

# Table des matières

<b>Introduction générale</b>	<b>1</b>
1 Situation . . . . .	1
2 Présentation du travail . . . . .	5
<hr/>	
<b>I Pour une théorie descriptive et formelle de la langue</b>	<b>7</b>
<b>1 Positionnement théorique</b>	<b>9</b>
1 Une approche non-générative . . . . .	10
2 Une grammaire non-présomptive . . . . .	12
3 Un traitement non-modulaire . . . . .	16
4 Une analyse multi-dimensionnelle . . . . .	22
<b>2 Architecture du formalisme</b>	<b>25</b>
1 Grammaire . . . . .	26
2 Spécification des propriétés . . . . .	36
3 Analyse . . . . .	45
<b>3 Notre modèle de grammaire</b>	<b>51</b>
1 Structuration des informations intrinsèques . . . . .	51
2 Propriétés et spécification . . . . .	57
<b>4 D'autres modèles en GP</b>	<b>71</b>
1 Une grammaire d'unification modulaire en GP . . . . .	72
2 Un modèle catégoriel procédural en GP . . . . .	90
3 Une Métagrammaire en GP . . . . .	111

---

<b>II</b>	<b>Éléments de grammaire du français</b>	<b>125</b>
	<b>Introduction</b>	<b>127</b>
1	Objectifs . . . . .	127
2	Notations simplifiées . . . . .	129
3	Organisation de la partie . . . . .	132
<b>5</b>	<b>Constructions de base</b>	<b>135</b>
1	Hierarchie générale des constructions . . . . .	135
2	Constructions de Base . . . . .	136
<b>6</b>	<b>Constructions nominales</b>	<b>145</b>
1	Noms construits (N) . . . . .	146
2	Syntagmes Nominaux (SN) . . . . .	148
<b>7</b>	<b>Constructions verbales</b>	<b>157</b>
1	Verbes construits (V) . . . . .	157
2	Syntagmes Verbaux (SV) . . . . .	180
<b>8</b>	<b>Constructions propositionnelles</b>	<b>193</b>
1	Propositions . . . . .	195
<b>9</b>	<b>Constructions de subordination</b>	<b>203</b>
1	Subordonnants (Sub) . . . . .	204
2	Conjoints de subordination (Conj-s) . . . . .	208
<b>10</b>	<b>Constructions adjectivales</b>	<b>215</b>
1	Adjectifs construits (Adj) . . . . .	216
<b>11</b>	<b>Entassements paradigmatiques</b>	<b>221</b>
1	Entassement Paradigmatique (EP) . . . . .	224
2	Coordination (EP-c) . . . . .	234
3	Disfluence (EP-c) . . . . .	241
4	Précisions des descriptions . . . . .	245
	<b>Conclusion</b>	<b>249</b>

---

<b>III Exemples d'analyses</b>	<b>253</b>
<b>Introduction</b>	<b>255</b>
1 Structure produite . . . . .	255
2 Analyse <i>vs.</i> Parsing . . . . .	256
<b>12 Un énoncé simple</b>	<b>259</b>
1 Début de l'analyse : Informations disponibles . . . . .	259
2 Premier niveau d'analyse . . . . .	260
3 Deuxième niveau . . . . .	263
4 Troisième niveau . . . . .	266
<b>13 Constructions verbales</b>	<b>269</b>
1 Différentes constructions verbales . . . . .	269
2 Sujet plein <i>vs.</i> sujet clitique . . . . .	281
3 Clitiques non-nominatifs . . . . .	284
<b>14 Coordinations</b>	<b>293</b>
1 Coordination de catégories différentes . . . . .	293
2 Coordination elliptique . . . . .	297
<b>Conclusion</b>	<b>305</b>
<hr/>	
<b>Conclusion générale</b>	<b>307</b>
1 Retour sur le développement . . . . .	307
2 Retour sur les objectifs . . . . .	310
<b>Perspectives</b>	<b>313</b>
3 Perspectives théoriques . . . . .	313
4 Perspectives grammaticales . . . . .	314
5 Perspectives formelles . . . . .	320
6 Perspectives informatiques . . . . .	324
<hr/>	
<b>Annexes</b>	<b>327</b>
<b>A La campagne d'évaluation des analyseurs syntaxiques EASY</b>	<b>327</b>
1 Les ressources utilisées . . . . .	327

*Table des matières*

---

2	Les parseurs . . . . .	329
3	La grammaire . . . . .	331
4	Interprétation des résultats . . . . .	347

---

<b>Table des figures</b>	<b>353</b>
--------------------------	------------

<b>Bibliographie</b>	<b>361</b>
----------------------	------------

# Introduction générale

Cette thèse propose des éléments de description de phénomènes syntaxiques du français, en référence à une théorie originale de la langue inspirée d’une synthèse des idées développées dans différentes approches, et représentée selon le formalisme des *Grammaires de Propriétés*.

## 1 Situation

Ce travail se situe à la fois dans une perspective très générale de développement d’une théorie de la langue (§1.1), et dans le cadre d’un travail de recherche d’équipe (§1.2).

### 1.1 Situation générale

**Linguistique formelle.** — On assiste actuellement à un développement tous azimuts des théories formelles de représentation de la langue. A l’aune de son cinquantenaire, la recherche dans le domaine foisonne de propositions aux caractéristiques variées, parfois éloignées mais souvent convergentes, issues de chemins différents, conçues à des fins diverses, et tentant à présent d’affirmer leurs relations autant qu’elles ont pu auparavant se concentrer sur ce qui les distinguait.

Cette variété se justifie d’abord par le fait que les **motivations sous-jacentes** aux modèles peuvent être d’une considérable variabilité, allant de la recherche des universaux linguistiques (p.ex. [Chomsky, 1957], [Prince & Smolensky, 1993]) à la volonté de reproduire l’utilisation effective de la langue en tant que mécanisme de construction de la signification (p.ex. [Muskens *et al.*, 1997], [Kempson *et al.*, 2001]), en tant qu’outil fonctionnel (p.ex. [Kay, 1979], [Dik, 1989], [Van Valin & LaPolla, 1997]) ou en tant que phénomène cognitif (p.ex. [Winograd, 1983], [Croft & Cruse, 2003]), en passant par un objectif de souplesse et de complétude de représentation (p.ex. [Gazdar *et al.*, 1985], [Pollard & Sag, 1994]) ou bien encore d’efficacité et de systématité (p.ex. [Chomsky, 1995], [Steedman, 1996]).

Elle se justifie également par les **domaines de recherche** qui peuvent en être à l’origine, tantôt linguistique (p.ex. [Sag *et al.*, 2003], [Pullum & Scholz, 2003]), tantôt logico-mathématique (p.ex. [Moortgat, 1989], [Kamp & Reyle, 1993]), tantôt informatique (p.ex. [Tugwell, 2000], [Baldridge & Kruijff, 2003]), bien sûr sans exclusion mutuelle.

Elle se justifie enfin par l'**ensemble de postulats** propre à chaque modèle, parmi lesquels peuvent compter l'universalité des propriétés linguistiques (p.ex. [Prince & Smolensky, 1993]), l'existence d'une capacité humaine à interpréter toute production (p.ex. [Kempson *et al.*, 2001]) ou à en juger de la qualité (p.ex. [Gazdar *et al.*, 1985]), la modularité des processus envisagés simultanément (p.ex. [Bresnan, 2001]) ou successivement (p.ex. [Mel'čuk, 1997]) ou bien leur interaction généralisée (p.ex. [Östman & Fried, 2005]), l'existence d'une compétence abstraite et sa distinction des productions linguistiques effectives (p.ex. [Chomsky, 1957]), la compositionnalité des informations linguistiques (p.ex. [Partee *et al.*, 1993]), ou encore la possibilité d'une structuration syntagmatique (p.ex. [Pollard & Sag, 1994]), dépendancielle (p.ex. [Mel'čuk, 1988]), ou uniquement procédurale (p.ex. [Chomsky, 1995], [Tugwell, 1998]).

**Linguistique descriptive.** — D'autre part se développent avec constance et régularité les descriptions de phénomènes langagiers, qui présentent eux aussi une notable variabilité.

Celle-ci se traduit tout d'abord par la **diversité des phénomènes** auxquels elle s'intéresse ; pour ne citer que des questions syntaxiques, p.ex. les dépendances non bornées, la coordination et la subordination (p.ex. [Bril & Rebuschi, 2005]), l'ordre des mots (p.ex. [Gerdes & Muller, 2004]), etc. A travers cette diversité des phénomènes décrits transparait une variabilité de la couverture des travaux, allant de l'étude très fine d'un mot, d'une structure donnée (*il y a* [Cappeau & Deulofeu, 2001], la construction "*what's X doing Y*" [Kay & Fillmore, 1999],...) à des grammaires complètes (p.ex. [Riegel *et al.*, 1999], [Le Goffic, 1993]) en passant par des descriptions de toute une partie de la langue (les relatives [Godard, 1988], la langue parlée [Blanche-Benveniste, 2000], la sémantique des noms [Flaux & Van de Velde, 2000],...).

Le nombre de descriptions des phénomènes linguistiques est également multiplié par la diversité des **approches théoriques** sous-jacentes à ces travaux. Certains par exemple s'appliquent à représenter la structure de constituants des énoncés (p.ex. [Hockett, 1958], [Bloomfield, 1961]) pendant que d'autres se concentrent sur les dépendances entre mots (p.ex. [Tesnière, 1959], [Mel'čuk, 1988]) ou encore sur les fonctions assurées par les diverses structurations linguistiques ([Dik, 1989], [Halliday, 1985]).

