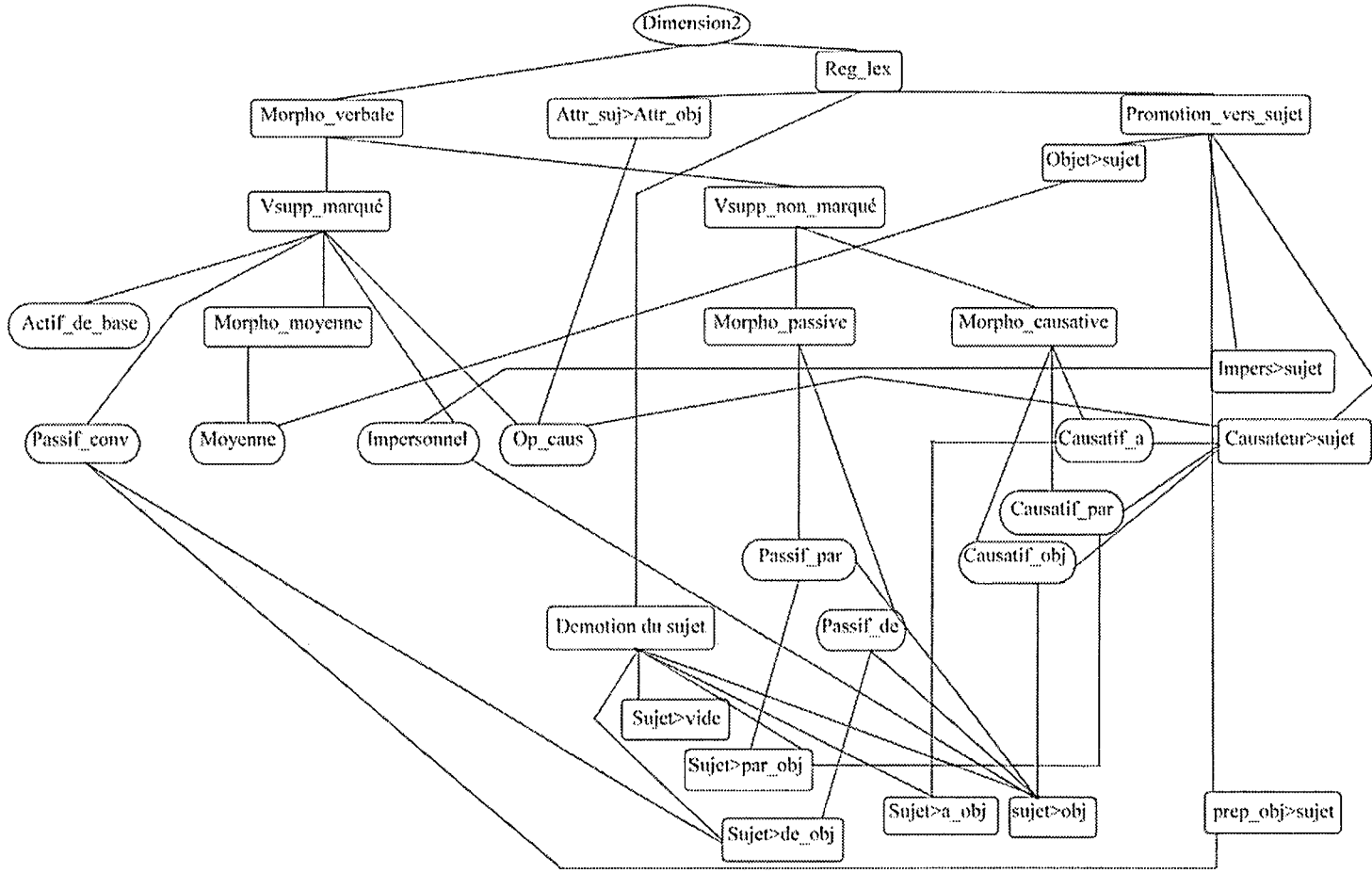
- Alors que la dimension 1 établit une sous-catégorisation catégorio-fonctionnelle, les règles lexicales ne concernent que les fonctions définies en dimension 1. Aucun travail sur les catégories établies en dimension 1 n'est effectué, ce qui amène le descripteur à réaliser des choix pas toujours motivés.
- Le travail effectué par la dimension 2 est double. Il y a d'une part des règles lexicales à décrire, et d'autre part, une ossature à dessiner. En général, les deux opérations sont liées, mais cela n'est pas toujours le cas.
- Enfin, les choix arborescents sont effectués en vue de la réalisation des arguments en dimension 3.

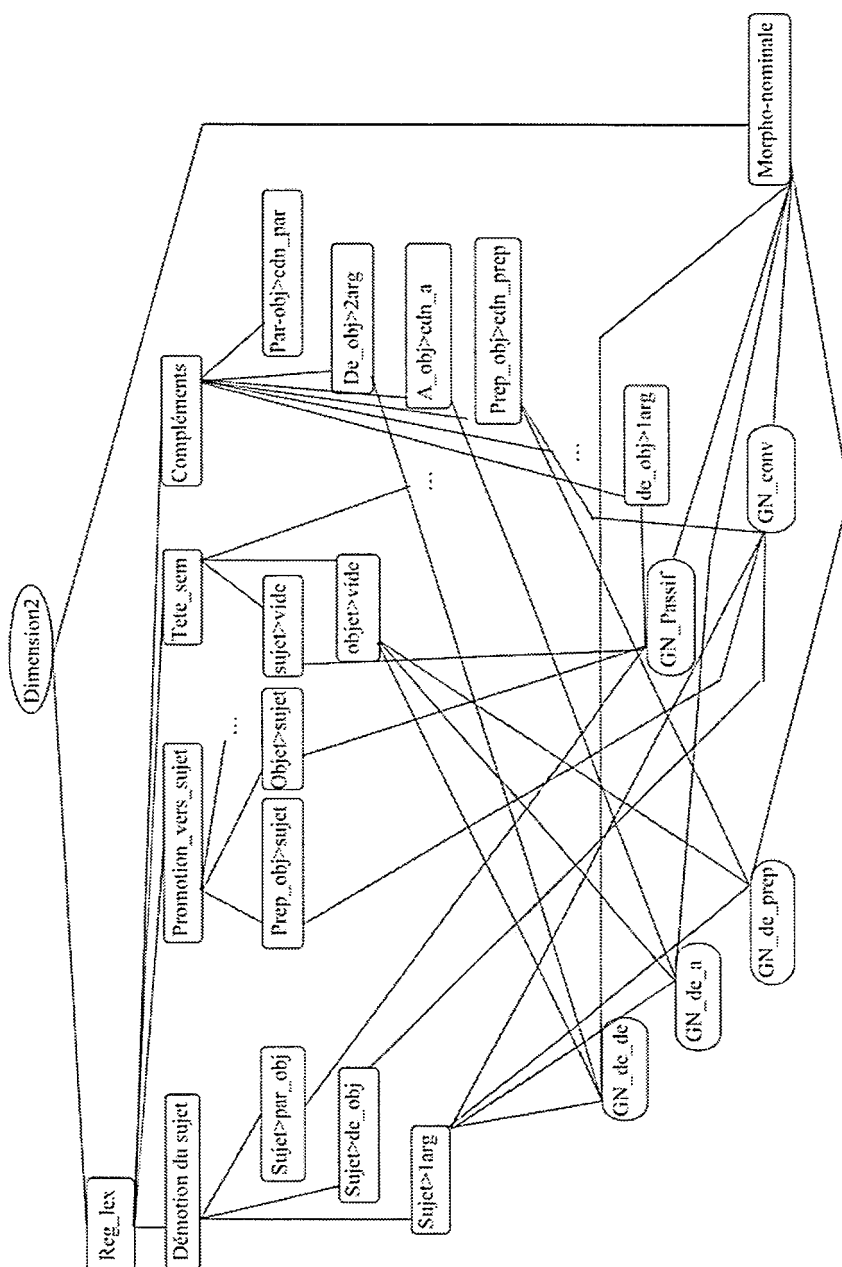
La proposition de [Vijay-Shanker et Schabes, 1992] établissait des règles lexicales extérieures à la hiérarchie et dessinait un nouvel arbre à partir d'un arbre déjà formé. Ici le cas est tout autre, puisqu'au stade de la dimension 2, aucun arbre complet n'est dessiné et que seule une ossature est vraiment décrite.

5.3.5 Vision graphique de la hiérarchie

Nous fournissons maintenant un bref récapitulatif des morphologies verbale et nominale de la dimension 2. Elles ont été simplifiées pour des raisons de lisibilité¹⁷. Le premier graphe d'héritage illustre la morphologie verbale, le second, la morphologie nominale.

¹⁷La morphologie verbale ne fait ainsi pas apparaître les classes concernant le réfléchi.

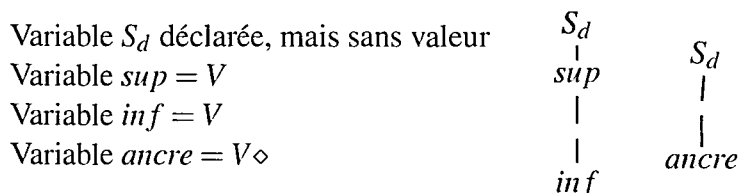




5.3.6 Principales différences avec l'implémentation de Candito

Par comparaison avec la hiérarchie de dimension 2 évoquée par [Candito, 1999a], [Candito, 1999b], notre hiérarchie semble très semblable (si l'on excepte bien entendu les phénomènes propres aux CVS) en terme de graphe d'héritage. Le contenu de nos classes, nous semble par contre beaucoup plus intuitif.

La classe (MORPHO-VERBALE) telle qu'elle est définie par Candito paraît de prime abord surprenante¹⁸ :



Aux vues de cette définition, plusieurs questions se posent alors :

- Pourquoi une classe censée représenter un phénomène le plus général possible, comporte-t-elle 3 éléments verbaux en son sein ?
- Ces éléments verbaux ont-ils tous la même valeur ?
- Que représentent exactement les constantes sup , inf et $ancre$?

Quelques éléments de réponse sont esquissés dans [Candito, 1999a], sans pour autant être totalement explicites. On peut en effet lire qu'il est nécessaire de "distinguer la notion d'ancre de la notion de tête syntaxique de l'arbre élémentaire". Cette remarque est tout à fait fondée, nous en avons déjà parlé. Ceci explique ainsi la bi-partition $ancre : sup/inf$, mais n'explique pas pourquoi un sup domine un inf ni quelle est véritablement la tête syntaxique.

Quelques pages plus loin, d'autres éléments de réponse sont apportés : la représentation de la MG du français "utilise plusieurs constantes verbales en relation de dominance large, qui nomment le même noeud dans le cas où le schème ne comporte pas de clitique". On peut donc affirmer que ces constantes sup et inf trouvent leur utilité lors de l'utilisation d'un clitique.

Ceci répond donc partiellement à nos deux dernières interrogations, mais pas à la première. De plus, il ne semble pas qu'une variable correspondant à une tête syntaxique existe réellement au sein de cette description. Est-il légitime de vouloir représenter dans une classe qui se veut la plus générale possible autant de phénomènes ? Nous répondons à cette question de manière négative. En effet, l'introduction d'un clitique n'est en soi nullement problématique. Si un clitique doit être utilisé, il amènera avec lui, la structure qui lui correspond¹⁹. De plus, la classe (MORPHO-VERBALE) lui prépare déjà une place d'accueil dans la mesure où elle spécifie bien qu'un verbe n'est pas forcément en relation de dominance immédiate avec un S .

¹⁸Cette définition varie d'ailleurs suivant les articles ou ouvrages dans lesquels elle est décrite. C'est pourquoi nous avons pris comme point de comparaison la hiérarchie effectivement implémentée par Candito.

¹⁹Rappelons que l'on n'a, en principe, pas de moyen de supprimer une constante qui a déjà été instanciée.

La représentation de la classe (MORPHO_VERBALE) de Candito semble donc fort différente de la nôtre. De manière générale, les descriptions de Candito semblent miser davantage sur la hiérarchisation de l'information que sur le contenu des classes. S'il est vrai que c'est certainement là ce qui représente l'invariant d'une Metagrammaire, la description des classes n'en demeure pas moins essentielle.

Dans sa conception ce chapitre tente de séparer clairement ce que l'on doit représenter de la manière dont il faut le représenter et établit donc plus clairement la différence entre une grammaire de bas niveau et une metagrammaire.

Nous avons donc étudié au cours de ce chapitre les éléments qui constitueront l'artère de la phrase ou du groupe nominal. Nous avons introduit de nouvelles fonctions syntaxiques, qui ne pouvaient être définies en dimension 1 (le complément d'agent, les compléments du nom et les 1er et 2nd argument) et avons jeté les bases de l'organisation hiérarchique de la Metagrammaire. Certains des choix que nous avons entrepris sont discutables d'un point de vue théorique, mais nous semblent préférables d'un point de vue pratique. Tout ce travail prépare la dimension 3, qui reste quant à elle, la plus fournie, sans être pour autant la plus difficile à décrire.

Chapitre 6

Dimensions 3 et 4 : la réalisation et l'ordonnancement des fonctions syntaxiques

Sommaire

6.1	Les réalisations canoniques	196
6.1.1	Le sujet	196
6.1.2	L'objet direct	199
6.1.3	Les a_objets, de_objets et prep_objets	199
6.1.4	Les compléments locatifs	201
6.1.5	L'attribut	202
6.2	Les réalisations non canoniques	205
6.2.1	Le sujet	205
6.2.2	Les clitiques	210
6.2.3	Les extractions	214
6.2.4	Accord du participe passé	221
6.3	Quelques remarques et exemples d'implémentation	223
6.3.1	La représentation des extractions	223
6.3.2	La représentation des clitiques	226
6.3.3	"Descente" du nom prédicatif	230
6.3.4	La représentation des syntagmes prépositionnels	231
6.3.5	Récapitulation générale : vision hiérarchisée	231
6.4	La dimension 4 : l'ordre des arguments réalisés	242

Au cours des chapitres précédents, nous avons annoncé les différentes fonctions syntaxiques que nous attendions. Certaines étaient présentes dès la dimension 1, d'autres n'apparaissaient qu'en dimension 2. Nous avons aussi défini les diverses armatures (les artères) des différents arbres

que nous voulions obtenir. Il nous reste maintenant une étape majeure : réaliser les différentes fonctions autour de ces armatures et éventuellement ordonner les différents arguments entre eux. Ceci est le rôle des dimensions 3 et 4. La dimension 3 est la plus riche en contenu, et la plus longue à décrire.

On s'intéressera tout d'abord aux différentes réalisations canoniques des diverses fonctions retenues, puis à leurs réalisations non-canoniques. Chaque section établit une présentation linguistique générale, puis éventuellement le traitement approprié en TAG. La dernière section revient plus spécifiquement sur quelques détails de l'implémentation.

6.1 Les réalisations canoniques

6.1.1 Le sujet

Considérations générales Le sujet peut se réaliser de manière canonique en français de plusieurs façons :

(206) Nominale : *Max* prend une douche

Phrastique : *Poser nu* fait scandale

Dans le cas des familles à nom prédicatif, le sujet nominal peut soit être un élément à substituer (207a), ou bien une tête sémantique (207b) :

(207) a. *Max* a un penchant pour la linguistique

b. *Le bruit* court que la gare sera détruite

Pour les sujets phrastiques, dans le cas d'une phrase tensée, le mode est obligatoirement au subjonctif :

(208) Subjonctif : Que *Max* vende ses parts de sa société fait scandale.

Indicatif : *Que *Max* vend ses parts de sa société fait scandale.

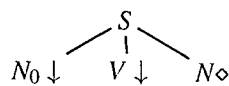
Dans le cas d'une infinitive, l'utilisation du complémenteur *de* marque un usage rare, et c'est pourquoi nous avons choisi de ne pas l'implémenter :

(209) a. *Poser nu* fait scandale

b. *De poser* nu fait scandale

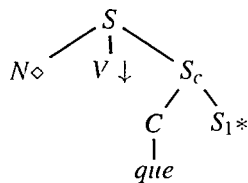
La représentation en TAG doit donc comporter plusieurs schémas d'arbres, suivant la nature de l'élément sujet. Le sujet nominal est un élément N à substituer ou une tête lexicale. Le sujet phrastique se réalise de 2 façons concurrentes. Il est également à substituer. Un trait concernant le $\langle mode \rangle$ est renseigné au niveau de S_0 et prend, selon les cas, la valeur d'infinitif (*inf*) ou de subjonctif (*subj*).

Les exemples ci-dessous illustrent ces divers points. Le premier arbre est issu de la famille $n0vN$, il représente un sujet nominal substitué. Le second arbre est issu de la famille $Nvs1$ et illustre l'ancre nominale sujet. Enfin, les 2 derniers arbres font tous partie de la famille $s0vN$ et représentent la réalisation du sujet phrastique.



$N_0 : \langle \text{fonc} \rangle = \text{subj}$

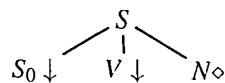
Arbre représentant un sujet nominal valable pour
Max prend une douche



Contraintes de non adjonction sur S_c et C

$N : \langle \text{fonc} \rangle = \text{subj}$

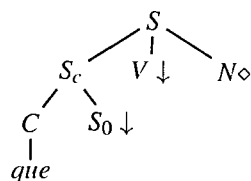
Arbre représentant une ancre nominale sujet valable pour
Le bruit court que la gare sera détruite



$S_0 : \langle \text{mode} \rangle = \text{inf}$

$S_0 : \langle \text{fonc} \rangle = \text{subj}$

Arbre représentant une infinitive sujet
valable pour *Poser nu fait scandale*



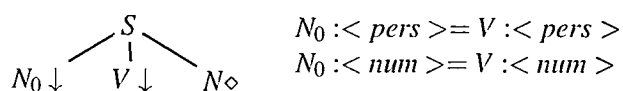
Contraintes de non adjonction sur S_c et C

$S_0 : \langle \text{mode} \rangle = \text{subj}$

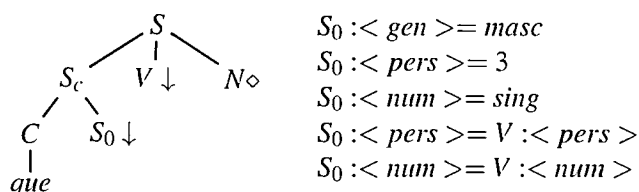
$S_c : \langle \text{fonc} \rangle = \text{subj}$

Arbre pour un sujet phrastique à complémenteur *que*
valable pour *Que Max pose nu fait scandale*

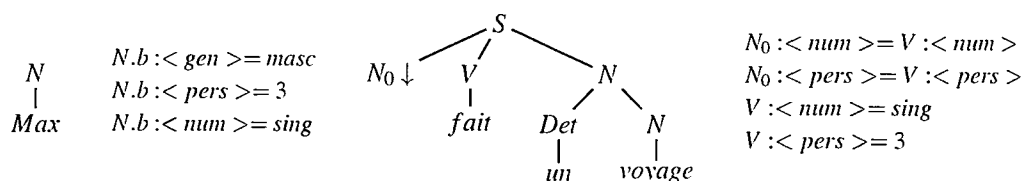
Gestion de l'accord L'accord Sujet/Verbe est réalisé par une équation d'accord sur les traits nombre *num* et personne *pers* du nœud sujet et du nœud verbe conjugué.



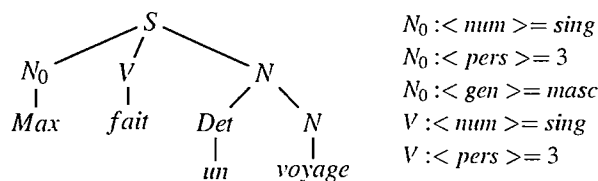
Pour les sujets phrastiques, leur nombre, leur genre et leur personne sont directement indiqués dans l'arbre :



L'exemple suivant illustre les phénomènes d'accord pour la phrase *Max fait un voyage* :



Résultat après substitution et unification :



Exemple d'accord Sujet/Verbe pour la phrase *Max fait un voyage*

6.1.2 L'objet direct

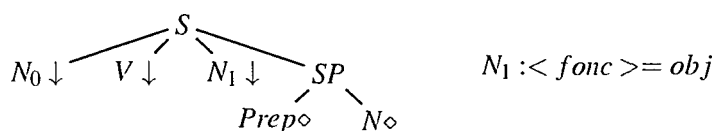
Tout comme le sujet nominal, l'objet direct nominal est représenté au sein de notre hiérarchie de diverses façons. Dans la majorité des cas, il s'agit d'un nom prédicatif renseigné comme une ancre (210), mais il peut également être représenté comme un élément à substituer (211).

- (210) CVS à l'actif : Max mène *une enquête* sur ce vol
Construction impersonnelle : Il court *le bruit* que la terre est ronde

- (211) Construction à attribut de l'objet : Max trouve *Léa* en colère



Arbre élémentaire permettant d'analyser
Max prend une douche



Arbre élémentaire permettant d'analyser
Max trouve Léa en colère

6.1.3 Les a_objets, de_objets et prep_objets

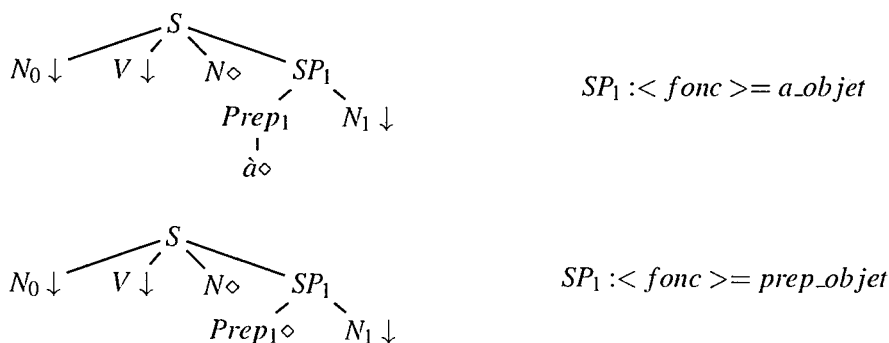
Nous avons vu que l'on différenciait les compléments prépositionnels suivant la nature de leur préposition, la pronominalisation du complément étant alors différente :

- (212) a. Max fait du chantage à *Luc* vs. Max *lui* fait du chantage
b. Max fait la censure *de cette page* vs. Max *en* fait la censure
c. Max commet un crime contre *Luc* vs. Max commet un crime *contre lui*

Dans ces exemples, les compléments de la préposition sont nominaux, mais ils peuvent également être phrastiques :

- (213) a. Max a une nette tendance à exagérer
b. Max a l'espoir de réussir
c. Max fait des efforts pour réussir

Le traitement de la préposition est fonction de la famille qui la contient. Ainsi, les familles en *aX* et *deX* renseignent à et *de* comme des co-têtes. Les familles en *pX* indiquent que la préposition est une co-tête.



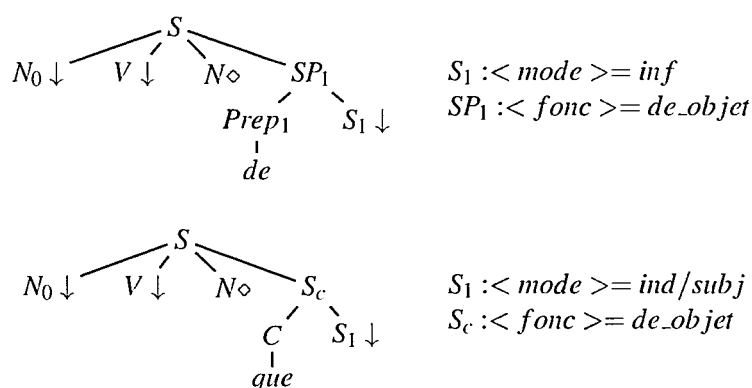
2 arbres à compléments prépositionnels permettant respectivement d'analyser les phrases
Max fait du chantage à Luc (1) et *Max commet un crime contre Luc* (2)

Les compléments d'agent en *par* suivent le même traitement que les compléments en *à* ou *de*. Le *par* est donc également une co-tête.

Les compléments de la préposition phrastiques sont quant à eux représentés de différentes façons suivant la manière dont ils se réalisent. Ils peuvent être des infinitives ou des complétives. Dans ce dernier cas, la tournure en prep ce que est parfois possible, mais est concurrencée par une construction directe, moins soutenue. C'est pourquoi nous ne représentons pas dans notre hiérarchie les constructions en prep ce que.

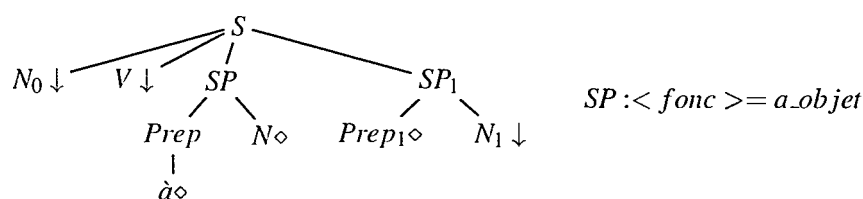
Les compléments phrastiques sont représentés par un nœud pied, ceci pour rendre compte des dépendances non bornées :

- (214) a. Max a l'espoir d'aller à Venise
 Venise où Max a l'espoir d'aller (est remplie de riches investisseurs)
- b. Max a l'espoir que tu ailles à Venise
 Venise où Max a l'espoir que tu ailles (est remplie de riches investisseurs)



Deux arbres permettant d'analyser les phrases
Max a l'espoir de réussir (1) et *Max a l'espoir que tous réussiront* (2)

Pour finir, notons que les compléments de la préposition peuvent également être des nœuds ancrés, lorsque le nom prédicatif fait partie d'un syntagme prépositionnel.



Un arbre élémentaire permettant d'analyser une phrase comme
Max procède à une enquête sur Luc

6.1.4 Les compléments locatifs

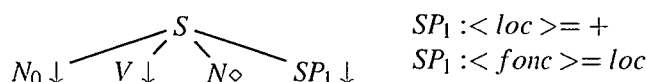
On distingue également les compléments prépositionnels à sens locatif dans la mesure où ils sont caractérisés par une pronominalisation, un relatif et un pronom interrogatif spécifique.

Pour le locatif de destination, on a ainsi les possibilités décrites par la série (215), alors que pour le locatif de provenance, les possibilités sont différentes (216) :

- (215) Max fait un voyage à Visan
 Max y fait un voyage
 Où Max fait-il un voyage ?
 Visan où Max a fait un voyage (a été la scène d'une émission populaire)

- (216) Max a fait une sortie remarquée de la salle de réunion
 Max en fait une sortie remarquée
 D'où Max a-t-il fait une sortie remarquée ?
 La salle de réunion d'où Max a fait une sortie remarquée (est maintenant fermée)

Le locatif, lorsqu'il n'est pas cliticisé, peut-être réalisé par un syntagme prépositionnel dont la préposition peut varier, ou bien par un adverbe locatif.

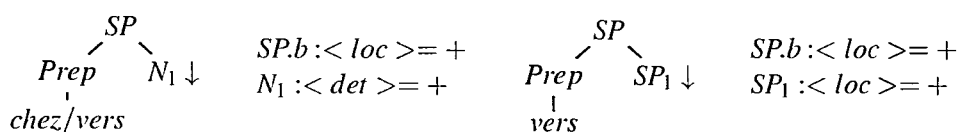


Arbre élémentaire permettant d'analyser
Max fait un voyage à Visan / ici

Le locatif de destination est réalisé par un syntagme prépositionnel introduit par *de*.

Suivant [Abeillé, 2002], nous considérons que les prépositions locatives, autonomes sémantiquement, sont, en tant qu'arguments, têtes d'arbres initiaux prépositionnels. La composition interne du SP varie selon la préposition : un SN (217a), un SP (217b) ou un adverbe locatif (217c), et ne dépend pas du verbe principal :

- (217) a. Max fait une sortie remarquée de la salle de réunion
 b. Max fait une sortie remarquée de chez Marie
 c. Max fait une sortie remarquée de là

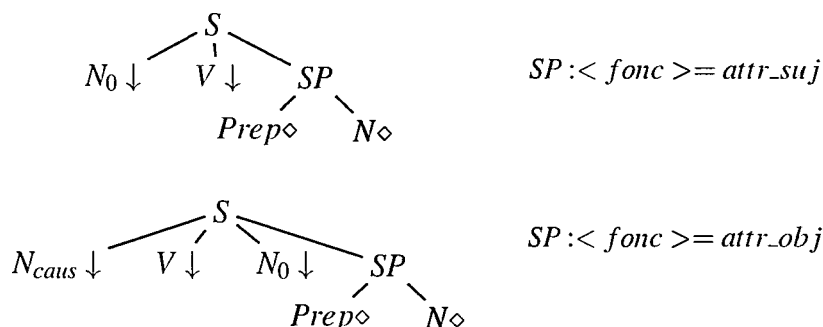


6.1.5 L'attribut

D'un point de vue strictement arborescent attribut du sujet et attribut de l'objet se représentent de la même façon : l'attribut reste un SP qu'il soit attribut du sujet ou attribut de l'objet.

- (218) a. Max est en colère
 b. Léa a trouvé Max en colère

Il n'y a pas d'accord à gérer pour les SP attributs avec le nom sujet (pour les attributs du sujet) ou objet direct (pour les attributs de l'objet).



Deux arbres élémentaires respectivement valables pour
Max est en colère (1) et *Léa met Max en colère* (2)

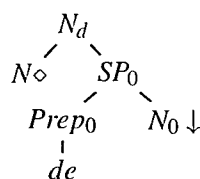
6.1.5.1 Les compléments du nom et les premier et second arguments

Ils nous permettront de rendre compte de la double analyse qui s'opère avec la plupart des CVS, mais également des compléments des groupes nominaux complexes. D'un point de vue strictement arborescent, ils ne diffèrent pas des compléments que l'on vient d'examiner dans cette section. Seuls les traits qu'ils portent, et en particulier leur fonction syntaxique diffèrent.

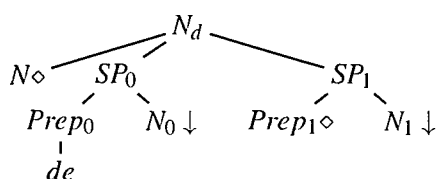
Pour éviter qu'une phrase non ambiguë comme *Max fait un transfert sur Luc* ne reçoive une double analyse (l'une considérant le complément prépositionnel comme un complément du verbe support, l'autre le considérant comme un complément du nom), on interdit la réalisation de complément du nom, lorsque la fonction objet du nom prédicatif se réalise de manière canonique. De manière analogue, lorsque les fonctions du nom prédicatif sont celles d'a_objet ou d'attribut du sujet, on interdit également la réalisation de complément du nom¹. De cette manière, on s'assure bien que seul le complément du verbe est considéré.

Dans le cadre du groupe nominal complexe, les compléments du nom sont appelés premier ou second argument lorsqu'ils correspondent à un syntagme prépositionnel introduit par *de*.

¹ Sauf bien entendu dans le cadre des familles n0vAN(dén1) et n0vEnN(dén1).

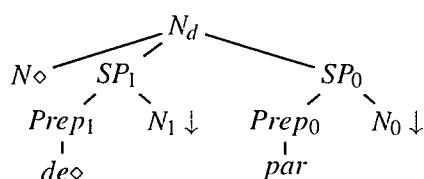


$SP_0 : < \text{fonc} > = 1arg$



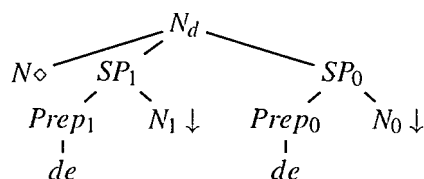
$SP_0 : < \text{fonc} > = 1arg$

$SP_1 : < \text{fonc} > = \text{cdn_prep}$



$SP_0 : < \text{fonc} > = \text{cdn_par}$

$SP_1 : < \text{fonc} > = 1arg$



$SP_0 : < \text{fonc} > = 1arg$

$SP_1 : < \text{fonc} > = 2arg$

Quatre arbres élémentaires permettant d'analyser respectivement

Le rêve de Max (1), *Le crime de Max contre Luc* (2)

L'achat de cet ouvrage par Max (3) et *Le portrait de Mona Lisa de Léonard* (4)

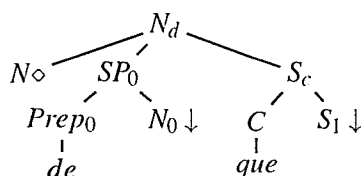
Lorsqu'ils sont compléments du nom, les compléments phrastiques apparaissent comme des îlots pour l'extraction :

(219) Max accepte l'hypothèse (de Luc) que Léa habite ici

*Là où Max accepte l'hypothèse (de Luc) que Léa habite

C'est pourquoi, on les considère comme des nœuds à substitution, et nom comme des nœuds pieds².

²On rappelle que lorsque les compléments phrastiques sont compléments du verbe support, ils autorisent l'extraction à longue distance :



Un arbre pour représenter *L'hypothèse de Luc que Léa habite ici*

Voilà qui clot les représentations canoniques des différentes fonctions syntaxiques que nous avons retenues. Comme le lecteur aura pu le remarquer, la représentation des groupes prépositionnels tient au sein de cette hiérarchie une place relativement importante. Nous allons maintenant nous intéresser aux représentations non canoniques de ces mêmes fonctions.

6.2 Les réalisations non canoniques

Nous allons tout d'abord nous tourner vers le sujet, qui peut être inversé ou omis, puis vers des réalisations qui concernent la majorité des fonctions syntaxiques : les réalisations comme clitiques et les phénomènes d'extraction (relatives, interrogatives et clivées).

6.2.1 Le sujet

6.2.1.1 Le sujet inversé

L'inversion du sujet est réputée être un phénomène complexe en français. Elle connaît trois variantes : l'inversion nominale (221a), l'inversion clitique (221b) et l'inversion complexe (221c).

- (221) a. Quand se répandra donc cette nouvelle ?
 b. Prendra-t-il une douche ?
 c. Max prendra-t-il une douche ?

-
- (220) Max fait l'hypothèse que Léa habite ici.
 Là où Max fait l'hypothèse que Léa habite

L'inversion nominale et l'inversion clitique ne présentent pas les mêmes contraintes [Kayne, 1972] : à la différence de l'inversion clitique, l'inversion nominale est impossible avec les verbes à objet direct postverbal réalisé (222), mais est possible dans une phrase enchâssée (223)

- (222) Quand viendra Max ?
 *Quand mangera Max son kiwi ?
 Quand mangeras-tu ton kiwi ?

- (223) a. Quand viendras-tu ?
 b. *Je me demande quand viendras-tu.
 c. Je me demande quand viendra Max.

L'inversion clitique et l'inversion complexe présentent les mêmes propriétés, sauf dans le cas de l'interrogative en *que* qui autorise l'inversion clitique (224a), mais pas l'inversion complexe (224b) :

- (224) a. Que fait-il ?
 b. *Que Max fait-il ?

Les inversions clitique et complexe n'ont lieu qu'en cas de contexte interrogatif en phrase principale, et sont réalisées immédiatement à droite du verbe qui porte les marques de flexion :

- (225) a. Quel voyage fait-il ?
 b. Quel voyage Max fait-il ?
 c. *Le voyage que Max fait-il ?
 d. *Le voyage que fait-il ?
 e. Quel voyage a-t-il fait ?
 f. *Quel voyage a fait-il ?