Certains, se basant sur l'opposition entre un système linguistique abstrait et un autre concret <sup>1</sup>, et associant à chacun une structure spécifique <sup>2</sup>, vont intégrer à leurs descriptions des mécanismes de passage de l'une à l'autre <sup>3</sup> et utiliser la grammaire pour représenter soit la structure abstraite (p.ex. [Guillaume, 1971], [Chomsky, 1993]) soit la concrète (p.ex. [Pollard & Sag, 1994], [Bresnan, 2001]).

Certains utilisent comme unités les parties du discours agrémentées d'informations de sous-catégorisation (p.ex. [Pollard & Sag, 1994]), d'informations fonctionnelles (p.ex. [Kaplan & Bresnan, 1982]) ou d'informations d'équivalence paradigmatique (p.ex. [van den Eynde & Mertens, 2003]), d'autres se basent sur une opposition entre *mots pleins* et *mots vides* (p.ex.

---

<sup>1</sup> *Langue et parole* chez [Saussure, 1916], *langue et dialogue* chez [Guillaume, 1971], *compétence et performance* chez [Chomsky, 1965].

<sup>2</sup> *Formes intérieure et extérieure* chez [Tesnière, 1959], *structures profonde et de surface* chez [Chomsky, 1957], *universel et singulier* chez [Guillaume, 1971].

<sup>3</sup> *Translation* chez [Tesnière, 1953], *transformation* chez [Harris, 1968], [Salkoff, 1979], *déplacement* chez [Chomsky, 1993], *promotion et dégradation* chez [Blake, 1990].

[Tesnière, 1953], [Bar-Hillel, 1964]).

Certains s'attachent à décrire toutes les productions attestées (p.ex. [Gross, 1975], [Berrendonner *et al.*, 1983], [Blanche-Benveniste, 2000]) là où d'autres se limitent volontairement aux seuls énoncés "bien formés" (p.ex. [Sag *et al.*, 2003], [Abeillé & Godard, 1999]), et d'autres encore se basent uniquement sur l'intuition du locuteur (p.ex. [Chomsky, 1965], [Abercrombie, 1983]).

Certains enfin, tentent de procurer à travers leurs travaux une théorie explicative de la langue (p.ex. [Guillaume, 1971], [Chomsky, 1965]) là où d'autres se bornent à décrire ce qu'ils observent (p.ex. [Blanche-Benveniste, 2000]).

**Pour une théorie descriptive et formelle de la langue.** — Linguistique formelle et linguistique descriptive sont fondamentalement interdépendantes. La première attend de la seconde sa matière première et la seconde, réciproquement, a besoin de la première pour être testée et validée <sup>4</sup>.

Cette thèse se situe à l'articulation des deux domaines. A la lumière d'une synthèse de certaines idées directrices, nous proposons un modèle dérivé d'une théorie originale, formalisé en *Grammaires de Propriétés* (ci-après GP, cf. par exemple [Blache, 2001] ou plus récemment [Blache, 2005]); ce qui a pour objectif à la fois de valider la théorie par sa généralisation en une grammaire et sa représentation formelle, et de valider le formalisme en l'employant à la représentation de ce modèle.

Comme on vient de le voir, les approches descriptives et formelles font l'objet d'une remarquable diversité et chaque théorie repose sur une combinaison complexe de postulats qui fait son unicité. Le fait d'utiliser GP pour représenter un nombre varié de phénomènes permettra, nous l'espérons, d'illustrer l'apport qu'il permet aux deux domaines, descriptif en offrant de représenter des théories qui n'avaient pas encore été formalisées, formel en permettant d'obtenir des traitements élégants, efficaces et jusqu'ici peu fréquents.

## 1.2 Développement des Grammaires de Propriétés

La recherche dont cette thèse est issue prend également place au sein du projet, plus précis, d'une équipe de recherche.

**La naissance d'un formalisme de TALN.** — Initialement apparues sous le nom de *Paradigme 5P* sous la plume de [Bès *et al.*, 1999], [Bès & Blache, 1999], les Grammaires de Propriétés <sup>5</sup> ont à leur origine répondu à une attente en Traitement Automatique des Langues Naturelles (ci-après TALN ou TAL). Il s'agissait alors de proposer un moyen d'allier robustesse de traitement et finesse d'analyse.

<sup>4</sup>C'est tout au moins ce qu'en disent [Chomsky, 1957, p. 112-113], [Pollard & Sag, 1994, p. 6-9], [Sag *et al.*, 2003].

<sup>5</sup>Cette appellation apparaît pour la première fois dans [Blache, 2000a].

A cette fin, les auteurs proposaient un système conçu de façon à être flexible et réutilisable, et au sein duquel on exprimait les connaissances grammaticales nécessaires sous une forme commune et nouvelle : celle des *propriétés*, contraintes exprimant des formulations atomiques des divers types de relations pouvant exister entre les objets linguistiques.

**Le développement d'un formalisme linguistique.** — Le formalisme a ensuite été appliqué à la description de phénomènes variés tels que les dépendances non bornées ([Blache, 2002]), l'interface entre les différents domaines d'analyse ([Blache & Hirst, 2001], [Blache & Meunier, 2004]), le traitement d'entrées tout-venant ([Blache & Azulay, 2002]), la multimodalité ([Guénot & Bellengier, 2004]) ou l'analyse de discours ([Bellengier & Priego-Valverde, 2004]).

Parallèlement à cela se sont développées plusieurs applications informatiques ([Blache & Balfourier, 2001], [Guénot & Vanrullen, 2003], [Dahl & Blache, 2004], [Vanrullen, 2005]), toutes centrées sur le mécanisme de satisfaction de contraintes. Celles-ci ont brillamment participé à la campagne nationale d'évaluation des analyseurs syntaxiques *EASY* ([Paroubek, 2005])<sup>6</sup>.

Certains des apports linguistiques du formalisme ont été mis en évidence au fil de ces travaux, comme la possibilité de tenir compte de la densité d'information linguistique ([Blache, 2004b]) ou de son hétérogénéité ([Blache, 2004a]), et de nouvelles perspectives se sont ouvertes sur les méthodes de développement de grammaire ([Blache *et al.*, 2003], [Guénot & Blache, 2005]) et de constitution de ressources annotées ([Blache & Guénot, 2002]). Le formalisme lui-même s'est affiné ([Vanrullen *et al.*, 2003]), et a trouvé une justification cognitive ([Blache, 2003b], [Blache *et al.*, 2006]).

**Eléments de grammaire du français.** — GP offre donc à présent un éventail d'applications variées et une assise applicative et cognitive. Cependant et contrairement à l'objectif de départ, mis à part la grammaire-jouet de [Blache, 2001] et la grammaire plate spécifiquement créée pour les besoins de la campagne *EASY* ([Balfourier *et al.*, 2005]), aucune grammaire complète n'avait été développée dans le formalisme. Notre travail, en tant que syntacticienne, a donc dû commencer par cela. Plutôt que de nous concentrer sur l'étude approfondie et la représentation formelle d'un phénomène donné, nous nous sommes attachée à proposer un ensemble de descriptions de finesse variable afin de constituer une première ressource grammaticale qui exploite réellement les capacités de GP, et puisse servir de base à toutes sortes de développements ultérieurs.

Cet ensemble de descriptions est encore loin de constituer une "grammaire" à proprement parler<sup>7</sup>, puisqu'une tâche d'une telle envergure nécessiterait bien plus que le temps imparti dans ce cadre précis ; néanmoins nous proposons ici un ensemble de représentations variées, toutes homogénéisées et intégrées en une même entité, ce qui constitue une nouveauté pour GP.

En outre, les descriptions que nous proposons sont inspirées de travaux d'écoles et de théories diverses, dont le mélange a fait émerger un modèle inédit dans ses propositions.

---

<sup>6</sup>Cf. Annexe A.

<sup>7</sup>Bien que nous la dénommions souvent ainsi par la suite, par raccourci.



## 2 Présentation du travail

### 2.1 Objectifs

A la lumière de ce que l'on vient d'exposer, voici les objectifs que nous nous fixons ici :

#### 2.1.1 Elaborer une théorie alternative

A travers l'illustration des extraits de grammaire que nous présentons ici, nous espérons montrer qu'il est possible d'envisager l'analyse linguistique comme un processus plus large et plus flexible que l'image qui en est donnée par les théories formalisées les plus courantes.

Nous souhaitons dans le même temps montrer que ce type de théorie est modélisable et directement formalisable, donc qu'elle est systématique et régulière. Nous espérons, à travers cela, proposer une ouverture de la linguistique formelle à des descriptions qui jusqu'à présent lui étaient difficilement accessibles.

#### 2.1.2 Elaborer un modèle mettant cette théorie en oeuvre

Nous introduisons un modèle de description grammaticale du français qui met en application la théorie proposée, et nous illustrons l'utilisation de ce modèle à travers la description de phénomènes syntaxiques du français.

Cet ensemble est relativement hétérogène dans la mesure où les descriptions ne sont pas toutes menées avec la même précision. Nous proposons de faire un tour général de grands types de constructions du français, et nous nous attardons sur quelques phénomènes qui ont retenu notre attention et auxquels on s'est intéressée de manière plus précise.

Toutefois, malgré la finesse inégale des descriptions, l'objet de cette grammaire est avant tout d'une part de proposer la première ressource GP de cette envergure, d'autre part d'apporter quelques solutions originales au traitement formel de la syntaxe.

#### 2.1.3 Formaliser ce modèle en *Grammaires de Propriétés*

Ce travail a demandé d'évaluer la capacité du formalisme à représenter de telles descriptions syntaxiques, et le choix de GP n'a pas été innocent sur ce point puisque ses spécificités ont rendu leur formalisation possible sans recours à des artifices *ad hoc*, ce qui n'aurait pas été le cas dans d'autres formalismes. Il constitue dans ce cadre une validation de la pertinence de GP en tant que formalisme linguistique.

Le fait de chercher à représenter une telle somme de descriptions en un ensemble homogène a fait en outre émerger un certain nombre de questionnements généraux sur la nature des

représentations formelles, et même plus généralement sur la nature des objets décrits dans une théorie linguistique.

## 2.2 Déroulement

Cette thèse s'organise en trois parties.

La **première partie** veut présenter la théorie sur laquelle on s'est basée, le formalisme que l'on a utilisé et la façon dont on a mis cela en oeuvre pour représenter notre grammaire. On abordera ces questions tout d'abord d'un point de vue très général, mettant en avant les quatre grandes positions constituant notre théorie : la non-générativité de l'approche, la non-modularisation de l'analyse, la non-lexicalisation de la grammaire et la multi-dimensionnalité des domaines (chapitre 1). On y verra les ponts qui se créent avec d'autres approches aux tenants et aboutissants très divers. Puis on introduira l'architecture du formalisme des *Grammaires de Propriétés* (chapitre 2), afin ensuite de présenter la façon dont nos positions théoriques sont mises en oeuvre, à travers une présentation de la modélisation de notre grammaire (chapitre 3). On verra au chapitre 4 que GP permet également de représenter des modèles dont les spécificités sont différentes des nôtres, ce qui permettra de comparer notre proposition à d'autres sur la base d'une représentation commune.

La **deuxième partie** consiste en la concrétisation de la première, à travers le développement d'éléments de grammaire du français. Ici l'on propose un ensemble de descriptions de constructions syntaxiques, variablement approfondies, et dont l'objectif est de procurer une base homogène et cohérente à des développements ultérieurs, tout en contenant quelques propositions nouvelles de description et de traitement de certains phénomènes. Chaque chapitre développe un type de construction de cette grammaire : le chapitre 5 traite des constructions de base, le chapitre 6 traite des constructions nominales, le chapitre 7 des constructions verbales, le chapitre 8 des constructions propositionnelles, le chapitre 9 des constructions de subordination, le chapitre 10 des constructions adjectivales, et enfin le chapitre 11 traite des entassements paradigmatiques (coordination et disfluences).

La **troisième partie** enfin offre quelques exemples de l'application de la grammaire précédente à des analyses. Après avoir expliqué le fonctionnement d'une analyse en GP, nous illustrons nos dires dans le chapitre 12 en menant pas à pas l'analyse d'un énoncé simple. Puis nous nous intéressons à des phénomènes qui ont retenu notre attention pendant le développement de la grammaire, et qui viennent montrer le traitement que notre description permet : nous présentons dans le chapitre 13 le traitement des clitiques dans les constructions verbales, et dans le chapitre 14 celui des coordinations.

En conclusion nous revenons sur les objectifs de cette thèse, et observons les perspectives que ce travail ouvre, à la fois pour le développement de GP, et plus généralement pour la théorie linguistique.

## Première partie

# Pour une théorie descriptive et formelle de la langue



# Chapitre 1

## Positionnement théorique

Le modèle que nous présentons en deuxième partie est basée sur deux choses : tout d’abord un **ensemble de positions théoriques** dont la conjonction fait l’originalité, et qui sont à l’origine du contenu des éléments de grammaire ; ensuite, un **formalisme** dans lequel nous pouvons représenter ce modèle, lui apportant régularité et validation de l’expression. Nous présentons dans ce chapitre les positions théoriques qui sont à l’origine de notre proposition. On peut la résumer à quatre positions-clés mises en évidence au fur et à mesure de nos recherches, et qui en résument les idées principales :

- La première position est celle du **non générativisme** (§1) : Le courant dans lequel nous nous inscrivons s’éloigne en de nombreux points des axiomes des théories génératives, notamment en ce qui concerne la définition d’une *grammaire*, et qui portent des implications profondes tant sur le contenu que sur la forme des descriptions.
- La deuxième position est celle de la **non-présomption** (§2) : On cherche à distinguer clairement et systématiquement, pour une unité linguistique donnée, les informations concernant sa *constitution* de celles concernant son *utilisation*. Refusant d’anticiper des relations qu’une unité entretiendra avec les autres au sein même de sa description interne, alors, par exemple, on ne représente pas d’informations syntaxiques dans un objet dont les constituants n’entretiennent pas de relations syntaxiques ; en cela, on se distingue des grammaires dites “lexicalisées”.
- La troisième position est celle de la **non modularité** (§3) : On se base ici sur l’idée selon laquelle l’information linguistique est *éparse et inconstante* dans les productions, et l’on cherche à en tenir compte dès le départ, sans postuler d’analyses séparées puis de mécanismes d’interfaçage entre les résultats, ni même de représentations distinctes des différents domaines liées par un mécanisme de liage.
- Enfin, la quatrième position est celle de la **multi-dimensionnalité** (§4) : On pose que la structuration linguistique <sup>8</sup> s’opère suivant deux axes, syntagmatique et paradigmatique. Cette remarque n’est pas aussi triviale qu’on pourrait le penser, puisqu’elle implique de

---

<sup>8</sup>Plus exactement, nous nous contentons ici de faire reposer notre idée de multi-dimensionnalité sur la syntaxe, bien que nous n’excluons pas qu’elle puisse s’appliquer à d’autres domaines.

prendre pleinement en compte ces deux dimensions dans l'analyse syntaxique, et non simplement la première comme c'est si souvent le cas.

L'organisation selon ces quatre axes a progressivement émergé de nos recherches tant du point de vue strictement théorique, que de celles portant sur le modèle ou sur le formalisme. Il s'agit ici de proposer une synthèse des positions que l'on tient par la suite, mais qui n'a évidemment pas vocation à constituer une théorie complète de la langue. Ici l'on n'aborde volontairement pas un certain nombre de points théoriques, tels que par exemple la question des universaux linguistiques, celle de la relation entre oral et écrit, etc. Nous nous concentrons exclusivement sur ce qui constitue un pré-requis à la lecture des parties suivantes.

## 1 Une approche non-générative

Selon les théories linguistiques génératives, une grammaire est un mécanisme permettant de *définir* une langue naturelle. L'idée sous-jacente est plus précisément de considérer la grammaire comme un *processus énumératif* à partir duquel on peut dériver strictement un langage, et par conséquent comme un mécanisme utilisé pour vérifier si une entrée donnée appartient ou non à une langue, en fonction de la possibilité ou non de construire une structure y correspondant. Mais cette conception est très restrictive pour plusieurs raisons.

**Filtrage des entrées.** — Premièrement, cela a des conséquences considérables sur la façon de représenter l'information linguistique, qui est alors exprimée de façon à *exclure* les occurrences considérées comme non-grammaticales. Ceci est particulièrement clair dans la *Théorie de l'Optimalité* (ci-après OT, [Prince & Smolensky, 1993]), dans laquelle les contraintes (considérées comme universelles) sont stipulées précisément dans cet objectif de filtrage. On retrouve le même mécanisme dans les *Métagrammaires* (ci-après MG, [Candito, 1999]) qui font usage de contraintes pour éliminer des arbres générés à partir de la grammaire ([Crabbé & Duchier, 2004]).

**Grammaticalité.** — Deuxièmement, considérer une grammaire comme un moyen de définir une langue se base sur une dichotomie claire entre productions grammaticales et non-grammaticales, alors même que [Chomsky, 1961] ou [Ziff, 1964] avaient déjà noté qu'une telle distinction ne correspondait pas à la réalité de la langue, pour un certain nombre de raisons. La linguistique de corpus (p.ex. [Deulofeu, 1982], [McEnery & Wilson, 1996], [Bilger, 2000], etc.) nous montre depuis des années et plus concrètement à quel point les productions peuvent être éloignées des descriptions générativistes, malgré la finesse vers laquelle certaines tendent, et ceci concerne tout aussi bien les productions orales qu'écrites.

**Couverture.** — De plus, la notion de grammaticalité nécessite une couverture totale de la langue : l'incertitude, l'ambiguïté, l'incomplétude, l'hétérogénéité n'ont pas leur place dans de telles approches. L'objectif (théorique) des analyses génératives est de reconstituer une structure englobant la totalité d'un énoncé, se basant sur le postulat de l'existence systématique d'une telle structure (d'où le "symbole initial" de ces grammaires). Or la quantité d'informa-

tions disponibles dans un énoncé est très variable (de même que la qualité et la finesse), et peut ne pas suffire pour reconstruire une structure générale.

Prenons un exemple concernant la syntaxe. Un énoncé tel que (1) présente une structuration syntaxique facilement identifiable, des informations en quantité suffisante pour en donner une analyse complète et sans conteste grammaticalement conforme.

- (1) Deux scientifiques travaillent actuellement sur un matériau qu'ils ont baptisé "claytronics", qui sera composé de nano-machines capables de s'organiser pour reproduire à distance en 3D, et via internet, n'importe quelle forme, y compris la vôtre.<sup>9</sup>

Toutefois les productions effectives s'éloignent rapidement de ce repère canonique, et peuvent réclamer un certain degré de *tolérance* ([Pullum & Scholz, 2003]) pour être analysées. On doit par exemple pouvoir traiter des phénomènes à l'intersection de plusieurs domaines, et dont la syntaxe à elle seule ne peut rendre compte (2).

- (2) les Anglais qui ont quand-même beaucoup d'humour euh les journaux anglais enfin les médias britanniques ont fini par répondre<sup>10</sup>

On doit également pouvoir analyser des énoncés qui présentent des écarts à la grammaire, qu'ils soient volontaires et acceptés (3a) ou non (3b), sans que lesdits écarts ne rendent la totalité de l'analyse impossible.