L'inversion nominale, quant à elle, peut avoir lieu dans un contexte d'extraction quelconque, le sujet se plaçant alors derrière le verbe (support, ou tête sémantique)

- (226) a. Le voyage que fait Max (est enrichissant)
 b. Quel voyage fait Max ?
 c. C'est ce voyage qu'a fait Max
 d. Le voyage qu'a fait Max (lui a beaucoup plu)
 e. *Le voyage qu'a Max fait (lui a beaucoup plu)

On l'a dit, avec les verbes pleins, l'inversion nominale est impossible si un objet canonique est réalisé :

- (227) *Léa à qui donne une fleur Max (est resplendissante)
 *Léa à qui donne Max une fleur (est resplendissante)

Mais avec les verbes supports, lorsque le nom prédicatif est en position d'objet, l'inversion du sujet semble possible (228), mais pas systématique (229) [Marandin, 2001] :

- (228) Madrid où fait escale le professeur (est splendide en cette saison)
 Le loup dont avait peur Max (a été capturé)
 Le matériel dont a besoin le professeur (vient d'arriver)

- (229) *Luc contre qui a commis un crime Max (est sain et sauf)
 *Luc à qui a fait du chantage Max (a porté plainte)

En fait, dès que le nom prédicatif n'est pas déterminé, l'inversion semble légitime, mais si une détermination est effectuée, l'acceptabilité des exemples n'est déjà plus si nette :

- (230) a. ?Madrid où fait une escale le professeur (est splendide en cette saison)
 b. ? ?Le loup dont avait une peur bleue Max (a été capturé)
 c. ? ?Le matériel dont avait un besoin impérieux le professeur (vient d'arriver)

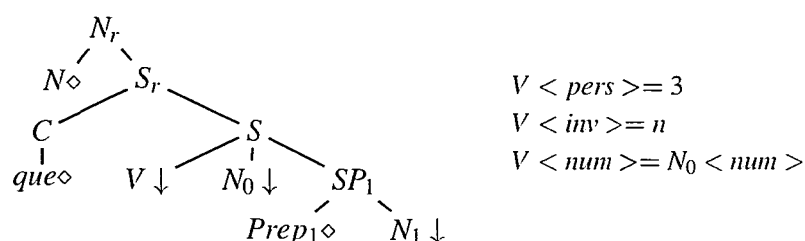
Comme le mentionnent [Bonami et Godard, 2001], le sujet nominal inversé possède différentes propriétés qui le différencient à la fois du sujet non inversé et de l'objet direct :

1. Contrairement au complément direct, et tout comme le sujet non inversé, le GN inversé peut lier le clitique *se* réfléchi :
 - *Paul est vraiment dans la lune : hier soir, il a voulu *se_i* présenter Marie_i.
 - Paul_i *s_i*'est présenté à Marie.
 - La personne à qui *s_i*'est présenté Paul_i l'a très bien accueilli.
2. Contrairement au complément direct, et tout comme le sujet inversé, le GN inversé ne peut pas être un indéfini ou un quantifieur lié à un *en* quantitatif :
 - Paul *en* a reçu beaucoup.
 - *Beaucoup *en* sont arrivés.
 - *Les élèves trouvent ce livre difficile ; mais nous pourrions étudier celui-ci qu'*en* ont déjà lu plusieurs.

3. Contrairement au complément direct, le GN inversé ne peut pas être un *de N* lié à un quantifieur flottant, entre l'auxiliaire et le participe :
 - *Le président a beaucoup lu de journaux, ces derniers jours.*
 - *L'événement qu'avaient relaté en détail beaucoup de journaux (est extraordinaire) vs. L'événement qu'avaient beaucoup en détail relaté de journaux (est extraordinaire).*
4. Contrairement au sujet non inversé et tout comme le complément direct, le GN inversé autorise l'extraction en *combien* :
 - *Combien viendront de ministres à cette cérémonie ?*
 - **Combien pensez-vous que de ministres viendront à cette cérémonie ?*
 - *Combien pensez-vous que viendront de ministres à cette cérémonie ?*
5. Contrairement au sujet non inversé, et tout comme le complément direct, le GN inversé peut prendre la forme *de N* en contexte négatif :
 - *Jean n'a jamais lu de livres de cet auteur.*
 - **D'enfants ne viennent plus jamais dans cette maison.*
 - *Une maison où ne viennent plus jamais d'enfants (est une maison triste).*
6. Parmi les quantifieurs "nus" qui sont acceptables comme sujets préverbaux (*un, deux, ..., tous, beaucoup, plusieurs, rien, tout*), seuls ceux qui sont aussi acceptables comme objets directs (*rien, tout*) peuvent être des GN inversés :
 - *Paul ne lit rien.*
 - *Une hypothèse que rien ne corrobore.*
 - *Une hypothèse que ne corrobore rien.*
 - **Paul lira tous.*
 - *Une hypothèse à laquelle tous s'intéressent.*
 - **Une hypothèse à laquelle s'intéressent tous.*
 - **Une hypothèse à laquelle se sont intéressés tous.*
7. Le GN inversé s'accorde en nombre, mais non en personne avec le verbe :
 - *L'appartement que Claire et Marie repeignent (est formidable). vs. *L'appartement que Claire et Marie peint (est formidable).*
 - *L'appartement que repeignent Claire et Marie (est formidable). vs. *L'appartement que peint Claire et Marie (est formidable).*
 - *L'appartement que Claire et toi habitez (est formidable).*
 - *L'appartement qu'habitaient Claire et toi (est formidable) vs. *L'appartement qu'habitez Claire et toi (est formidable).*

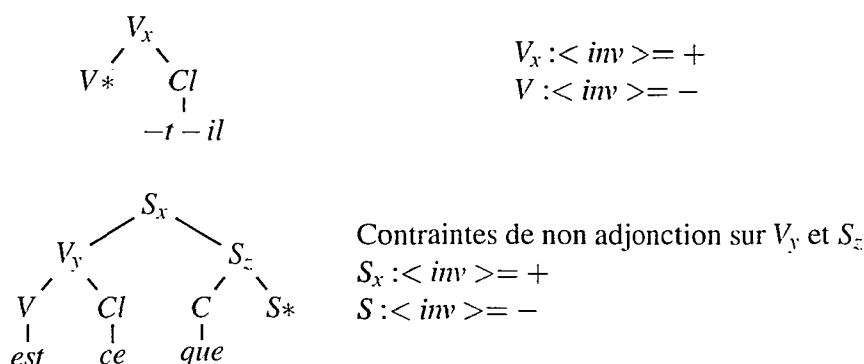
Le test du "en quantitatif" permet définitivement d'accorder le statut de sujet à ce type de GN inversé en contexte d'extraction.

La représentation en TAG que nous effectuons est réalisée suivant [Abeillé, 2002]. Aussi, le sujet nominal inversé se place-t-il à droite du verbe support. Les inversions clitique et complexe sont implémentées par adjonction d'un arbre auxiliaire représentant la particule interrogative à droite du verbe portant les marques flexionnelles. L'adjonction de cette particule ou de l'expression *est-ce-que* est forcée en contexte interrogatif par une chaîne de traits coindexés (les traits $\langle inv \rangle$ et $\langle princ \rangle$ [Abeillé, 2002]).



Un arbre élémentaire pour *Le crime que commet Max contre Luc*

Les arbres suivants sont utilisés pour représenter les inversions clitique et complexe ainsi que la construction en *est-ce que* [Crabbé, 2005] :



6.2.1.2 Les sujets sans forme réalisée

Bien que nous ne traitions pas de la question de l'élipse, il y a cependant deux cas de non réalisation du sujet à considérer : le sujet non réalisé des infinitives (231) et des impératives (232).

- (231) Max veut faire un voyage
Max persuade Luc de faire un voyage

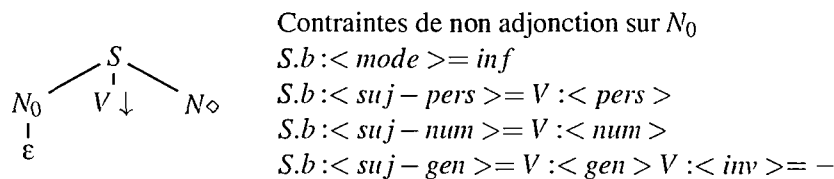
(232) Fais donc ce voyage !

L'infinitive évite l'expression du sujet ou de l'objet redondant avec le contrôleur dans la phrase principale :

(233) *Max veut Max faire un voyage

*Max persuade Luc de Luc faire un voyage

Les arbres TAG positionnent un nœud vide pour le sujet et renseignent les traits permettant de gérer le contrôle du sujet³.



Arbre initial pour représenter le sujet omis des infinitives

Max veut prendre une douche

Pour l'impérative, on considère également qu'il y a sujet omis, suivant en cela [Candito, 1999a, Abeillé, 2002]. La représentation du sujet omis et des phrases à l'impérative est donc sensiblement semblable, sauf pour le placement des clitiques, qui lui est différent :

(234) a. Max veut lui faire du chantage

b. Fais-lui du chantage

6.2.2 Les clitiques

En français, le paradigme des pronoms personnels peut être scindé en deux parties : d'une part, les formes fortes (ou pronoms forts), qui ont une distribution semblable à celle des groupes nominaux, et d'autre part les formes faibles, dites clitiques (aussi appelées particules préverbaux, formes conjointes ou pronoms atones). [Kayne, 1975] a soigneusement décrit les propriétés de surface de ces dernières : les clitiques ne peuvent être ni accentués, ni modifiés, ni focalisés, ni coordonnés, et sont attachés à un verbe. Ils apparaissent en position préverbale lorsque le verbe n'est pas un impératif positif, et après le verbe dans le cas contraire (voir également par exemple [Delaveau et Kerleroux, 1985, Vergnaud, 1985, Miller, 1992, Zribi-Hertz, 1994] ou encore [Heap et Robergue, 2001] pour un historique de l'étude de la cliticisation).

³On a par exemple *Léa veut être invitée*, mais pas **Léa veut être invité*.

Diverses caractéristiques permettent de distinguer le clitique sujet des autres clitiques compléments. Ainsi, contrairement aux autres clitiques compléments, le clitique sujet a une portée large en cas de coordination :

(235) Clitique sujet : il viendra demain et apportera du champagne.

Clitique complément : Max nous apportera du champagne et donnera des livres $\neq$ Max nous apportera du champagne et nous donnera des livres.

De plus, alors que tous les autres clitiques compléments doivent être post-verbaux en même temps, l'inversion du clitique sujet est compatible avec des clitiques compléments préverbaux [Abeillé, 2002] :

(236) Donne-le moi !

*Me donne-le !

* Le donne moi !

Me le donneras-tu ?

[Miller et Sag, 1997], à la suite de [Miller, 1992], ont proposé de considérer les clitiques comme des affixes sur la base de différents critères établis par [Zwicky et Pullum, 1983] :

- **Degré de sélection par rapport à l'hôte** : les clitiques pronominaux s'attachent toujours à un verbe, même s'ils dépendent initialement d'un adjectif ou d'un nom.
- **Coordination** : le clitique objet doit être répété lors de la coordination de 2 verbes, et ne peut pas avoir de portée large dans le cas d'une coordination :

(237) a. Jean le voit et le suit

b. *Jean le voit et suit

- **Rigidité de l'ordre des clitiques** qui ne dépend pas de critères syntaxiques. Par exemple, l'ordre accusatif/datif, dépend de la personne, et non du cas du clitique :

(238) Max *la_{acc}* *lui_{dat}* donne

Max *me_{dat}* *la_{acc}* donne

- **Existence de "trous" dans la combinatoire** : toutes les combinaisons de clitiques ne sont pas autorisées.

(239) a. Il le lui a présenté

b. *Il me t'/lui a présenté

c. Il m'a présenté à toi/elle

De manière analogue, la première personne du clitique nominatif en position inversée n'est pas acceptable avec la plupart des verbes :

(240) a. *Viens-je ?

b. *Cours-je ?

- **Idiosyncrasies morphologiques** On trouve des irrégularités notoires dans la combinatoire. L'exemple le plus communément cité est celui du verbe *aller* :

(241) a. Pierre y va

*Pierre va

b. Pierre ira

Pierre y ira

- **Les affixes sont sujets aux règles phonologiques** : dans certains contextes, les clitiques sont sujets à une liaison obligatoire. (242) se prononce ainsi obligatoirement [ãna].

(242) Marie en a

En se basant sur ces observations [Miller, 1992, Miller et Sag, 1997] proposent une formalisation en HPSG : l'affixation est réalisée dans le lexique à l'aide d'un principe d'affixation. Pour l'interface syntaxe sémantique, ils considèrent le verbe comme un foncteur dont un des arguments a déjà été consommé par le clitique.

Représenter l'hypothèse affixale en TAG soulève des difficultés. Il faudrait déjà modifier l'architecture des parseurs. Un traitement morphologique des clitiques affecte en effet la sous-catégorisation des mots. Dès lors, ce ne seraient plus des lemmes qui seraient associés à des familles, mais des formes fléchies, ce qui demanderait de revoir l'architecture d'un lexique à la XTAG.

Un autre argument est apporté par [Crabbé, 2005]. En effet, la construction sémantique proposée par [Miller, 1992, Miller et Sag, 1997] supprime toute possibilité de calculer la sémantique sur l'arbre de dérivation : "le langage de construction sémantique qu'ils présupposent leur permet de reconstituer la sémantique pour des cas du type *Le garçon l'a vue* où *l'* est bien affixé à *avoir* mais dépendant de *voir*. Ce qui nécessite de définir la forme *l'a* comme un foncteur qui consomme comme argument le foncteur *voir*. La composition est faite de telle sorte que ce soit bien le foncteur *l'a* qui consomme l'objet de *voir*".

La solution morphologique, même si elle a de fait été présentée par [Candito, 1999a] dans sa thèse, n'a pas été effectivement implémentée. Candito reste elle-même très vague sur la réalisation à adopter à l'aide d'une métagrammaire⁴. La représentation syntaxique que nous présentons a également été avancée par [Abeillé, 2002].

⁴En tous cas, la solution morphologique ne dispense pas la création de classes spéciales dans la métagrammaire.

Dans l'optique d'un traitement en syntaxe, il faut remarquer que les clitiques, à la différence des pronoms forts, reçoivent une catégorie particulière, ce qui les empêche de se coordonner avec des groupes nominaux et prépositionnels, et de prendre des modificateurs. De plus, la séquence formée par les clitiques et le verbe est amenée à réaliser un sous-constituant, en vue de ses propriétés phonétiques et prosodiques (liaison obligatoire), et du fait de l'interdiction d'un modificateur entre clitiques et verbe.

Les clitiques sont considérés comme des nœuds à substituer dans les arbres verbaux. En dimension 3, ils se réalisent de manière indépendante. Un niveau pour le clitique sujet est créé, et un niveau supplémentaire pour tous les clitiques compléments également. C'est en fait la dimension 4 qui se chargera de les placer correctement et d'interdire les placements incorrects.

La solution syntaxique a l'avantage d'être plus directement lisible et intuitive. Elle souffre cependant elle aussi de divers problèmes : d'une part, elle augmente potentiellement le nombre d'arbres d'une famille, puisqu'à un arbre traitant les clitiques morphologiquement, peut en correspondre plusieurs traitant les clitiques syntaxiquement. Ainsi, à un verbe plein, portant des clitiques sujet, accusatif et datif, peut correspondre 3 arbres rendant ainsi compte des différents ordres possibles :

- Il le(acc) lui(dat) donne.
- Il nous(dat) le(acc, def) donne.
- Il nous(dat) en(acc, indef) donne.

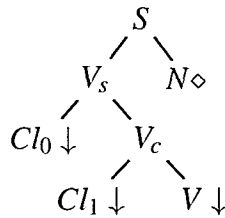
D'autre part, elle ne permet pas de gérer efficacement les interactions avec l'auxiliaire. Dans le cas où les clitiques sont préverbaux, aucun problème ne se pose : l'adjonction de l'auxiliaire se fait entre le clitique et le verbe. Mais si les clitiques sont postverbaux, il faudrait alors pouvoir distinguer forme simple et forme composée, ce qui n'est bien entendu pas possible, par soucis d'homogénéité^{5 6}.

- Prends-tu une douche ?
- As-tu pris une douche ?

On fournit ci-après un arbre illustrant le traitement syntaxique des clitiques, pour la phrase *Il lui fait du chantage* :

⁵Candito signale également ce type de problème avec les verbes à montée.

⁶La solution pratique à ce problème est d'adjoindre le clitique (cf. [Abeillé, 1991])



Un arbre illustrant le traitement des clitiques en syntaxe
Il lui fait du chantage

Le fait que nous ayons choisi en dimension 1 de ne pas sous-catégoriser les groupes prédicatifs par des formes clitiques interdira toute réalisation comme clitique.

6.2.3 Les extractions

Nous allons maintenant terminer notre panorama des alternances de réalisation en présentant les phénomènes d'extraction. 3 phénomènes vont ici nous intéresser : la question (243a), la relative (243b) et la construction clivée (243c).

- (243) a. Quel crime Max a-t-il commis ?
 b. Le crime qu'il a commis (était sans aucun sens)
 c. C'est un crime qu'il a encore commis

De manière générale, les phénomènes d'extraction se caractérisent par la position de leur argument extrait, antéposé à la phrase, et par la possibilité pour ces mêmes arguments d'être éloignés à distance non bornée du prédicat dont ils dépendent.

6.2.3.1 Les questions

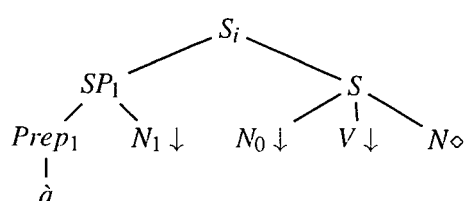
Les questions représentent un cas typique de phénomène d'extraction : l'élément interrogé est placé en tête de phrase et peut être réalisé, sauf pour le sujet interrogé, à longue distance de son hôte :

- (244) a. Quel individu a commis cette odieuse agression ?
 *Quel individu crois-tu qu'a commis cette odieuse agression ?
 b. Quel crime a donc commis cet individu ?
 Quel crime crois-tu qu'a donc commis cet individu ?

Pour les familles à nom prédicatif, les questions sur les compléments non prédicatifs sont réalisées, comme dans le cas des verbes pleins, par des arbres spécifiques à N ou SP antéposé.

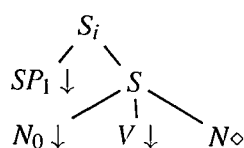
Suivant, [Candito, 1999a, Abeillé, 2002], nous considérons les pronoms interrogatifs (*qui*, *que*, *lequel*) comme des noms à substituer, portant les traits $\langle det \rangle = +$ et $\langle wh \rangle = +$.

Par exemple, les arbres ci-après illustrent la représentation de questions portant sur des noms ordinaires (a-objet et locatif interrogés). Le premier arbre permet par exemple d'analyser la phrase *A qui Max fait-il du chantage ?*, le second, *Où Max fait-il un pèlerinage ?*



Contrainte de non adjonction sur S_i

$S_i.b : \langle princ \rangle = +$
 $S_i.b : \langle wh \rangle = +$
 $S_i.b : \langle inv \rangle = x$
 $S_i.b : \langle princ \rangle = x$
 $S.t : \langle inv \rangle = x$
 $S.b : \langle wh \rangle = -$
 $N_1 : \langle wh \rangle = +$
 $N_0 : \langle wh \rangle = -$



Contrainte de non adjonction sur S_i

$S_i.b : \langle princ \rangle = +$
 $S_i.b : \langle wh \rangle = +$
 $S_i.b : \langle inv \rangle = x$
 $S_i.b : \langle princ \rangle = x$
 $S.t : \langle inv \rangle = x$
 $S.b : \langle wh \rangle = -$
 $N_0 : \langle wh \rangle = -$
 $SP_1 : \langle wh \rangle = +$
 $SP_1 : \langle loc \rangle = +$

Deux arbres permettant respectivement d'analyser
A qui Max fait-il du chantage ? et *Où Max fait-il un voyage ?*

Le questionnement portant sur le nom prédicatif est tout différent : comme le nom prédicatif est tête lexicale de la structure, il ne peut être remplacé par un pronom interrogatif :

(245) a. *Que fait Max ?

Une enquête sur la population

b. *Que fait Max sur la population ?

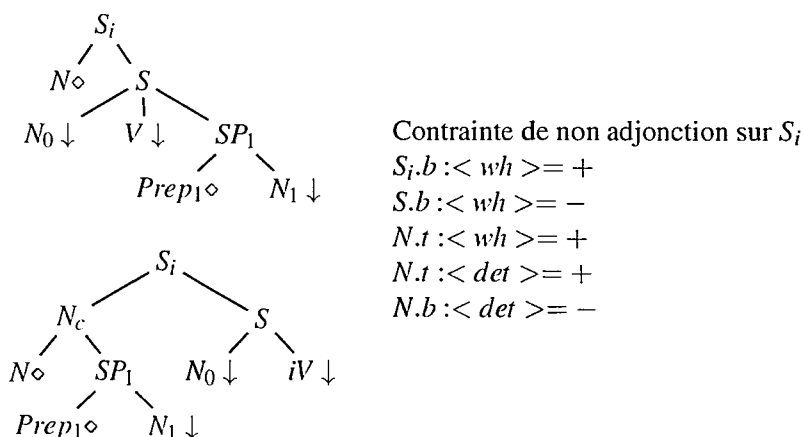
Une enquête

La seule question qui puisse être autorisée est celle où le nom prédicatif est accompagné d'un déterminant⁷ :

- (247) a. Quelle enquête Max fait-il donc ?
 b. Quelle enquête sur la population Max fait-il donc ?

Pour forcer l'adjonction du déterminant, on fait porter en top du nom prédicatif, le trait $\langle det \rangle = +$, et en bottom, le trait $\langle det \rangle = -$. Seul *quel* pourra alors s'y adjoindre, le nom prédicatif portant également en top le trait $\langle wh \rangle = +$.

Les deux arbres suivants illustrent l'interrogation sur le nom prédicatif objet :



Deux arbres permettant respectivement d'analyser
Quelle enquête Max fait-il sur la population ? et
Quelle enquête sur la population Max fait-il ?

⁷Ceci est en fait une approximation. On a vu au chapitre 3 que la question portant sur le nom prédicatif pouvait être possible dès lors que le verbe permettait de déterminer l'objet interrogé :

- (246) a. Qu'a commis Max ?
 b. Un crime odieux
 c. A quoi a procédé Max ?
 d. A une opération délicate

Lorsque la question sur le groupe prédicatif n'est pas possible comme dans (248) , on note au niveau du lexique du nom prédicatif le trait $N.t : < wh > = -$.

- (248) *Dans quelle ignorance Max est-il ?
 *Dans quelle ignorance de cet événement Max est-il ?

6.2.3.2 Les relatives

Traditionnellement, les relatives sont décrites comme des propositions subordonnées qui sont introduites par un pronom relatif :

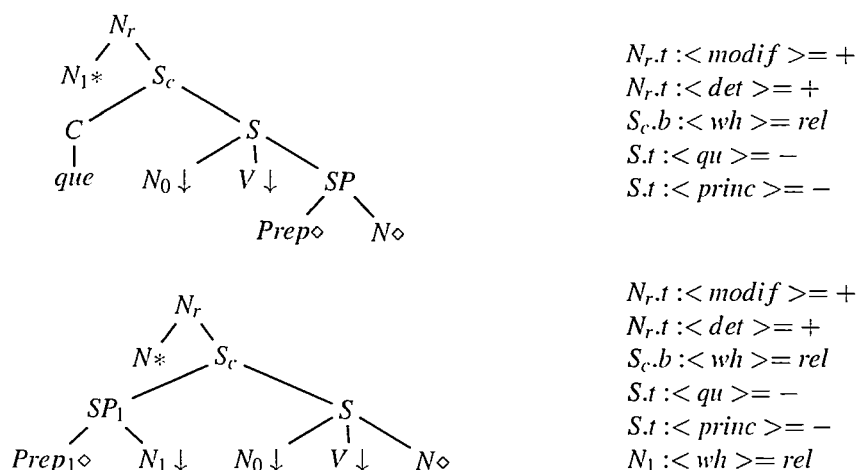
- simple : *qui, que, dont, où*
- composé : *lequel* qui s'amalgame avec les prépositions *à* et *de* pour former *auquel* et *duquel*, et qui varie en genre et en nombre.

Mais, selon [Kayne, 1976] et [Godard, 1988], on peut distinguer parmi les pronoms relatifs 2 types de comportement. Certains sont des complémenteurs, qui n'ont ni genre, ni nombre, ni trait sémantique, et se trouvent toujours en position complémenteur ; d'autres sont des mots *qu-* et peuvent varier en genre et en nombre, renvoyer à des animés et occuper des positions variées dans la phrase.

On distingue donc *qui, que, et dont* des autres introducteurs qui sont des mots *qu-*. *qui* a un statut particulier car, il est suivant les cas, complémenteur ou mot *qu-* : le *qui* de la relativisation du sujet est insensible au trait +/- animé, alors que le *qui* qui apparaît dans la relativisation du complément d'une préposition implique le trait + animé.

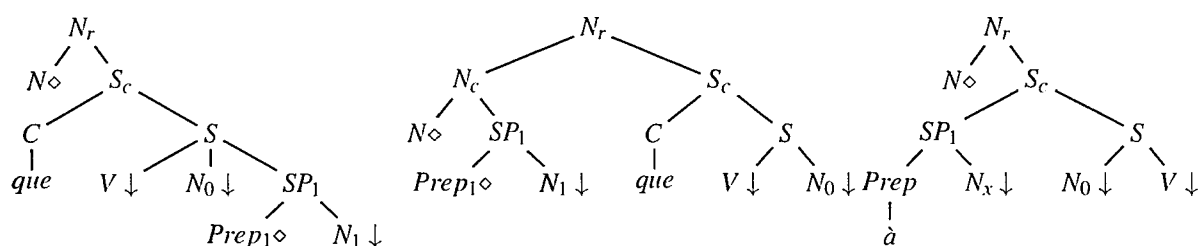
Pour les familles à nom prédicatif, plusieurs cas se présentent. Pour les compléments non prédicatifs, nous suivons [Candito, 1999a, Abeillé, 2002]. Dans cette optique, les relatives sont représentées par des arbres auxiliaires auxquels il manque un argument totalement ou partiellement identifié à l'antécédent de la relative. Les mots *qu-* ancrent des arbres de catégorie N, portant le trait *wh* et ayant pour valeur *rel*. Les complémenteurs sont représentés directement dans l'arbre comme des co-têtes.

Les arbres fournis ci-après illustrent la représentation de quelques relatives. Le premier arbre illustre un sujet relativisé, le suivant un complément prépositionnel relativisé.



Deux arbres auxiliaires permettant d'analyser les groupes nominaux
Le voisin que Max met en colère (1) et *Le voisin contre lequel le voleur commet un crime* (2)

Tout comme pour les questions, le cas de la relativisation du complément prédicatif est tout autre. En effet, lorsque le nom prédicatif appartient à l'antécédent de la relative, celui est une tête lexicale. L'arbre n'est donc plus un arbre auxiliaire, mais un arbre initial.



Pour certains groupes prédicatifs prépositionnels (ceux introduits par *en*), un changement de préposition peut être observé. C'est pourquoi, les familles en *etreEnN* renseignent les deux prépositions.

- (249) Max est en colère
 Max est dans une colère noire

*La colère en laquelle Max est (semble destructrice)
 La colère dans laquelle Max est (semble destructrice)

6.2.3.3 Les clivées

Une phrase comme *C'est le chaton qui a mangé tes lacets* est ambiguë entre une interprétation déictique (présentatif + relative) et une interprétation contrastive, caractéristique des clivées.

Dans l'interprétation déictique, le chaton qui a mangé les lacets est opposé par le locuteur à d'autres chatons ; *qui a mangé tes lacets* est une précision permettant alors de différencier un chaton particulier parmi d'autres. Dans l'interprétation contrastive, on ne se trouve plus dans cette situation, on oppose un chaton à d'autres êtres qui n'ont pas mangé de lacets (Réponse à la question *Qui a mangé tes lacets ?*).

Selon [Vikner, 1972], on peut distinguer trois parties dans une phrase clivée :

- le présentatif constitué par *c'est*
- le focus (ou foyer), placé entre le présentatif et *qui* ou *que*
- la phrase tronquée : ce qui suit *qui* ou *que*

D'un premier abord, on pourrait être tenté d'effectuer un rapprochement entre clivées et relatives. Leurs propriétés syntaxiques diffèrent cependant. [Moreau, 1976] note que le foyer d'une clivée est moins contraint que l'antécédent d'une relative. En particulier, il peut être un syntagme prépositionnel⁸ ou infinitif :

- C'est à Léa que Max donne un baiser
- C'est poser nu qui fait scandale

En outre, elle note également que les relatives restrictives n'acceptent pas comme antécédent un nom propre, un pronom personnel ou démonstratif, ou une complétive, alors que les clivées n'ont pas cette restriction au niveau de leur foyer :

- C'est Max qui m'amuse vs. *J'ai donné un cadeau à Max qui m'amuse
- C'est elle qui est gentille vs. *J'ai acheté des fleurs à elle qui est gentille
- C'est celui-là qui m'amuse vs. *Celui-là qui m'amuse me regarde avec insistance
- C'est qu'il ne soit pas encore là qui m'inquiète vs. *Qu'il ne soit pas encore là qui est évident m'inquiète

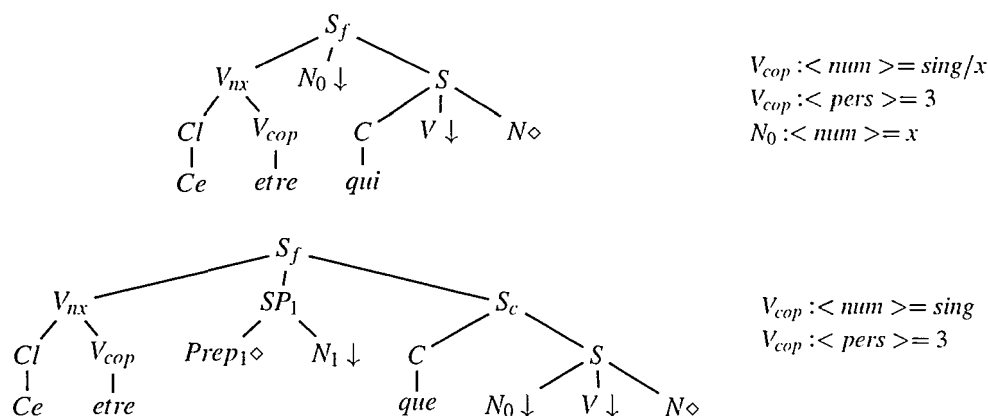
Le verbe *être* du présentatif ne connaît pas de variation de personne, mais en connaît en nombre et en temps : on a *est, sera, était...* lorsque l'élément à droite de *être* est singulier, et *sont, seront, étaient...*, lorsque cet élément est au pluriel⁹.

La représentation des phrases clivées en TAG suit celle proposée par [Candito, 1999a, Abeillé, 2002]. Les phrases clivées sont donc considérées comme des arbres élémentaires à têtes multiples (*ce, être*¹⁰ et le complémenteur lexicalisé sont des co-têtes fonctionnelles). L'accord en nombre pour *être* est noté par une disjonction : au singulier, l'accord est réalisé avec *ce*, au pluriel, l'accord est ou non réalisé avec le foyer nominal de la clivée.

⁸Il existe une construction concurrente dans laquelle seul le syntagme nominal complément de la préposition est extrait : *C'est Léa à qui Max donne un baiser*. Ce type de construction est considéré comme agrammatical par certains locuteurs [Moreau, 1976], mais peut être utilisé occasionnellement dans un style formel [Jones, 1996].

⁹L'accord en nombre ne se fait cependant pas lorsque le présentatif est suivi de *nous* ou *vous*. De plus, l'accord n'est pas toujours réalisé, même lorsque *être* est suivi d'un syntagme pluriel autre que *nous* ou *vous*.

¹⁰Le verbe copule n'est pas une co-tête dans l'implémentation de Candito. Il doit être substitué. Aucune explication n'est fournie pour expliquer cette divergence pratique.



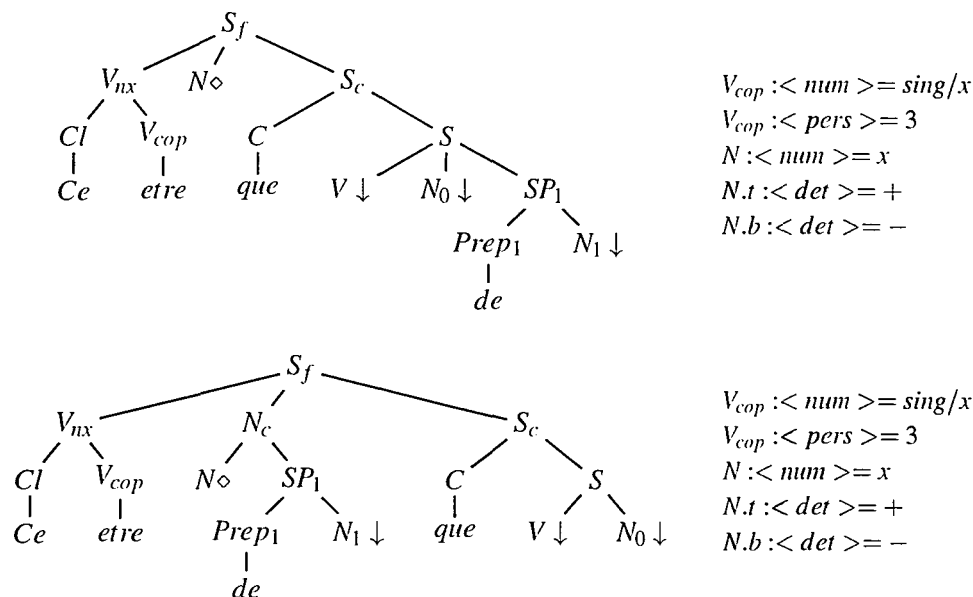
2 arbres permettant respectivement d'analyser les phrases

C'est Max qui prend une douche (1) et *C'est contre Luc que Max commet un crime* (2)

Outre les clivées des arguments libres, on prévoit également pour les familles à nom prédicatif des structures pour l'extraction du groupe prédicatif. Le foyer de la clivée porte toujours en top le trait $\langle det \rangle = +$ et en bottom le trait $\langle det \rangle = -$, ce qui force l'adjonction du déterminant. On a donc (250a), tout en excluant (250b).

(250) a. C'est une peur énorme du loup que Max a

b. *C'est peur du loup que Max a



Deux arbres permettant d'analyser respectivement les phrases

C'est une censure que fait Max de cette page (1) et

C'est une censure de cette page que fait Max (2)

Pour les familles en *etreEnN*, on observe là encore, une alternance dans l'usage de la préposition. Lorsque le nom prédicatif n'admet pas de déterminant, la préposition *en* est utilisée, mais lorsque le nom prédicatif est déterminé, la préposition utilisée est *dans*.

- (251) a. C'est en colère qu'il s'est mis
 b. C'est dans une colère noire qu'il s'est mis

6.2.4 Accord du participe passé

Jusqu'à présent, nous avons représenté différents phénomènes d'extraction, en laissant de côté l'accord du participe passé. Cet accord a la réputation d'être un phénomène complexe, et parfois controversé. [Grevisse et Goose, 1993] fournissent ainsi quelques 16 pages de règles permettant de décrire cet accord. D'autres comme [Tesnière, 1959] semblent signaler que la suppression de cet accord se fera un jour ou l'autre :

Aussi bien l'usage actuel est-il purement livresque, et aujourd'hui l'accord préconisé par la grammaire ne se fait plus même dans la langue parlée courante des personnes cultivées. On dit sans sourciller : *la lettre que j'ai écrit*. La règle est morte de complications... (Tesnière, 581 :24)

Sans passer d'un extrême à l'autre, on peut néanmoins tenter de fournir quelques généralisations. Nous nous appuyons en cela sur la grammaire pédagogique de [Dubois et Lagane, 1973] :

- Les participes passés non réfléchis employés avec *être* s'accordent avec leur sujet :
 - Les feuilles mortes sont tombées hier
 - *Les feuilles mortes sont tombé hier
- Sauf en cas d'emploi avec un impersonnel, ou un clitique *en*, le participe s'accorde avec son objet direct cliticisé ou extrait :
 - Les accidents qu'il est arrivé ce jour-là (ont été meurtriers).
 - *Les accidents qu'il est arrivés ce jour-là (ont été meurtriers).
 - (J'ai cueilli des fraises dans le jardin et) j'en ai mangé.
 - *(J'ai cueilli des fraises dans le jardin et) j'en ai mangées.
 - Les chocolats que Max a offerts à Léa (semblent délicieux).
 - *Les chocolats que Max a offert à Léa (semblent délicieux).

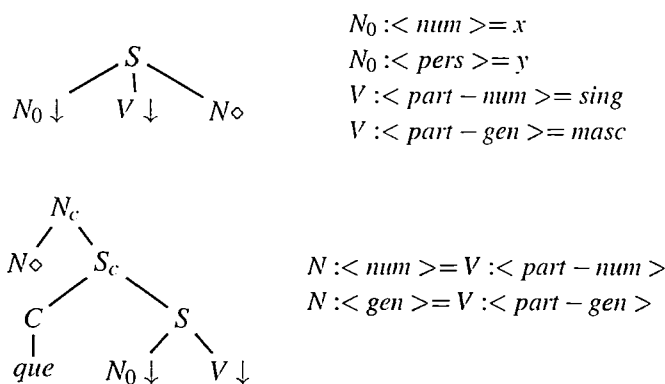
- Max (a acheté des chocolats et) les a offerts à Léa.
- *Max (a acheté des chocolats et) les a offert à Léa.

- La voiture que Max s'est offerte (est splendide).
- *La voiture que Max s'est offert (est splendide).

- Sa voiture, Max se l'est offerte (en Janvier).
- *Sa voiture, Max se l'est offert (en Janvier).

- Dans tous les autres cas, le participe passé est masculin singulier :
- Max a offert une voiture à son père.
- Max s'est offert une splendide voiture.
- Les filles ont dormi dans la chambre.

Pour la **représentation en TAG**, nous suivons [Candito, 1999a] qui utilise les traits *part – num* et *part – gen* pour réaliser l'accord du participe passé¹¹. Suivant les cas, ces traits prendront les valeurs respectives de *masc* et *sing*, ou la valeur des traits de *num* et *gen* de N_0 , de $N_\diamond$, de N_1 (extraits) ou de Cl_1 . Les auxiliaires identifient les traits *part – num* et *part – gen* avec le nombre et le genre du participe auquel ils s'adjoignent.



Deux exemples d'accord du participe passé

Sachant maintenant comment représenter les phénomènes qui nous intéressent, nous allons pouvoir décrire l'implémentation réalisée.

¹¹ On distingue les traits *part – num* et *part – gen* des traits *num* et *gen* car les mêmes arbres peuvent être ancrés par des formes finies du verbe. Dans ce cas, seuls ces derniers sont pertinents.

6.3 Quelques remarques et exemples d'implémentation

Bien entendu, nous ne pourrions pas être exhaustif. La dimension 3 est en effet relativement complexe et comporte énormément de classes. Nous allons donc nous attacher à la description de certains phénomènes qui nous paraissent les plus représentatifs de l'usage que l'on peut faire du compilateur, renvoyant le lecteur à [Candito, 1999a] pour des détails complémentaires.

6.3.1 La représentation des extractions

Les phénomènes d'extraction constituent un cas particulièrement intéressant pour la description de la dimension 3. En effet, ils permettent de partager l'information de manière très pertinente. Nous allons décrire ces particularités avec la présentation des clivées.

Dans tous les cas, le présentatif est invariant, et peut donc être immédiatement décrit dans la classe la plus générale représentant l'extraction.

La classe (CLIVAGE) hérite de la classe (EXTRACTIONS)	
Elle renseigne le bout d'arbre représentant le présentatif	
Syntaxe : (CLIVAGE (EXTRACTIONS) (init-constants '(((S_f) :cat ("S") :ind "f") ((V_nx) :cat ("V") :ind "nx") ((Cl_ce) :cat ("Cl") :ind "x") ((ce) :cat ("ce")) ((v_cop) :cat ("V"))))) (init-simple-desc '(S_f (V_nx (Cl_ce (ce)) (V_cop))))))	
A l'issue de cette description on a donc créé le bout d'arbre : $ \begin{array}{c} S_f = S_f \\ \\ V_nx = V_nx \\ \swarrow \quad \searrow \\ Cl_ce = Cl \quad V_cop = V_cop \\ \qquad \qquad \\ Ce \qquad \qquad \text{etre} \end{array} $	

Ensuite, deux actions sont possibles : soit c'est le sujet qui est clivé, et alors la forme de la clivée est en *C'est ... qui*, soit c'est un complément qui est clivé et alors la forme de la clivée est en *C'est ... que*.

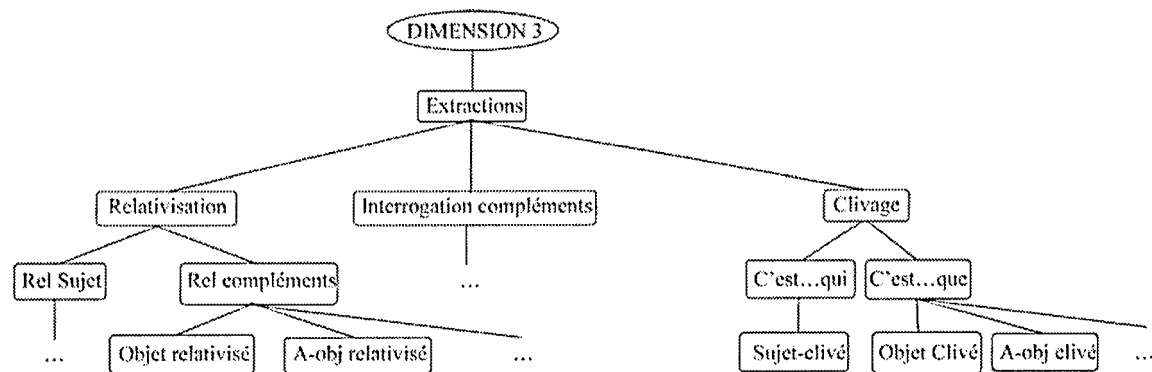
La description du clivage est donc alors naturellement scindée en 2 classes concurrentes.

La classe (C'est...qui) hérite de la classe (CLIVAGE) et précise le complémentateur
Syntaxe : (CEST...QUI (CLIVAGE) (init-constants '(((foyer_suj) :cat ("N" "S") :function sujet) ((Comp_qui) :cat ("C") :constraints "NA") ((qui) :cat ("qui")))) (init-simple-desc '(S_f (V_nx) (foyer_suj) (S_d (Comp_qui (qui)))))) (init-description '((S_d nil) ((Comp_qui nil)) ((vmarq t))))) Après héritage, le bout d'arbre de la classe est donc :

La classe (CEST...QUE) subit un traitement analogue, si ce n'est qu'elle ne fournit pas de fonction au foyer de la clivée, et qu'elle introduit un niveau S intermédiaire.

La classe (C'est...que) hérite de la (CLIVAGE) et précise le complémentateur
Syntaxe : (CEST...QUE (CLIVAGE) (init-constants '(((foyer) : cat ("N" "S")) ((Comp_que) :cat ("C") :constraints "NA") ((que) :cat ("que")) ((S_c) :cat ("S") :ind "c" :constraints "NA")) (init-simple-desc '(S_f (V_nx) (foyer) (S_c (Comp_que (que)) (S_d))))) Après héritage, le bout d'arbre de la classe est donc :

Le graphe d'héritage décrit page suivante permet de représenter l'ensemble des phénomènes d'extraction pris en compte. Outre, les phénomènes de clivage décrits ci-avant, il intègre donc les relatives et les interrogatives.

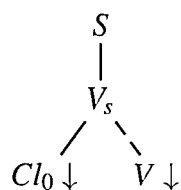


6.3.2 La représentation des clitiques

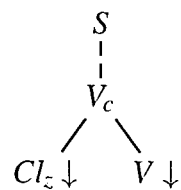
La représentation des clitiques en dimension 3 doit préparer l'ordonnement en colonne de [Perlmutter, 1970]. La réalisation du clitique pour une fonction particulière n'est donc pas unique et dépend de la colonne dans laquelle il se situe.

Position	I	II	III	IV	V	VI
	je	me	le	lui	y	en
	tu	te	la	leur		
	il	se	les			
	elle	nous				
	on	vous				
	nous					
	vous					
	ils					
	elles					

Pour les éléments de la colonne I, la réalisation du clitique sujet est décrite ainsi :



Les éléments des autres colonnes, sont représentés un niveau plus bas¹² :



Le placement des éléments de la colonne I par rapport aux éléments des autres colonnes est donc réalisé naturellement, grâce au verbe qui permet de placer V_c sous V_s .

Pour les clitics compléments, on les représente en dimension 3, indépendamment les uns des autres, mais en indiquant à quelle colonne ils appartiennent.

¹²L'indice z du clitique représente la valeur de l'argument.

La déclaration des colonnes se fait de la façon suivante :

```
(CLITIQUES
(DIMENSION3)
)

(CLITIQUES_COMPLEMENTS
(CLITIQUES)
)

(COL2
(CLITIQUES_COMPLEMENTS
(init-constants '(((col2) :cat ("Cl"))))
)

(COL3
(CLITIQUES_COMPLEMENTS
(init-constants '(((col3) :cat ("Cl"))))
)

(COL4
(CLITIQUES_COMPLEMENTS
(init-constants '(((col4) :cat ("Cl"))))
)

(COL5
(CLITIQUES_COMPLEMENTS
(init-constants '(((col5) :cat ("Cl"))))
)

(COL6
(CLITIQUES_COMPLEMENTS
(init-constants '(((col6) :cat ("Cl"))))
)
```

Il reste ensuite à réaliser chaque clitique suivant la fonction et la colonne à laquelle il appartient. Par exemple, pour les clitiques objets :

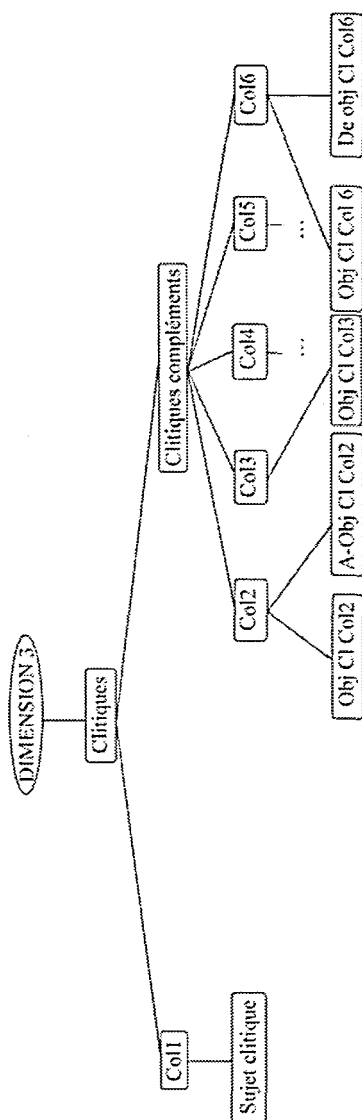
```
(OBJ-CLT-COL2
(FONCTION.OBJET COL2)
(init-constants '(((objet col2) :function objet :cat ("Cl") :qual subst :ind "1")
                  ((sous_comp_verb) :cat ("V") :ind "c"))))
(init-description '((S_d nil) ((sous_compl_verb t))))
(init-description '((sous_compl_verb nil) ((col2 nil)) ((vmarq nil))))
)

(OBJ-CLT-COL3
(FONCTION.OBJET COL3)
(init-constants '(((objet col3) :function objet :cat ("Cl") :qual subst :ind "1")
                  ((sous_comp_verb) :cat ("V") :ind "c"))))
(init-description '((S_d nil) ((sous_compl_verb t))))
(init-description '((sous_compl_verb nil) ((col3 nil)) ((vmarq nil))))
)

(OBJ-CLT-COL6
(FONCTION.OBJET COL2)
(init-constants '(((objet col6) :function objet :cat ("Cl") :qual subst :ind "1")
                  ((sous_comp_verb) :cat ("V") :ind "c"))))
(init-description '((S_d nil) ((sous_compl_verb t))))
(init-description '((sous_compl_verb nil) ((col6 nil)) ((vmarq nil))))
)
```

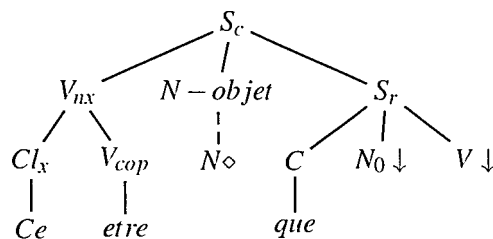
A ce stade, chaque clitique est donc associé à une fonction et à une colonne. Il ne restera donc plus qu'à réaliser l'ordonnement en dimension 4.

La figure fournie page suivante décrit le graphe d'héritage complet des clitiques préverbaux.

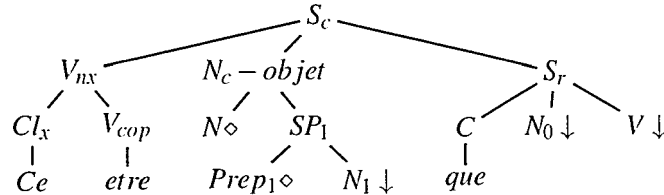


6.3.3 "Descente" du nom prédicatif

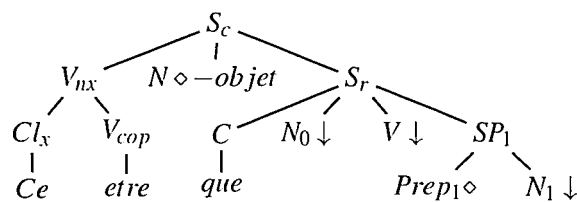
Pour traiter la "descente" du nom prédicatif dans un syntagme, on utilise une relation sous-spécifiée (voir le chapitre introductif sur la Métagrammaire) entre un N qui porte une fonction et le nom prédicatif ancre, qui n'en porte pas. On crée ainsi par exemple une relation sous-spécifiée entre un *N* objet et un nom prédicatif ancre. Lorsqu'un syntagme prépositionnel est complément du nom, ce SP se place à la droite du nom prédicatif, sous le nom objet. En revanche, le lien est ramené à 0, lorsque la descente n'est pas possible : le *N* objet et le nom prédicatif sont alors identifiés comme un seul nœud. Nous fournissons ci-après un exemple pour un objet clivé. Suivant les situations, cet arbre permet la génération de plusieurs résultats :



Contenu de la classe objet clivé pour une ancre



Résultat après croisement de cette classe avec un complément du nom prépositionnel



Résultat après croisement de cette classe avec un complément du verbe prépositionnel

6.3.4 La représentation des syntagmes prépositionnels

Il semble évident que les syntagmes prépositionnels, mis à part leur co-ancre éventuelle, se représentent globalement tous de la même façon. Le grammairien pourrait donc être tenté de créer une classe générale, convenant à tous les syntagmes prépositionnels. Malheureusement, la Métagrammaire traite les variables de manière globale. Ce qui implique qu'on ne peut réaliser un argument prépositionnel qu'une seule fois, car si l'on souhaite réaliser 2 SP, ces 2 SP auront les mêmes noms de variables et l'unification échouera, et ce même si ces SP ont des fonctions syntaxiques différentes. Les variables de chaque SP se doivent donc d'être différentes. Et on ne peut donc proposer un traitement général.

6.3.5 Récapitulation générale : vision hiérarchisée

Nous fournissons ci-après les représentations hiérarchisées du sujet, de l'objet, de l'a_objet, du de_objet et des compléments du nom.

Pour les fonctions qui peuvent réaliser conjointement un nœud à substituer ou un nœud ancre, on a différencié certaines classes suivant la valeur du trait *qual* porté par l'argument. En effet, les classes où l'argument est à substituer sont généralement plus simples (les liens ne sont pas sous spécifiés entre un père et son fils). De plus, la représentation de certains phénomènes syntaxiques n'est pas forcément la même¹³. Les classes terminales qui ne mentionnent pas ce renseignement au niveau de leur nom ont, en principe, un argument à substituer.

Ces représentations correspondent à un agencement théorique. Il se peut que la hiérarchie effectivement implémentée soit parfois différente. Néanmoins, ces représentations permettent d'envisager la lourde tâche du grammairien descripteur de la hiérarchie...

6.3.5.1 Représentations du sujet

On rappelle ici que l'on n'a pas de classe particulière pour un sujet interrogé. L'interrogation est réalisée par la même classe que le sujet nominal en position canonique. La classe du sujet clitique inversé ne crée pas un argument clitique à substituer. Elle assure juste qu'un sujet nominal soit réalisé de manière vide¹⁴. Une intervention au niveau des croisements de certaines classes (contraintes négatives et positives) est par ailleurs nécessaire :

- La réalisation de l'impératif n'est pas possible dans un contexte d'extraction
- La réalisation du sujet clitique inversé n'est pas possible avec une relative ou une clivée
- La réalisation du sujet nominal inversé n'est possible que dans un contexte d'extraction et lorsque l'objet n'est pas réalisé de façon canonique.

¹³Par exemple, dans le cas d'une relativisation, on a selon les cas, des arbres auxiliaires ou des arbres initiaux.

¹⁴Rappelons que l'inversion du clitique sujet est réalisée par adjonction.

Pour aider le lecteur à identifier chaque classe, nous fournissons ci-dessous quelques exemples :

(252) a. Suj-phr-can-ss-comp : *Poser nu* fait scandale

b. Suj-phr-can-subj : *Que Max soit venu à la réunion* a fait scandale

c. Suj-nom-can-subst : *Cet apprenti* a commis une erreur

d. Suj-nom-can-ancr : *Une erreur* a été commise

e. Préverbal (sujet clitique) : *Il* a fait une erreur

f. Inversé (sujet clitique) : *Comment-il* une erreur ?

g. Suj-nom-inv-subst : *C'est* une erreur que fait *cet apprenti*

h. Suj-nom-inv-ancr : *C'est* contre Luc qu'est commis *un crime*

i. Suj-comp-subst (sujet relativisé avec complémenteur) : *L'apprenti qui* a commis cette erreur est rouge de honte

j. Suj-comp-ancr (sujet relativisé avec complémenteur) : *Le crime qui* a été commis hier est ignoble

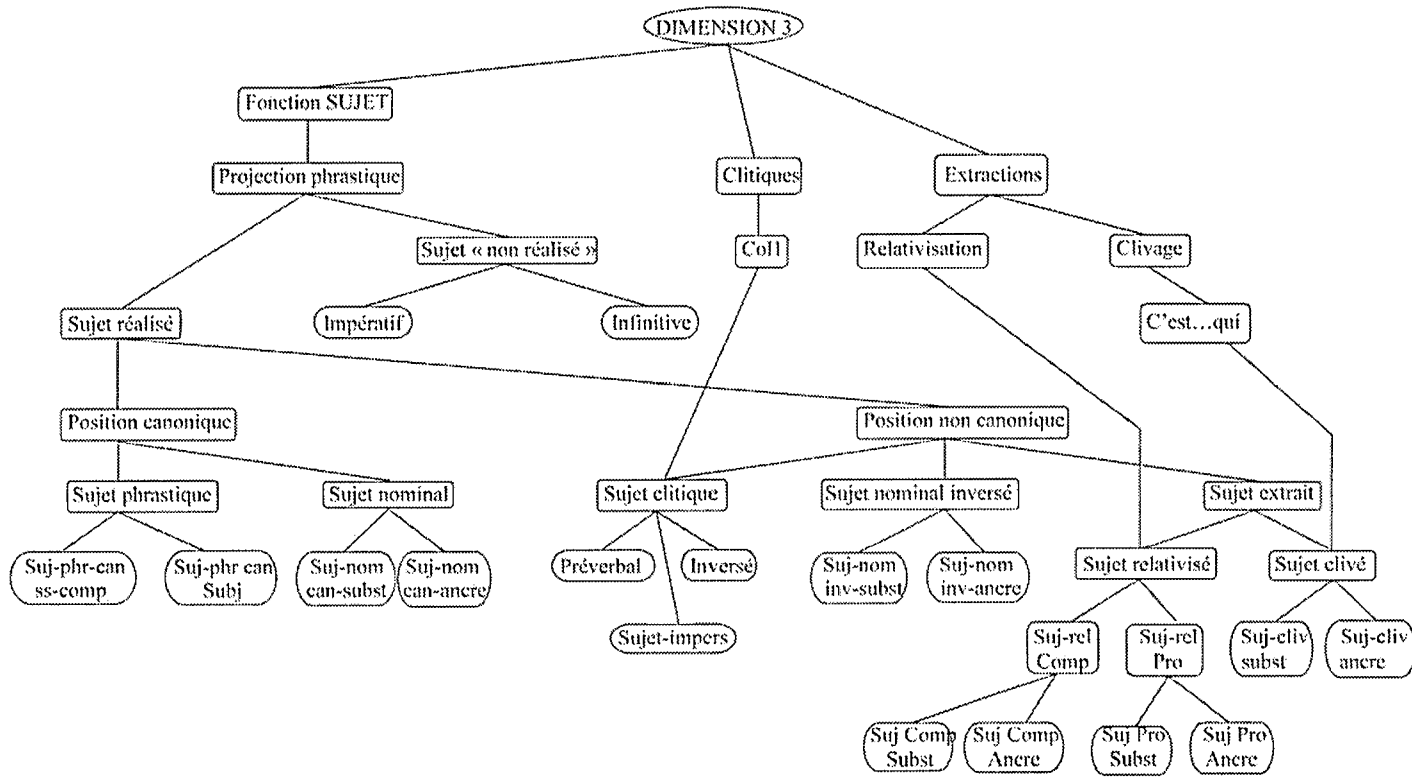
k. Suj-pro-subst (sujet relativisé avec mot qu-) : Il faut en référer à *cet apprenti lequel* a commis une grave erreur

l. Suj-pro-ancr (sujet relativisé avec mot qu-) : *Cette erreur laquelle* a été commise par cet apprenti devrait être sanctionnée

m. Suj-cliv-subst : *C'est cet apprenti qui* a commis une erreur

n. Suj-cliv-ancr : *C'est un crime qui* est commis contre Luc

o. Sujet-impers : *Il* est commis un crime contre Luc

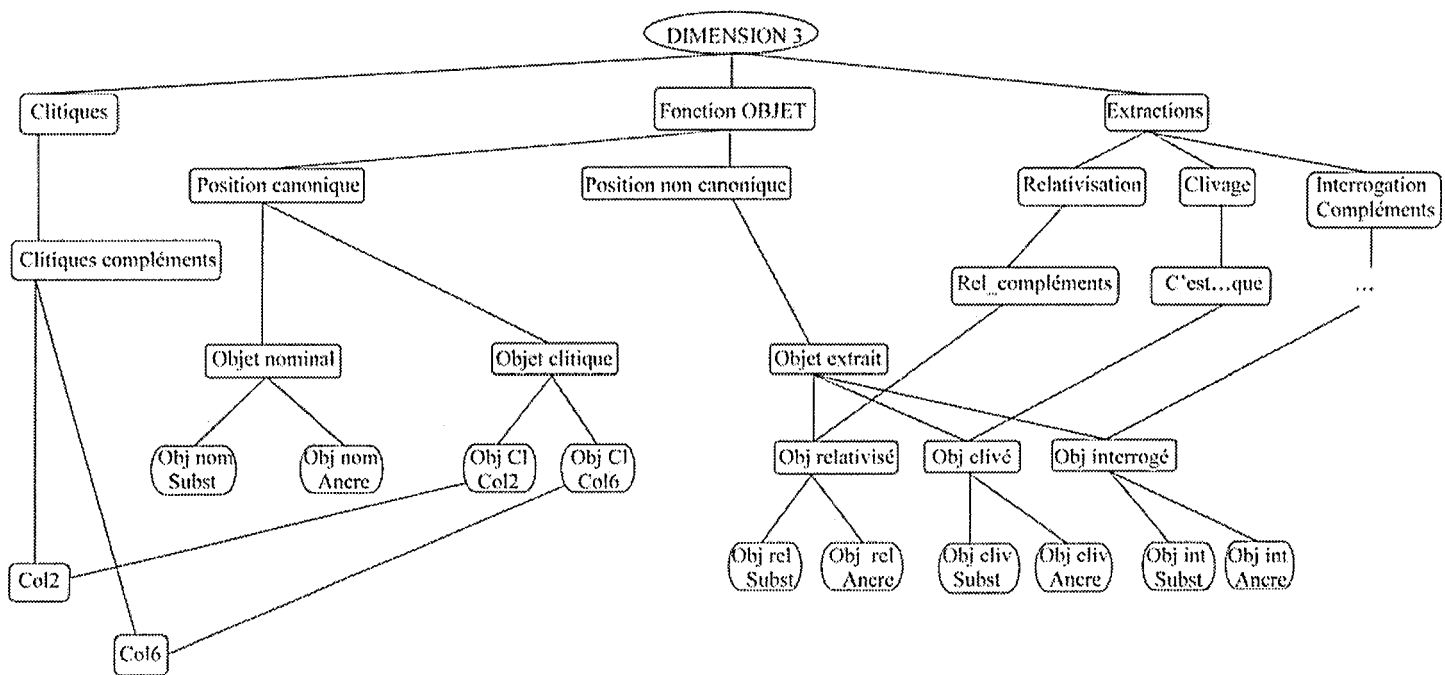


6.3.5.2 Représentations de l'objet

Les familles à nom prédicatif objet excluent naturellement la réalisation de l'objet en tant que clitique, car la fonction objet peut uniquement être portée par le nom. Les autres familles n'ont pas cette restriction.

Le croisement de la réalisation objet du nom prédicatif (en position canonique) avec la réalisation des compléments du nom est par ailleurs interdit.

- (253) a. Obj-nom-subst : Max met *le voisin* en colère
- b. Obj-nom-ancr : L'apprenti a fait *une erreur*
- c. Obj-cl-col2 : Max *nous* met en colère
- d. Obj-cl-col3 : Max *le* met en colère
- e. Obj-cl-col6 : Max *en* met en colère
- f. Obj-rel-subst : *Le voisin que* Max met en colère
- g. Obj-rel-ancr : *L'erreur que* l'apprenti a faite est irréparable
- h. Obj-cliv-subst : *C'est le voisin que* Max a mis en colère
- i. Obj-cliv-ancr : *C'est une erreur irréparable que* l'apprenti a faite
- j. Obj-int-subst : *Qui* Max a-t-il mis en colère
- k. Obj-int-ancr : *Quelle erreur* l'apprenti a-t-il donc commise ?



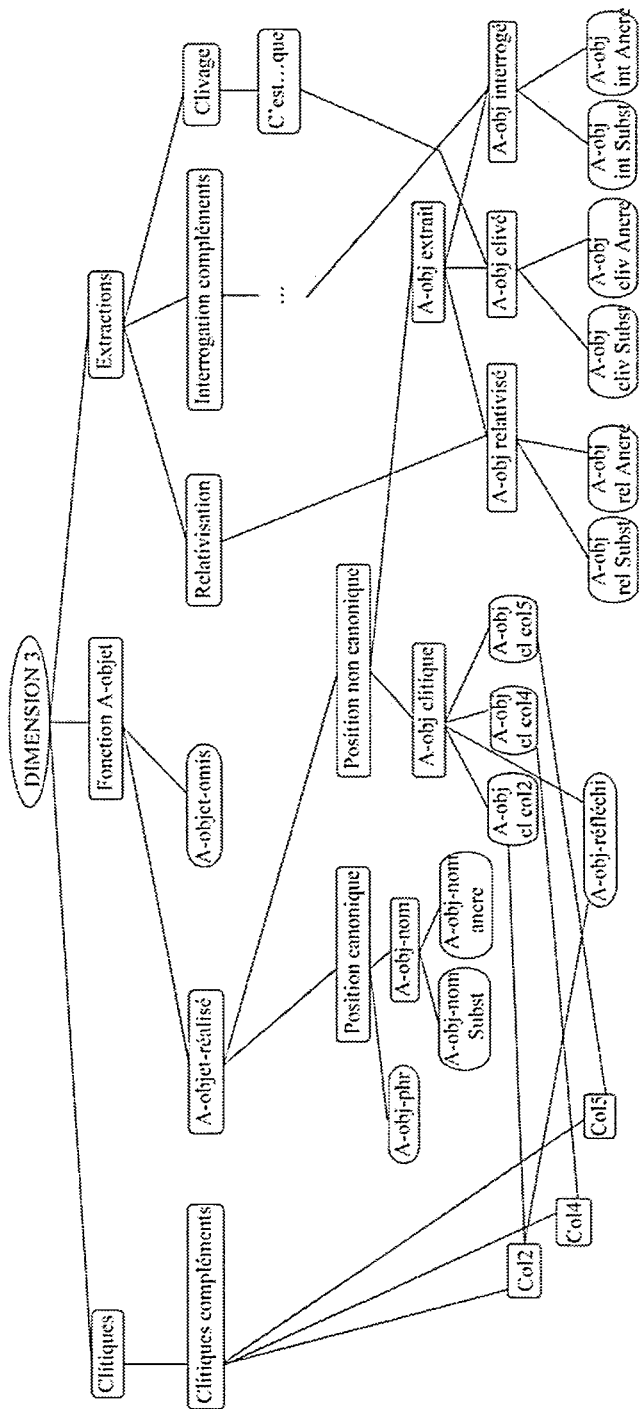
6.3.5.3 Représentation des compléments prépositionnels

Les compléments prépositionnels, qu'ils soient introduits par *à*, *de*, *par* ou une autre préposition, ont une représentation hiérarchique relativement semblable. Seules quelques modifications doivent être effectuées. Malheureusement, comme on l'a vu, il n'est pas possible de décrire un comportement global. Il est donc impossible d'éliminer la grande part de redondance dans la description de ces compléments.

Nous n'entrerons pas dans le détail de toutes ces réalisations. Nous fournissons simplement le graphe d'héritage pour les fonctions de *a_objet* et de *de_objet*.

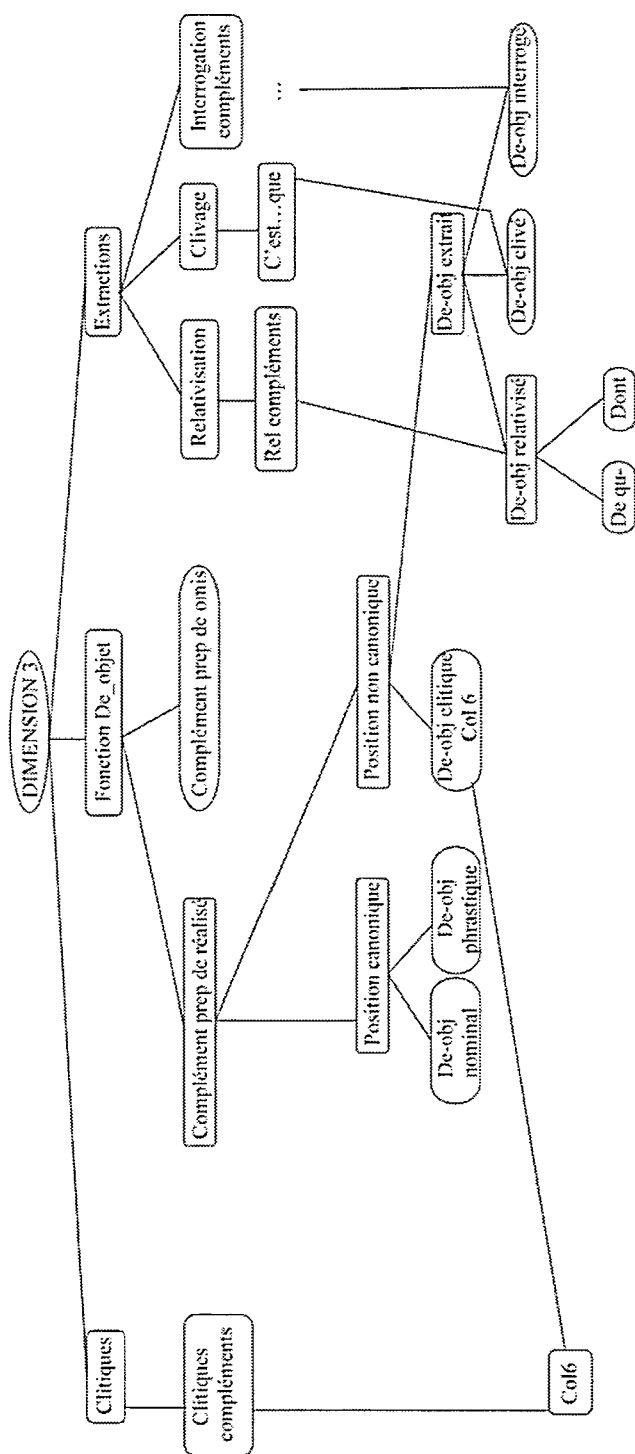
Les *a_objet* peuvent représenter des noms prédicatifs. Dans ce cas, ils ne peuvent bien entendu, pas être omis. Pour la même raison que pour l'objet prédicatif, l'objet indirect cliticisé est également naturellement rejeté.

- (254) a. *a_objet-phr* : Max a de la peine à *se décider*
- b. *a_objet-nom-subst* : Max a fait du chantage à *Luc*
- c. *a_objet-nom-ancr* : Max procède à *l'aménagement de sa chambre*
- d. *a_objet-cl-col2* : Max *nous* fait du chantage
- e. *a_objet-reflechi* : Max *se* donne une gifle
- f. *a_objet-col4* : Max *lui* fait du chantage
- g. *a_objet-col5* : Max *y* fait référence
- h. *a_objet-rel-subst* : *L'apprenti à qui* le directeur fait des reproches sera renvoyé
- i. *a_objet-rel-ancr* : *L'analyse à laquelle* Max fait référence est astucieuse
- j. *a_objet-cliv-subst* : *C'est à cet apprenti que* le directeur fait des reproches
- k. *a_objet-cliv-ancr* : *C'est à des aménagements que* Max a procédé
- l. *a_objet-int-subst* : A qui Max fait-il référence ?
- m. *a_objet-int-ancr* : A quels aménagements Max a-t-il procédé ?



Les de_objets

- (255) a. de-obj-nominal : Max a peur *du loup*
- b. de-obj-phr : Max a peur *de devoir partir*
- c. de-obj-clitique-col6 : Max *en* a peur
- d. de-qu-(de-obj-relativisé) : *Les clichés desquels* le proviseur fera l'agrandissement recevront un prix
- e. dont (de-obj-relativise) : *Les clichés dont* le proviseur fera l'agrandissement recevront un prix
- f. de-obj-clive : *C'est du loup que* Max a peur
- g. de-obj-interrogé : *De quoi* Max a-t-il peur ?



Les compléments du nom

(256) a. Det : Max procède à *son* aménagement

b. De-nominal : C'est à l'aménagement *de sa chambre* que procède Max

c. De-phrastique : C'est un besoin énorme *de partir* qu'a Max

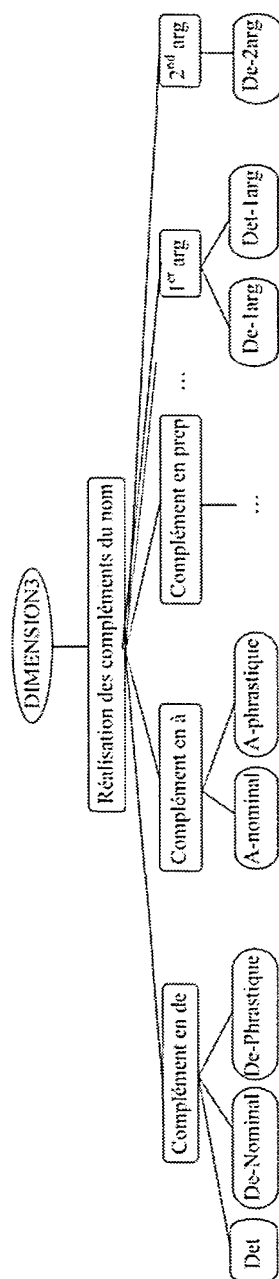
d. A-nominal : C'est une remarque à *l'apprenti* que le directeur a faite

e. A-phrastique : C'est une peine énorme à *se décider* qu'a Max

f. de-1arg : Le portrait *de Léonard* de Mona-Lisa est étonnant

g. det-1arg : *Son* portrait de Mona-Lisa est étonnant

h. de-2arg : Le portrait de Léonard *de Mona Lisa* est étonnant



6.4 La dimension 4 : l'ordre des arguments réalisés

Il nous reste maintenant à ordonner les différents clitiques entre eux suivant en cela le classement original proposé par [Perlmutter, 1970]. La dimension 3 a préparé la réalisation des clitiques et s'est chargée de réaliser chaque clitique suivant la fonction et la colonne dans laquelle il doit apparaître.