- (3) a. T'occupe!  
b. c'est parcequ'on est 3 à bosser sur le site et que le troisième à casser son portable alors la com passe mal!<sup>11</sup>

Enfin, on doit pouvoir traiter les énoncés qui ne contiennent quasiment pas d'informations syntaxiques (cf. p.ex. [Deulofeu, 2005]) (et dont le manque d'informations de ce domaine est souvent compensé par une augmentation des informations d'autres domaines, prosodie notamment pour l'oral) (4).

- (4) a. lundi lavage mardi repassage mercredi repos<sup>12</sup>  
b. il y a mon frère sa moto le guidon eh ben complètement naze quoi<sup>13</sup>

**Une ressource descriptive.** — Pour ces raisons, nous pensons qu'une grammaire ne doit pas être restreinte à être un mécanisme définitoire : la question fondamentale n'est pas de savoir si un énoncé appartient ou non à une langue mais bien d'extraire l'information de cet énoncé. Nous nous basons donc sur une conception plus générale selon laquelle une grammaire est une **ressource** réellement **descriptive** contenant toutes les informations à partir desquelles l'analyse (et l'interprétation) d'une entrée (réellement produite) peut être construite, quelle que soit sa forme.

<sup>9</sup>fr.sci.philo, 21 septembre 2006.

<sup>10</sup>[Blanche-Benveniste *et al.*, 1990].

<sup>11</sup><http://www.onpeutlefaire.com/forum/index.php?showtopic=6164>, 26 juillet 2006.

<sup>12</sup>[Mertens, 1993].

<sup>13</sup>[Cappeau & Deulofeu, 2001], d'après A. Cuiloli.

Il s’agit d’un point de vue partagé par des approches très diverses, telles que *Dynamic Syntax* (DS, [Tugwell, 2000], [Kempson *et al.*, 2001]), les *Grammaires de Construction* (CxG, [Goldberg, 1995]), ou l’*Approche Pronominale* (AP, [Blanche-Benveniste *et al.*, 1990]).

**Non-dérivationalité.** — De même que *Role and Reference Grammar* (RRG, [Van Valin & LaPolla, 1997], ou [Bentley, 2003] sur ce point), CxG ([Kay & Fillmore, 1999]), *Lexical-Functional Grammar* (LFG, [Bresnan, 2001]) ou *Word Grammar* ([Hudson, 1984]), l’approche que l’on propose est **non-dérivationnelle** : nous ne faisons pas usage de la notion de “déplacement” / “mouvement” (*movement*) du *Programme Minimaliste* ni de celle de “promotion” vs. “dégradation” (*demotion*) de la *Grammaire Relationnelle* (RG, [Blake, 1990]).

**Pas de structure initiale.** — En outre et comme on l’a vu précédemment, dans la plupart des cas les autres approches considèrent que tout énoncé de la langue doit pouvoir faire l’objet d’une structure l’englobant complètement :

*A candidate sentence is licensed as a sentence of the language if and only if there exists in the grammar of that language a set of constructions which can be combined in such a way as to produce a representation of that sentence.* ([Kay, 1995, pp. 172-173], à propos de CxG)

Nous refusons de considérer que la complétude d’une structure puisse être une condition *sine qua non* à la grammaticalité d’un énoncé<sup>14</sup>, et notre modèle ne nécessite donc pas une couverture complète et permet de construire si cela se justifie des **structures partielles**, non-connectées, selon les données fournies (et la couverture de la grammaire).

## 2 Une grammaire non-présomptive

La raisonnement qui nous a conduit à adopter cette position repose sur une réflexion à propos des grammaires dites “lexicalisées”. Nous commençons par présenter ceci avant de généraliser notre propos à la structuration de l’information.

### 2.1 Lexicalisation et non-lexicalisation

#### 2.1.1 Grammaires dites “lexicalisées”

**Lexique et grammaire dans les modèles linguistiques formels.** — La plupart des modèles de description linguistique courants font partie de ce qu’on appelle les “grammaires lexicalisées”. Des grammaires telles que TAG ou HPSG sont caractérisées par l’intégration, au

---

<sup>14</sup>Ne serait-ce que parce que pour que ce soit le cas, il faudrait pouvoir envisager qu’une grammaire puisse être exhaustive ; or nous pensons que ce point n’est qu’un objectif vers lequel le développement de grammaire tend, mais sans pouvoir l’atteindre concrètement. La conséquence de cela est que quoi qu’il en soit, une grammaire ne pouvant réellement pas être tout à fait complète, on ne peut la considérer comme un outil de jugement d’appartenance d’énoncés à la langue.



sein même de la représentation d’une unité lexicale (au même titre que celle d’un syntagme), à la fois d’informations qui définissent l’objet en tant que partie du système *lexical* de la langue, et d’informations qui le définissent en tant que partie du système *grammatical* de la langue. Pour preuve de la distinction nécessaire des deux systèmes pourtant, le recours dans ces modèles à des règles de dérivation lexicale pour obtenir les variantes grammaticales possibles de chaque objet lexical.

D’après [Kahane, 2006], la justification de l’approche lexicaliste tient au fait que la représentation syntagmatique masque les relations (parfois distantes entre des objets de profondeurs variables) entre mots (ou morphèmes) :

*(...) the recourse to phrasal description masks what combinations of words are involved in extraction. From a theoretical point of view, we think that purely linguistic data might be introduced in word descriptions and that phrasal schemata might only encode how subphrases combine together.*

Selon ce point de vue, la grammaire est réduite à une procédure de formation de syntagmes (ou d’autres objets ou relations), i.e. à un mécanisme de représentation des informations linguistiques qui, elles, sont toutes rassemblées dans le lexique <sup>15</sup>. Pourtant la façon dont les objets se combinent fait bien l’objet de la syntaxe, et il s’agit donc d’information linguistique au même titre que les autres informations auxquelles l’auteur se réfère.

**Lexiques procéduraux.** — Dans les approches dites “lexicalistes”, on mêle donc dans une seule classe d’objets (les objets lexicaux) des informations concernant deux systèmes, certes interdépendants, néanmoins distincts. A tel point que dans des formalismes tels que celui des *Grammaires Catégorielles* (p.ex. [Steedman, 1996]) – qui se définit lui-même comme “extrêmement lexicalisé” (*extreme lexicalization*, [Kruijff, 2004]) –, le *Programme Minimaliste* comme on l’a vu, ou encore *Dynamic Syntax*, ce sont les objets du lexique eux-mêmes, définis sous la forme de procédures d’analyse, qui contiennent toutes les informations grammaticales des modèles. En effet dans ce cas il n’existe pas de grammaire distincte, toute ou partie, du lexique lui-même (pas de description d’autres objets), mis à part quelques principes généraux qui régissent les combinaisons possibles et qui permettent donc de dériver à partir de ces unités “lexicales” la structure linguistique d’un énoncé sans passer par des descriptions de syntagmes (ni de dépendances syntaxiques). Les informations proprement lexicales des unités sont exprimées de manière *implicite* à l’intérieur même de leur description procédurale (qui est un ensemble d’informations explicitement grammaticales puisqu’elles expriment des procédures de construction de structure). Ici le lexique est véritablement la grammaire, et *vice versa*, et les deux sont confondus comme une seule et même entité.

**Pas d’isolation des informations.** — Toutes ces grammaires, comme l’a en partie montré [Kahane, 2004], ont en outre ceci de commun que les objets qu’elles manipulent sont interprétables comme étant soit des “ressources” soit des “besoins”, soit des ensembles “ressources-besoins” (où chaque partie de l’unité peut être l’un ou l’autre). En revanche il est impossible de ne prendre en considération dans ces modèles qu’une partie des “ressources” et/ou des

<sup>15</sup>L’auteur évoque la question des constructions dans lesquelles l’information linguistique ne peut être rapportée à un mot, mais il n’explique pas clairement comment ces types d’objets pourraient être traités élégamment et logiquement par une telle grammaire lexicalisée.

“besoins” des unités, séparément du reste ; en d’autres termes, il est impossible de manipuler des sous-ensembles variables (d’où la possibilité de les représenter sous forme de systèmes polarisés, dans lesquels chaque unité a une “charge” donnée, définie dans le lexique, et dont la neutralisation revient à la prise en considération, c’est-à-dire à l’analyse de l’unité).

**Grammaires lexicalisées ou lexiques grammaticalisés ?.** — On peut penser qu’il est assez curieux d’avoir nommé “grammaire lexicalisée” l’approche consistant à intégrer au lexique des informations grammaticales, puisque littéralement il s’agirait plutôt de “lexiques grammaticalisés”. Comme on l’a dit dans les divers exemples précédents, c’est pourtant le lexique qui fait office de grammaire dans ces approches, que ce soit totalement (approches totalement lexicalisées comme CG) ou en partie (approches semi-lexicalisées comme HPSG qui combine un lexique grammaticalisé et un ensemble de principes et de mécanismes de constitution de syntagmes).

Si l’on accepte cette remarque, alors à l’inverse, la position que l’on adopte ici correspondrait plus rigoureusement à ce que pourrait être une “grammaire lexicalisée” en un sens littéral, puisqu’au contraire de ces approches, on intègre véritablement le lexique dans la grammaire et non le contraire. Toutefois, pour éviter la confusion inévitable, nous nous gardons bien de qualifier notre approche de “lexicalisée”, et au contraire, afin de montrer notre opposition aux grammaires habituellement appelées “lexicalisées” nous appelons notre position “non-lexicalisation”.

### 2.1.2 La non-lexicalisation de notre modèle

Pour notre part, nous postulons que ce n’est pas toujours l’unité lexicale considérée dans sa globalité, en tant que partie du discours, qui contient les informations sur ses relations avec le reste de l’énoncé. En effet, les études menées par exemple sur ce qui est appelé la “non-compositionnalité” de nombreuses constructions linguistiques que l’on ne saurait réduire aux seuls phénomènes de figement <sup>16</sup> (p.ex. [Kay & Fillmore, 1999], [Kracht, 2005a], etc.) montrent bien qu’il est parfois nécessaire de pouvoir prendre en considération uniquement une partie de l’ensemble d’informations constituant une unité afin de l’intégrer à la construction d’une structure plus grande. Par exemple, on fera usage de l’ensemble des informations contenues dans l’unité lexicale *terre* pour l’intégrer en tant que constituant de syntagme dans (5a), alors qu’on devra en éliminer une certaine partie dans l’analyse de (5b) qui est une expression semi-figée dans laquelle le morphème ne conserve pas la totalité de son contenu sémantique ainsi que sa capacité à recevoir des compléments ; enfin, dans (5c) où l’unité est employée dans une expression figée (un “*mot complexe non construit*” selon l’expression de [Corbin, 1987]), *terre* perd complètement son statut de morphème et donc toute une partie de son statut.

- (5) a. Voici la **terre** sous l’épanchement des premiers feux de l’aube <sup>17</sup>  
b. Que faisaient les techniciens qui nous avait informé la veille aux infos du tremblement de **terre** de magnitude 8.5 en pleine mer ? <sup>18</sup>

---

<sup>16</sup>Et quand bien même on ne considérerait que les phénomènes de figement comme des constructions dites non-compositionnelles, ceci demande aux approches dites “lexicalisées” pour les prendre en compte de traiter ceci comme une exception à leur règle lexicaliste, donc comme une limite à leur modèle.

<sup>17</sup>fr.lettres.ecriture - 10 nov 2002.

<sup>18</sup>fr.soc.religion - 29 décembre 2004.

c. Le tatin de pommes de terre, c'est ça ? <sup>19</sup>

Pourtant, dans les trois cas une partie (variable et plus ou moins large) de l'unité reste stable, ce qui justifie selon nous de pouvoir séparer ces informations-ci de ce qui concerne l'emploi de l'unité en tant que constituant de structure plus grande.

C'est là l'une des raisons pour lesquelles on tient à ne pas présumer de l'utilisation d'un objet au sein même de sa description. Conséquence de cela, on cherche à effectuer une distinction <sup>20</sup> claire entre le système lexical et le système grammatical, sans avoir à passer par des règles qui ont pour charge de modifier le contenu de base des unités lexicales (i.e., de changer la "grammaire" contenue dans le "lexique") <sup>21</sup>. Selon ce postulat les informations grammaticales (concernant les relations entre objets d'un énoncé) peuvent être exprimées de manière régulière en fonction d'un ensemble d'informations dont la constitution est variable à volonté, pouvant contenir des caractéristiques provenant de domaines différents (morphologie, syntaxe, sémantique, prosodie, pragmatique). Cela signifie que ces ensembles manipulés dans la grammaire ne correspondent pas forcément toujours à des unités lexicales, mais peuvent pointer des objets plus larges ou plus fins, en fonction des descriptions effectuées.

La frontière que nous posons ici a pour conséquence que de nombreuses informations de sous-catégorisation que l'on trouve habituellement mentionnées dans le lexique, font dans notre modèle l'objet de constructions grammaticales. Nous faisons par exemple une différence de traitement entre valence et transitivité. Nous représentons la valence dans les unités lexicales parce qu'il s'agit d'une information de sémantique lexicale, stable pour chacune de ces entrées, mais nous représentons la transitivité dans la grammaire parce qu'elle dépend des occurrences produites (Cf. chapitre 7 pour une discussion à ce propos).

## 2.2 Une grammaire non-présomptive

D'un point de vue plus général, cette position concernant la "lexicalisation" des grammaires est en réalité l'une des manifestations d'un choix selon lequel on refuse de faire figurer dans une construction donnée des informations portant sur la structure (de quelque domaine que ce soit) dans laquelle elle est susceptible d'entrer. Concernant le lexique, cela signifie qu'on ne fait pas figurer d'informations syntaxiques dans les unités lexicales (ce sont les constructions syntaxiques qui pourront contenir des restrictions plus ou moins fines et pouvant mentionner des informations lexicales), mais ceci ne s'applique pas uniquement à la relation entre "lexique" et "grammaire".

---

<sup>19</sup>fr.rec.cuisine - 28 jan 2000.

<sup>20</sup>Soulignons que nous parlons ici de distinction théorique, et non formelle.

<sup>21</sup>On cherche également à ne pas avoir recours à ce type de règles parce que cela nous demanderait d'utiliser un mécanisme différent de ceux que l'on utilise déjà.

### 3 Un traitement non-modulaire

#### 3.1 Deux acceptions de la “modularité”

Dans tous les cas, il s’agit d’isoler des parties d’un ensemble (et de convenir de méthodes pour les combiner par la suite) ; mais suivant la nature de l’ensemble que l’on considère, il s’ensuit deux utilisations possibles de la notion : représentation modulaire de la structuration des objets, ou analyse modulaire.

##### 3.1.1 Modularité des objets

**Représentations modulaires.** — Selon un premier point de vue, la modularisation consiste à scinder le contenu des *objets linguistiques* en sous-parties. C’est typiquement le cas de la *Grammaire Lexicale-Fonctionnelle* (LFG, [Kaplan & Bresnan, 1982], [Bresnan, 2001]), où l’analyse se base sur une ressource grammaticale unique qui contient toutes les informations nécessaires à la construction d’une double représentation (au moins : *c-structure* et *f-structure*), lesquelles informations sont représentées au sein des mêmes objets selon des conventions différentes, et mises en relation entre elles *via* un mécanisme de “projection” (*linking*) issue de la théorie du même nom (*linking theory*, ou *lexical mapping theory*, [Levin, 1986], [Bresnan & Zaenen, 1990]). (fig. 1 <sup>22</sup>).

```

la :      Det, (↑Genre)=fém, (↑Num)=sing, (↑Def)=déf

souris :  N, (↑Genre)=fém, (↑Num)=sing, (↑Pred)='souris'

donne :   V, (↑Mod)=ind, (↑Temps)=prés, (↑Suj Num)=sing,
          (↑Suj Pers)=3, (↑Pred)='donner,<Suj,Obj,A-Obj>'

```

FIG. 1 – Représentation modulaire des unités en LFG.

Les *Grammaires d’Adjonction d’Arbres* (TAG, [Joshi *et al.*, 1975], [Joshi & Schabès, 1997]) utilisent elles aussi la théorie de la projection : quand elles ne distinguent pas entre “arbre dérivé” (structure syntaxique) et “arbre de dérivation” (structure de combinaison des arbres) ([Vijay-Shanker, 1987]) pour calculer à partir de cette dernière les représentations sémantiques, leurs unités grammaticales sont représentées sous une forme modulaire constituée d’une “dimension <sup>23</sup> syntaxique” (un arbre élémentaire) et une “dimension sémantique” (une prédication élémentaire) <sup>24</sup> ([Gardent & Kallmeyer, 2003]) (fig. 2 <sup>25</sup>).

On peut rapprocher cette modularisation des informations liées entre elles par un mécanisme de mise en correspondance, de la *rule-to-rule hypothesis* de [Montague, 1974] selon laquelle à

<sup>22</sup>Tirée de [Vilnat, 2005].

<sup>23</sup>Le terme de “dimension” ici fait référence à un niveau ; il est utilisé de la même façon en *Multi-Modal Combinatory Categorical Grammars* ([Kruijff, 2004]).

<sup>24</sup>Le nombre de “dimensions” adoptées par la suite dépend des théories ; [Crabbé, 2005a] p.ex. en définit trois : un niveau sémantique, un niveau syntaxique et un niveau d’interface entre les deux.

<sup>25</sup>Tirée de [Crabbé, 2005b]



et à cette fin adoptent une représentation typiquement procédurale des unités (fig. 4 <sup>27</sup>).

```

IF      ?Ty(e) ‘holds for’ description {  $\langle \uparrow_0 \rangle Tn(a), ?Ty(e), \diamond \}$ ,
THEN   add Fo(John), Ty(e) and  $\llbracket \downarrow \rrbracket \perp$  to the description,
ELSE   ABORT.

```

FIG. 4 – Représentation procédurale des objets linguistiques en DS.

Ici la description de l’unité (en l’occurrence la fig. 4 représente une entrée pour *John*) consiste à dire : au moment où je lis *John* dans mon entrée à analyser, SI au niveau de l’arbre où je suis à ce moment-là <sup>28</sup> je vois la description  $\{ \langle \uparrow_0 \rangle Tn(a), ?Ty(e), \diamond \}$  <sup>29</sup> ALORS j’ajoute à cette description les informations *Fo(John)* et *Ty(e)* <sup>30</sup>, SINON je ne fais RIEN. On n’a ici aucune mention du domaine linguistique d’appartenance des informations données dans les descriptions, on ne fait que les manipuler, en en ajoutant et en se déplaçant dans la structure que l’on construit.

D’une manière comparable, les unités des *Grammaires Minimalistes* (fig. 5 <sup>31</sup>) sont l’objet d’une représentation en termes de configuration, c’est-à-dire que leur contenu et leur ordre, certes ne font pas référence à des actions à effectuer, mais toutefois dépendent directement du mécanisme de parsing (qui lit, et donc utilise, ces informations dans un ordre donné).