Il reste donc à ordonner les arguments représentant les éléments des colonnes II, III, IV, V, et VI. Cet ordre est défini en Dimension 4 en s'appuyant sur le codage réalisé. Par exemple, pour stipuler que les clitiques de colonne II, se réalisent à gauche de ceux de la colonne III, on écrira :

```
(II<III
(DIMENSION4)
(init-simple-desc '(sous_compl_verb (col2) (col3)))
)
```

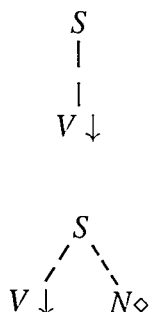
On s'assure également que les croisements ne sont assurés que lorsque les deux éléments sont effectivement présents à l'aide d'un croisement contraint comme :

- (II < III) est utilisée quand les classes (COL2) et (COL3) sont croisées.
- (II < V) est utilisée quand les classes (COL2) et (COL5) sont croisées.
- (II < VI) est utilisée quand les classes (COL2) et (COL6) sont croisées.
- (III < IV) est utilisée quand les classes (COL3) et (COL4) sont croisées.
- (III < V) est utilisée quand les classes (COL3) et (COL5) sont croisées.
- (III < VI) est utilisée quand les classes (COL3) et (COL6) sont croisées.
- (IV < V) est utilisée quand les classes (COL4) et (COL5) sont croisées.
- (IV < VI) est utilisée quand les classes (COL4) et (COL6) sont croisées.
- (V < VI) est utilisée quand les classes (COL5) et (COL6) sont croisées.

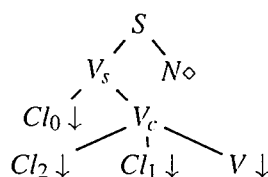
Ce qui s'écrira plus formellement :

```
(defparameter *constrained-crossings* '(
(II<III          COL2          COL3)
(II<V            COL2          COL5)
(II<VI           COL2          COL6)
(III<IV          COL3          COL4)
(III<V           COL3          COL5)
(III<VI          COL3          COL6)
(IV<V            COL4          COL5)
(IV<VI           COL4          COL6)
(V<VI            COL5          COL6)
))
```

Constatons maintenant l'utilité de ces règles et observons les arbres réalisés après le croisement des classes (ACTIF), (OBJET), (SUJ-CLT), (DE-OBJ-CLT) et (A-OBJ-CLT2)¹⁵. Les classes représentant les clitiques ont été définies ci-avant. Il nous reste à rappeler les classes (ACTIF) et (OBJET).



Sans utiliser de règle de dimension 4, on obtiendrait après croisement 2 arbres, dont l'un serait faux, puisque l'ordre entre les deux clitiques compléments n'est pas libre. La règle de la classe ($IV < VI$) nous permet de remédier à ce problème. Et l'on obtiendra ainsi l'arbre correct désiré. C'est-à-dire :



Le seul arbre élémentaire possible pour représenter *Il lui en fait le récit*

Notons que si l'on souhaite également rendre compte du placement des clitiques pour les verbes à l'impératif, il faut définir un nouvel ordre : les clitiques sont immédiatement postposés au verbe et l'ordre est alors $III < II - IV < V < VI$, avec les mêmes restrictions combinatoires que celles des formes anteposées. On définit alors de nouvelles classes d'ordre, qui seront uniquement valables pour un impératif.

¹⁵La classe représentant le nom prédicatif objet n'est pas représentée de cette façon au sein de la hiérarchie. Cela n'a cependant pas d'importance pour la démonstration. La classe représentant l'actif est également plus simple que celle effectivement implémentée.

On a donc présenté les dimensions 3 et 4 nécessaires à la représentation des familles à nom prédicatif. Cette implémentation constitue, pour une part, une réimplémentation de la hiérarchie de Candito, dans un cadre plus strict et plus proche du concept de Métagrammaire défini dans sa thèse. Si l'on souhaitait utiliser cette hiérarchie pour générer les familles des verbes pleins, il faudrait néanmoins y ajouter des éléments phrastiques objets qui ne sont pas définis dans le cadre de nos familles.

Cette présentation hiérarchique et parfois morcelée des données, peut amener à penser que leur implémentation est une chose aisée et naturelle. Or, cela n'est pas tout à fait le cas. En effet, toute la présentation adoptée laisse à penser qu'une hiérarchie s'écrit en une seule fois, en un seul bloc. Il est pourtant bien évident, que l'écriture se fait de manière incrémentale : on commence par traiter un phénomène, puis d'autres, tout en vérifiant que les nouvelles données implémentées ne viennent pas gêner les représentations existantes. C'est ici que le travail le plus laborieux est effectué, car bien souvent, des liens que l'on avait renseignés comme directs, doivent en fait être sous-spécifiés, amenant donc à redéfinir au moins de manière partielle, le travail antérieur effectué¹⁶.

La systématisation des croisements, et donc l'élargissement de la couverture d'une grammaire se fait au prix d'une surveillance accrue de la bonne formation des structures générées. Il s'agit là d'une tâche supplémentaire que la description manuelle ne laissait envisager que très sommairement. Jusqu'à présent ces aspects ont été très peu mis en valeur. Nous espérons les avoir mis au moins partiellement en lumière.

¹⁶De plus lorsque les résultats obtenus diffèrent de ceux attendus, on ne sait pas forcément si c'est l'implémentation de la hiérarchie qui demande à être précisée, ou bien si c'est le compilateur qui ne se comporte pas de manière appropriée. En effet, les choix d'implémentation du compilateur n'ont pas été discutés par Candito, ce qui laisse une grande part d'incertitude. Notre expérimentation du compilateur s'est donc faite de manière très empirique.

Troisième partie

Expériences en corpus

Chapitre 7

Expériences sur corpus et codage lexical

Sommaire

7.1	Le corpus de Paris 7	248
7.2	Préférences lexicales	250
7.2.1	Recherches sur le verbe <i>faire</i>	250
7.2.2	Recherches sur le verbe <i>avoir</i>	253
7.2.3	Recherches sur le verbe <i>commettre</i>	255
7.2.4	Recherches sur le verbe <i>procéder</i>	257
7.2.5	Recherches sur le verbe <i>courir</i>	258
7.3	Extraction de noms prédicatifs	260
7.3.1	Les noms à 2 compléments	260
7.3.2	Autres noms prédicatifs	263
7.4	Expérimentations avec le parseur d'A. Sarkar	265
7.4.1	Analyses sur des phrases test	265
7.4.2	Analyses sur phrases de corpus	273
7.4.3	Raisons des rejets ou des suranalyses	279
7.4.4	Récapitulatif	280

Le présent chapitre a pour but de présenter les expérimentations que nous avons menées sur le corpus de Paris 7 [Abeillé et al., 2003] concernant les constructions à verbe support et les groupes nominaux complexes. Après avoir brièvement introduit le corpus, nous examinons plusieurs verbes afin de déterminer si des préférences lexicales sont observables, puis nous tâchons d'extraire du corpus une liste de noms prédicatifs les plus productifs. Nous terminons nos observations avec une expérimentation à l'aide du parseur d'Anoop Sarkar [Sarkar, 2000] de la grammaire que nous avons décrite durant les chapitres précédents.

7.1 Le corpus de Paris 7

Nous allons utiliser pour nos expérimentations le corpus le Monde développé à l'Université Paris 7. Il s'agit d'un corpus journalistique comportant environ 1 million de mots (soit environ 875 000 mots sans ponctuation), représentant 32 000 phrases, couvrant une période allant de 1989 à 1993. Ce corpus embrasse de nombreux domaines. En particulier, les pages culturelles, économiques et sportives n'ont pas été exclues.

Les textes du corpus constituent une image d'un français écrit contemporain, sans pour autant être représentatifs des différentes productions écrites du monde francophone. Ils ne sont pas particulièrement marqués par un niveau de langue, excluent en général le discours oral ou encore le dialogue.

D'un point de vue théorique, leurs auteurs ont souhaité que ce corpus puisse servir aussi bien à des informaticiens, qu'à des linguistes ou des psycholinguistes. A la différence d'autres projets (Penn Treebank, Prague Dependency Treebank), aucune théorie syntaxique n'a été fidèlement appliquée, permettant ainsi une représentation relativement consensuelle, traduisible dans différents cadres théoriques.

Le corpus a subi plusieurs phases de développement, chaque étape ayant été suivie d'une validation humaine manuelle systématique :

- Dans un premier temps, un étiquetage morpho-syntaxique a été effectué. C'est à cette étape qu'ont été réalisés, la segmentation en mots et en phrase, la désambiguïsation catégorielle et morphologique, le regroupement des mots composés et la lemmatisation [Clément, 2001].
- Puis est venue une phase plus syntaxique, comprenant d'abord le marquage en constituants puis l'assignation de fonctions grammaticales à ces constituants [Abeillé et al., 2003, Abeillé et Barrier, 2004].

Bien entendu, il n'est pas question pour nous de présenter ici les choix linguistiques opérés. Nous renvoyons le lecteur aux différents guides des annotateurs qui ont été renseignés durant toutes les phases du projet [Abeillé et Clément, 1997, Abeillé et al., 2002, Abeillé, 2004], pour qu'il puisse également se faire une idée plus précise de l'encodage réalisé.

Nous noterons cependant que le guide des annotations en constituants renseigne le comportement à adopter avec les CVS à plusieurs reprises :

1. "On a fait l'expérience qu'une lecture suffisamment attentive permettait toujours de désambiguer les phrases du corpus (...) Il ne faut pas hésiter à faire appel à des connaissances encyclopédiques (...)
Dans certains cas, l'interprétation est exactement la même et on a le choix entre deux découpages. Si les tests syntaxiques marchent aussi bien pour les

deux (...), on opte pour celui qui comporte le moins d'enchâssement (principe du rattachement minimum). C'est le cas par exemple pour la plupart des constructions à verbe support, on peut inclure ou non le PP dans le NP, sans différence d'interprétation, on choisit donc de ne pas l'inclure." (Abeillé et al., 2003, p.8)

2. "Dans certains cas on peut trouver un verbe support suivi d'un nom et d'un complément prépositionnel : *avoir peur de, avoir envie de*, etc.

On ne compose pas le verbe parce que le nom peut former un syntagme plus complexe (*avoir une peur bleue de, avoir la plus grande envie de*) et parce qu'on peut le déplacer (*la peur que j'ai eue*), ce qui montre que ce type de construction n'est pas figé." (Abeillé et al., 2003, p.23)

En pratique, ces recommandations n'ont pas toujours été suivies, et il subsiste encore au sein du corpus quelques erreurs d'annotations.

```
<SENT nb="8957">
  <NP>
    <w lemma="un" ei="Dfs" ee="D-ind-fs" cat="D" subcat="ind" mph="fs">une</w>
    <w compound="yes" lemma="opération commando" ei="NCfs" ee="N-C-fs" cat="N" subcat="C" mph="fs">
      <w catint="N">opération</w>
      <w catint="N">commando</w>
    </w>
  </NP>
  <VN>
    <w lemma="avoir" ei="VP3s" ee="V--P3s" cat="V" subcat="" mph="P3s">a</w>
    <w lemma="être" ei="VKms" ee="V--Kms" cat="V" subcat="" mph="Kms">été</w>
    <w lemma="mener" ei="VKfs" ee="V--Kfs" cat="V" subcat="" mph="Kfs">menée</w>
  </VN>
  <PP>
    <w lemma="par" ei="P" ee="P" cat="P">par</w>
    <NP>
      <w lemma="le" ei="Dms" ee="D-def-ms" cat="D" subcat="def" mph="ms">l'</w>
      <w lemma="ex-" ei="PREF" ee="PREF" cat="PREF">ex-</w>
      <w lemma="FLNC" ei="NPs" ee="N-P-s" cat="N" subcat="P" mph="s">FLNC</w>
    </NP>
  </PP>
  <PP>
    <w lemma="contre" ei="P" ee="P" cat="P">contre</w>
    <NP>
      <w lemma="un" ei="Dfp" ee="D--fp" cat="D" subcat="ind" mph="fp">des</w>
      <w lemma="installation" ei="NCfp" ee="N-C-fp" cat="N" subcat="C" mph="fp">installations</w>
    <AP>
      <w lemma="touristique" ei="Afp" ee="A-qual-fp" cat="A" subcat="qual" mph="fp">touristiques</w>
    </AP>
  </NP>
  </PP>
  <PP>
    <w lemma="en" ei="P" ee="P" cat="P">en</w>
    <NP>
      <w lemma="Corse" ei="NPfs" ee="N-P-fs" cat="N" subcat="P" mph="fs">Corse</w>
    </NP>
  </PP>
</SENT>
```

Exemple d'annotation (simplifié) du corpus après les phases de tagging et de shallow parsing

Les interrogations que nous avons effectuées portent sur une version intermédiaire du corpus. A l'époque où nous avons lancé nos investigations, les annotations fonctionnelles n'avaient en effet pas encore fait l'objet d'une validation systématique. Nous nous appuyons donc exclusivement sur la segmentation du corpus. Les résultats des requêtes ont été soigneusement analysés (avec suppression des requêtes erronées) et classés.

Notre interrogation a porté sur 40 fichiers, représentant environ les deux tiers du corpus.

7.2 Préférences lexicales

Nos premières expériences ont porté sur l'observation des verbes *faire*, *avoir*, *commettre*, *procéder* et *courir*. Le but était de déterminer si des préférences lexicales étaient observables, et en particulier de savoir si l'emploi d'un verbe comme support était productif ou au contraire marginal.

7.2.1 Recherches sur le verbe *faire*

Nous avons commencé nos interrogations du corpus avec le verbe *faire*. Ce dernier présente en effet de nombreux emplois (les dictionnaires de langue en sont le reflet), qu'il est parfois difficile de classer.

Dans la majorité des cas, le verbe *faire* est utilisé avec un complément nominal. Néanmoins, il est possible de relever d'autres emplois :

1. faire que :

À cette proposition - formulée par Jean Poperen, et que soutiennent tous les autres courants à l'exception de Lionel Jospin et Louis Mermaz, - on a objecté qu'elle favoriserait l'activisme militant des petites minorités aux dépens des grands courants majoritaires, dont la densité même fait que les militants sont plus difficiles à déplacer.

2. faire pour :

Ce théâtre-là est fait pour être vu, et lu.

3. faire avec sujet impersonnel :

Comme en octobre dernier, le blocus a repris et "il fait froid, écrivent les Izvestia, dans les appartements d'Erevan".

Et bien entendu, le faire causatif, très largement représenté :

Troisième - et sans doute principal - motif de colère des universitaires : à vouloir brusquer les choses, le ministère risque de faire voler en éclats les délicats équilibres de la communauté universitaire.

De manière générale, les emplois de *faire* se découpent donc en 3 classes, dont la répartition est la suivante :

Faire + complément nominal	47%
Faire causatif	39%
Autres emplois de faire	14%
Occurrences de <i>faire</i>	1743

Ce sont les emplois du verbe *faire* suivi d'un complément nominal qui nous ont préférentiellement intéressé. Le découpage des différents emplois en classes a été relativement aisé, la majorité des emplois étant bien représentée. On a ainsi globalement retrouvé le classement proposé par [Giry-Schneider, 1986]¹, à ceci près que nous avons distingué plus finement les emplois de *faire* en rapport avec un nom prédicatif :

1. faire verbe plein :

Nous passons d'un métier où nous faisons des produits de luxe à une activité où nous travaillerons dans des conditions comparables à l'automobile ou l'électronique grand public.

2. faire verbe figé :

Que le ministre de l'équipement, qui avait longtemps fait figure de " dauphin " de Mr Mauroy, ait jugé utile de se faire représenter personnellement à cette négociation confirme qu'il entend agir de façon autonome par rapport au maire de Lille.

3. faire verbe attributif :

Le ministre des affaires étrangères n'a sûrement pas eu tort d'ajouter qu'une " nouvelle approche " du problème avait été définie à partir de la proposition australienne de faire de l'ONU le véritable pivot d'un règlement.

4. faire opérateur causatif sur avoir :

En France, la tendance élitiste leur fait du mal.

5. faire verbe support :

C'est peut-être ce problème qui explique qu'aucune annonce n'avait été encore faite au public jeudi en fin de matinée.

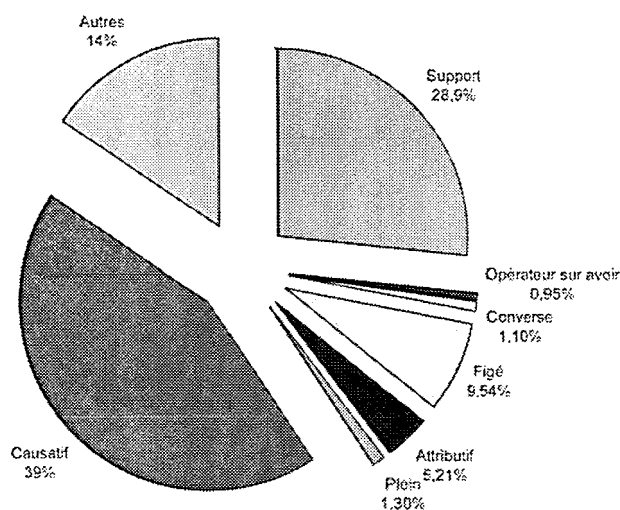
6. faire support converse :

¹Ce classement avait été établi en examinant les 82 emplois du verbe *faire* dans le *Litttré*. Giry-Schneider avait alors remarqué que les emplois proposés pouvaient en fait se réduire à moins d'une dizaine.

Les déficits n'ont jamais fait peur au comité d'établissement du CIB, l'ex-usine Renault de Billancourt.

On a fait rentrer dans les constructions converses des exemples du type *faire peur* dans la mesure où ils n'ajoutent pas d'argument à la structure de base en *avoir* et ne permettent pas la formation du groupe nominal. L'opérateur causatif sur *avoir* ne permet également pas la formation du groupe nominal, mais ajoute, quant à lui, un argument à la structure de base en *avoir*.

Les emplois de causatif et de support sont majoritaires. Les emplois en tant que verbe plein sont peu représentés. La répartition du verbe est illustrée par le diagramme suivant :



Répartition du verbe *faire*

Les 30 premiers noms prédicatifs utilisés avec le verbe *faire* sont les suivants² (classement par ordre alphabétique) :

allusion à	analyse de	arbitrage entre/sur
bilan de	bond	calcul
choix de	communication à	concession à
constat	déclaration à	demande
démonstration de	effort sur/pour	entrée dans
erreur de	évaluation	expérience de
geste envers	offre à	place à
promesse à	proposition à	ravage dans
référence à	tentative	travail
usage de	visite à	voyage

²On a indiqué derrière le nom la préposition qui peut être utilisée avec celui-ci. Ce recensement de l'usage des prépositions n'est bien entendu pas exhaustif : il n'est que le reflet de l'usage qui est effectivement fait en corpus.

Les substantifs *effort* et *proposition* sont les plus utilisés au sein du corpus avec le support *faire*.

7.2.2 Recherches sur le verbe *avoir*

L'étude du verbe *avoir* s'est révélée plus difficile que celle du verbe *faire*. *Avoir* présentait en effet de nombreux emplois qu'il fallait également distinguer de l'emploi de l'auxiliaire. Plus de 12000 occurrences ont été trouvées, dont plus de 91% constituaient des auxiliaires de temps. Parmi les quelques 9% restants, certains emplois ne nous intéressaient pas, puisqu'ils ne faisaient pas intervenir de groupe nominal complément :

Elle s'est faite par un travail de recherche sur le texte et par une réflexion sur la responsabilité de l'acteur, directement engagé par le texte qu'il a à dire, et bien sûr par une relativisation du rôle du metteur en scène, sanctifié durant les années 70".

Au final, nous avons étudié 1047 occurrences du verbe *avoir*. L'observation des résultats de la recherche nous a permis d'effectuer le découpage suivant :

1. avoir verbe plein :

sens de posséder : *"Si j'ai une bagnole, je serai respecté dans le quartier", dit-il.*

sens d'obtenir : *Il a un revenu mensuel de 13 000 f, dont 900 F d'allocations familiales.*

2. avoir verbe figé :

Les citoyens savent que la qualité de la vie quotidienne dépend pour une part importante de la qualité des services publics auxquels ils ont affaire journellement.

3. avoir verbe support :

Accueillant les représentants des cinq, le ministre français des affaires étrangères leur a dit qu'ils détenaient "des responsabilités particulières dans le maintien de la paix" et que "certains ont une influence importante sur les protagonistes" du conflit.

4. avoir support converse :

Il avait la confiance de tout le monde et connaissait personnellement chacun des ouvriers.

5. avoir dans une tournure en *il y a* :

Hormis le Danemark et la RFA - où la situation "doit être examinée en tenant compte des migrations en provenance de la RDA et d'autres pays d'Europe orientale" - il n'y a pas eu de détérioration du marché du travail.

6. avoir verbe attributif :

Le conducteur du véhicule de presse a eu le bras cassé lors de l'accident.

avoir pour : *La faiblesse souvent dénoncée de leurs primes doit, a priori, avoir pour conséquence une grande proximité entre leur traitement indiciaire et le salaire qu'ils reçoivent réellement.*

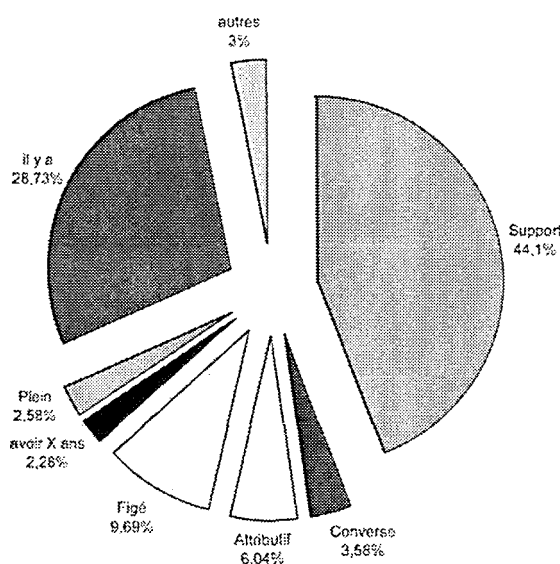
avoir quelque chose de : *Le phénomène n'a rien de surprenant, les grilles de classification établies par les conventions collectives le rendaient même inéluctable.*

7. avoir au sens d'être agé de :

Si l'on s'en tient à la répartition par âge, 57% des actionnaires ont plus de 55 ans alors que les 15-34 ans ne sont seulement que 15% à détenir des titres.

Un même nom pouvait entrer dans plusieurs classes. C'est le cas par exemple de *raison* qui, suivant les cas, pouvait être classé avec le *avoir* figé, ou le *avoir* support, ou bien le cas de *confiance* qui pouvait entrer dans une construction converse ou à support :

- avoir figé : *Je pense qu'il a eu raison.*
- avoir support : *Aucun pouvoir civil n'aime se mettre dans la main des pouvoirs militaire et policier et Mr Gorbatchev a toutes les raisons d'y être particulièrement réticent dans la mesure où il danse en permanence au-dessus d'un précipice dans lequel beaucoup de gens souhaitent le voir tomber.*
- avoir converse : *Il avait la confiance de tout le monde et connaissait personnellement chacun des ouvriers.*
- avoir support : *Vos déclarations les plus récentes montrent bien que vous n'avez plus confiance en Jacques Chirac ?*



Répartition du verbe *avoir* + SN

Les 30 premiers noms prédicatifs dégagés sont les suivants (classement par ordre alphabétique) :

accès à	besoin de	capacité à/de
confiance dans/en	connaissance de	conséquence sur
défaut	effet sur	efficacité
entretien avec sur	envie de	expérience de
idée de	importance	incidence sur
intention de	lien avec	opinion sur
peur de	portée	pouvoir de
problème avec	raison de	recours à
regret	répercussion sur	rôle dans
sentiment de	tendance à	volonté de

Là encore, comme dans le cas du support *faire*, on observe une utilisation préférentielle de certains noms avec leur support *avoir*. Les mots les plus fréquemment utilisés sont les noms *besoin* et *effet*.

7.2.3 Recherches sur le verbe *commettre*

Nous nous sommes également intéressés au verbe *commettre* et avons donc extrait toutes les phrases comportant ce verbe. Avec cette requête, nous pensions préférentiellement obtenir l'emploi du verbe comme support et c'est effectivement ce qui s'est passé.

Nous n'avons relevé aucune occurrence du verbe qui renvoie au fait de *désigner officiellement une personne ou une autorité à laquelle un tribunal confère une mission*. Ce qui peut sembler troublant au sein d'un corpus journalistique. En tous cas, nous ne trouvons pas d'explication qui pourrait justifier cette absence.

Parmi toutes les occurrences relevées, seule une n'est pas le reflet d'un support :

L'irréparable avait été commis.

Les 39 autres emplois du verbe *commettre* sont ceux de support et sont donc tous en relation avec un nom prédicatif :

abus	acte	attentat
coup d'Etat	crime	délit
entorse à	erreur dans	faute
hold-up contre	infraction	irrégularité
massacre	meurtre	parcours
violence		

La plupart de ces noms renvoie à un acte répréhensible ou mauvais ou à une erreur :

- *Les policiers de la Crim qui s'intéressent aussi à des attentats à la boule de pétanque ("une boule de pétanque utilisée comme réceptacle de la matière explosive")*, commis par des immigrés, ont bon espoir d'aboutir.
- L'erreur qu'ont commis tous les partis politiques, *cela a été de laisser Mr Le Pen récupérer un certain nombre de patrie, de nation, d'identité nationale, de défense de la culture française, etc. - qui devraient être partagées et défendues par tout le monde.*

Mais si le nom ne comporte pas sémantiquement cette valeur, il peut alors être modifié :

- *Le suédois a déclaré qu'il était responsable d'autres actes terroristes commis à l'étranger...*

Ceci n'est cependant pas obligatoire :

- *Le drame que je vis, la carence, c'est que la justice ne prend pas cet acte qui a été commis.*

Dans ce dernier cas, l'*acte* a cependant été décrit au cours de l'article.

Parfois, l'emploi de *commettre* peut sembler un peu forcé :

- *Bien que, formellement, le ministère des finances demeure fidèle à ses engagements et n'ait pas eu recours aux bons d'Etat destinés à financer le déficit, il a néanmoins dû commettre une entorse au sacro-saint principe de rigueur budgétaire en émettant à nouveau des bons de financement de la construction.*

L'emploi qui nous semble le plus représentatif de cet emploi forcé est celui de *commettre* en rapport avec *parcours* :

- *Et de façon presque paradoxale, ce sont les plus brillants, ceux qui commettent des parcours presque sans fautes, qui sont le plus souvent montrés du doigt..*

Alors qu'un "parcours presque sans fautes" véhicule une notion positive, il est tout de même employé avec "commettre". Mais c'est justement ce type de parcours qui suscite de la part de certains jalousie et médisance. Ainsi, la notion positive se transforme-t-elle en notion négative... On peut penser qu'un emploi moins marqué aurait eu recours au verbe support *faire*.

La tournure passive est le plus souvent utilisée (62% des entrées). Le pourcentage restant étant réservé à la tournure active (équivalent d'une classe sans redistribution en dimension 2). Aucune construction moyenne n'est trouvée, ni même d'impersonnel ou de causatif. La prédisposition de *commettre* a une utilisation passive s'explique facilement. L'auteur d'une action répréhensible n'est en effet pas toujours immédiatement connu. C'est le cas par exemple d'un attentant dont la revendication peut tarder, ou d'un crime dont le coupable n'est en général identifié que plusieurs mois après une enquête.

Nous avons ensuite cherché à savoir si les noms prädicatifs que nous avons relevés pouvaient s'employer avec un autre support. Malheureusement, notre recherche s'est révélée peu intéressante. Seul le mot *erreur* a pu être retrouvé avec un autre support : *faire*. Nous pensons que le corpus était trop peu fourni et qu'un corpus plus conséquent nous aurait permis d'obtenir

de meilleurs résultats. En particulier, nous nous attendions à trouver des supports en rapport avec *massacre* ou *attentat* comme *perpétrer* ou *accomplir*.

7.2.4 Recherches sur le verbe *procéder*

Jusqu'à présent, nous avons examiné des verbes plaçant le nom prédicatif en position d'objet direct. Nous nous sommes donc également attaché à l'étude du verbe *procéder*. Ce dernier est peu présent au sein du corpus (53 occurrences). Ses emplois sont majoritairement ceux de support (49 occurrences). Néanmoins, un emploi comme intransitif est également relevable :

1. *procéder* intransitif (au sens d'agir) :

Dans la droite ligne de la position qui avait été adoptée par le SPD Ouest-allemand, lors de son congrès de décembre à Berlin-Ouest, ce texte n'en souligne pas moins d'abord qu'il s'agit de procéder par étapes, ensuite que "toutes ces étapes doivent être intégrées dans le processus d'unification européenne".

2. *procéder* à support :

Mais, comme l'ensemble de ses concurrents, le groupe procède à des désinvestissements non stratégiques pour compenser les manques à gagner.

Aucun des emplois de *procéder* ne fait apparaître d'impersonnel passif. Ce type de passif est réputé difficile à observer, voire impossible³.

La majeure partie des compléments prépositionnels du nom prédicatif sont introduits par *de* (33 occurrences). Les seuls autres exemples du corpus faisant intervenir d'autres prépositions (1 fois *dans* et 2 fois *avec*) sont les suivants :

Car la capitale fédérale a toujours payé son tribut aux gens de robe, confortablement installés au congrès (...), d'où ils procèdent parfois à quelque incursion dans le monde des affaires avant de regagner le Capitol Hill.

Nous avons alors pris le parti de procéder à des négociations bilatérales avec RVI.

Procéder est le seul verbe support que nous ayons observé en rapport avec des noms prédicatifs coordonnés :

De nombreux pays de l'Europe de l'Est n'ont pas mis en œuvre des réformes et n'ont pas procédé à une restructuration et à des changements.

³Une recherche sur internet par *google* portant sur la séquence "il a été procédé à" fait apparaître que ce type de passif est tout de même envisageable. Néanmoins, l'absence du complément d'agent est régulière. Les exemples avec complément d'agent sont beaucoup plus marginaux, mais les textes de type juridique peuvent y avoir recours :

Il a été procédé à ces enregistrements le 1er juillet 2000 par Rémi D.

Les noms prédicatifs qui sélectionnent le support *procéder à* sont les suivants (classement par ordre alphabétique) :

ablation de	acquisition	ajustement de
appel d'offres	audit	augmentation
changement	classement de	comparaison
consultation	dépôt de	désinvestissement
échange de	étude	examen de
gommage de	hausse de	incursion dans
licenciement de	liquidation de	mariage de
modification de	montage de	négociation avec
opération	ouverture de	perestroïka
recrutement	réduction de	rééquipement
réexamen de	refonte de	relance
renouvellement de	réorganisation	restructuration
revalorisation de	sélection	tirage
travaux	visite de	

De cette liste, on observera que des noms étrangers, comme *perestroïka*, peuvent également être considérés comme prédicatifs :

Dès 1986, Mr Mikhaïl Gorbatchev avait pressé les communistes vietnamiens de procéder à leur propre perestroïka.

7.2.5 Recherches sur le verbe *courir*

Le dernier verbe que nous ayons examiné est le verbe *courir*. Il est peu présent au sein du corpus (40 occurrences) mais a des emplois très diversifiés :

1. *courir* intransitif :

sans complément : *Dans la chaleur des moteurs courent des silhouettes chargées de sacs qui croisent des motos et des autos, dont personne ne peut affirmer si elles sont, ou non, encore inscrites en course.*

avec complément locatif : *Pour compléter cette impression d'emprisonnement, des canalisations courent au fond de la mer, prêtes à amener le futur pétrole des plates-formes avoisinantes dans les cales du bateau avant que d'autres pétroliers ne viennent plus tard en prendre livraison.*

2. *courir à* :

Cependant, à recoller des morceaux de-ci de-là, sans vision globale, sans perception réelle de la vie quotidienne des acteurs du système éducatif, sans projets d'avenir, on court à une crise sans précédent.

3. courir + infinitive :

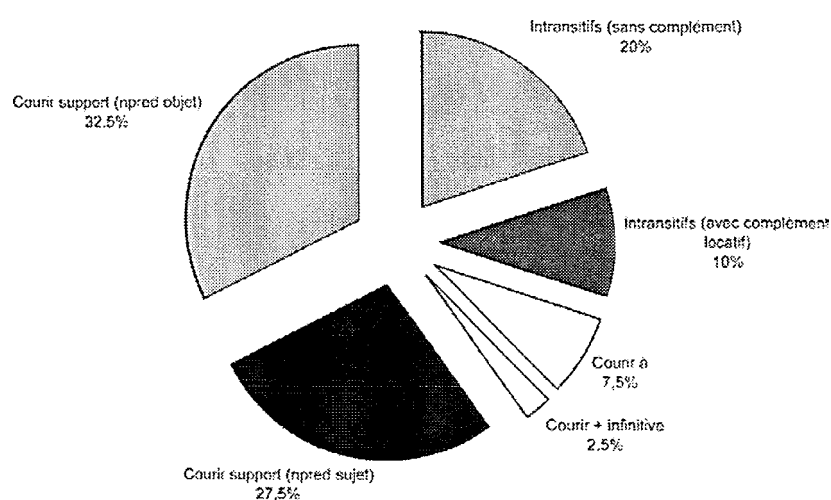
Quand mes gosses s'entaillaient le doigt ou le genou bien profond, je les plan-tais là immédiatement et je courais m' enfermer dans les toilettes en hurlant...

4. courir support :

nom prédicatif sujet : Une rumeur court, *ces temps-ci : le gouvernement italien chercherait par tous les moyens à écarter "Ripa" de Bruxelles, où il est jugé encombrant.*

nom prédicatif objet : L'adjudant court le risque d'être dégradé, *rattrapé puis dépassé par l' évolution technique, mais il peut tout aussi bien être promu lieutenant pour répondre aux besoins d'une armée qui manque de cadres de terrain.*

A eux seuls, les emplois de courir comme intransitif ou support représentent plus de 70% des entrées. Les autres emplois (et en particulier celui de courir+infinitif) sont beaucoup plus marginaux. *Rumeur* et *Bruit* sont les deux seuls noms prédicatifs sujets utilisés avec le support (l'emploi de *rumeur* est majoritaire avec 80% des entrées). *Risque* est le seul nom prédicatif objet. La répartition est la suivante :



Répartition du verbe *courir*

De cette étude, on peut affirmer que l'emploi d'un verbe comme support, même s'il n'est pas toujours majoritaire est très important. En particulier, il est bien souvent supérieur à celui de son équivalent plein. Bien entendu, il conviendrait d'examiner d'autres verbes pour déterminer si cette constatation est globalement généralisable. Néanmoins, l'étude de verbe comme *faire* ou *avoir* constitue déjà un bon indice, car ces verbes présentent des emplois variés en nombre relativement important.

7.3 Extraction de noms prédicatifs

Nos précédentes recherches nous ont permis de relever l'emploi de certains noms prédicatifs en rapport avec un verbe support particulier. Nous voudrions maintenant déterminer quels sont les noms prédicatifs les plus fréquemment utilisés au sein du corpus.

Pour cela, nous allons d'une part rechercher des groupes nominaux relevant d'une configuration particulière, et d'autre part observer les noms les plus fréquents du corpus.

7.3.1 Les noms à 2 compléments

Aussi allons-nous dans un premier temps nous intéresser aux groupes prépositionnels que les groupes nominaux contiennent. Il ne nous semble pas opportun de relever les groupes nominaux ne contenant qu'un groupe prépositionnel, puisque leurs têtes peuvent être des noms concrets. Si nous voulons avoir le plus de chance d'observer des noms prédicatifs, il nous faut récupérer les groupes nominaux qui contiennent au moins deux groupes prépositionnels.

Cette requête nous permet d'obtenir 1798 emplois, reflets potentiels de groupes nominaux prédicatifs. Un classement sur les noms par fréquence fait néanmoins vite ressortir que certains groupes doivent être éliminés. C'est le cas des groupes qui ont pour tête un nom ayant rapport à des chiffres comme millions, milliards ou %.

Certains groupes ne contiennent en fait qu'un seul groupe prépositionnel complément, les autres étant des modificateurs :

- l'odeur de l'eau de Javel dans les couloirs
- un dépôt de sperme à la suite d'un cancer du testicule
- la réunion du comité central de décembre

De tels groupes sont très fréquents et représentent près de 40% des occurrences dégagées.

D'autres groupes sont mal représentés, et devraient, de fait, contenir des groupes prépositionnels enchassés :

- l'autoconservation par congélation de sperme
- Le courrier des lecteurs de l'hebdomadaire Révolution

Enfin, certains compléments sont inclus dans le groupe nominal, alors qu'ils devraient être considérés comme des compléments du verbe :

Fournissant une nouvelle arme aux constructeurs pour le maintien de leur monopole de distribution, le texte reconnaît toutefois une profession en plein essor.

Si l'on élimine ces emplois déviants, il reste alors 968 groupes nominaux. Leurs 90 premiers représentants les plus fréquents sont les suivants (classement par ordre alphabétique) :

abandon par de / de à	accès de à	accord sur avec
accroissement de de	achat par de	acquisition par de
aide de à	annonce par de / de sur	arrivée de sur
augmentation de de	besoin de de	capacité de à
cession de à	conflit de avec	conseil de à
construction par de	création par de	départ de de
dépôt par de / de à	désignation de à	dévaluation par de
développement de en	échange de avec	effort de sur
émergence de dans	envoi par de / de dans / à de	entrée de dans
exonération par de	extension de sur	fuite de dans
importance de sur	importation de à	indexation de sur
instauration par de	intervention de dans	introduction de à
lancement par de	maintien de dans	manque de de
mise de sur	nomination de à	ouverture de à
partage de en	participation de à	passage de à
perte de sur	place de dans	présence de dans
présentation par de / de à	prévision de sur	prise par de
production par de	progression de dans	projet de pour
promotion par de	proposition de sur	publication par de
rachat par de à	rapport de sur	recours de à
recrutement par de	recul de de	redistribution de à
réduction de sur	réévaluation par de	regroupement de en / de autour
relation de avec	relèvement par de	remise par de / de à
rencontre de avec	renouvellement par de	répartition de dans
répression de contre	reprise par de	responsabilité de dans
retour de à	retrait de de	signature par de
signe de à	sortie de de	soutien par de / de à
suppression de de	transfert de à	transformation de en
utilisation par de	vente par de / de à	venue de à
versement de à	visite de à	voyage de en

L'observation de ces noms suggère que les noms à interprétation passive se trouvent fréquemment en corpus. C'est en fait le classement alphabétique qui donne cette impression. Les groupes nominaux ayant des compléments en *par* ont en effet peu d'occurrences (de l'ordre de 3 ou 2 par nom prédicatif), alors que les groupes nominaux à interprétation active sont beaucoup plus nombreux.

Bien souvent, l'apparition de tous les compléments n'est pas simultanée :

Un refus de la part des ministres concernés et des "eurocrates de Bruxelles" revient

pour cette organisation (...) à "une stratégie de renoncement dans la finalité et d'abandon aux Etats-Unis du contrôle de l'arme alimentaire".

Mais il reste possible d'observer des groupes nominaux comportant 3 compléments :

Le rachat par Air France à la société Chargeurs SA d'UTA (qui lui donne également le contrôle d'Air Inter) dote la France d'un groupe aérien unique, à l'instar de nombreux pays européens.

En outre, la répartition des compléments prépositionnels laisse apparaître des préférences de placement⁴. On observe ainsi qu'en général, le premier complément prépositionnel est introduit par *de*. Les groupes nominaux comportant deux compléments en *de* sont également très bien représentés. Il est plus rare d'observer des groupes nominaux sans complément prépositionnel en *de*, mais cela ne demeure pas impossible.

Placement des compléments prépositionnels	Pourcentage
$SP_{de} SP_{de}$	19,3 %
$SP_{de} SP_{à}$	23,6 %
$SP_{à} SP_{de}$	9,8 %
$SP_{de} SP_{prep}$	33,7 %
$SP_{prep} SP_{de}$	12,5 %
$SP_{prep} SP_{prep}$	0,6 %
Autres placements (>2 compléments)	0,5 %

Une étude plus spécifique des données, confirme le fait que les emplois de groupes nominaux dits "passifs" sont très minoritaires. Parmi les 968 groupes nominaux relevés, seuls 92 font apparaître un complément prépositionnel en *par*⁵.

Deux emplois nous ont frappé, car ils paraissent peu conventionnels :

1. *Elles avaient pour origine* des propos au Nouvel Observateur par M. Abel Farnoux.
2. *Le gouvernement a réaffirmé son ouverture (...) allant dans le sens (...) d'une meilleure maîtrise de leur destin* par les Corses eux-mêmes.

⁴On n'a retenu ici que 3 types de compléments prépositionnels : ceux introduits par *de*, ceux introduits par *à* et ceux introduits par une préposition différente de *de* ou *à*. On n'a pas cherché à repérer les compléments prépositionnels à valeur locative, ni les compléments en *par*.

⁵Bien entendu, il peut subsister d'autres groupes nominaux "passifs", les compléments prépositionnels n'étant pas systématiquement présents

Le premier emploi est à notre sens plutôt déviant, même si facilement compréhensible. Le second illustre un emploi passif du nom prédicatif *maîtrise*, pourtant généralement exclu. L'ensemble des noms prédicatifs dégagés est le suivant (classement par ordre alphabétique - les noms prédicatifs déjà présents dans la liste précédente sont en italique) :

<i>abandon</i> par de	absorption par de	<i>achat</i> par de
<i>acquisition</i> par de	administration par de	adoption par de
<i>annonce</i> par de	annulation par de	approbation par de
blocage par de	compensation par de	condamnation par de
consultation par de	couverture par de	<i>création</i> par de
découverte par de	<i>dépôt</i> par de	élaboration par de
entreprise	<i>envoi</i> par de	évacuation par de
fondation par de	gestion par de	illustration par de
indemnisation par de	livraison par de	maîtrise par de
mise en place par de	organisation par de	perception par de
poursuite par de	pénétration par de	<i>présentation</i> par de
prise en charge par de	propos par à	<i>rachat</i> par de à
réalisation par de	recherche par de	règlement par de
régulation par de	rejet par de	remise à plat par de
remise en cause par de	renforcement par de	<i>reprise</i> par de
<i>signature</i> par de	<i>utilisation</i> par de	<i>vente</i> par de

Là encore, on observe un placement préférentiel du complément prépositionnel en *de* en première position (60 % des entrées des groupes nominaux passifs).

7.3.2 Autres noms prédicatifs

Cette première recherche nous a d'ores et déjà permis de relever un nombre important de noms prédicatifs. Elle présente cependant un défaut, puisqu'elle ne prend en compte que les nominaux ayant 2 compléments. Une seconde recherche, cette fois plus générale, nous permettra de pallier ce problème. Nous allons examiner les 200 premiers noms les plus fréquents du corpus de manière à observer si les noms prédicatifs y sont bien représentés.

D'emblée, les noms les plus fréquents semblent très souvent liés au domaine de l'économie. On retrouve ainsi des mots comme *franc*, *milliard*, *million*, *prix*, *dollar* ou *économie*... Les jours de la semaine, et les mois de l'année sont également très présents, puisque tous font partie de la liste des 200 noms.

Reste que cette liste contient de nombreux noms prédicatifs. Bien entendu, hors de tout contexte, il est parfois difficile de déterminer si un nom est vraiment prédicatif. Comment différencier par

exemple un *effet* personnel, d'un *effet* sur une substance ou une personne, le *droit*, discipline, du *droit* sur quelque chose ?

Les 72 noms dégagés sont fournis dans la liste ci-dessous. Les noms en gras désignent les noms déjà mis en valeur par la recherche portant sur les groupes nominaux comportant 2 compléments. Bien entendu, cela ne signifie pas que les noms restants sont forcément des noms n'ayant qu'un seul complément :

accord	action	activité
affaire	aide	augmentation
avenir	baisse	besoin
condition	conseil	conséquence
consommation	contrat	contrôle
création	croissance	débat
décision	demande	développement
difficulté	droit	effet
effort	élection	émission
évolution	exportation	fin
formation	gestion	hausse
idée	image	intérêt
investissement	mesure	mouvement
négociation	objectif	opération
opposition	organisation	participation
perte	place	plan
position	pouvoir	problème
production	programme	progression
projet	proposition	question
raison	rapport	recherche
réduction	réforme	reprise
résultat	réunion	risque
rôle	service	situation
succès	travail	vente

Au terme de ces sections, nos diverses expérimentations ont donc confirmé la présence importante de verbes supports dans un corpus de type journalistique. Quant à l'extraction de noms prédicatifs, elle a été facilitée par l'encodage du corpus qui nous a donc évité des dépouillements longs et fastidieux.

Bien entendu, les résultats présentés ne sauraient prétendre à une quelconque représentativité. En effet, rien ne permet d'affirmer que le corpus soit de taille suffisante, ou que le style journalistique n'ait pas d'influence sur les résultats. Néanmoins ces recherches confirment l'intuition que l'on se fait de l'importance des verbes supports et des noms prédicatifs. Il conviendra de confirmer ou affiner nos constatations en effectuant nos recherches sur la totalité du corpus.

7.4 Expérimentations avec le parseur d'A. Sarkar

Pour finir, nous allons expérimenter la grammaire que nous avons décrite avec la Métagrammaire. Nous ne parlons pas d'évaluation dans la mesure où nous ne disposons pas d'un corpus de phrases tests suffisant. Ce qui nous importe ici, c'est de savoir que des phrases (produites ou non par nos soins) pourront être analysées correctement. En particulier, parmi les analyses obtenues, nous allons vérifier que l'analyse pertinente est bien produite par l'analyseur. Nous allons effectuer nos analyses en deux temps. Tout d'abord, nous allons examiner quelques phrases tests, au vocabulaire peu important, mais suffisamment représentatives de notre grammaire. Ensuite, nous tâcherons de reprendre et d'étendre les analyses que [Werli, 2001] avait obtenues, lors des phases précédentes d'évaluation de la grammaire.

Pour expérimenter notre grammaire, nous allons utiliser le parseur développé par [Sarkar, 2000]. Ce parseur, implémenté en ANSI C, utilise une variante de l'algorithme de [Van Noord, 1994], de type ascendant et head-corner. Les fichiers, initialement développés au format XTAG, sont convertis automatiquement pour répondre au format du nouveau parseur. Notons que nous n'effectuerons pas d'évaluation en termes d'efficacité de traitement. L'objet de notre expérimentation n'est en effet pas l'analyseur en lui-même, mais la grammaire TAG du français.

7.4.1 Analyses sur des phrases test

Nous avons pensé en premier lieu utiliser le corpus de phrases tests TSNLP. Malheureusement, ce dernier ne référençait pas explicitement de phrases à verbe support. Il nous a donc fallu observer les phrases grammaticales et déterminer quelles étaient celles qui faisaient apparaître des CVS. Peu sont pertinentes pour notre propos, car peu permettent de tester uniquement le phénomène qui nous intéresse. Au final, nous avons retenu trois phrases, qui ne font malheureusement pas apparaître de complément prépositionnel. Ces phrases ont été analysées avec succès et nous avons obtenu la bonne analyse. Le nombre d'analyses est indiqué entre parenthèses :

- L'ingénieur pose une question (2)
- La question est posée (2)
- Une question est posée par l'ingénieur (5)

Nous avons ensuite continué nos analyses sur des phrases fabriquées par nos soins. Là encore, tout comme pour TSNLP, le vocabulaire utilisé est succinct, mais renseigné dans les fichiers des lexiques. Ainsi, l'analyseur n'a pas à utiliser le fichier de défauts qui assigne à un nom inconnu un arbre ou une famille particulière. Il s'agit donc là d'un environnement idéal pour tester la pertinence de notre grammaire.

Les phrases que nous avons retenues sont les suivantes (le nombre d'analyses obtenues est renseigné entre parenthèses) :

- L'ingénieur commet une agression contre la standardiste (2)
- Ce sont des baisers que l'ingénieur donne à la standardiste (24)
- Une gaffe est faite par la standardiste (4)
- Il est commis un crime contre l'ingénieur (3)
- Le photographe prend une photo de ces enfants (7)

Toutes ces phrases font intervenir des constructions à verbe support. La dernière est ambiguë, entre une interprétation à verbe plein (Le photographe prend un objet à quelqu'un) et une interprétation à verbe support (Le photographe réalise une photo de quelqu'un).

Les résultats obtenus sont très satisfaisants : toutes les phrases sont correctement analysées. Les phrases utilisant le support *commettre* ne font appel qu'aux familles à nom prédicatif (en effet, le lexique ne renseigne pas le verbe *commettre* en tant que verbe plein). Les phrases renfermant les supports *donner*, *faire* et *prendre* peuvent également être analysées avec les familles des verbes pleins. C'est ce qui explique des résultats plus importants (en termes de nombre d'analyses).

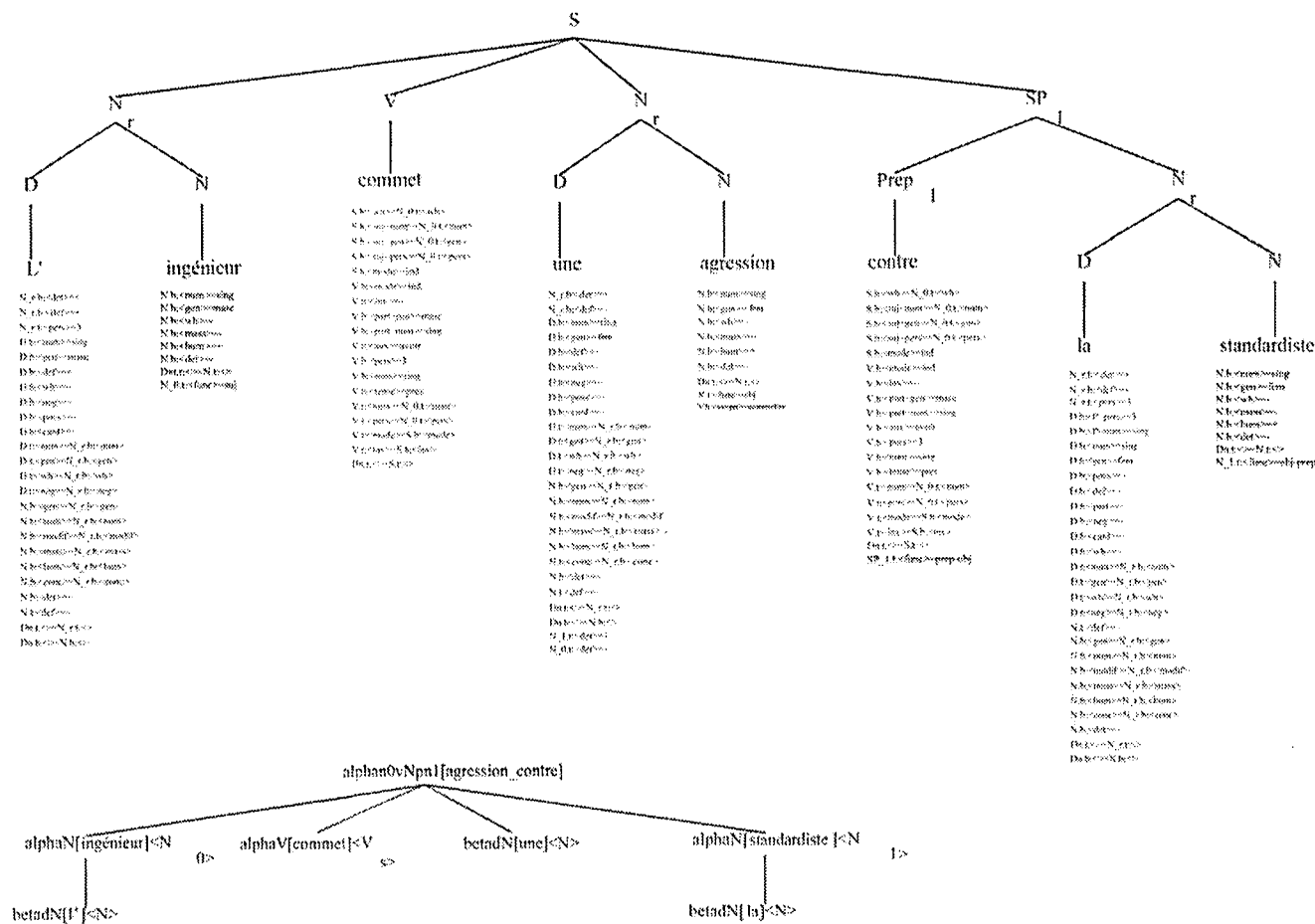
Nous fournissons maintenant les arbres dérivés et de dérivation de chaque phrase⁶. Ces arbres ont été dessinés par le parseur en postscript⁷. Les traits sont indiqués sous chaque nœud⁸

Nous tâcherons d'analyser les raisons de la surgénération des analyses dans la sous-section 7.4.3

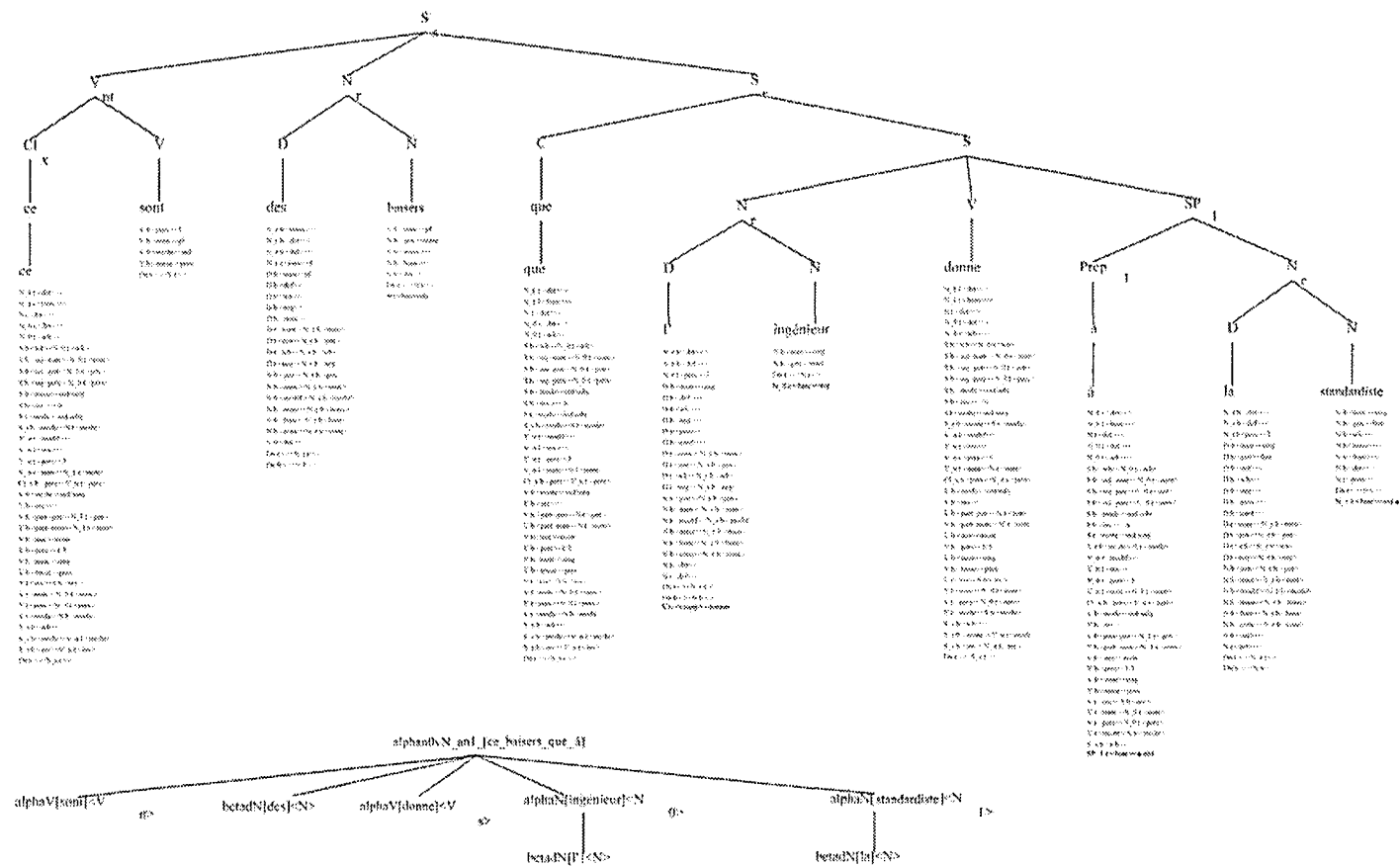
⁶On notera un problème d'affichage des co-têtes dans les arbres élémentaires. Ces co-têtes sont automatiquement dupliquées. Ce problème n'est pas dû à l'extension de la grammaire. La grammaire des verbes pleins se comporte de la même façon avec les co-têtes de ses arbres élémentaires.

⁷Ces exemples sont les sorties effectives du parseur. Seule la présentation des arbres dérivé et de dérivation varie ici. Nous les avons regroupés sur une même page afin de faciliter la lecture.

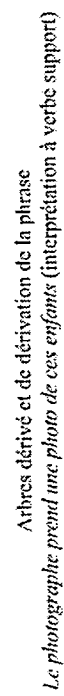
⁸Cette mise en page a été mise en place par Stéphanie Werli. Elle évite que les traits ne soient affichés comme une simple collection les uns à la suite des autres. Cependant, même si les traits sont indiqués sous un nœud, cela ne signifie pas pour autant qu'il concerne effectivement ce nœud. La lecture est néanmoins facilitée. Nous avons conscience que les traits ne sont pas toujours très lisibles. Ils donnent néanmoins une idée du codage effectif de la grammaire et permettent d'entrevoir les éléments que le parseur a à manier.

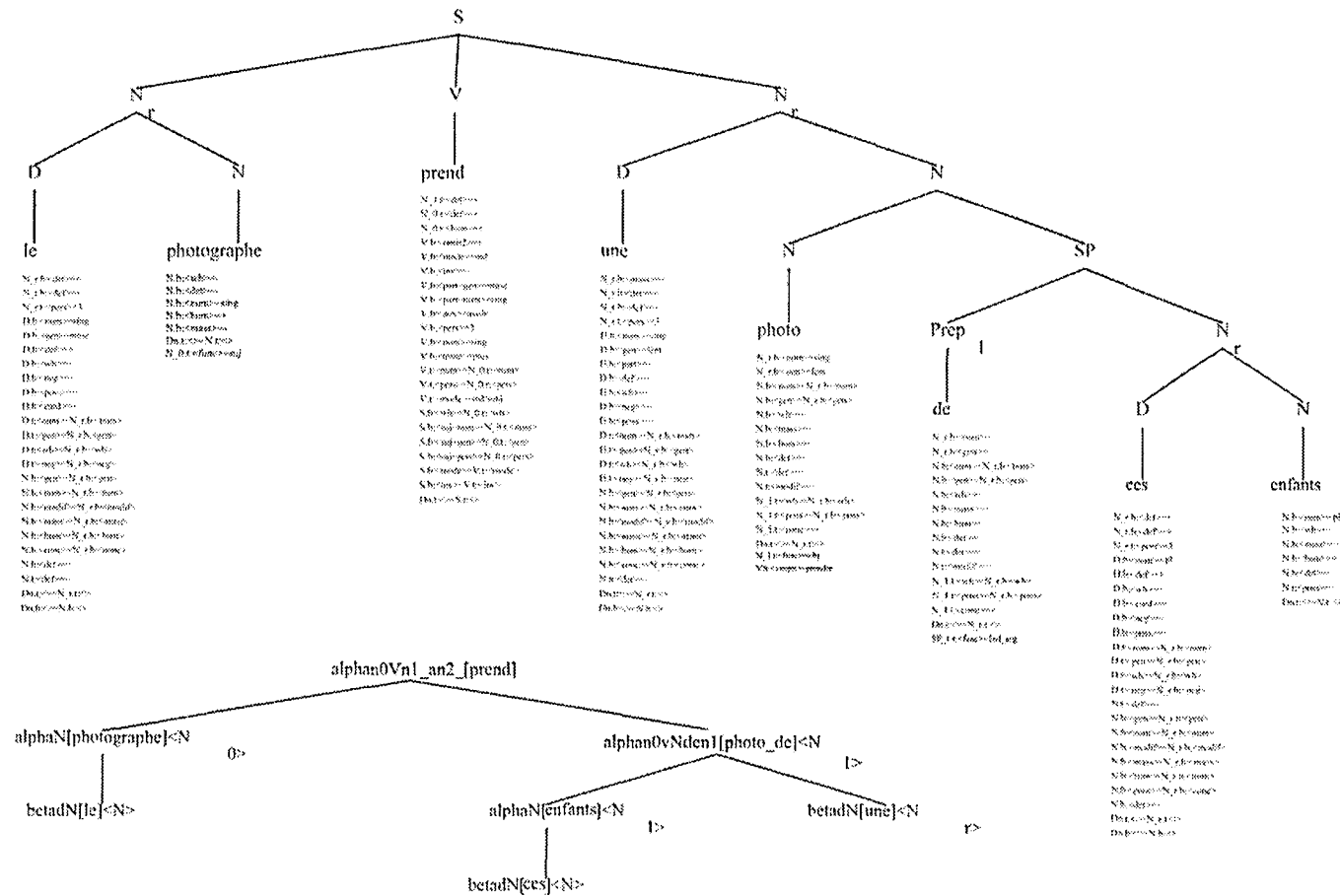


Arbres dérivé et de dérivation de la phrase
L'ingénieur commet une agression contre la standardiste



Arbres dérivé et de dérivation de la phrase
Ce sont des bakers que l'ingénieur donne à la standardiste





Arbres dérivé et de dérivation de la phrase
Le photographe prend une photo de ces enfants (interprétation à verbe plein)

7.4.2 Analyses sur phrases de corpus

L'analyse sur phrases de corpus s'avère quant à elle plus délicate. Nous avons tout d'abord observé les phrases utilisées par S. Werli lors de son évaluation afin de déterminer lesquelles faisaient apparaître des constructions à verbe support. Ces phrases comportaient moins de 15 mots, et étaient toutes extraites du corpus *Le Monde*. Sur plus de 300 phrases, seule une vingtaine fait apparaître des CVS. Seule une partie de ces phrases avait été analysée (phrases 3, 4, 5 et 9). D'autres n'avaient reçu aucune analyse (phrases 12 et 13). Enfin, aucune mention n'avait été faite pour les trois quarts restants, qui ont donc été vraisemblablement écartés. Les analyses proposées étaient forcément erronées puisqu'aucune famille à nom prédicatif n'avait été décrite à l'époque. Le verbe support était donc considéré arbitrairement comme un verbe plein. Le nom prédicatif était, quant à lui, considéré comme un argument du verbe.

Nous avons donc réanalysé ces phrases, en effectuant tout d'abord un prétraitement (désamalgames des articles et des prépositions, ensemble Nom propre/ Nom commun considéré comme un seul mot, suppression de la ponctuation...). Nous n'avons pas vérifié que la grammaire comportait dans son lexique l'ensemble des mots de chaque phrase⁹, mais avons renseigné chaque couple nom prédicatif / verbe support. Le nombre d'analyses obtenues est indiqué entre parenthèses :

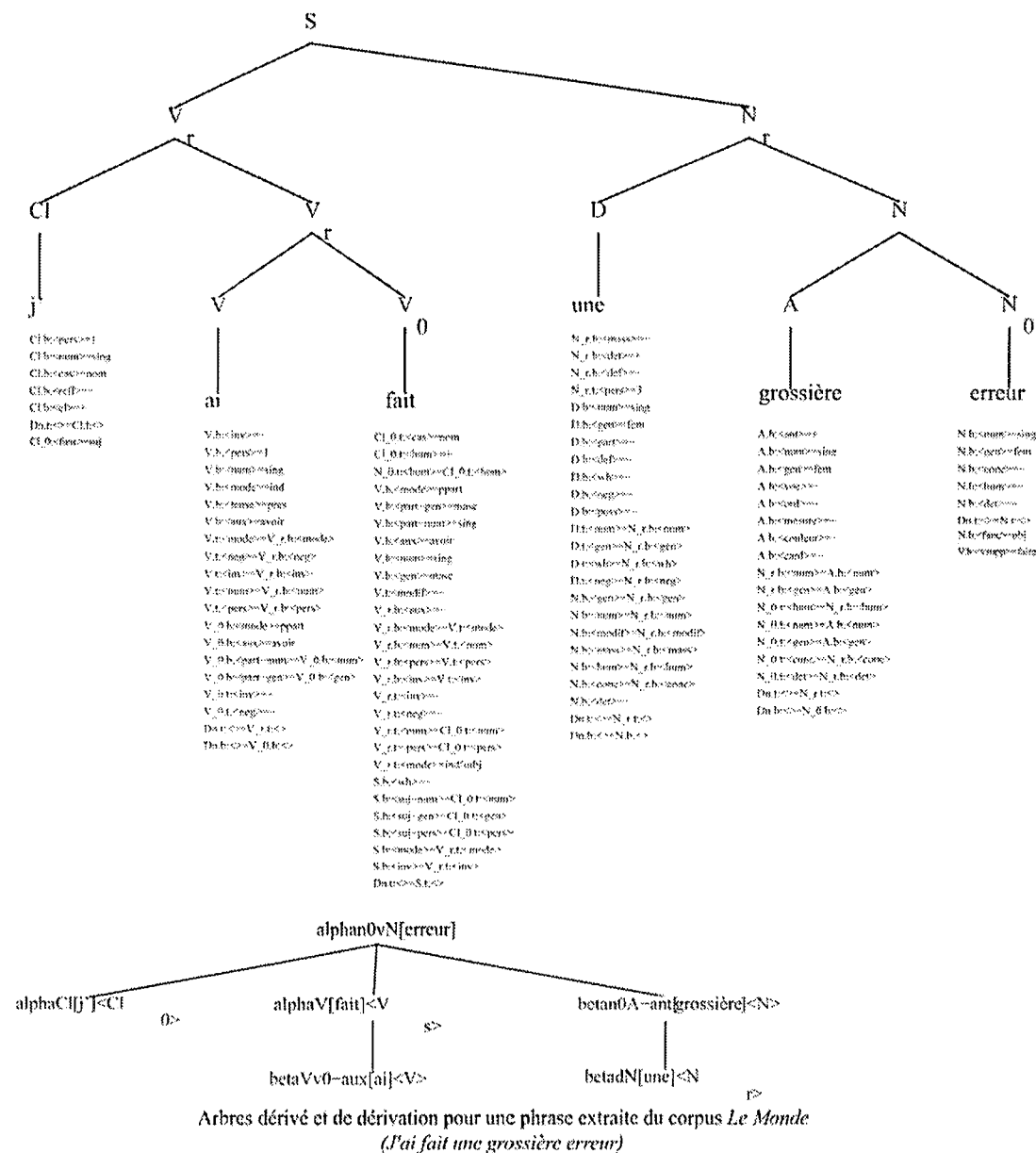
1. Une initiative diplomatique en vue d'apaiser les esprits a été prise le 13 Janvier à Lomé (0)
2. Le contrat d'importation des ordures helvétiques a été conclu en bonne et due forme (0)
3. Certains se sont posé la question (1)
4. Il a porté des coups, un peu partout, sur son corps (12)
5. J'ai fait une grossière erreur (3)
6. Nous en avons la volonté (3)
7. D'autres entreprises n'ont pas cette souplesse (0)
8. Mais ces développements ont des incidences limitées sur le chiffre d'affaires (0)
9. Il n'avait aucun préjugé défavorable envers l'armée (12)
10. Les études qu'il propose ont été mises en œuvre par Robert Badinter (pb)

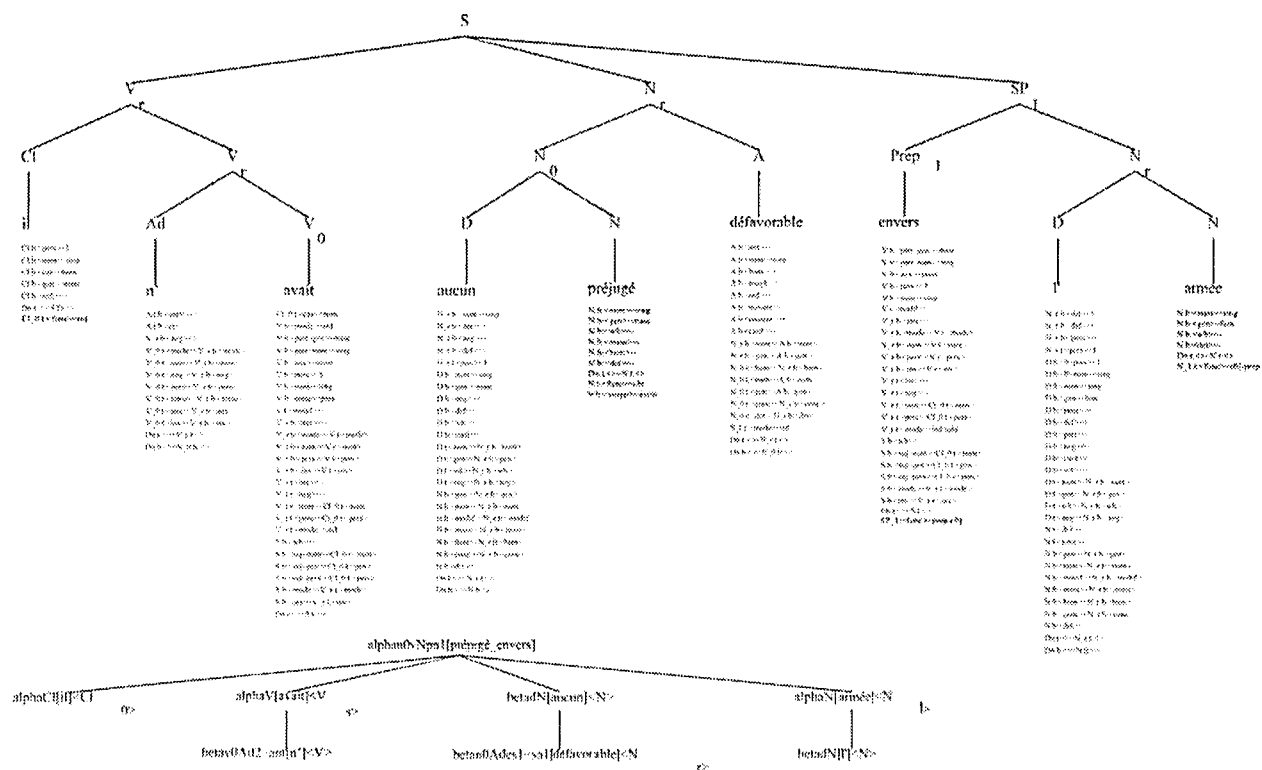
⁹Le parseur dispose d'un système de défaut qui attribue à un mot inconnu une famille ou un arbre.

11. Sauf que, dans l'original, le professeur avait un cancer (5)
12. Il m'a fallu en trouver qui aient une certaine innocence (pb)
13. Je suis persuadé que le choix est déjà fait, prétend-il (pb)
14. Lui aussi il a ses chances (0)
15. Le festival a fait faillite, personne n'a été payé (pb)
16. On n'en recense qu'une dizaine dont aucune n'a pris de dimension judiciaire (21)
17. Il avait de graves ennuis de santé (4)
18. MM Pasqua et Seguin pourraient également effectuer quelques déplacements en province (0)
19. Aujourd'hui, dit-elle, nous avons besoin de perspectives (0)
20. Depuis la fin du mois de novembre, des contacts avaient été noués (17)

Les résultats que nous avons, pour notre part, obtenus sont moins satisfaisants que précédemment. Certaines phrases ne reçoivent aucune analyse¹⁰. D'autres, sont analysées mais ne fournissent pas la bonne analyse. Pour le tiers restant, des analyses correctes sont fournies parmi d'autres. Notons au passage que 4 phrases (les phrases 10, 12, 13 et 15) posent problème au parseur qui s'arrête brusquement en renvoyant une erreur d'exécution. Ces résultats permettent tout de même de confirmer que l'extension de la grammaire permet de prendre en compte des phrases qui jusqu'alors étaient rejetées.