$\epsilon ::= Refl12 +k T$	$\epsilon ::= Acc3 +k T$	$\epsilon ::= Dat3 +k T$	$\epsilon ::= v +k T$
$se ::= Acc3 +F Refl12$	$se ::= Dat3 +F Refl12$	$se ::= v +F Refl12$	
$le ::= Dat3 +G Acc3$	$le ::= v +G Acc3$		$lui ::= v +F Dat3$
$\epsilon ::= vacc \leq =D v$	$\epsilon ::= vdat \leq =D +k vacc$	$\epsilon ::= V \leq =p vdat$	
$montrera::V$	$donne::V$	$\epsilon ::= > V =D +k =D v$	
$\epsilon ::= P \leq =p$	$a ::= D +k P$	$\epsilon ::= p -F$	
$Jean::D -k$	$Marie::D -k$	$roi::N$	$livre::N$
$le ::= N D -k$	$\epsilon ::= D -k -F$	$\epsilon ::= D -k -G$	$\epsilon ::= T C$

FIG. 5 – Représentation procédurale des objets linguistiques dans une Grammaire Minimaliste.

Dans la figure 5 par exemple, on voit que chaque unité lexicale est multipliée par le nombre de configurations possibles lors de son analyse : à chaque ligne de l’extrait correspond une forme lexicale, à laquelle on fait correspondre un ensemble d’informations dont le contenu et l’ordre correspond à une suite différente d’actions à effectuer et d’informations à ajouter dans la structure que l’on est en train de construire.

<sup>27</sup>Tirée de [Kempson *et al.*, 2001].

<sup>28</sup>En Dynamic Syntax l’analyse consiste à construire un arbre et le compléter progressivement en se déplaçant sur ses branches. Une entrée lexicale peut donc aussi bien ajouter de l’information à un noeud que commander le déplacement du pointeur dans l’arbre.

<sup>29</sup>Cette description signifie à peu près “Je suis le premier argument ( $\langle \uparrow_0 \rangle$ ) du noeud initial ( $Tn(a)$ ), j’existe mais je ne sais pas à quoi je corresponds dans l’énoncé ( $?Ty(e)$ ), et je porte le pointeur ( $\diamond$ )”.

<sup>30</sup>“Je suis de la forme *John*” (*Fo(John)*) et “Je sais à qui je corresponds dans l’énoncé” (*Ty(e)*).

<sup>31</sup>Il s’agit d’un extrait du lexique proposé dans [Amblard, 2006] pour la reconnaissance des clitics.

### 3.1.2 Modularité des analyses

La seconde acception possible de la modularité consiste en le découpage, non pas des unités de la grammaire, mais des analyses. C'est typiquement le cas de la *Théorie Sens Texte* (TST, [Mel'čuk, 1997]) au sein de laquelle on distingue quatre niveaux standards (sémantique, puis syntaxique, morphologique et phonologique de surface) et trois niveaux intermédiaires (syntaxique, morphologique et phonologique profonds) (fig. 6 <sup>32</sup>). Ces niveaux sont ordonnés et leur succession fournit autant d'analyses et de représentations différentes, aux extrémités desquelles se trouvent le *Sens* (la signification, l'interprétation) d'une part et le *Texte* (la production graphique ou phonique) d'autre part.

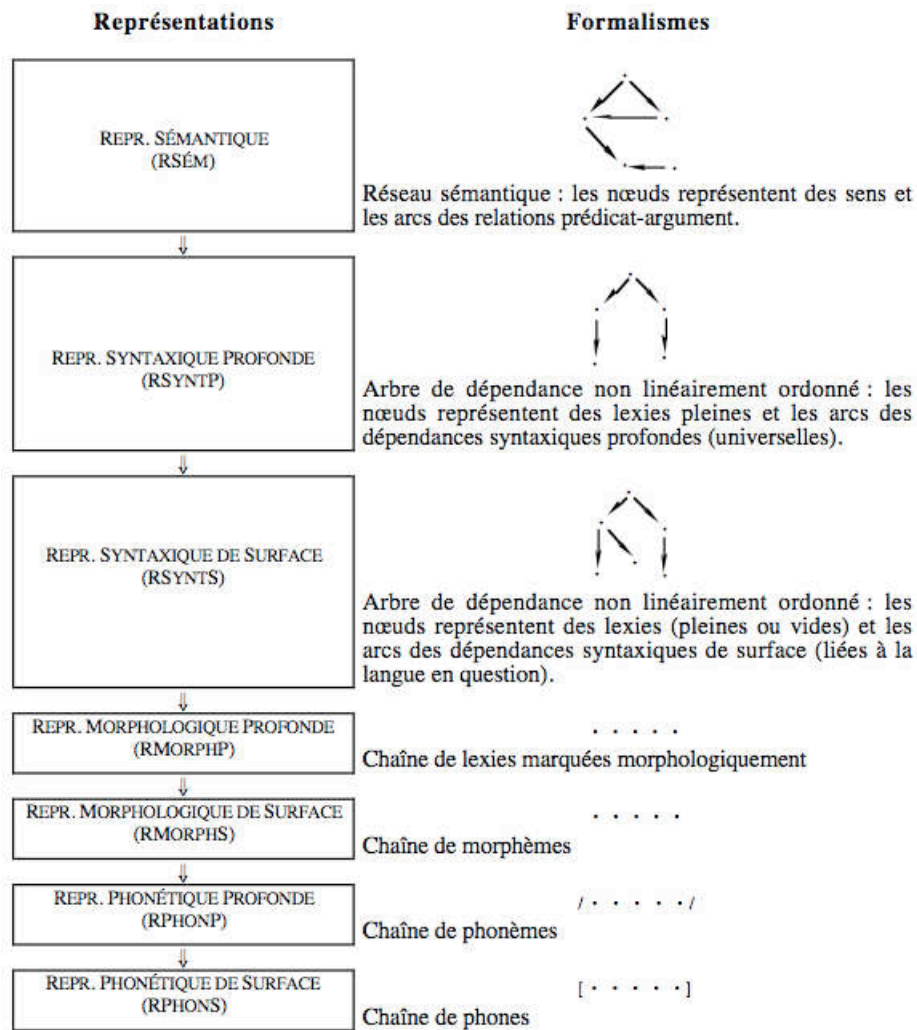


FIG. 6 – Représentation modulaire des analyses en TST.

C'est le cas aussi de la conception de TAG que l'on a évoquée précédemment, selon laquelle on calcule les représentations sémantiques des énoncés à partir de l'arbre de dérivation obtenu

<sup>32</sup>Tirée de [Polguère, 1998].

lors de l’analyse syntaxique (p.ex. [Schabès & Shieber, 1994], [Candito & Kahane, 1998], [Kallmeyer, 2002]).

C’est également le cas du système FIPS ([Laenzlinger & Wehrli, 1991]) qui utilise une grammaire du français inspirée du modèle *Gouvernement et Liage*. Cet analyseur, entièrement étudié pour permettre l’implémentation de la théorie chomskyenne, est constitué d’un ensemble ordonné de composants (modules) qui répondent chacun à des processus particuliers, allant des structures lexicales aux structures sémantico-pragmatiques (cf. fig. 7).

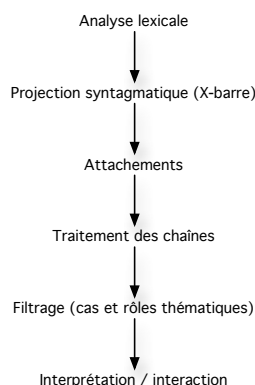


FIG. 7 – Représentation modulaire des analyses en GB dans FIPS.

## 3.2 La non-modularité de notre modèle

De ce qui précède nous retenons une distinction entre ce que l’on appelle “modèles modulaires” et qui font une distinction de notation entre les domaines, en faisant des systèmes autonomes, et ce que l’on appelle les “modèles non-modulaires” qui, même s’ils peuvent séparer les informations des différents domaines, les représentent toutefois en une même entité si bien qu’il est possible d’envisager des interactions de toutes sortes, y compris de bas niveau, entre les domaines, de même que d’envisager une représentation avec partage d’information entre domaines. Dans le premier groupe on fait figurer LFG, TAG et la TST ; dans le second, HPSG, DS et GM. Nous nous plaçons pour notre part aux côtés des seconds.

Plus précisément, en reprenant ces interprétations possibles de la modularité, on peut dire que notre modèle adopte une représentation modulaire de ses objets à la manière de HPSG, mais les rassemble au sein d’une grammaire non-modulaire (§3.2.1), qui effectue des analyses non-modulaires (cf. §3.2.2).

### 3.2.1 Grammaire non-modulaire

Dans notre conception, une grammaire est un ensemble d’associations d’informations indiquant les interactions entre différents objets. L’idée est de rassembler plusieurs sources d’informa-



tion sitôt qu'elles ont une conséquence remarquable sur (la construction de) l'interprétation, quelles que soient leur nature et leur importance. Cette conception est celle des *Grammaires de Construction* (sur ce point, cf. p.ex. [Kay & Fillmore, 1999]), qui a été intégrée à notre propre cadre de travail ([Blache, 2004a]).

**Représentation homogène.** — Contrairement aux approches qui font usage de la *linking theory*, en GP la totalité des informations que nous représentons l'est sous une unique forme, stable et homogène : quel que soit leur domaine d'appartenance, nous utilisons une unique formalisation, ce qui permet à ces informations d'être mises en relation dès leur expression dans la grammaire, et sans recours à l'expression d'un mode d'interfaçage spécifique. Cela signifie également que nous n'introduisons **pas de hiérarchie** entre les différents domaines d'information, à la différence des approches telles que la TST. En outre, cela permet également de ne pas avoir à préciser le domaine d'appartenance d'une information, et donc de pouvoir préciser tout aussi directement des relations existant entre des éléments de domaines différents (la simultanéité d'un geste et un son dans une analyse multimodale p.ex.).