¹⁰Deux cas se présentent : soit le parseur indique qu'il n'a pas trouvé d'analyse valable, soit le parseur "ne répond plus" (on a considéré qu'aucune analyse n'était valable après 10 minutes d'attente sur un Pentium IV, 2,4 GHz)



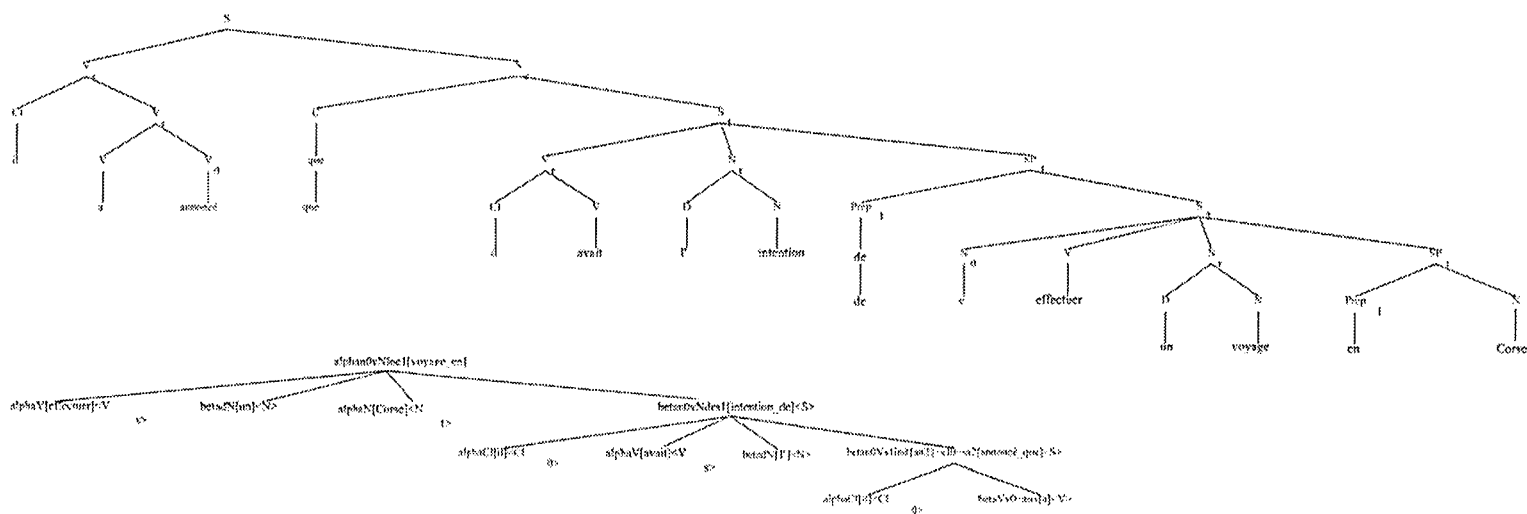


Arbres dérivé et de dérivation pour une phrase extraite du corpus *Le Monde*
(Il n'avait aucun préjugé défavorable envers l'armée)

Dans le but d'étendre le corpus de test et de consolider les résultats obtenus, nous avons ajouté à ce corpus d'autres phrases. Tout comme précédemment, nous avons choisi des phrases assez courtes, sans nous soucier des constructions syntaxiques qui les composent. Ces phrases ont en grande partie été extraites des données qui nous ont servi de base pour effectuer les statistiques concernant la répartition d'un verbe supposé support. On ne s'étonnera donc pas ici que le verbe support varie peu. Les phrases retenues sont les suivantes :

- 21 Et encore, j'ai de la veine, je suis pas trop atteinte (pb)
- 22 Mais George Borrow n'a peur de rien (pb)
- 23 Il a annoncé qu'il a l'intention d'effectuer un voyage en Corse (6)
- 24 Cette conférence avait deux objectifs principaux (4)
- 25 Séguin a eu le coup de foudre pour Pasqua (4)
- 26 La France a un problème avec son logement (2)
- 27 Le choix a vite été fait (2)
- 28 Nous faisons des efforts pour résorber son déséquilibre commercial (32)
- 29 Les autres grands constructeurs ont fait la même proposition (4)
- 30 Les déréglementations seront faites avec toute prudence usuelle (1)
- 31 Car il paraît que ça fait des ravages chez les gamins (pb)
- 32 Avant leur week-end à Rome, les dirigeants allemands font escale à Paris (17)
- 33 Pour y faire face, le gouvernement compte faire largement appel aux collectivités locales (pb)
- 34 Elle ne sera pas condamnée lorsque l'infraction aura été commise par un employé (16)
- 35 L'URSS a déjà procédé à une réduction de son aide et en annonce de nouvelles (28)
- 36 Nous procédons alors à une pré-sélection (2)
- 37 Voilà le rêve ambigu caressé par la Fédération (0)
- 38 Mr Sautier n'a cependant pas perdu tout espoir (pb)
- 39 A chargeurs, on confirmait que des négociations avec Air France étaient menées (50)
- 40 La première attaque est menée par la Royal Air Force (1)

On remarquera encore ici que les résultats ne sont pas aussi bons que sur des phrases fabriquées par nos soins. Néanmoins, on observera également que des phrases à la structure relativement complexe, comme la phrase (23) peuvent être analysées correctement. On trouvera sur la page suivante les arbres dérivé et de dérivation de cette phrase.



Arbres dérivé et de dérivation pour une phrase extraite du corpus *Le Monde*
(Il a annoncé qu'il avait l'intention d'effectuer un voyage en Corse)

7.4.3 Raisons des rejets ou des suranalyses

Il est bien sûr difficile d'expliquer totalement les raisons de rejet ou de suranalyse de toutes ces phrases. D'une part, la grammaire, en constante évolution peut contenir des erreurs de cohérence au niveau des traits, des arbres ou du lexique (mot mal orthographié, mot associé à une mauvaise famille). D'autre part, le parseur peut lui aussi être source d'erreurs.

Néanmoins, nous relevons de nos expérimentations, que les phrases fabriquées qui ne contiennent ni modifieur, ni incise, sont mieux analysées. Une analyse sur corpus taggé aurait certainement permis d'obtenir de meilleurs résultats et aurait par exemple évité les homonymies (par exemple, "donne" peut être un verbe plein, un verbe support, ou un nom). Malheureusement, il semble que cette démarche soit infructueuse dans l'état actuel du parseur. Tout comme S. Werli, nous avons relevé que les analyses comportant des arbres à co-têtes sont systématiquement rejetées dans ce cas...

Nous pouvons cependant apporter quelques raisons pour les suranalyses. Ce sont principalement les groupes prépositionnels qui posent problème :

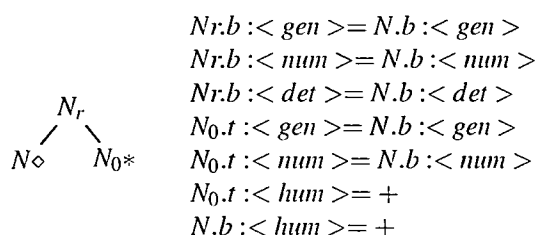
- Ils peuvent en effet être considérés comme des compléments ou des modifieurs.
- Lorsqu'un complément prépositionnel n'est pas présent, il n'est également pas possible de déterminer quelle famille doit être utilisée. Ainsi, par exemple, en plus de son emploi de support, prendre est aussi concerné par les familles à verbe plein $n0Vn1(an2)$ et $n0vn1(loc2)$. En l'absence du complément prépositionnel, il n'y a aucun moyen de déterminer quelle est la famille pertinente.
- Dans le même ordre d'idée, on peut également signaler que les homonymies peuvent être problématiques. On pensera ainsi par exemple au verbe *faire* qui en plus de son de support, présente également celui de verbe plein (sens de fabriquer).

Concernant les raisons de rejet ou de mauvaise analyse, on peut avancer les hypothèses suivantes :

- On l'a déjà dit, les phrases à incise, ou à modifieurs multiples sont sources de problèmes
- Les phrases juxtaposées (parataxe) ou coordonnées sont également problématiques. On a ainsi vérifié que si chaque sous-phrase autonome était analysée séparément, les résultats seraient pertinents.
- Certains phénomènes syntaxiques ne sont pas encore couverts. On peut ainsi signaler les relatives extraposées, ou bien encore les phrases à présentatifs (voici, voilà)

Enfin, on notera que des arbres spécifiques devraient être ajoutés. Pour notre évaluation,

nous avons considéré les séquences renseignant des titres, des noms propres ou des compagnies comme un seul mot, mais il serait souhaitable de prendre en compte de nouveaux arbres, comme par exemple le suivant :



On notera pour finir que les problèmes mentionnés dans cette partie ne sont pas spécifiques aux familles à nom prédicatif. L'ajout des nouvelles familles ne fait que mettre en valeur un phénomène déjà existant.

7.4.4 Récapitulatif

Il n'est guère évident de vouloir établir une comparaison avec les évaluations ultérieures. Nous ne disposons pas en effet de corpus de tests totalement identiques. En 2001, [Werli, 2001] indiquait un nombre moyen d'analyses par phrase de 4,9. Antérieurement, [Candito, 1999a] indiquait que la grammaire qu'elle avait développée fournissait 2,96 analyses par phrase. Les chiffres des évaluations de 1996 étaient quant à eux de 1,5. Pour notre part, nous obtenons sur nos corpus de tests, un nombre moyen de 6,1 analyses¹¹. Ces chiffres confirment le fait que plus la couverture de la grammaire s'accroît, plus le nombre d'analyses est important.

Même si les résultats demeurent perfectibles, notre grammaire a tout de même permis le passage de phrases qui jusqu'alors ne recevaient pas d'analyse. Ces résultats sont donc encourageants pour les études futures¹².

¹¹ Il est d'ailleurs amusant de constater que le nombre moyen d'analyses ne varie pas entre le corpus de phrases fabriquées et le corpus de phrases. Ceci est certainement dû au fait que toutes les phrases du corpus n'ont pas pu être parsées. Si tel avait pu être le cas, on peut penser que l'on aurait obtenu un nombre moyen d'analyses beaucoup plus important en corpus.

¹² Le fait que l'on obtienne sur les phrases tests et les phrases en corpus un nombre moyen d'analyses identique est un phénomène surprenant, mais absolument pas représentatif. Si l'on ne prend en charge que les phrases qui ont été effectivement acceptées par le parseur, le nombre moyen d'analyses pour les phrases en corpus monte alors à 7,4.

	Phrases tests	Phrases en corpus	Toutes phrases confondues
Phrases à analyser	8	40	48
Phrases acceptées par le compilateur	8 (100 %)	32 (80 %)	40 (83,4 %)
Phrases recevant au moins une analyse	8 (100 %)	23 (57,5 %)	31 (64,6 %)
Phrases recevant au moins une analyse valide	8 (100 %)	18 (45 %)	26 (54,2 %)
Nombre moyen d'analyses par phrase	6,1	6,1	6,1

Les tests que nous avons entrepris nous ont permis d'expérimenter notre grammaire sur deux points : nous avons testé la capacité de la grammaire à reconnaître et à analyser correctement une phrase grammaticale, et avons également déterminé son taux moyen d'ambiguïté.

Conclusion

Nous voici maintenant arrivé au terme de notre travail. Il ne nous reste plus qu'à présenter nos conclusions.

Notre étude avait pour but la description et la formalisation des CVS à l'aide d'une métagrammaire. Pour cela, nous les avons considérées tant d'un point de vue lexical, que sémantique et syntaxique et ne nous sommes pas limité à des CVS qui placent le nom prédicatif en position d'objet direct. En examinant la plupart des critères généralement mis en avant pour identifier une CVS, nous avons pu constater qu'ils souffraient d'exceptions. Cela n'est pas nouveau en soi, mais la remise en cause de la double analyse mise en avant par [Gross, 1976] l'est un peu plus.

Syntaxiquement, les CVS présentent un comportement relativement uniforme et régulier, mais ce qui frappe lorsqu'on les observe c'est le caractère parfois déroutant de leurs compléments prépositionnels ou phrastiques. Certains doivent être uniquement considérés comme des compléments du verbe support, d'autres uniquement comme des compléments du nom, alors que d'autres acceptent les deux possibilités.

Le mode de représentation que constitue la Métagrammaire s'avère appréciable, tant il permet d'éviter la redondance de l'information, que de systématiser des croisements, rendant ainsi la couverture de la grammaire exhaustive. Nous avons pu juger de sa pertinence en prenant en compte des phénomènes qui jusqu'alors n'avaient pas été implémentés. Si la description des CVS ne nous a pas posé de problèmes de développement majeurs¹³, elle nous a néanmoins amené à reconsidérer certains aspects du formalisme. En particulier, l'ajout d'une quatrième dimension s'est avéré rapidement nécessaire pour ordonner les arguments syntaxiques entre eux.

Nous avons également redéfini la notion d'artère, initialement mise en avant par

¹³Les principaux problèmes que nous avons rencontrés concernent davantage le compilateur que le formalisme lui-même. Les descriptions hiérarchiques sont relativement intuitives.

[Candito, 1999a]. Si cette dernière en avait une vision plutôt sémantique, nous en avons quant à nous une vision résolument syntaxique. Cela nous paraît d'autant plus approprié que les fonctions utilisées par la métagrammaire sont des fonctions grammaticales et que la métagrammaire se veut un formalisme de description syntaxique. Les informations sémantiques que le descripteur peut vouloir ajouter, ne doivent pas, selon nous, se situer au niveau méta, mais au niveau de la grammaire de bas niveau, c'est-à-dire au niveau du contenu des classes de la métagrammaire.

A ce stade, il nous semble pertinent de souligner certains problèmes laissés ouverts et d'envisager les perspectives que ce travail peut amener. Le formalisme que nous avons utilisé a été testé sur la grammaire d'arbres adjoints. Rien ne permet d'affirmer qu'il soit suffisant ou adapté à la description d'autres grammaires de bas niveau. Il semble désormais acquis qu'une hiérarchie se décrit de façon unique, et que c'est au compilateur de prendre en charge les réalisations effectives vers les grammaires de bas niveau, mais très peu de travaux ont véritablement été menés en ce sens. A notre connaissance, seule l'étude de [Crabbé, 2005] soulève le problème. Gageons que dans un avenir proche de nouvelles réponses pourront éclairer le sujet.

Un second point qui nous semble intéressant d'explorer est celui de la réutilisabilité des classes utilisées pour décrire les CVS. Nous aurions pu choisir de reprendre la hiérarchie développée par [Candito, 1999a] pour implémenter nos familles à nom prédicatif. Malheureusement, la tâche a vite semblé impossible, tant les choix d'implémentation paraissaient parfois obscurs et différents de ceux présentés dans sa thèse. Nous n'avons donc pas pu étudier la question de la réutilisabilité d'une classe. Tout ce que nous pouvons dire, c'est qu'il semble que la description du graphe d'héritage semble être un phénomène invariant. A ce titre, la hiérarchie que nous décrivons, pour les phénomènes communs à ceux des verbes pleins, semble globalement semblable.

Un autre point qui nous semble important de mentionner concerne les interactions de la métagrammaire avec d'autres domaines que ceux qu'elle traite. Comment, par exemple, le formalisme doit-il se comporter avec les composantes sémantiques de la grammaire qu'il ne traite pas ?

Pour conclure, nous espérons que les idées qui émergent de ce travail pourront contribuer à la description de grammaire à plus large couverture et également à des analyses linguistiques plus approfondies de certains phénomènes syntaxiques. Nous avons nous-même contribué à l'enrichissement de la grammaire TAG du français et espérons que d'autres prendront le relai. La tâche n'est certes pas toujours évidente (car il faut alors s'inscrire dans une suite et comprendre les analyses précédentes qui ont été avancées), mais elle en vaut vraiment la peine.

L'évaluation que nous avons réalisée, montre d'ailleurs que la prise en compte de nou-

veaux phénomènes a permis d'analyser des phrases qui jusqu'alors étaient rejetées par le parseur. Bien entendu, elle montre également que la couverture de la grammaire résiste mal à l'analyse sur corpus non restreints. Si la grammaire que nous avons développée est davantage le reflet de la compétence, il n'en demeure pas moins que malgré la complexité et la précision des analyses qu'elle renferme, il est à l'heure actuelle difficile de prétendre analyser automatiquement et exhaustivement une langue naturelle comme le français.

Références

Abeillé, A. (1988). Light verb constructions and extraction out of NP in tree adjoining grammar. In *Papers from the 24th Regional Meeting of the Chicago Linguistic Society*, Chicago.

Abeillé, A. (1991). *Une grammaire lexicalisée d'arbres adjoints pour le français*. Thèse de Doctorat, Université Paris 7.

Abeillé, A. (1993). *Les nouvelles syntaxes*. Armand Colin, Paris.

Abeillé, A. (1996a). Corpus et syntaxe : l'apport de l'informatique linguistique. *Revue française de linguistique appliquée*.

Abeillé, A. (1996b). Fonction ou position objet ? *Le Gré des langues*, 11.

Abeillé, A. (1997). Fonction ou position objet ? *Le Gré des langues*, 12.

Abeillé, A. (2002). *Une grammaire électronique du français*. CNRS Editions, Paris.

Abeillé, A. (2004). Guide des annotations fonctionnelles. Documentation interne. Université Paris 7, disponible sur <http://www.llf.cnrs.fr/guide-fonctions.new.pdf>.

Abeillé, A. et Barrier, N. (2004). Enriching a french treebank. In *4th international conference on language resources and evaluation LREC*, Lisbon.

Abeillé, A., Barrier, N., et Barrier, S. (2001). La grammaire FTAG. Documentation interne, Université Paris 7.

Abeillé, A. et Candito, M.-H. (2000). FTAG : A lexicalized tree adjoining grammar for french. In Abeillé, A. et Rambow, O., (eds), *Tree Adjoining Grammars formalisms, linguistic analysis and processing*. CSLI Publications, Stanford.

Abeillé, A. et Clément, L. (1997). Annotation morpho-syntaxique : les mots simples et composés. Documentation interne. Université Paris 7, disponible sur <http://www.llf.cnrs.fr/Abeille/guide-morpho-synt.02.pdf>.

Abeillé, A., Clément, L., et Toussenet, F. (2003). Building a treebank for french. In Abeillé,

A., (ed), *Building and Using Parsed Corpora*, volume 20 de *Text, Speech and Language Technology*. Kluwer Academic Publishers, Dordrecht.

Abeillé, A. et Godard, D. (1996). La complémentation des auxiliaires français. In Abeillé, A. et Godard, D., (eds), *Nouveaux raisonnements syntaxiques*, volume 122 de *Langages*. Larousse, Paris.

Abeillé, A. et Godard, D. (1999). French word order and lexical weight. In Borsley, R. D., (ed), *The nature and function Syntactic categories*, volume 32 de *Syntax and Semantics*. Academic Press, New-York.

Abeillé, A. et Godard, D. (2000). Varieties of ESSE in romance languages. In Flickinger D. et Kathol, A., (ed), *Proceedings of the 7th International HPSG Conference*, UC Berkeley.

Abeillé, A. et Godard, D. (2004). De la légèreté en syntaxe. *BSL*, XCIX/1.

Abeillé, A., Godard, D., et Sag, I. (1998). Two kinds of composition in french complex predicates. In Hinrichs, E. et Kathol, A. et Nakazawa, T., (eds), *Complex Predicates in non transformational Syntax*, volume 30 de *Syntax and Semantics*. New-York Press.

Abeillé, A. et Rambow, O., (eds) (2000). *Tree Adjoining Grammars*. CSLI Publications, Stanford.

Abeillé, A., Toussenet, F., et Cheradame, M. (2002). Annotations en constituants, guide pour les correcteurs. Documentation interne. Université Paris 7, disponible sur <http://www.llf.cnrs.fr/fr/Abeille/guide-annot.pdf>.

Alonso Ramos, M. (1998). *Etude sémantico-syntaxique des constructions à verbe support*. Thèse de Doctorat, Université de Montréal.

Alonso Ramos, M. (2001). Constructions à verbe support dans les langues SOV. *Bulletin de la Société de linguistique de Paris*, 96-1.

Alsina, A. (1993). *Predicate Composition : a theory of syntactic function alternations*. Thèse de Doctorat, Stanford University.

Amiot, D., De Mulder, W., Flaux, N., et Tenchéa, M., (eds) (1999). *Fonctions syntaxiques et rôles sémantiques*. Cahiers scientifiques de l'université d'Artois. Artois Presses Université.

Angelis, A. d. (1989). Nominalizations with the italian support verbe *avere*. In *Lingvisticae Investigationes*, volume 13-2. John Benjamins, Amsterdam.

Anscombre, J.-C. (1986). Article zéro, termes de masse et représentation d'événements en français contemporain. In David, J. et Kleiber, G., (eds), *Déterminants : syntaxe et sémantique*, volume XI de *Recherchers linguistiques*. Klincksieck, Paris.

Références

- Anscombre, J.-C. (1991). La détermination zéro : quelques propriétés. In Anscombre, J.-C., (ed), *Absence de détermination et déterminant zéro*, volume 102 de *Langages*. Larousse, Paris.
- Aït-Mokhtar, S., Hagège, C., et Sandor, A. (2003). Problèmes d'intersubjectivité dans l'évaluation des analyseurs syntaxiques. In *Actes du colloque TALN 2003 - Tome 2*, Batz-sur-Mer.
- Bae, S.-M. (2002). *Le dictionnaire électronique des séquences nominales figées en coréen et de leurs formes fléchies*. Thèse de Doctorat, Université de Marne-La-Vallée.
- Bally, C. (1932). *Linguistique générale et linguistique française*. Francke, Berne. Troisième édition, conforme à la deuxième, 1950.
- Baron, I. (1998). Objet effectué et constructions nominales. In *Les objets : relations grammaticales et rôles sémantiques*, volume 35 de *Travaux de linguistique*. Duculot, Bruxelles.
- Barrier, N. (2002). Une métagrammaire pour les adjectifs du français. In *Actes du colloque TALN 2002*, Nancy.
- Barrier, N., Barrier, S., et Kynion, A. (2000). Lexik : a maintenance tool for FTAG. In *Proceedings of TAG+5*, Paris.
- Barrier, S. (1999). Classification et repérage de valence verbale : expériences avec FTAG. Mémoire de D.E.A., Université Paris 7.
- Barrier, S. et Barrier, N. (2003). Une métagrammaire pour les noms prédicatifs du français. In *Actes du colloque TALN 2003*, Batz-sur-Mer.
- Barrier, S. et Barrier, N. (2004). Metagrammars : a new implementation for FTAG. In *TAG+7 Proceedings*, Vancouver.
- Batoux, D. (1999). Les verbes supports. In Touratier, C., (ed), *La grammaticalisation - la terminologie*, volume 18 de *Travaux*. Publications de l'Université de Provence. Cercle linguistique d'Aix en Provence.
- Blanche-Benveniste, C. (2003). Les formes grammaticales de réalisation des sujets. *BFDL - Le sujet*. Actes augmentés de quelques articles du colloque *Le sujet* organisé en Septembre 2001 à l'Université de Provence.
- Blinkenberg, A. (1960). *Le problème de la transitivité en français moderne*. Munksgaard, Copenhagen.
- Bonami, O. (1999). *Les constructions du verbe : le cas des groupes prépositionnels argumentaux*. Thèse de Doctorat, Université Paris 7.

- Bonami, O. et Godard, D. (2001). Inversion du sujet, constituance et ordre des mots. *Cahier Jean-Claude Milner*. Editions Verdier.
- Boonen, D. (2001). *Le prédicat adjectival en FTAG*. Mémoire de DEA. Université Paris 7.
- Boons, J.-P., Guillet, A., et Leclère, C. (1976). *La structure des phrases simples en français*. Droz, Genève.
- Bresnan, J. (2001). *Lexical Functional Syntax*. Blackwell Textbooks in Linguistics. Blackwell, Oxford.
- Bresnan, J. et Kaplan, R. (1982). Lexical functional grammar : a formalism for grammatical representations. In Bresnan, J., (ed), *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge.
- Brinton, L. J. et Akimoto, M., (eds) (1999). *Collocational and idiomatic Aspects of Composite Predicates in the History of English*, volume 47 de *Studies in Language Companion Series*. John Benjamins, Amsterdam.
- Brunot, F. (1926). *La pensée et la langue. Méthodes, principes et plan d'une théorie nouvelle du langage appliquée au français*. Masson, Paris.
- Buvet, P.-A. (2003). La possessivation dans les constructions à support. In Conenna, M. et Laporte, E., (eds), *Grammaires et lexiques comparés : actes du colloque*, volume 26-1 de *Linguisticae Investigationes*. John Benjamins, Amsterdam.
- Candito, M.-H. (1996). A principle-based hierarchical representation of LTAGs. In *Actes de la conférence COLING96*, Copenhague.
- Candito, M.-H. (1999a). *Représentation modulaire et paramétrable de grammaires électroniques lexicalisées. Application au français et à l'italien*. Thèse de Doctorat, Université Paris 7.
- Candito, M.-H. (1999b). Un outil multilingue de construction semi-automatique de grammaires d'arbres adjoints. application au français et à l'italien. *T.A.L.*, 40.
- Carlson, G. N. (1977). Amount relatives. *Language*, 53 :3.
- Cattell, R. (1984). *Composite Predicates in English*, volume 17 de *Syntax and Semantics*. Academic Press, New-York.
- Caviola, F. (1992). Quelques nominalisations en français F et en français Q : éléments d'une syntaxe des Nprocessus. In Labelle, J. et Leclère, C., (eds), *Lexiques-grammaires comparés en français*, volume 17 de *Linguisticae investigationes Supplementa*. John Benjamins, Amsterdam.

- Chaurand, J. (1991). Verbes supports et emploi ou absence de l'article. In Anscombre, J.-C., (ed), *Absence de détermination et déterminant zéro*, volume 102 de *Langages*. Larousse, Paris.
- Chen, Z. (1990). The chinese vocables TEACH and TEACHING : two families of lexical entries for an explanatory combinatorial dictionary of chinese. In *Meaning-Text-Theory : Linguistics, Lexicography and Implications*. University of Ottawa Press. Steele, J. (ed.).
- Chomsky, N. (1957). *Structures syntaxiques*. Editions du Seuil. Trad. fr. 1969.
- Chomsky, N. (1965). *Aspects de la théorie syntaxique*. Editions du Seuil. Trad. fr. 1971.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- Clément, L. (2001). *Construction et exploitation d'un corpus syntaxiquement annoté pour le français*. Thèse de Doctorat, université Paris 7.
- Clément, L. et Kinyon, A. (2003a). Generating parallel multilingual LFG-TAG grammars with a metagrammar. In *Proceedings of Formal Grammars 03*, Vienne.
- Clément, L. et Kinyon, A. (2003b). Generating LFGs with a metagrammar. In *Proceedings of the LFG03 conference*, Saragota Springs, NY.
- Corblin, F. (1988). Savoir la grammaire et faire de la grammaire. *De la grammaire scientifique à la grammaire scolaire*. Huot (ed).
- Crabbé, B. (2005). *Représentation informatique de grammaires fortement lexicalisées : application à la grammaire d'arbres adjoints*. Thèse de Doctorat, Université de Nancy 2, Nancy.
- Crabbé, B., Gaiffe, B., et Roussanaly, A. (2004). Représentation et gestion de grammaires d'arbres adjoints lexicalisées. In Villemonte de la Clergerie, E. et Rajman, M., (eds), *Evolutions en analyse syntaxique*, volume 44-3 de *Traitement automatique des langues*. Hermès, Cachan.
- Curat, H. (1982). *La locution verbale en français moderne*. Presses de l'Université de Laval, Québec.
- Curme, G. O. (1935). *A Grammar of the English Language, Voll : Parts of Speech*. Heath, Boston : D.C.
- Daladier, A. (1978). *Problèmes d'analyse d'un type de nominalisation en français et de certains groupes nominaux complexes*. Thèse de Doctorat, Université Paris 7.
- Daladier, A. (1996). Le role des verbes supports dans un système de conjugaison nominale et l'existence d'une voie nominale en français. In Ibrahim, A. H., (ed), *Les supports*, volume 121 de *Langages*. Larousse, Paris.

- Dalrymple, M. (2001). *Lexical Functional Grammar*, volume 34 de *Syntax and Semantics*. Academic Press, New-York.
- Dalrymple, M. et Lødrup, H. (2000). The grammatical functions of complement clauses. In *Proceedings of the LFG00 Conference*.
- Danlos, L. (1980). *Représentations d'informations linguistiques : constructions "N être Prép X"*. Thèse de Doctorat, Université Paris 7.
- Danlos, L. (1988). Les expressions figées contruites avec le verbe support être prép. In Danlos, L., (ed), *Les expressions figées*, volume 90 de *Langages*. Larousse, Paris.
- Danlos, L. (1992). Support verb constructions : linguistics properties, represenation, translation. In *Journal of French Linguistic Study*, volume 2. N°1 CUP.
- Danlos, L. (1994). Coder des informations monolingues sur les noms pour éviter des règles bilingues sensibles au contexte. In Ladmiral, J.-R., (ed), *Le traducteur et l'ordinateur*, volume 116 de *Langages*. Larousse, Paris.
- Delaveau, A. et Kerleroux, F. (1985). *Problèmes et exercices de syntaxe française*. Armand Colin, Paris.
- Dixon, R. (1991). *A new approach to English Grammar; on Semantic Principles*. Clarendon, Oxford.
- Dubois, J. et Lagane, R. (1973). *La nouvelle grammaire du français*. Larousse, Paris.
- Dupont, N. (1991). *Les clitiques attributs en français*. Presses universitaires de Lyon. De Gaulmyn, M.-M. et Rémi-Giraud, S. (eds).
- Emorine, M. (1992). *Formalisation syntaxique et sémantique des constructions à verbes supports en français et en en espagnol dans une grammaire catégorielle d'unification*. Thèse de Doctorat, Université Clermont-Ferrand 2, Clermont-Ferrand.
- Erbach, G. et Krenn, B. (1994). Idioms and support verb constructions in HPSG. In Nerbonne, J., Netter, K., et Pollard, C., (eds), *German Grammar in HPSG*, volume 46 de *CSLI Lecture Notes*. CSLI Publications, Stanford.
- Fellbaum, C. et Zribi-Hertz, A. (1989). La construction moyenne en français et en anglais : étude de syntaxe et de sémantique comparées. volume 18 de *Recherches linguistiques de Vincennes*.
- Fontenelle, T. (1997). *Turning a bilingual dictionary into a lexical-semantic database*. Lexicographica Series Maior 79. Max Niemeyer Verlag, Tübingen.
- Gaätone, D. (1981). Les "locutions verbales" : pour quoi faire ? *Revue romane*, 16,1-2.

- Gaatone, D. (1998a). *Le passif en français*. Duculot, Liège.
- Gaatone, D. (1998b). L'objet direct comme notion formelle dans la formulation des règles syntaxiques. In *Les objets : relations grammaticales et rôles linguistiques*, volume 35 de *Travaux de linguistique*. Duculot, Bruxelles.
- Gaiffe, B., Crabbé, B., et Roussanaly, A. (2002). A new metagrammar compiler. In *Proceedings of TAG+6*, Venise.
- Gardès-Tamine, J. (2003). La place du sujet en français. *BFDL - Le sujet*. Actes augmentés de quelques articles du colloque *Le sujet* organisé en Septembre 2001 à l'Université de Provence.
- Gerdes, K. (2002). *Topologie et grammaires formelles de l'allemand*. Thèse de Doctorat, Université Paris 7.
- Girard, G. (2003). La notion de sujet - une notion à définir. *BFDL - Le sujet*. Actes augmentés de quelques articles du colloque *Le sujet* organisé en Septembre 2001 à l'Université de Provence.
- Girard, G. (2004). La notion de sujet et la notion de complément. *Cercles*, 9.
- Giry-Schneider, J. (1976). Formes passives à sujet sans déterminant. In *Recherches linguistiques*, volume 2. Université de Paris 8.
- Giry-Schneider, J. (1978). *Les nominalisations en français*. Droz, Genève-Paris.
- Giry-Schneider, J. (1981). Les compléments nominaux du verbe *dire*. In Gross, M., (ed), *Formes syntaxiques et prédicats sémantiques*, volume 63 de *Langages*. Larousse, Paris.
- Giry-Schneider, J. (1986). Les noms construits avec faire : compléments ou prédicats ? In Gross, G. et Vivès, R., (eds), *Syntaxe des noms*, volume 69 de *Langue française*. Larousse, Paris.
- Giry-Schneider, J. (1987). *Les prédicats nominaux en français*. Droz, Genève-Paris.
- Giry-Schneider, J. (1991). L'article zéro dans le lexique-grammaire des noms prédicatifs. In Anscombre, J.-C., (ed), *Absence de déterminant et déterminant zéro*, volume 102 de *Langages*. Larousse, Paris.
- Giry-Schneider, J. (1996). La notion de modifieur obligatoire dans les phrases à verbe support *avoir*. In Ibrahim, A. H., (ed), *Les supports*, volume 121 de *Langages*. Larousse, Paris.
- Godard, D. (1988). *La syntaxe des relatives en français*. CNRS Editions, Paris.
- Godard, G. (1992). Extraction out of NP in french. *Natural language and linguistic Theory*, 10.

- Gougenheim, G. (1971). Une catégorie lexico-grammaticale : les locutions verbales. *Etudes de linguistique appliquée*, 2. Didier.
- Gracia, L. (1986). *la teoria tematica*. Thèse de Doctorat, Universitat Autònoma de Barcelona, Barcelone.
- Grevisse, M. et Goose, A. (1993). *Le Bon Usage*. De Boeck - Duculot, Liège. 13ème édition.
- Grewe, K. (1993). Une analyse sémantique et syntaxique des phrases à verbes supports de l'allemand et du français. *Acta Linguistica Hungarica : An International Journal of Linguistics*, 41.1-4.
- Grimshaw, J. (1982). Subcategorization and grammatical relations. In Bresnan, J., (ed), *Subjects and other subjects : Proceedings of the Harvard conference on the representation of grammatical relations*. Indiana University Linguistics Club, Bloomington.
- Grimshaw, J. (1990). *Argument Structure*, volume 18 de *Linguistic Inquiry Monograph*. MIT Press, Cambridge.
- Grimshaw, J. et Mester, A. (1988). Light verbs and θ -marking. *Linguistic Inquiry*, 19-2.
- Gross, G. (1989). *Les constructions converses du français*. Droz, Genève-Paris.
- Gross, G. (1993a). Les passifs nominaux. In Gross, G., (ed), *Sur le passif*, volume 109 de *Langages*. Larousse, Paris.
- Gross, G. (1993b). Trois applications de la notion de verbe support. volume 59 de *L'information grammaticale*.
- Gross, G. (1995). Une sémantique nouvelle pour la traduction automatique : les classes d'objets. *La tribune des Industries de la Langue et de l'Information électronique*, 17-18-19.
- Gross, G. (1996a). *Les expressions figées en français*. Ophrys, Paris.
- Gross, G. (1996b). Pour une typologie des prédicats nominaux. In *Prédication, assertion, information, Actes du colloque d'Uppsala en linguistique française*, Uppsala.
- Gross, G. (1996c). Prédicats nominaux et compatibilité aspectuelle. In Ibrahim, A. H., (ed), *Les supports*, volume 121 de *Langages*. Larousse, Paris.
- Gross, G. et Kiefer, F. (1995). La structure événementielle des substantifs. volume 30/1-2 de *Folia Linguistica*. Mouton de Gruyter, Berlin.
- Gross, G. et Valli, A. (1991). Déterminant zéro et verbes supports en moyen français et en français moderne. In Anscombre, J.-C., (ed), *Absence de déterminant et déterminant zéro*, volume 102 de *Langages*. Larousse, Paris.