Toutefois, permettre une interaction de bas niveau entre les différents domaines opérant au sein même des descriptions grammaticales, ne signifie pas que toutes ces informations sont confondues en un seul et même système (au sens de [Saussure, 1916]). En effet, mêler les différents domaines en une unique ressource requiert avant tout d'être à même de poser une différence claire entre chacun d'eux, afin d'en cerner les besoins propres, les mécanismes spécifiques et les caractéristiques remarquables <sup>33</sup>. Aussi nous concentrons-nous ici uniquement sur les caractéristiques strictement syntaxiques des phénomènes étudiés, parce que notre participation à l'élaboration de cette grammaire se situe précisément à ce niveau. Par ailleurs ce sera une occasion de montrer que même pour effectuer une analyse syntaxique (*i.e.*, pour obtenir une structure syntaxique en guise de résultat d'analyse), l'ensemble des informations syntaxiques, toutes fines qu'elles puissent être, ne sont bien souvent pas suffisantes.

Ce n'est donc pas la même conception de la non-modularité que celles des théories fonctionnelles (comme p.ex. [Givón, 1984], la *Cognitive Grammar* de [Langacker, 1987], la *Systemic Functional Grammar* de [Halliday, 1985], la *Functional Grammar* de [Dik, 1989], etc.), qui certes associent à chaque objet linguistique un couple forme-sens tout comme c'est le cas ici, mais ne représentent pas explicitement les informations linguistiques constituant les objets; ils préfèrent à cela représenter une sorte de synthèse fonctionnelle de ces informations.

### 3.2.2 Analyse non-modulaire

**Monostratilité.** — Les résultats d'analyse que l'on construit se présentent sous une forme monostratale : toutes les informations que l'on peut extraire des énoncés donnent lieu à une structure unique les rassemblant. En outre, de la même façon que dans la grammaire, toutes ces informations sont représentées sous une forme homogène, ce qui permet la même flexibilité d'expression.

<sup>33</sup>Par exemple, nous rejoignons [Kracht, 2005b] sur ce point, qui explique notamment que les notions d'indice, d'ordre et de multiplicité ne font pas partie de la sémantique.

Ceci nous permet de ne pas sur-analyser de domaine, en ne nous demandant pas d’y ajouter des règles de structuration pour tenir compte des cas où il manque des informations de ce niveau précis afin de pouvoir malgré tout produire une structure pour ce niveau, laquelle structure devra servir d’entrée à l’analyse du niveau supérieur dans la hiérarchie.

## 4 Une analyse multi-dimensionnelle

Depuis [Saussure, 1916] il est admis que l’analyse linguistique s’articule selon deux axes, syntagmatique et paradigmatique. Or l’immense majorité des théories linguistiques ne considèrent que le premier de ces axes pour élaborer leurs grammaires (c’est notamment ce que constate [Deulofeu, 2003]). Ce que nous proposons, à travers l’idée de multi-dimensionnalité de l’analyse, est de proposer un traitement qui tienne compte des deux axes à la fois.

Ceci remet en question une grande partie des éléments qui sont souvent considérés comme des invariants de l’analyse syntaxique (notamment), et pour en mesurer les conséquences il convient de revenir un peu plus précisément sur la notion de paradigme. [Dubois *et al.*, 1994] en dit la chose suivante :

*(...) un paradigme est constitué par l’ensemble des unités entretenant entre elles un rapport virtuel de substituabilité. F. de Saussure retient surtout le caractère virtuel de ces paradigmes. En effet, la réalisation d’un terme (= sa formulation dans l’énoncé) exclut la réalisation concomitante des autres termes. À côté des rapports in præsentia, les phénomènes du langage impliquent également des rapports in absentia, virtuels. On dira ainsi que les unités a, b, c, ...n appartiennent au même paradigme si elles sont susceptibles de se substituer les unes aux autres dans le même cadre typique (syntagme, phrase, morphème).*

Si effectivement la mise en relation d’objets sur l’axe paradigmatique restait virtuelle dans les productions alors on n’aurait pas besoin d’en tenir compte dans une analyse. Mais il arrive pourtant que plusieurs objets occupent le même paradigme, comme cela a été montré par Claire Blanche-Benveniste (cf. p.ex. [Blanche-Benveniste, 1987]), et il est donc nécessaire de pouvoir le représenter et en tenir compte. Voilà pourquoi nous posons que la représentation des relations paradigmatiques revêt la même importance que celle des relations syntaxiques dans notre grammaire, et voilà pourquoi nous posons que notre grammaire doit être multi-dimensionnelle.

D’un certain point de vue, cette position rejoint celles de [Meillet, 1924] opposant “*syntaxe de rection*” et “*syntaxe de parataxe*”, de [Bally, 1965] opposant “*syntaxe liée*” (ou “*dictum*”) et “*syntaxe segmentée*” (ou “*modus*”), de [Perrot, 1994] opposant “*syntaxe de l’énoncé comme phrase*” et “*comme message*”, de [Blanche-Benveniste, 1997] opposant “*microsyntaxe*” et “*macro-syntaxe*”, ou encore de [Deulofeu, 2006] suivant cette dernière et opposant “*détachement*” et “*rattachement*”. La différence existant entre ces approches (et notamment l’approche macro-syntaxique) et la nôtre est que les deux dimensions sur lesquelles nous nous reposons ne sont pas les mêmes : là où [Deulofeu, 2003] p.ex. oppose **deux “modules”**<sup>34</sup> **syntactiques**,

---

<sup>34</sup>“Module” est ici employé au sens de [Nølke, 1994], et se rapproche à peu près de ce que l’on appelle ici “domaine”, i.e. à un ensemble de phénomènes que l’on peut étudier en utilisant un ensemble de méthodes de description homogène.

l'un micro (relations syntaxiques de rection) et l'autre macro (relations syntaxiques au-delà de la rection), nous opposons **deux “dimensions” syntaxiques**, l'une syntagmatique (qui rassemble, en fait, une partie des relations micro- et macrosyntaxiques sous l'ensemble des relations hypotactiques) et l'autre paradigmaticque (qui rassemble l'autre partie, i.e. les relations paratactiques). L'un des points essentiels avec lesquels la séparation entre notre proposition et l'approche macrosyntaxique est notable, c'est ce qui concerne la relation entre les domaines d'analyse : là où la macrosyntaxe va user d'indices p.ex. prosodiques pour justifier d'une structuration (macro)syntaxique, nous pourrions utiliser cet indice pour proposer une construction grammaticale, mais sans nécessairement rapporter celui-ci à une dimension syntaxique. Autrement dit, nous posons que dans (6),

(6) lundi lavage mardi repassage mercredi repos

*lundi lavage* et *mardi repassage*, p.ex., n'entretiennent simplement pas de relation syntaxique (ni micro, ni macro).

Dans notre proposition, nous posons que les deux dimensions, quoique coexistantes et intégrées à la grammaire sur un pied d'égalité (ce qui signifie que la grammaire contiendra une structure syntagmatique et une structure paradigmaticque), n'entretiennent pas les mêmes relations ; ainsi, les constructions syntagmatiques pourront intégrer des constituants paradigmaticques et inversement, mais un objet donné pourra être constituant à la fois d'une structure syntagmatique et d'une structure paradigmaticque. Par exemple, dans (7), *je* et *Pierre* sont les constituants immédiats d'une construction paradigmaticque ; parallèlement à cela, *je te vois* constituera un syntagme verbal.

(7) je te vois et Pierre aussi

### Résumé

Dans ce chapitre nous avons présenté les idées directrices de notre position théorique, en nous basant sur une comparaison avec celles d'un certain nombre d'autres approches théoriques. De ce bilan, nous avons dégagé les quatre idées suivantes, dont la conjonction fait l'originalité de notre proposition :

- Le **non générativisme** (§1) : Le courant dans lequel nous nous inscrivons s'éloigne en de nombreux points des axiomes des théories génératives, notamment en ce qui concerne la définition même d'une *grammaire*, et qui portent des implications profondes tant sur le contenu que sur la forme des descriptions.
- La **non-présomption** (§2) : On cherche à distinguer clairement et systématiquement, pour une unité linguistique donnée, les informations concernant sa *constitution* de celles concernant son *utilisation*. Refusant d'anticiper des relations qu'une unité entretiendra avec les autres au sein même de sa description interne, alors, par exemple, on ne représente pas d'informations syntaxiques dans un objet dont les constituants n'entretiennent pas eux-mêmes de relations syntaxiques ; en cela, on se distingue des grammaires dites "lexicalisées".
- La **non modularité** (§3) : On se base ici sur l'idée selon laquelle l'information linguistique est *éparse et inconstante* dans les productions, et l'on cherche à en tenir compte dès le départ, sans postuler d'analyses séparées puis de mécanismes d'interfaçage entre les résultats, ni même de représentations distinctes des différents domaines associées par un mécanisme de liage.
- La **multi-dimensionnalité** (§4) : On pose que la structuration linguistique s'opère suivant deux axes, syntagmatique et paradigmatique. Ceci implique de prendre pleinement en compte ces deux dimensions dans l'analyse, et non simplement la première comme c'est si souvent le cas.

## Chapitre 2

# Architecture du formalisme

Les grands principes de l'architecture de GP ont été largement introduits dans [Blache, 2001]. Ils sont initialement issus d'un ensemble de postulats théoriques et d'objectifs d'applications au TALN. Cependant une exploration linguistique fine du formalisme était nécessaire. C'est ainsi qu'en commençant à manipuler l'esquisse de grammaire existante <sup>35</sup> pour la tester et l'améliorer, nous avons vu émerger un certain nombre de caractéristiques du modèle qui à notre connaissance n'avaient pas été soulignées auparavant (ceci étant sans doute dû à la différence de perspective). Nous avons alors tenté de formaliser divers types de représentation, différentes sortes de relations possibles, afin de vérifier si ce qui nous apparaissait conservait sa pertinence linguistique, tout en restant satisfaisant du point de vue du modèle. A l'issue de ceci nous étions capable de présenter le formalisme dans une formulation destinée à montrer comment un linguiste peut représenter les différents types d'informations sous forme de GP. C'est ce que nous proposons dans ce chapitre. Cette présentation tâche de mettre en évidence les caractéristiques notables du point de vue du développement de grammaire, c'est-à-dire à la fois l'architecture générale du formalisme telle que proposée par Philippe Blache <sup>36</sup> dans ses diverses publications à ce propos, ainsi que les compléments apportés et les aspects mis en évidence à partir de manipulations diverses <sup>37</sup>.