- Gross, M. (1968). *Grammaire transformationnelle du français - Syntaxe du verbe*. Larousse, Paris. Réédition Cantilène 1986.
- Gross, M. (1969). Remarques sur la notion d'objet direct en français. In Lagane, R. et Pinchon, J., (eds), *La syntaxe*, volume 1 de *Langue française*. Larousse, Paris.
- Gross, M. (1976). Sur quelques groupes nominaux complexes. In Chevalier, J.-C. et Gross, M., (eds), *Méthodes en grammaire française*. Klincksieck.
- Gross, M. (1977). *Grammaire transformationnelle du français - Syntaxe du nom*. Larousse, Paris.
- Gross, M. (1981). Les bases empiriques de la notion de prédicat sémantique. In Gross, M., (ed), *Formes syntaxiques et prédicats sémantiques*, volume 63 de *Langages*. Larousse, Paris.
- Gross, M. (1982). Une classification des phrases "figées" du français. In *Revue Québécoise de Linguistique*, volume 11-2. Presses de l'Université du Québec, Montréal.
- Gross, M. (1986). Les nominalisations d'expressions figées. In Gross, G. et Vivès, R., (eds), *Syntaxe des noms*, volume 69 de *Langue française*. Larousse, Paris.
- Gross, M. (1988). Les limites de la phrase figée. In Danlos, L., (ed), *Les expressions figées*, volume 90 de *Langages*. Larousse, Paris.
- Gross, M. (1998). La fonction sémantique des verbes supports. In Lamiroy, B., (ed), *Le lexique-grammaire*, volume 37 de *Travaux de linguistique*. Duculot, Bruxelles.
- Guenther, F. et Blanco, X. (2004). Multi-lexemic expressions : an overview. In Leclère, C., Laporte, E., Piot, M., et Silberstein, M., (eds), *Lexique, Syntaxe et Lexique-Grammaire*, volume 24 de *Linguisticae investigationes Supplementa*. John Benjamins, Amsterdam.
- Guimier, C., (ed) (1993). *1001 circonstants*. Presses universitaires de Caen.
- Han, C.-H. et Rambow, O. (2000). The sino-korean light verb construction and lexical argument structure. In *Proceedings of TAG+5*, Paris.
- Hancock, V. (2002). L'emploi des constructions clivées dans le français parlé. In *XV Skandinaviske romanistkongress*, Oslo.
- Harris, S. Z. (1968). *Mathematical Structure of Languages*. Wiley-Interscience, New-York.
- Harris, S. Z. (1976). *Notes du cours de syntaxe*. Editions du Seuil, Paris.
- Hausmann, F. (1979). Un dictionnaire de collocations est-il possible ? *Travaux de littérature et de linguistique de l'Université de Strasbourg*.

Heap, D. et Robergue, Y. (2001). Cliticisation et théorie syntaxique, 1971-2001. *Revue québécoise de linguistique*, 30-1.

Ibrahim, A. H. (2002a). Les verbes supports en arabe. *Bulletin de la Société de Linguistique de Paris*, 97-2.

Ibrahim, A. H. (2002b). Maurice Gross : une refondation de la linguistique au crible de l'analyse automatique. In *Actes de la 9ème conférence Annuelle sur le Traitement Automatique des Langues Naturelles*, Nancy.

Jayaseelan, K. A. (1988). Complex predicates and θ -theory. In Wilkins, W., (ed), *Syntax and Semantics - Thematic relations*, volume 21. Academic Press, New-York.

Jespersen, O. (1942). *A modern English Grammar on Historical Principles*. George Allen & Unwin.

Jones, M. A. (1996). *Foundations of French Syntax*. Cambridge University Press.

Joshi, A., Levy, L., et Takahashi, M. (1975). Tree adjunct grammars. *Journal of the Computer and System Sciences*, 10-1.

Kayne, R. (1972). Subject inversion in french interrogatives. In Casagrande, J. et Saciuk, B., (eds), *Generative studies in Romance languages*. Newbury House, Rowley.

Kayne, R. S. (1975). *French syntax. The Transformational Cycle*. The MIT Press, Cambridge. Traduction française aux Editions du Seuil, 1977.

Kayne, R. S. (1976). French relative *que*. *Current studies in Romance languages*. F. Hensey and M. Lujan (eds) - Georgetown University Press.

Keenan, E. (1976). Towards a universal definition of "subject". In Li, C. N., (ed), *Subject and Topic*. Academic Press, New-York.

Kinyon, A. (2000). Hypertags. In *Coling-00*, Sarrebrück.

Kjaersgaard, P. S. (2002). Fonctions syntaxiques - définitions et problèmes. In *XV Skandinaviske romanistkongress*, Oslo.

Kupferman, L. (1991). L'aspect du groupe nominal et l'extraction de *en*. *Le français moderne*, LIX-2.

Kupferman, L. (1996). Les génitifs : gouvernement d'antécédent et gouvernement thématique. In Kupferman, L., (ed), *Un bien grand mot : de. De la préposition au mode de quantification*, volume 109 de *Langue française*. Larousse, Paris.

- Kupferman, L. (2004). *Le mot "de" : domaines prépositionnels et domaines quantificationnels*. Duculot, Bruxelles.
- Kuroda, Y. (1968). English relativization and certain related problems. In *Language*, volume 44. Linguistic Society of America, New-York.
- La Fauci, N. (1984). Le seuil sémantique : verbes à complétives et constructions à verbe support. In Guillet, A. et La Fauci, N., (eds), *Lexique-Grammaire des langues romanes*, volume 9 de *Lingvisticae investigationes Supplementa*. John Benjamins, Amsterdam.
- Labelle, J. (1974). *Etudes de constructions avec opérateur AVOIR (nominalisations et extensions)*. Thèse de Doctorat, Université Paris 7.
- Labelle, J. (1984a). Le prédicat nominal avec support avoir. In Guillet, A. et La Fauci, N., (eds), *Lexique-Grammaire des langues romanes*, volume 9 de *Lingvisticae investigationes Supplementa*. John Benjamins, Amsterdam.
- Labelle, J. (1984b). Le prédicat nominal avec support avoir. In Guillet, A. et La Fauci, N., (eds), *Lexique-Grammaire des Langues romanes*, volume 9 de *Lingvisticae Investigationes Supplementa*. John Benjamins, Amsterdam.
- Lagae, V. (1998). En quantitatif : pronom lié à la fonction objet ou à une position ? In *Les objets : relations grammaticales et rôles sémantiques*, volume 35 de *Travaux de linguistique*. Duculot, Bruxelles.
- Lagane, R. (1969). Problèmes de définition : le sujet. In Lagane, R. et Pinchon, J., (eds), *La syntaxe*, volume 1 de *Langue française*. Larousse, Paris.
- Langer, S. (2005). A linguistic test battery for support verb constructions. In Gross, G. et Pontonx, S., (eds), *Nouvel état des lieux*, volume 27-2 de *Lingvisticae Investigationes Supplementa*. John Benjamins, Amsterdam.
- Lazard, G. (1994). *L'actance*. Linguistique nouvelle. PUF, Paris.
- Lazard, G. (2003). Le sujet en perspective interlinguistique. *BFDL - Le sujet*. Actes augmentés de quelques articles du colloque *Le sujet* organisé en Septembre 2001 à l'Université de Provence.
- Le Goffic, P. (1993). *Grammaire de la phrase française*. Hachette Supérieur, Paris.
- Le Pesant, D. et Mathieu-Colas, M. (1998). Introduction aux classes d'objets. In Le Pesant, D. et Mathieu-Colas, M., (eds), *Les classes d'objets*, volume 131 de *Langages*. Larousse, Paris.
- Leclère, C. (1971). Remarques sur les substantifs opérateurs. In Gross, M. et Stéfanini, J., (eds), *Syntaxe transformationnelle du français*, volume 11 de *Langue française*. Larousse, Paris.

- Lees, R. B. (1960). *The Grammar of English Nominalizations*. Mouton, La Haye.
- Lindberg, L. (1898). *Les locutions verbales figées dans la langue française*. Thèse pour le doctorat. Faculté des lettres d'Upsal, Suède.
- Marandin, J.-M. (1997). Dans le titre se trouve le sujet. Thèse d'habilitation. Université Paris 7.
- Marandin, J.-M. (2001). Unaccusative inversion in french. *Going Romance 1999 :selected papers*. John Benjamins.
- Marchello-Nizia, C. (1996). Les verbes supports en diachronie : le cas du français. In Ibrahim, H., (ed), *Les supports*, volume 121 de *Langages*. Larousse, Paris.
- Matsumoto, M. (1999). Composite predicates in middle english. In Brinton, L. J. et Akimoto, M., (eds), *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*, volume 47 de *Studies in Language Companion Series*. John Benjamins, Amsterdam.
- Matsumoto, Y. (1996). *Complex Predicates in Japanese - A Syntactic and Semantic Study of the Notion 'Word'*. CSLI & Kurocio, Stanford.
- Mel'čuk, I. (1984). Un nouveau type de dictionnaire : le dictionnaire explicatif et combinatoire du français contemporain. In Mel'čuk, I., Arbatchewsky-Jumarie, N., Elnitsky, L., Iordanskaja, L., et Lessard, A., (eds), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*. Presses de l'Université de Montréal, Montréal.
- Mel'čuk, I. (1988). *Dependency Syntax : theory and Practice*. State University of New-York Press, Albany.
- Mel'čuk, I. (1997). Vers une linguistique sens-texte. leçon inaugurale. Collège de France.
- Mel'čuk, I., Clas, A., et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot, Louvain-la-Neuve.
- Meunier, A. (1984). Une construction à prédicat nominal : Luc a l'audace de mentir à léa. In Guillet, A. et La Fauci, N., (eds), *Lexique-Grammaire des langues romanes*, volume 9 de *Lingvisticae investigationes Supplementa*. John Benjamins, Amsterdam.
- Miller, P. (1992). *Clitics and Constituents in Phrase Structure Grammar*. Garland, New-York.
- Miller, P. et Monachesi, P. (2003). Les pronoms clitiques dans les langues romanes. *Les langues romanes, problèmes de la phrase simple*. CNRS Editions, Paris, Godard D. (ed.).
- Miller, P. et Sag, I. A. (1997). French clitic movement without clitics or movement. *Natural Language and Linguistic Theory*.
- Milner, J.-C. (1982). *Ordres et raisons de langue*. Editions du Seuil, Paris.

- Milner, J.-C. (1986). Introduction à un traitement du passif. Publication de l'ERA 942. Université Paris 7.
- Moignet, G. (1974). *Etude de psychosystématique française*. Klincksieck, Paris.
- Moreau, M. L. (1976). *C'est. Etude de syntaxe transformationnelle*. Université de Mons.
- Muller, C., (ed) (2001). *Clitiques et cliticisation*. Honoré Champion. Textes réunis par Muller, Actes du colloque de Bordeaux, Octobre 1998.
- Namer, F. (1998). *Support verb constructions*, chapitre 8. European Commission.
- Namer, F. et Schmidt, P. (1997). Support verb constructions in a typed structure framework. *Verbum*, 19.
- Negroni-Peyre, D. (1978). Nominalisation par être en et réflexivation. In *Lingvisticae Investigationes*, volume 2-1. John Benjamins, Amsterdam.
- Nølke, H. (1983). Quelques réflexions sur la structure sémantique des phrases clivées en français moderne. *Modèles linguistiques*, 5.
- Ogata, K. (1982). Note sur les verbes supports en japonais. Mémoire d'université de Paris 7.
- Parmentier, Y. (2003). Méta-grammaires et calcul sémantique pour les grammaires d'arbres adjoints. Projet de fin d'Etudes DEA, Loria, Nancy.
- Pasero, R., Royauté, J., et Sabatier, P. (2004). Sur la syntaxe et la sémantique des groupes nominaux à tête prédicative. volume 27-1 de *Lingvisticae Investigationes*. John Benjamins, Amsterdam.
- Perlmutter, D. M. (1970). Surface structure constraints in syntax. *Linguistic Inquiry*, 1.
- Perlmutter, D. M. (1971). *Deep and surface structure constraints in Syntax*. Holt, Rinehart & Winston, New-York.
- Polenz, P. v. (1963). *Funktionsverben im heutigen Deutsch. Sprache in der rationalisierten Welt*. Schwann, Düsseldorf.
- Ranchhod, E. (1989). Lexique-grammaire du portugais : prédicats nominaux supportés par *estar*. In *Lingvisticae Investigationes*, volume 13-2. John Benjamins, Amsterdam.
- Rappaport, M. et Levin, B. (1988). What to do with θ -roles. In Wilkins, W., (ed), *Thematic relations*, volume 21 de *Syntax and Semantics*. Academic Press, New-York.
- Reuther, T. (1996). On dictionary entries for support verbs : The cases of russian VESTI, PROVODIT' and PROIZVODIT'. In Wanner, L., (ed), *Lexical Functions in Lexicography and*

Natural Language Processing, volume 31 de *Studies in Language Companion Series*. John Benjamins, Amsterdam.

Riegel, M., Pellat, J.-C., et Rioul, R. (1996). *Grammaire méthodique du français*. Linguistique nouvelle. PUF, Paris.

Rivière-Papon, N. (1996). *Repères temporels et repères énonciatifs : la construction impersonnelle, le pronominal et le participe passé adjectival en français*. Thèse à la carte. Presses universitaires du Septentrion, Lille.

Rémi-Giraud, S. et Roman, A., (eds) (1998). *Autour du circonstant*. Linguistique et sémiologie. Presses universitaires de Lyon.

Rogers, J. et Vijay-Shanker, K. (1994). Obtaining trees from their descriptions : An application to tree-adjointing grammars. *Computational Intelligence*, 10.

Rohrer, C. (1967). Definition of locutions verbales. *French review*, 41-3.

Rooryck, J. et Van den Bussche, H. (1988). Les attributs en français et leur intégration phrasique. *Travaux de linguistique*, 16. Duculot.

Royauté, J. et Sabatier, P. (2004). Contraintes d'ordre dans les groupes nominaux prédicatifs. In *Actes des journées de la Syntaxe*, Bordeaux.

Ruwet, N. (1972). *Théories syntaxiques et syntaxe du français*. Editions du Seuil, Paris.

Salkoff, M. (1990). Automatic translation of support verb constructions. In *Proceedings of COLING-90*, Helsinki.

Samvelian, P. (1995). *Les nominalisation en français : arguments syntaxiques et actants sémantiques*. Thèse de Doctorat, Université Paris 7.

Sarkar, A. (2000). Practical experiments in parsing using tree adjoining grammars. In *TAG+5 Proceedings*.

Serbat, G. (1981). *Cas et fonctions*. Linguistique nouvelle. PUF, Paris.

Simatos, I. (1997). Expression idiomatique, expression figée, prédicat complexe. In Fradin, B. et Marandin, J., (eds), *Mot et grammaires*. Didier Erudition, Paris.

Stein, G. (1991). The phrasal verb type 'to have a look' in modern english. *International Review of Applied Linguistics in Language Teaching*, 29.

Tasmowski, L. (1984). ?*lui faire téléphoner quelqu'un d'autre : une stratégie ? In *Linguisticae Investigationes VIII*.

- Tesnière, L. (1959). *Eléments de syntaxe structurale*. Klincksieck, Paris.
- Touratier, C. (2003). Sujet et syntaxe ? *BFDL - Le sujet*. Actes augmentés de quelques articles du colloque *Le sujet* organisé en Septembre 2001 à l'Université de Provence.
- Uchida, Y. et Nakayama, M. (1993). Japanese verbal noun constructions. In *Linguistics*, volume 31.
- Van Noord, G. (1994). Head-corner parsing for TAG. *Computational Intelligence*, 10.
- Vergnaud, J.-R. (1985). *Dépendances et niveaux de représentation en syntaxe*, volume 13 de *Lingvisticae Investigationes Supplementa*. John Benjamins, Amsterdam.
- Vijay-Shanker, K. (1992). Using descriptions of trees in a tree adjoining grammar. *Association for Computational Linguistics*, 18.
- Vijay-Shanker, K. et Schabes, Y. (1992). Structure sharing in lexicalized tree adjoining grammar. In *Proceedings of COLING-92*.
- Vikner, C. (1972). Quelques réflexions sur les phrases clivées en français moderne. In *V skandinaviske romanistkongress*, Turku.
- Villemonte de la Clergerie, E. et Thomasset, F. (2005). Comment obtenir plus des métagrammaires. In *Actes du colloque TALN 2005*, Dourdan.
- Vivès, R. (1983). *Avoir, prendre, perdre : constructions à verbe support et extensions aspectuelles*. Thèse de Doctorat, Université Paris 8.
- Vivès, R. (1993). La prédication nominale et l'analyse par verbe support. volume 59 de *L'information grammaticale*.
- Vivès, R. (2004a). Alternances casuelles, verbes supports et quantification. In Gross, G. et de Pontoux, S., (eds), *Verbes supports : nouvel état des lieux*, volume 27-2 de *Lingvisticae Investigationes*. John Benjamins, Amsterdam.
- Vivès, R. (2004b). Une grille d'analyse pour les prédicats nominaux. In Leclère, C., Laporte, E., et Piot, M. et Silberstein, M., (eds), *Syntaxe, Lexique et Lexique-Grammaire. Hommage à Maurice Gross*, volume 24 de *Lingvisticae Investigationes Supplementa*. John Benjamins, Amsterdam.
- Wagner, R. L. et Pinchon, J. (1962). *Grammaire du français classique et moderne*. Hachette Supérieur, Paris. Nouvelle édition 1991.
- Werli, S. (2001). *Développement et évaluation d'une grammaire TAG pour le français*. Mémoire de DEA. Université Paris 7.

- Wierzbicka, A. (1982). Why can you *have a drink* when you can't **have an eat* ? *Language*, 58.
- Wilmet, M. (1998). *Grammaire critique du français*. Hachette Supérieur, Paris. 2^o édition.
- Xia, F. (2001). *Automatic grammar generation from two different perspectives*. Thèse de Doctorat, University of Pennsylvania.
- XTAG, G. (2001). A lexicalized tree adjoining grammar for english. disponible sur internet à l'adresse <http://www.cis.upenn.edu/xtag/tech-report/rech-report.html>.
- Zholkovsky, A. et Mel'čuk, I. (1967). Sur la synthèse sémantique. *T.A. Informations*, 2.
- Zinglé, H. et Brobeck-Zinglé, M.-L. (2003). *Dictionnaire combinatoire du français*. La Maison du Dictionnaire.
- Zribi-Hertz, A. (1994). La syntaxe des clitiques nominatifs en français standard et en français avancé. *Travaux de linguistique et de philologie*, 32. Klincksieck.
- Zubizarreta, M.-L. (1987). *Levels of Representation in the Lexicon and in the Syntax*. Foris, Dordrecht.
- Zwicky, A. et Pullum, G. (1983). Cliticisation vs. inflection : English *n't*. *Language*, 59.

Annexes

Annexe A

Hiérarchie exemple des familles à nom prédicatif

Nous présentons dans cette annexe une mini-hiérarchie permettant de représenter 2 familles à nom prédicatif du français. Cette hiérarchie a été testée sous Clisp 2.27 et permet de générer 26 arbres pour la famille n0vN et 112 pour la famille n0vNpn1.

Elle traite en dimension 2 :

- Les actifs de base (sans redistribution) : L’aimant exerce une attraction sur le fer
- Le passif : Une attraction est exercée par l’aimant sur le fer
- Le passif impersonnel : Il est exercé une attraction sur le fer
- Le passif conversif : Le fer subit l’attraction de l’aimant
- Les constructions moyennes : Une attraction s’exerce sur le fer

La dimension 3, bien que paraissant ici très développée est encore très rudimentaire, mais elle permet de représenter les arguments réalisés de manière canonique, ainsi que les phénomènes d’extraction.

Afin de ne pas alourdir inutilement cette hiérarchie, l’ordonnancement des arguments sera présenté dans l’annexe suivante.

```
#####  
;# Hiérarchie de démonstration pour les familles à nom #  
;# prédicatif du français #  
#####  
;# Familles décrites : #  
;# n0vN - Max prend une douche #  
;# n0vNpn1 - Max commet un crime contre Luc #  
#####
```

```

#####
;# Auteur : Barrier Sébastien - LLF                                     #
;#          sbarrier@linguist.jussieu.fr                               #
;#####
;# Testée sous Clisp 2.27                                             #
;#####

(in-package :user)

(setq *families* '(
    "n0vN"
    "n0vNpn1") )

(setq *redistributions* '(
    ACTIF-DE-BASE
    PASSIF
    PASSIF-IMPERSONNEL
    PASSIF-CONV
    MOYENNE) )

(setq *sujets* '((
    SUJ-NOM-CAN-SUBST
    SUJ-NOM-CAN-ANCRE
    PREVERBAL                ;(Sujet clitique préverbal)
    SUJ-NOM-INV-SUBST
    SUJ-NOM-INV-ANCRE
    SUJ-COMP-SUBST
    SUJ-COMP-ANCRE
    SUJ-CLIV-SUBST
    SUJ-CLIV-ANCRE
    SUJET-IMPERS
    ):function sujet) )

(setq *objets* '((
    OBJ-NOM-SUBST
    OBJ-NOM-ANCRE
    OBJ-REL-SUBST
    OBJ-REL-ANCRE
    OBJ-CLIV-SUBST
    OBJ-CLIV-ANCRE
    OBJ-INT-SUBST
    OBJ-INT-ANCRE
    ):function objet) )

(setq *prepobjs* '((

```

```

PREP-OBJ-CAN
PREP-OBJ-CAN2
PREP-OBJ-REL
PREP-OBJ-CLIV
PREP-OBJ-INT
):function prep-obj) )

(setq *deobjs* ' (
  DE-OBJ-NOM-CAN
  DE-OBJ-NOM-CAN2
  DONT
  DE-OBJ-CLIVE
  DE-OBJ-INT
):function de-obj) )

(setq *parobjs* ' (
  PAR-OBJ-OMIS
  PAR-OBJ
):function parobj) )

(setq *realizations* ' (
  *sujets*
  *objets*
  *prepobjs*
  *deobjs*
  *parobjs*) )

(defparameter *inheritance-constraints* ' (
  (SUJ-NOM-INVERSE EXTRACTIONS)
  (PASSIF-IMPERSONNEL PAR-OBJ-OMIS) ) )

(defparameter *neg-inheritance-constraints* ' (
  (SUJ-NOM-INVERSE OBJET-NOM)
  (PREP-OBJ-CAN2 OBJET-NOM)
  (DE-OBJ-NOM-CAN2 OBJET-NOM) ) )

(defparameter *constrained-crossings* ' (
  )
)

(defparameter *disjunctive-nodes* '(EXTRACTIONS)
)

(DIMENSION1
(TOP) )

```

```

(DIMENSION2
  (TOP) )

(DIMENSION3
  (TOP) )

(DIMENSION4
  (TOP) )

; #####
; # Classes de dimension 1 #
; #####

(SUJET-INITIAL
  (DIMENSION1)
  (init-constants '(((arg0) :cat ("N" "Cl" "S" "Det") :function sujet))) )

(SUJET-NOMINAL
  (SUJET-INITIAL)
  (init-constants '(((arg0) :cat ("N" "Cl" "Det") ))) )

(SUJ-N-SUBST
  (SUJET-NOMINAL)
  (init-constants '(((arg0) :ind "0" :qual subst ))) )

(SUJ-N-HEAD
  (SUJET-NOMINAL)
  (init-constants '(((arg0) :qual head))) )

(OBJET-INITIAL
  (DIMENSION1)
  (init-constants '(((arg1) :cat ("N" "Cl" "S") :function objet))) )

(OBJET-NOMINAL
  (OBJET-INITIAL)
  (init-constants '(((arg1) :cat ("N" "Cl")))) )

(OBJ-N-SUBST
  (OBJET-NOMINAL)
  (init-constants '(((arg1) :ind "1" :qual subst))) )

(OBJ-N-HEAD
  (OBJET-NOMINAL)

```

```

(init-constants '(((arg1) :cat ("N") :qual head))) )

(PREP-OBJ-INITIAL
 (DIMENSION1)
 (init-constants '(((arg2) :cat ("N" "S") :function prep-obj))) )

(PREP-OBJ-NOMINAL
 (PREP-OBJ-INITIAL)
 (init-constants '(((arg2) :cat ("N") :qual subst :ind"1")))) )

(DE-OBJ-INITIAL
 (DIMENSION1)
 (init-constants '(((arg2) :cat ("N" "Cl" "S") :function de-obj))) )

(DE-PHR
 (DE-OBJ-INITIAL)
 (init-constants '(((arg2) :cat ("S") :qual subst))) )

(DE-NOM
 (DE-OBJ-INITIAL)
 (init-constants '(((arg2) :cat ("N" "Cl") :qual subst))) )

(n0vNpn1
 (SUJ-N-SUBST OBJ-N-HEAD PREP-OBJ-NOMINAL)
 (init-name "n0vNpn1-" :direction :left) )

(n0vN
 (SUJ-N-SUBST OBJ-N-HEAD)
 (init-name "n0vN-" :direction :left) )

; Fin de déclaration des classes de la dimension 1

;#####
;# Classes de dimension 2 #
;#####

(MORPHO-VERBALE
 (DIMENSION2)
 (init-constants '(((S_d) :cat ("S"))
                    ((vsupp) :cat ("V") :qual subst)
                    ((vmarq) :cat ("V") )))
 (init-description '((S_d nil) ((vmarq t))))
 (init-description '((S_d nil) ((vsupp t)))) )

```

```

(VSUPP-NON-MARQUE
  (MORPHO-VERBALE)
  (init-constants '(((vmarq) :qual subst)))
  (init-description '((S_d nil) ((vmarq t)) ((vsupp nil)))) )

(VSUPP-MARQUE
  (MORPHO-VERBALE)
  (init-constants '(((vmarq vsupp)))) )

(ACTIF-DE-BASE
  (VSUPP-MARQUE)
  (init-name "Actif-" :direction :left) )

(REG_LEX
  (DIMENSION2) )

(DEM-SUJ
  (REG_LEX) )

(SUJ>PAROBJ
  (DEM-SUJ)
  (revise-function 'sujet 'parobj) )

(SUJ>DEOBJ
  (DEM-SUJ)
  (revise-function 'sujet 'de-obj) )

(SUJ>VIDE
  (DEM-SUJ)
  (revise-function 'sujet 'empty) )

(PROM-SUJ
  (REG_LEX) )

(OBJ>SUJ
  (PROM-SUJ)
  (revise-function 'objet 'sujet) )

(PREPOBJ>SUJ
  (PROM-SUJ)
  (revise-function 'prep-obj 'sujet) )

(PASSIF
  (OBJ>SUJ SUJ>PAROBJ VSUPP-NON-MARQUE)
  (init-equations '(, (cr-eq vsupp "t" "mode" "ppart")

```

```

, (cr-eq vsupp "b" "mode" "ppart")
, (cr-eq vsupp "b" "passive" "+")
, (cr-eq vsupp "t" "neg" "-")
, (cr-eq vsupp "t" "inv" "-")
, (cr-eq vsupp "t" "neg" "-")
, (cr-eq vmarg "t" "inv" "-")
, (cr-eq vmarg "b" "aux" "avoir")
, (cr-eq vmarg "b" "caus" "-")
, (cr-eq vmarg "b" "cop" "+"))
(init-constants '(((vmarg) :ind "cop")))
(init-name "Passif-" :direction :left) )

(PASSIF-IMPERSONNEL
  (SUJ>PAROBJ VSUPP-NON-MARQUE)
  (init-constants '(((impers) :cat ("il") :function sujet)
                    ((vmarg) :ind "cop"))))
  (init-name "Passif-impersonnel-" :direction :left) )

(PASSIF-CONV
  (PREPOBJ>SUJ SUJ>DEOBJ VSUPP-MARQUE) )

(MORPHO-MOYENNE
  (VSUPP-MARQUE)
  (init-constants '(((refl) :cat ("Cl"))
                    ((se) :cat ("se") :qual no_qual)
                    ((Vc) :cat ("V") :ind "c")
                    ))
  (init-equations '(, (cr-eq refl "b" "pers" "3"))))
  (init-equations-masquantes '(, (cr-eq vsupp () "aux" "etre"))))
  (init-description '((S_d nil) ((Vc t))))
  (init-simple-desc '(Vc (refl) (vmarg)))
  (init-simple-desc '(refl (se)))
  (init-name "Moyenne-" :direction :left) )

(MOYENNE
  (MORPHO-MOYENNE OBJ>SUJ SUJ>VIDE) )

;Fin de déclaration des classes de la dimension 2

;#####
;# Classes de dimension 3 #
;#####

(FONCTION-SUJET
  (DIMENSION3)

```

```
(init-constants '(((sujet) :cat ("N" "Cl" "S") :function sujet)))
)
```

```
(PROJECTION-PHRASTIQUE
  (FONCTION-SUJET) )
```

```
(SUJET-REALISE
  (PROJECTION-PHRASTIQUE)
  (init-equations '(, (cr-eq sujet "t" "num" vmarq "t" "num")
                      , (cr-eq sujet "t" "gen" vmarq "t" "gen")
                      , (cr-eq sujet "t" "pers" vmarq "t" "pers")
                      , (cr-eq sujet "t" "num" S_d "b" "suj-num")
                      , (cr-eq sujet "t" "gen" S_d "b" "suj-gen")
                      , (cr-eq sujet "t" "pers" S_d "b" "suj-pers")
                      , (cr-eq S_d "b" "mode" "ind/subj")))) )
```

```
(SUJET-CANONIQUE
  (SUJET-REALISE)
  (init-equations '(, (cr-eq S_d "b" "wh" sujet () "wh")
                      , (cr-eq sujet "t" "fonc" "sujet")))) )
```

```
(SUJET-NOM
  (SUJET-CANONIQUE) )
```

```
(SUJ-NOM-CAN-SUBST
  (SUJET-NOM)
  (init-constants '(((sujet) :cat ("N") :qual subst)))
  (init-description '((S_d nil) ((sujet nil)) ((vmarq t)))) )
```

```
(SUJ-NOM-CAN-ANCRE
  (SUJET-NOM)
  (init-constants '(((ancre) :cat ("N") :qual head)
                    ((sujet) :cat ("N") :qual head)))
  (init-description '((S_d nil) ((sujet nil)) ((vmarq t)))) )
  (init-description '((sujet nil) ((ancre t)))) )
```

```
(SUJET-NON-CANO
  (SUJET-REALISE) )
```

```
(SUJ-NOM-INVERSE
  (SUJET-NON-CANO) )
```

```
(SUJ-NOM-INV-SUBST
  (SUJ-NOM-INVERSE)
  (init-constants '(((sujet) :cat ("N") :qual subst)))
```

```

(init-description '((S_d nil) ((vsupp t)) ((sujet nil)))) )

(SUJ-NOM-INV-ANCRE
  (SUJ-NOM-INVERSE)
  (init-constants '(((ancre) :cat ("N") :qual head)
                    ((sujet) :cat ("N") :qual head)))
  (init-description '((S_d nil) ((vsupp t)) ((sujet nil))))
  (init-description '((sujet nil) ((ancre t)))) )

(FONCTION-OBJET
  (DIMENSION3)
  (init-constants '(((objet) :cat ("N" "Cl" "S") :function objet))) )

(OBJET-CANONIQUE
  (FONCTION-OBJET) )

(OBJET-NOM
  (OBJET-CANONIQUE)
  (init-constants '(((objet) :cat ("N" "Cl"))))
  (init-equations '(, (cr-eq objet "t" "fonc" "objet"))) )

(OBJ-NOM-SUBST
  (OBJET-NOM)
  (init-constants '(((objet) :qual subst)))
  (init-description '((S_d nil) ((vsupp t)) ((objet nil)))) )

(OBJ-NOM-ANCRE
  (OBJET-NOM)
  (init-constants '(((ancre) :cat ("N") :qual head)
                    ((objet) :cat ("N") :qual head)))
  (init-description '((S_d nil) ((vsupp t)) ((objet nil))))
  (init-description '((objet nil) ((ancre t)))) )

(FONCTION-PREP-OBJ
  (DIMENSION3)
  (init-constants '(((prepobjet) :cat ("N" "S") :function prep-obj))) )

(PREP-OBJ-NOM
  (FONCTION-PREP-OBJ)
  (init-constants '(((prepobjet) :cat ("N") :qual subst))) )

(PREP-OBJ-CAN
  (PREP-OBJ-NOM)
  (init-constants-with-index 'prep-obj '(((sp_prep) :cat ("SP"))

```

```

((prep) :cat ("Prep") :qual co_head)))
(init-equations '(, (cr-eq prep-obj "t" "fonc" "prep-obj")))
(init-simple-desc '(sp_prep (prep) (prepobjet)))
(init-description '((S_d nil) ((vsupp t)) ((sp_prep nil)))) )

(PREP-OBJ-CAN2
  (PREP-OBJ-NOM)
  (init-constants-with-index 'prep-obj '(((sp_prep) :cat ("SP"))
                                           ((prep) :cat ("Prep") :qual co_head)))
  (init-constants '(((racine) :cat ("N") :ind "c"))))
  (init-simple-desc '(sp_prep (prep) (prepobjet)))
  (init-simple-desc '(racine (ancree) (sp_prep))) )

(FONCTION-DE-OBJ
  (DIMENSION3) )

(DE-OBJ-NOM
  (FONCTION-DE-OBJ)
  (init-constants '(((deobjet) :cat ("N" "Cl") :qual subst)))) )

(DE-OBJ-NOM-CAN
  (DE-OBJ-NOM)
  (init-constants-with-index 'de-obj '(((sp_de) :cat ("SP"))
                                         ((prep_de) :cat ("Prep"))
                                         ((deobjet))))
  (init-equations '(, (cr-eq sp_de "t" "fonc" "de-obj")))
  (init-constants '(((lex_prep_de) :cat ("de") :qual no_qual)))
  (init-simple-desc '(sp_de (prep_de) (deobjet)))
  (init-simple-desc '(prep_de (lex_prep_de)))
  (init-description '((S_d nil) ((vsupp t)) ((sp_de nil)))) )

(DE-OBJ-NOM-CAN2
  (DE-OBJ-NOM)
  (init-constants-with-index 'de-obj '(((sp_de) :cat ("SP"))
                                         ((prep_de) :cat ("Prep"))
                                         ((deobjet))))
  (init-constants '(((lex_prep_de) :cat ("de") :qual no_qual)))
  (init-constants '(((racine) :cat ("N") :ind "c"))))
  (init-simple-desc '(sp_de (prep_de) (deobjet)))
  (init-simple-desc '(prep_de (lex_prep_de)))
  (init-simple-desc '(racine (ancree) (sp_de))) )

(FONCTION-PAR-OBJ
  (DIMENSION3) )

```

```

(PAR-OBJ-OMIS
  (FONCTION-PAR-OBJ)
  (init-constants '(((non_realise_physiquement) :function par-obj))) )

(PAR-OBJ
  (FONCTION-PAR-OBJ)
  (init-constants-with-index 'parObj '(((nom) :cat ("N") :function parObj )
                                         ((SPpar) :cat ("SP") )
                                         ((Preppar) :cat ("Prep")))))
  (init-equations '(, (cr-eq SPpar "t" "fonc" "par-obj")))
  (init-constants '(((lexPrep_par) :cat ("par") :qual no_qual)))
  (init-description '((S_d nil) ((vsupp t)) ((SPpar nil))))
  (init-simple-desc '(SPpar (Preppar (lexPrep_par)) (nom))) )

(CLITIQUES
  (DIMENSION3) )

(COL1
  (CLITIQUES) )

(PREVERBAL
  (COL1)
  (init-constants '(((supsv) :cat ("V") :ind "s" :qual no_qual)
                    ((suj_cl) :cat ("Cl") :function sujet)))
  (init-equations '(, (cr-eq suj_cl () "cas" "nom")
                    , (cr-eq supsv "b" "aux" "-")
                    , (cr-eq supsv "t" "inv" "-")
                    , (cr-eq supsv "t" "neg" "-")
                    , (cr-eq S_d "b" "wh" "-")))
  (init-simple-desc '(S_d (supsv)))
  (init-description '((supsv nil) ((suj_cl nil)) ((vmarg t)))) )

(SUJET-IMPERS
  (COL1)
  (init-constants '(((supsv) :cat ("V") :ind "s" :qual no_qual)
                    ((suj_cl) :cat ("Cl"))
                    ((il) :cat ("il") :function sujet :qual no_qual)))
  (init-equations '(, (cr-eq suj_cl () "cas" "nom")
                    , (cr-eq supsv "b" "aux" "-")
                    , (cr-eq supsv "t" "inv" "-")
                    , (cr-eq supsv "t" "neg" "-")
                    , (cr-eq S_d "b" "wh" "-")))
  (init-simple-desc '(S_d (supsv)))
  (init-description '((supsv nil) ((suj_cl nil)) ((vmarg t)))) )

```

```

(init-simple-desc '(su_j_cl (il))) )

(EXTRACTIONS
  (DIMENSION3) )

(OBJ-EXTRAIT
  (FONCTION-OBJET)
  (init-equations-masquantes `((, (cr-eq vmarq "b" "part-num" objet () "num")
                                   , (cr-eq vmarq "b" "part-gen" objet () "gen")))) )

(CLIVAGE
  (EXTRACTIONS)
  (init-constants '(((Sf) :cat ("S") :ind "f")
                    ((Vnx) :cat ("V") :ind "nx")
                    ((Cl_ce) :cat ("Cl") :ind "x")
                    ((ce) :cat ("ce") :qual no_qual)
                    ((v_cop) :cat ("V"))
                    ((etre) :cat ("etre") :qual no_qual)))
  (init-equations `((, (cr-eq V_cop () "modif" "-")
                      , (cr-eq V_cop () "inv" "-")
                      , (cr-eq V_cop () "pers" "3")
                      , (cr-eq V_cop () "pers" Cl_ce "b" "pers")
                      , (cr-eq Sf "b" "wh" "-")
                      , (cr-eq Sf "b" "mode" V_cop "t" "mode")
                      , (cr-eq Sf "b" "inv" V_cop "t" "inv")
                      , (cr-eq S_d "t" "mode" "ind/subj"))))
  (init-simple-desc '(Sf (Vnx (Cl_ce (ce)) (v_cop (etre))))) )

(CEST...QUI
  (CLIVAGE)
  (init-constants '(((foyer_suj) :cat ("N" "S") :function sujet)
                    ((comp_qui) :cat ("C") :constraints "NA")
                    ((qui) :cat ("qui") :qual no_qual)))
  (init-simple-desc '(Sf (Vnx) (foyer_suj) (S_d (comp_qui (qui)))))
  (init-description '((S_d nil) ((comp_qui nil)) ((vmarq t)))) )

(SUJET-CLIVE
  (CEST...QUI) )

(SUJ-CLIV-SUBST
  (SUJET-CLIVE)
  (init-equations `((, (cr-eq V_cop () "num" foyer_suj () "num")))
  (init-constants '(((foyer_suj) :cat ("N") :qual subst)))) )

```

```

(SUJ-CLIV-ANCRE
  (SUJET-CLIVE)
  (init-equations `((cr-eq V_cop () "num" foyer_suj "t" "num")))
  (init-constants '(((foyer_suj) :cat ("N") :qual head)
                     ((ancre) :cat ("N") :qual head)))
  (init-description '((foyer_suj nil) ((ancre t)))) )

(CEST...QUE
  (CLIVAGE)
  (init-constants '(((foyer) :cat ("N" "S" "SP"))
                     ((comp_que) :cat ("C") :constraints "NA")
                     ((que) :cat ("que") :qual no_qual)
                     ((S_c) :cat ("S") :ind "c" :constraints "NA"))))
  (init-equations `((cr-eq S_d "b" "wh" "-")
                     , (cr-eq S_c "b" "mode" S_d "t" "mode")
                     , (cr-eq V_cop () "num" foyer "t" "num")))
  (init-simple-desc '(Sf (Vnx) (foyer) (S_c (comp_que (que)) (S_d)))) )

(OBJET_CLIVE
  (CEST...QUE)
  (init-constants '(((foyer) :cat ("N" "S") :function objet)))) )

(OBJ-CLIV-SUBST
  (OBJET_CLIVE)
  (init-constants '(((foyer) :cat ("N") :qual subst)))) )

(OBJ-CLIV-ANCRE
  (OBJET_CLIVE)
  (init-constants '(((foyer) :cat ("N") :qual head)
                     ((ancre) :cat ("N") :qual head)))
  (init-description '((foyer nil) ((ancre t)))) )

(PREP-OBJ-NOM-NON-CAN
  (PREP-OBJ-NOM) )

(PREP-OBJ-EXTRAIT
  (PREP-OBJ-NOM-NON-CAN) )

(PREP-OBJ-CLIV
  (PREP-OBJ-EXTRAIT CEST...QUE)
  (init-constants-with-index 'prep-obj '(((foyer) :cat ("SP"))
                                           ((prep) :cat ("Prep") :qual co_head)
                                           ((prepobjet) :cat ("N") :qual subst)))
  (init-simple-desc '(foyer (prep) (prepobjet)))) )

```

```

(DE-OBJ-NOM-NON-CAN
  (DE-OBJ-NOM) )

(DE-OBJ-EXTRAIT
  (DE-OBJ-NOM-NON-CAN) )

(DE-OBJ-CLIVE
  (DE-OBJ-EXTRAIT CEST...QUE)
  (init-constants-with-index 'de-obj '(((foyer) :cat ("SP"))
                                ((prep_de) :cat ("Prep"))
                                ((deobjet) :cat ("N") :qual subst)))
  (init-constants '(((lexprep_de) :cat ("de") :qual no_qual)))
  (init-simple-desc '(foyer (prep_de) (deobjet)))
  (init-simple-desc '(prep_de (lexprep_de))) )

(RELATIVISATION
  (EXTRACTIONS) )

(SUJ-RELATIVISE
  (RELATIVISATION)
  (init-constants '(((root) :cat ("N") :ind "r")
                    ((rel) :cat ("N") :function sujet)
                    ((comp) :cat ("C"))
                    ((lexComp) :cat ("qui") :qual no_qual)))
  (init-equations `((cr-eq S_d "t" "mode" "ind/subj")
                    , (cr-eq S_d "t" "inv" "-")
                    , (cr-eq root "b" "modif" "+")
                    , (cr-eq root "t" "det" "+")))
  (init-description '(((root nil) ((rel t)) ((S_d nil))))
  (init-description '(((S_d nil) ((comp nil)) ((vmarq t))))
  (init-simple-desc '(comp (lexcomp))) )

(SUJ-COMP-SUBST
  (SUJ-RELATIVISE)
  (existing-constant '(((rel) :cat ("N") :qual subst)))
  (revise-qual 'foot 'rel)
  (init-equations `((cr-eq root "b" "num" rel () "num")
                    , (cr-eq root "b" "gen" rel () "gen")
                    , (cr-eq root "b" "det" rel () "det")
                    , (cr-eq root "b" "conc" rel () "conc")
                    , (cr-eq root "b" "hum" rel () "hum")
                    , (cr-eq rel () "wh" "-")))
  (init-simple-desc '(root (rel) (S_d))) )

(SUJ-COMP-ANCRE

```

```

(SUJ-RELATIVISE)
(init-constants '(((rel) :cat ("N") :qual head)
                  ((ancre) :cat ("N") :ind "" :qual head)))
(init-equations '(, (cr-eq root "b" "num" rel () "num")
                   , (cr-eq root "b" "gen" rel () "gen")
                   , (cr-eq root "b" "det" rel () "det")
                   , (cr-eq root "b" "conc" rel () "conc")
                   , (cr-eq root "b" "hum" rel () "hum")
                   , (cr-eq rel () "wh" "-"))
(init-description '(((rel nil) ((ancre t)))) )

(REL-COMPLEMENTS
 (RELATIVISATION)
 (init-constants '(((root) :cat ("N") :ind "r")
                   ((rel) :cat ("N"))
                   ((S_r) :cat ("S") :ind "r"))))
 (init-equations '(, (cr-eq S_d "b" "wh" "-")
                   , (cr-eq S_r "b" "mode" S_d "t" "mode")
                   , (cr-eq S_r "t" "inv" "-/n")
                   , (cr-eq root "b" "num" rel () "num")
                   , (cr-eq root "b" "gen" rel () "gen")
                   , (cr-eq root "b" "det" rel () "det")
                   , (cr-eq root "b" "conc" rel () "conc")
                   , (cr-eq root "b" "hum" rel () "hum")
                   , (cr-eq rel () "wh" "-"))
 (init-description '(((root nil) ((rel t)) ((S_r nil))))
 (init-simple-desc '(S_r (S_d))) )

(OBJ-REL-SUBST
 (REL-COMPLEMENTS)
 (existing-constant '(((rel) :cat ("N") :qual subst)))
 (init-constants '(((rel) :function objet)
                   ((comp) :cat ("C"))
                   ((lexcomp) :cat ("que") :qual no_qual)))
 (revise-qual 'foot 'rel)
 (init-simple-desc '(root (rel) (S_r)))
 (init-simple-desc '(S_r (comp (lexcomp)) (S_d))) )

(OBJ-REL-ANCRE
 (REL-COMPLEMENTS)
 (init-constants '(((rel) :cat ("N") :function objet :qual head)
                   ((ancre) :cat ("N") :ind "" :qual head)
                   ((comp) :cat ("C"))
                   ((lexcomp) :cat ("que") :qual no_qual)))
 (init-simple-desc '(S_r (comp (lexcomp)) (S_d)))

```

```

(init-description '((rel nil) ((ancre t)))) )

(PREP-OBJ-REL
  (REL-COMPLEMENTS)
  (init-constants-with-index 'prep-obj '(((sp) :cat ("SP"))
                                           ((prep) :cat ("Prep") :qual co_head)
                                           ((nrel) :function prep-obj :cat ("N")))))
  (init-constants '(((rel) :qual foot)))
  (init-simple-desc '(root (rel) (S_r (sp) (S_d))))
  (init-simple-desc '(sp (prep) (nrel))) )

(DONT
  (REL-COMPLEMENTS)
  (init-constants-with-index 'de-obj '(((rel) :cat ("N") :function de-obj)))
  (init-constants '(((comp) :cat ("C"))
                    ((lexcomp) :cat ("dont") :qual no_qual)))
  (revise-qual 'foot 'rel)
  (init-simple-desc '(root (rel) (S_r)))
  (init-simple-desc '(S_r (comp (lexcomp)) (S_d))) )

(QUESTIONS
  (EXTRACTIONS)
  (init-constants '(((S_r) :cat ("S") :ind "r")
                    ((quest))))
  (init-equations '(, (cr-eq S_r "b" "wh" "+")
                     , (cr-eq S_d "b" "wh" "-")
                     , (cr-eq S_r "b" "mode" S_d "t" "mode")
                     , (cr-eq S_r "b" "inv" S_d "t" "inv")
                     , (cr-eq quest () "wh" "+")))
  (init-description '((S_r nil) ((quest t)) ((S_d nil)))) )

(OBJ-INT-SUBST
  (QUESTIONS)
  (init-constants '(((quest) :cat ("N") :qual subst :function objet)))
  (init-simple-desc '(S_r (quest) (S_d))) )

(OBJ-INT-ANCRE
  (QUESTIONS)
  (init-constants '(((quest) :cat ("N") :function objet)
                    ((ancre) :cat ("N") :ind "" :qual head)))
  (init-description '((quest nil) ((ancre t)))) )

(PREP-OBJ-INT
  (QUESTIONS)
  (init-constants-with-index 'prep-obj '(((quest) :cat ("N") :function prep-obj)
                                           ((prep) :cat ("Prep") :qual co_head)
                                           ((nrel) :function prep-obj :cat ("N")))))

```

```

                ((SP) :cat ("SP"))
                ((prep) :cat ("Prep") :qual co_head)))
(init-simple-desc '(S_r (SP)))
(init-simple-desc '(SP (prep) (quest))) )

(DE-OBJ-INT
 (QUESTIONS)
 (init-constants-with-index 'de-obj '(((quest) :cat ("N") :function de-obj)
                                     ((SP_de) :cat ("SP"))
                                     ((prep_de) :cat ("Prep")))))
 (init-constants '(((lexprep_de) :cat ("de") :qual no_qual)))
 (init-simple-desc '(S_r (SP_de)))
 (init-simple-desc '(SP_de (prep_de) (quest)))
 (init-simple-desc '(prep_de (lexprep_de))) )

```


Annexe B

Implémentation des problèmes d'ordre

B.1 La solution de K. Gerdes

L'implémentation suivante reprend la proposition de [Gerdes, 2002], dans le langage décrit par [Candito, 1999a] pour ordonner un objet direct et un objet indirect (l'objet direct précédant l'objet indirect). On notera qu'elle permet d'ordonner de la même façon le *SP* locatif à substituer.

Pour ne pas alourdir inutilement cette description qui permet de générer 2 arbres pour la famille $n0Vn1$, et 4 arbres pour les familles $n0Vn1_pn2_$ et $n0Vn1_loc2_$, les descriptions d'arbres ont été simplifiées et ne comportent aucune information sur les traits.

```
(setq *families* '("n0Vn1" "n0Vn1_pn2_" "n0Vn1_loc2_"))

(setq *redistributions* '(ACTIF))

(setq *sujets* '((SUJ-NOM) :function sujet))
(setq *objs* '((OBJ-OMIS OBJ-NOM) :function objet))
(setq *pobjs* '((PREP-OBJ-OMIS PREP-OBJ-NOM-COHEAD-REALISE) :function pObj))
(setq *locs* '((LOC-OMIS LOC-REALISE) :function locat))

(setq *realizations* '(*sujets* *objs* *pobjs* *locs*))

(defparameter *inheritance-constraints* '())
(defparameter *neg-inheritance-constraints* '())
(defparameter *disjunctive-nodes* '())
(defparameter *constrained-crossings* '())

; Sous catégorisation initiale
; Dimension 1
```

; Définition d'un nom prédicatif

```
(DIMENSION1
  (TOP))
```

```
(SUJET_INITIAL
  (DIMENSION1)
  (init-constants '(((arg0) :cat ("N") :ind "0" :function sujet))))
```

```
(OBJ-INITIAL
  (DIMENSION1)
  (init-constants '(((arg1) :cat ("N") :ind "1" :function objet))))
```

```
(PREP-OBJ-INITIAL
  (DIMENSION1)
  (init-constants '(((arg2) :cat ("N") :ind "2" :function pObj))))
```

```
(LOCATIF-INITIAL
  (DIMENSION1)
  (init-constants '(((arg2) :cat ("SP") :ind "2" :function locat))))
```

; Classes terminales de la dimension 1
; 1 argument réalisé

```
(n0Vn1
  (SUJET_INITIAL OBJ-INITIAL)
  (init-constants '(((arg0) :cat ("N"))))
  (init-name "n0Vn1" :direction :left))
```

; 2 arguments réalisés

```
(n0Vn1_pn2_
  (n0Vn1 PREP-OBJ-INITIAL)
  (init-constants '(((arg2) :cat ("N") :function pObj)))
  (init-name "(pn2)" :direction :right))
```

```
(n0Vn1_loc2_
  (n0Vn1 LOCATIF-INITIAL)
  (init-constants '(((arg2) :cat ("SP") :function locat)))
  (init-name "(loc2)" :direction :right))
```

; Fin classes terminales dimension 1

; Redistribution

; Dimension 2

```
(DIMENSION2
  (TOP))
```

```
(MORPHO-VERBALE
  (DIMENSION2)
  (init-constants '(((S_d) :cat ("S"))
                    ((verbe) :cat ("V") )))
  (init-description '((S_d nil) ((verbe t))))))
```

```
(ACTIF
  (MORPHO-VERBALE)
  (init-name "Actif-" :direction :left))
```

; Réalisation
; Dimension 3

```
(DIMENSION3
  (TOP))
```

```
(FONCTION-SUJET
  (DIMENSION3))
```

```
(SUJ-NOM
  (FONCTION-SUJET)
  (init-constants '(((sujet) :cat ("N") :function sujet)
                    ((verbe) :cat ("V") )))
  (init-description '((S_d nil) ((sujet t)) ((verbe t))))))
```

```
(FONCTION-OBJET
  (DIMENSION3))
```

```
(OBJ-OMIS
  (FONCTION-OBJET)
  (init-name "n1-omis" :direction :right))
```

```
(OBJ-NOM
  (FONCTION-OBJET)
  (init-constants '(((suparg2) :qual no_qual)))
; L'objet attend potentiellement un autre complément à sa droite.
  (init-constants '(((objet) :cat ("N") :ind "1" :qual subst)))
  (init-description '((S_d nil) ((verbe t)) ((objet t)) ((suparg2 t))))))
```

;Syntagme prépositionnel

```

(FONCTION-POBJ
  (DIMENSION3))

(PREP-OBJ-OMIS
  (FONCTION-POBJ)
  (init-name "-sp2-omis" :direction :right))

(PREP-OBJ-NOM-ou-PHR-REALISE
  (FONCTION-POBJ)
  (init-constants '(((suparg2) :cat ("SP") :ind "2")
                    ((verbe) :cat ("V")) ))
  (init-simple-desc '(S_d (verbe) (suparg2))))

(PREP-OBJ-NOM-REALISE
  (PREP-OBJ-NOM-ou-PHR-REALISE)
  (init-constants '(((arg2) :cat ("N") :ind "2")
                    ((parg2) :cat ("Prep") :ind "2"))))
  (init-simple-desc '(suparg2 (parg2) (arg2)))
  (init-name "-sp2-nom" :direction :right))

(PREP-OBJ-NOM-COHEAD-REALISE
  (PREP-OBJ-NOM-REALISE)
  (init-constants '(((parg2) :qual co_head))))

(LOCATIF
  (DIMENSION3))

(LOC-OMIS
  (LOCATIF)
  (init-name "-locatif_omis" :direction :right))

(LOC-REALISE
  (LOCATIF PREP-OBJ-NOM-ou-PHR-REALISE)
  (revise-qual 'subst 'suparg2)
  ; Révision du qual Suparg2 est à substituer dans le cas d'un Locatif
  (init-constants '(((suparg2) :cat ("SP") :ind "2" :qual subst)))
  (init-simple-desc '(S_d (verbe) (suparg2))))

```

B.2 Notre proposition

Nous reprenons ici le même exemple, mais en ôtant les notions d'ordre entre compléments de la dimension 3. Celles-ci sont renseignées, à part, en dimension 4. Le croisement se fait

alors par contraintes de croisement. Si l'avantage de notre méthode n'est pas flagrant ici, on imagine cependant facilement le gain de temps en maintenance lors de l'écriture d'un fichier volumineux, permettant la représentation de toutes les familles d'une hiérarchie...

Nous ne reprenons pas ici les classes de dimension 1 et de dimension 2 qui restent bien évidemment inchangées par rapport à l'exemple précédent. Nous renseignons donc simplement les classes de dimension 3 et de dimension 4.

```
(setq *families* '("n0Vn1" "n0Vn1_pn2_" "n0Vn1_loc2_"))

(setq *redistributions* '(ACTIF))

(setq *sujets* '((SUJ-NOM) :function sujet))
(setq *objs* '((OBJ-OMIS OBJ-NOM) :function objet))
(setq *pobjs* '((PREP-OBJ-OMIS PREP-OBJ-NOM-COHEAD-REALISE) :function pObj))
(setq *locs* '((LOC-OMIS LOC-REALISE) :function locat))

(setq *realizations* '(*sujets* *objs* *pobjs* *locs*))

(defparameter *inheritance-constraints* '())
(defparameter *neg-inheritance-constraints* '())
(defparameter *disjunctive-nodes* '())

(defparameter *constrained-crossings* '(
(N<SP      OBJ-NOM      PREP-OBJ-NOM-ou-PHR-REALISE)))

...

(DIMENSION3
  (TOP))

(FONCTION-SUJET
  (DIMENSION3))

(SUJ-NOM
  (FONCTION-SUJET)
  (init-constants '(((sujet) :cat ("N") :function sujet)
                    ((verbe) :cat ("V") )))
  (init-description '((S_d nil) ((sujet t)) ((verbe t)))))

(FONCTION-OBJET
  (DIMENSION3))

(OBJ-OMIS
  (FONCTION-OBJET))
```

```

  (init-name "n1-omis" :direction :right))

(OBJ-NOM
  (FONCTION-OBJET)
  (init-constants '(((objet) :cat ("N") :ind "1" :qual subst)))
  (init-description '((S_d nil) ((verbe t)) ((objet t)))))

;Syntagme prépositionnel

(FONCTION-POBJ
  (DIMENSION3))

(PREP-OBJ-OMIS
  (FONCTION-POBJ)
  (init-name "-sp2-omis" :direction :right))

(PREP-OBJ-NOM-ou-PHR-REALISE
  (FONCTION-POBJ)
  (init-constants '(((suparg2) :cat ("SP") :ind "2")
                    ((verbe) :cat ("V")) ))
  (init-simple-desc '(S_d (verbe) (suparg2))))

(PREP-OBJ-NOM-REALISE
  (PREP-OBJ-NOM-ou-PHR-REALISE)
  (init-constants '(((arg2) :cat ("N") :ind "2")
                    ((parg2) :cat ("Prep") :ind "2"))
  (init-simple-desc '(suparg2 (parg2) (arg2)))
  (init-name "-sp2-nom" :direction :right))

(PREP-OBJ-NOM-COHEAD-REALISE
  (PREP-OBJ-NOM-REALISE)
  (init-constants '(((parg2) :qual co_head)))))

(LOCATIF
  (DIMENSION3))

(LOC-OMIS
  (LOCATIF)
  (init-name "-locatif_omis" :direction :right))

(LOC-REALISE
  (LOCATIF PREP-OBJ-NOM-ou-PHR-REALISE)
  (init-constants '(((suparg2) :cat ("SP") :ind "2" :qual subst)))
  (init-simple-desc '(S_d (verbe) (suparg2))))

```

```
;Placement des constituants
```

```
(DIMENSION4  
  (TOP))
```

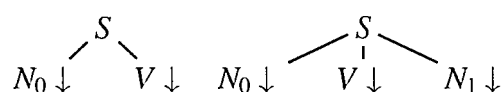
```
(PLACEMENT_DES_CONSTITUANTS  
  (DIMENSION4))
```

```
(N<SP  
  (PLACEMENT_DES_CONSTITUANTS)  
  (init-simple-desc '(S_d (objet) (suparg2))))
```

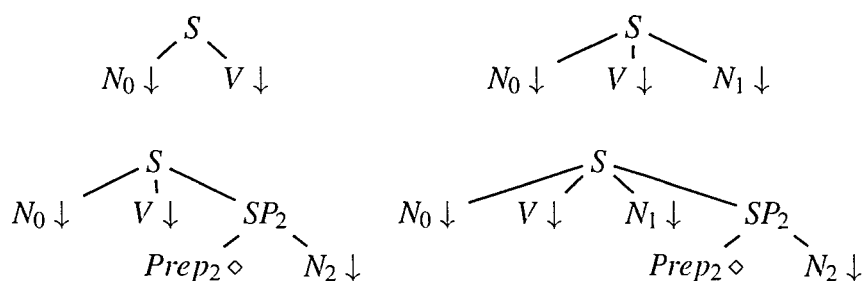
B.3 Les arbres générés

Les arbres suivants ont été générés à l'aide des 2 fichiers de description. Chaque fichier a généré les mêmes arbres en nombre identique. Il est clair que l'exemple choisi n'est pas juste linguistiquement parlant (tout du moins pour le français), mais il permet d'illustrer le fonctionnement de la Métagrammaire, tout en donnant une idée de la tâche à accomplir pour générer quelques arbres. Rassurons le lecteur, ce sont les premiers arbres qui demandent le plus de travail.

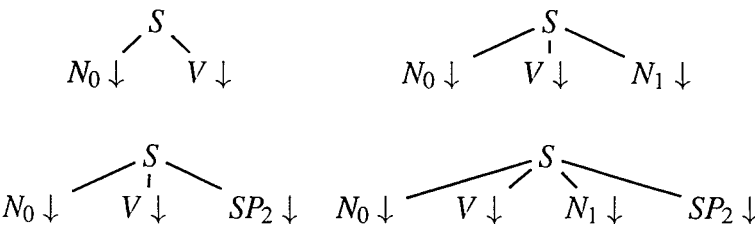
B.3.1 Famille n0Vn1



B.3.2 Famille n0Vn1(pn2)



B.3.3 Famille n0Vn1(loc2)



Annexe C

Extrait des lexiques syntaxiques

Nous fournissons dans cette annexe un extrait des lexiques syntaxiques développés pour les familles à nom prédicatif.

Nous avons préférentiellement renseigné les noms prédicatifs qui apparaissaient le plus fréquemment en corpus. Pour le choix du verbe support, le corpus n'étant pas forcément exhaustif, nous nous sommes appuyé sur notre intuition, et sur les listes déjà existantes constituées entre autres par [Giry-Schneider, 1978, Giry-Schneider, 1987, Gross, 1989] et [Zinglé et Brobeck-Zinglé, 2003]. Le travail reste à ce stade très perfectible, mais nous avons tenté de renseigner autant que faire se peut les informations concernant les différentes variantes de support, les converses et les opérateurs lexicaux.

```
#####  
;# Lexique syntaxique des noms prédicatifs #  
#####  
;# Contact : Sébastien Barrier - LLF      #  
;#           sbarrier@linguist.jussieu.fr #  
#####  
;# Pour les traits :                      #  
;# vsup indique le verbe support          #  
;# vconv indique le conversif             #  
;# vcaus indique l'opérateur causatif     #  
#####
```

INDEX: analyse/1

ENTRY: analyse

POS: N

FAM: n0vNden1

FS: V:<vsup>=faire, V:<vconv>=subir

;EX: Max fait une analyse de ce texte

INDEX: analyse/1
ENTRY: analyse
POS: N
FAM: n0vANden1
FS: V:<vsup>=procéder, V:<vconv>=subir
;EX: Max procède à l'analyse du terrain

INDEX: arbitrage/1
ENTRY: arbitrage sur
POS: N Prepl
FAM: n0vNpn1
FS: V:<vsup>=faire, V:<vconv>=subir
;EX: Max fait un arbitrage sur cette situation

INDEX: arbitrage/1
ENTRY: arbitrage
POS: N
FAM: n0vNden1
FS: V:<vsup>=faire
;EX: Max fait un arbitrage de la situation

INDEX: arbitrage/1
ENTRY: arbitrage sur
POS: N Prepl
FAM: n0vANpn1
FS: V:<vsup>=procéder
;EX: Max procède à un arbitrage sur cette situation

INDEX: arbitrage/1
ENTRY: arbitrage
POS: N
FAM: n0vANden1
FS: V:<vsup>=procéder
;EX: Max procède à un arbitrage de la situation

INDEX: attentat/1
ENTRY: attentat contre
POS: N Prepl
FAM: n0vNpn1
FS: V:<vsup>=commettre/réaliser/perpétrer, V:<vconv>=essayer
;EX: Les bandits ont commis un attentat contre la Préfecture

INDEX: bilan/1
ENTRY: bilan

POS: N
 FAM: n0vNden1
 FS: V:<vsup>=dresser/effectuer/faire
 ;EX: Max dresse le bilan de son action

INDEX: bilan/1
 ENTRY: bilan
 POS: N
 FAM: n0vANDen1
 FS: V:<vsup>=procéder
 ;EX: Max procède au bilan de sa vie

INDEX: communication/1
 ENTRY: communication
 POS: N
 FAM: n0vNan1
 FS: V:<vsup>=faire
 ;EX: Max fait une communication à cette assemblée

INDEX: concession/1
 ENTRY: concession
 POS: N
 FAM: n0vNan1
 FS: V:<vsup>=faire, V:<vconv>=obtenir
 ;EX: Max a fait énormément de concessions à Luc

INDEX: confiance/1
 ENTRY: confiance en
 POS: N Prep1
 FAM: n0vNpn1
 FS: V:<vsup>=avoir, V:<vconv>=avoir/inspirer, V:<vcaus>=donner,
 #PASSIF-
 ;EX: Max a confiance en Luc

INDEX: confiance/1
 ENTRY: confiance pour
 POS: N Prep1
 FAM: n0vNpn1
 FS: V:<vsup>=ressentir, V:<vconv>=avoir/inspirer, V:<vcaus>=donner,
 #PASSIF-
 ;EX: Max ressent de la confiance pour Luc

INDEX: connaissance/1
 ENTRY: connaissance
 POS: N

FAM: n0vNden1

FS: V:<vsup>=avoir/prendre, #PASSIF-

;EX: Max a eu connaissance de cet événement

INDEX: déclaration/1

ENTRY: déclaration

POS: N

FAM: n0vNan1

FS: V:<vsup>=adresser/faire

;EX: Max fait une déclaration à cette assemblée

INDEX: demande/1

ENTRY: demande

POS: N

FAM: n0vNden1an2

FS: V:<vsup>=adresser/effectuer/faire/présenter/soumettre

;EX: Max effectue une demande de crédit au guichetier

INDEX: demande/1

ENTRY: demande auprès

POS: N Prep2

FAM: n0vNden1pn2

FS: V:<vsup>=adresser/effectuer/faire/présenter/soumettre

;EX: Max effectue une demande de crédit auprès de la banque

INDEX: démonstration/1

ENTRY: démonstration

POS: N

FAM: n0vNden1an2

FS: V:<vsup>=faire

;EX: Max fait la démonstration de sa force à Léa

INDEX: effet/1

ENTRY: effet sur

POS: N Prep1

FAM: n0vNpn1

FS: V:<vsup>=avoir/faire

;EX: Ce médicament a un effet surprenant sur Max

INDEX: entorse/1

ENTRY: entorse

POS: N

FAM: n0vNan1

FS: V:<vsup>=commettre/faire

;EX: Max commet une entorse au règlement

INDEX: entorse/1
 ENTRY: entorse envers
 POS: N Prep1
 FAM: n0vNpn1
 FS: V:<vsup>=commettre/faire
 ;EX: Max commet une entorse envers le règlement

INDEX: entretien/1
 ENTRY: entretien
 POS: N
 FAM: n0vNan1
 FS: V:<vsup>=accorder, V:<vconv>=obtenir
 ;EX: Max accorde un entretien au journaliste

INDEX: entretien/1
 ENTRY: entretien avec
 POS: N Prep1
 FAM: n0vNpn1
 FS: V:<vsup>=avoir, V:<conv>=obtenir,
 #PASSIF-
 ;EX: Max a un entretien avec la direction

INDEX: massacre/1
 ENTRY: massacre
 POS: N
 FAM: n0vN
 FS: V:<vsup>=faire/commettre
 ;EX: Les barbares ont commis un massacre

INDEX: massacre/1
 ENTRY: massacre
 POS: N
 FAM: n0vANden1
 FS: V:<vsup>=procéder
 ;EX: Les barbares ont procédé au massacre de la population

INDEX: pouvoir/1
 ENTRY: pouvoir sur
 POS: N Prep1
 FAM: n0vNpn1
 FS: V:<vsup>=avoir, V:<vcaus>=donner, #PASSIF-
 ;EX: Max a le pouvoir sur la population

INDEX: promesse/1

ENTRY: promesse

POS: N

FAM: n0vNan1

FS: V:<vsup>=donner/faire, V:<vconv>=avoir

;EX: Max fait une promesse à Léa

INDEX: regret/1

ENTRY: regret

POS: N

FAM: n0vN

FS: V:<vsup>=avoir/éprouver/ressentir, #PASSIF-

;EX: Max éprouve des regrets

INDEX: regret/1

ENTRY: regret

POS: N

FAM: n0vNdes1

FS: V:<vsup>=avoir, #PASSIF-

;EX: Max a le regret d'annoncer son départ

INDEX: rôle/1

ENTRY: rôle dans

POS: N Prep1

FAM: n0vNpn1

FS: V:<vsup>=avoir/jouer

;EX: Max joue un rôle dans la société

INDEX: visite/1

ENTRY: visite

POS: N

FAM: n0vNan1

FS: V:<vsup>=faire/rendre, V:<vconv>=recevoir

;EX: Max fait une visite éclair à Léa

INDEX: volonté/1

ENTRY: volonté

POS: N

FAM: n0vNdes1

FS: V:<vsup>=avoir/manifeste, V:<vcaus>=donner

;EX: Max manifeste la volonté de combattre les barbares

Annexe D

Catégories et traits utilisés par la grammaire TAG du français

Nous fournissons dans cette annexe l'ensemble des catégories et traits utilisés par la grammaire TAG du français. Tous n'ont pas été décrits ici : outre l'exposé des chapitres précédents, cette annexe reprend donc également les travaux de [Abeillé, 1991, Abeillé, 2002, Candito, 1999a, Barrier, 1999, Barrier et al., 2000, Boonen, 2001, Barrier, 2002] et [Barrier et Barrier, 2003].

D.1 Catégories utilisées

- A : groupe adjectival ou adjectif
- Adv : groupe adverbial ou adverbe
- C : complémenteur
- Cl : clitique
- Conj : conjonction de coordination
- Det : déterminant
- N : Groupe nominal, nom ou pronom
- Prep : préposition
- S : phrase
- SP : syntagme prépositionnel
- V : verbe

D.2 Traits utilisés

<i>Attribut</i>	<i>Valeur</i>	<i>Interprétation</i>	<i>Noeuds concernés</i>
ant	+, -	Anteposé	A, Adv, Prep
aux	avoir, etre, -	Sélection de l'auxiliaire	V
card	+, -	Cardinal	A, Det, N
cas	acc, dat, gen, loc, nom	Cas	V, Cl
cl	obj, suj, -	"Montée" des clitiques	V
compar	+, -	Forme comparative	A, Adv
conc	+, -	Concret	N
cop	+, -	Verbe attributif	V
def	+, -	Défini	N
degre	+, -	gradable	A, Adv
el	+, -	forme élidée	Det, A, Cl, C, Prep
facile	+, -	Constructions tough	A, V
fonc	a_objet, attr_suj, attr_obj, de_loc, de_obj, loc, objet, par_obj, prep_obj, sujet, obj_a, obj_de obj_par, obj_prep, cdn_a, cdn_de, cdn_par cdn_prep, larg, 2arg	Fonction grammaticale	A, Cl, Det, N, S, SP
gen	fem, masc	Genre	A, Det, N, V, S
hum	+, -	Humain	N
inv	+, -, n	Inversion du sujet	V, S
loc	+, -	Locatif	Adv, N, SP
man	+, -	Manière	Adv, N, SP
mode	imp, inf, ind, subj, part, ppart	Mode	V, S
modif	+, -	Modifié	A, Adv, N, SP, V
mass	+, -	Massique	N
neg	+, -	Négativé	A, Adv, Conj, Det, N, S, V
num	pl, sing	Nombre	A, Det, N, V, S
ord	+, -	Ordinal	A, Det, N
pers	1, 2, 3	Personne	N, S, V
partitif	+, -	Partitif	Det, N
part-gen	fem, masc	Genre du participe passé	V
part-num	pl, sing	Nombre du participe passé	V
passif	de, par, -	Passif	V, N
poss	+, -	Possessif	Det

<i>Attribut</i>	<i>Valeur</i>	<i>Interprétation</i>	<i>Noeuds concernés</i>
psycho	+, -	Nom psychologique	N
princ	+, -	Principale	V
processif	+, -	Processif	N
P-num	pl, sing	Nombre du possesseur	Det
P-pers	1, 2, 3	Personne du possesseur	Det
quant	+, -	Quantifieur	Det, N
refl	+, -	Réfléchi	Cl, V
suj-gen	fem, masc	Genre du contrôleur	S, V
suj-num	pl, sing	Nombre du contrôleur	S, V
suj-pers	1, 2, 3	Personne du contrôleur	S, V
temp	+, -	Temporel	Adv, N, SP
tense	cond, fut, imparf, past pres	Temps	V, S
vcaus	donner, mettre...	Forme de l'opérateur causatif	V
vconv	exercer, subir...	Forme du support converse	V
vsup	avoir, donner, faire...	Forme du verbe support	V
wh	+, -, rel	Interrogatif ou relatif	Det, N, V, S, SP