**Architecture générale.** — (fig. 8) D'une manière générale, GP est constituée de deux ressources de base, une **grammaire** (§1) et une **spécification des propriétés** (§2). Ces deux ressources permettent d'effectuer l'analyse d'énoncés, en produisant des **caractérisations** (§3) selon un mécanisme de satisfaction de contraintes.

Un couple grammaire - spécification des propriétés constitue un module d'analyse. Cela signifie que dans une perspective non-modulaire comme la nôtre il ne figurera qu'un seul couple, et

---

<sup>35</sup>Celle proposée dans [Blache, 2001].

<sup>36</sup>Ainsi que Gabriel Bès dans les premières présentations du formalisme, cf. par exemple [Bès & Blache, 1999].

<sup>37</sup>Cela signifie donc que cette présentation vient en complément des autres coexistantes. Cette remarque est valable également pour les développements les plus récents du formalisme, dont on n'a pu tenir compte pour des raisons de chronologie (on pense notamment à [Blache *et al.*, 2006]), mais qui ne s'opposent pas à ce que l'on présente ici.

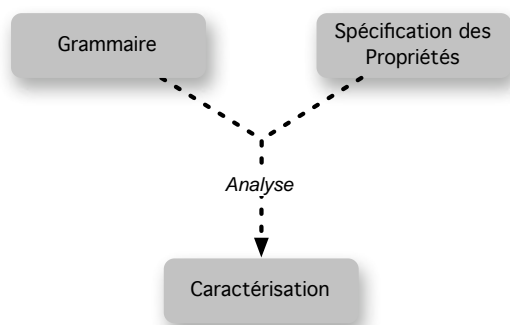


FIG. 8 – Architecture générale du formalisme.

dans une perspective modulaire il pourra exister autant de couples que de modules <sup>38</sup>.

## 1 Grammaire

La grammaire est la ressource qui contient les descriptions des objets linguistiques. Ces objets qui la constituent sont nommés **constructions** (§1.1), et chacune de ces constructions est elle-même constituée d'un **identifiant** (§1.2), d'**informations intrinsèques** (§1.3) et **extrinsèques** (§1.4) (fig. 9).

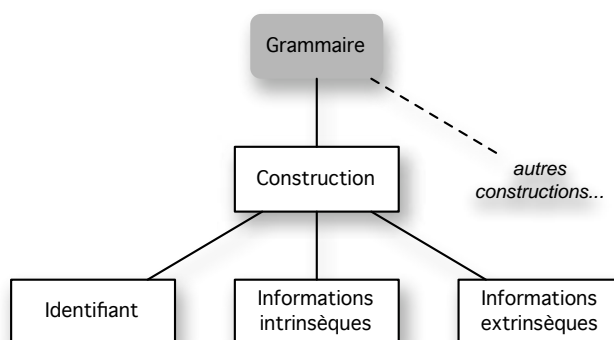


FIG. 9 – Architecture de la grammaire.

<sup>38</sup>C'est par exemple le cas de la *Grammaire d'Unification Sens-Texte* (GUST) que l'on a représentée au chapitre 4 §1 : ce modèle présente une série de grammaires qui se succèdent les unes aux autres, et en GP cela se traduit par autant de couples grammaire - spécification qu'il y a d'analyses successives dans la GUST.

## 1.1 Constructions

### 1.1.1 Définition

Chaque objet manipulé dans notre grammaire est appelé **construction** au sens de CxG (cf. par exemple [Goldberg, 1995]), c'est-à-dire que chaque objet est considéré comme un “patron” plus ou moins complexe, plus ou moins général, faisant référence à un ensemble de caractéristiques pouvant être mixte, et qui constitue avec les autres un réseau complexe de descriptions linguistiques.

Cela signifie plusieurs choses. Tout d'abord, chaque construction est le reflet d'une **régularité structurelle** observée, exprimée sous la forme d'un ensemble d'indices remarquables permettant de postuler la définition d'un objet linguistique donné, constitué des éléments mis en relation au sein de cette construction.

L'ensemble d'indices utilisé peut être **plus ou moins précis**, sa précision dépendant de la finesse de la structure au sein du système grammatical. Par exemple on peut facilement envisager qu'une construction très générale telle qu'un SV, dont la réalisation linguistique est d'une considérable variabilité, sera définie par un ensemble d'indices nettement moins précis que la construction décrivant une construction verbale très figée telle que *étant donné* ou *ce faisant*. La construction très générale du SV devra en effet contenir toutes, mais uniquement les caractéristiques que l'on peut retrouver dans tous les SV possibles quelle que soit leur forme, alors que les constructions figées pourront bénéficier de descriptions d'autant plus fines que leur forme est contrainte.

Enfin, cet ensemble d'indices caractérisant une construction peut être **mixte**. Cela signifie que les constructions ne sont pas uniquement des objets répondant à des indices syntaxiques et/ou lexicaux, mais peuvent également répondre à des caractéristiques sémantiques, prosodiques ou autres, suivant le type d'informations que l'on souhaite observer, et qui nous sont accessibles lors de l'analyse.

Concrètement, cela ne signifie pas que toutes les constructions doivent systématiquement recevoir une définition à la fois syntaxique et sémantique et prosodique même si certaines d'entre elles ne présentent pas d'indices pertinents dans l'un ou l'autre domaine ; cela signifie que par exemple, si l'on a accès aux informations prosodiques et que l'on observe qu'une construction donnée est remarquable par son contour intonatif régulier, alors on pourra faire figurer cette information dans la description de ladite construction, sans nécessiter pour autant que le contour intonatif soit pris en considération dans tous les autres cas.

**Héritage.** — Les constructions, dans une grammaire, sont liées entre elles par un réseau d'héritage. En effet, dans la mesure où les constructions peuvent être plus ou moins précises et que plus elles sont fines, plus elles reçoivent de contraintes, on peut les organiser entre elles en permettant à chacune d'elles d'hériter des caractéristiques d'une ou plusieurs autres constructions, plus générales. Pour reprendre notre exemple précédent, on peut imaginer que les constructions verbales figées *ce faisant* ou *étant donné* héritent (chacune séparément) de la construction qui décrit les SV, ce qui signifie qu'elles récupèrent les contraintes générales

du SV, auxquelles elles ajoutent d'autres contraintes qui leur sont propres.

**Informations intrinsèques et extrinsèques.** — Chaque construction que nous manipulons dans notre grammaire rassemble d'après ce qu'on vient de voir un ensemble d'informations, pertinentes à une régularité structurelle observée. Cette régularité structurelle est elle-même constituée de deux types d'informations, que nous appelons informations intrinsèques et informations extrinsèques <sup>39</sup>, et qui s'opposent par leur nature.

Si l'on s'appuie sur la définition donnée par le TLF-i <sup>40</sup>, est **intrinsèque** ce qui

*[...] est inhérent, indépendamment de tous les facteurs extérieurs [...] qui appartient à un objet [...] en lui-même et non dans ses relations à l'autre.*

En se référant à [Lalande, 1972], on peut ajouter qu'

*[...] une chose est dite en particulier avoir une “valeur intrinsèque” [...] si elle possède cette valeur par sa propre nature, et non pas en tant qu'elle est le signe ou le moyen d'autre chose.*

A l'inverse, dans le TLF-i sera considéré comme **extrinsèque** ce qui

*[...] est extérieur à l'objet que l'on considère, qui ne lui appartient pas mais dépend des circonstances, des faits accessoires.*

Appliquant cette opposition aux caractéristiques d'une construction, on considérera comme **intrinsèques** les informations qui sont propres à la construction en tant qu'elle-même, comme par exemple le fait qu'elle porte un genre ou un mode ou la façon dont est construite sa signification, parce que ces caractéristiques ne dépendent pas des objets qui constituent la construction mais de sa nature propre. A l'opposé, on considérera comme **extrinsèques** les informations qui concernent la mise en relation des objets entrant dans la constitution d'une instance de la construction dans une production donnée. On se place dans une perspective où les constituants et la construction, chacun en tant qu'objets, entretiennent une forme spécifique de relation, qui peut se réaliser sous la forme de diverses occurrences possibles de la construction (et de ses constituants donnés dans un énoncé donné). Par exemple, le fait qu'un accord en genre entre un nom et un adjectif soit l'une des caractéristiques notables d'un SN (entre autres) est extrinsèque au SN puisque cela concerne les qualités de ses constituants.

L'une des conséquences notables de cette distinction est que les caractéristiques intrinsèques sont parfaitement **régulières** aux constructions (un SN a toujours un genre, un SV a toujours un mode), alors que les caractéristiques extrinsèques sont **variables** en fonction des occurrences (un certain SN peut être introduit grâce à un ensemble donné d'indices dont par exemple l'accord en genre entre nom et adjectif, alors qu'une autre instance de SN peut être introduite sans que cet indice ne soit relevé, par exemple si ce dernier SN ne contient pas d'adjectif). Pour être plus précise, ce ne sont pas les caractéristiques extrinsèques elles-mêmes

---

<sup>39</sup>Cette terminologie est introduite dans [Guénol, 2005a]; on trouve chez d'autres auteurs à ce même propos mention d'une opposition entre informations “statiques” et “dynamiques”, entre “traits” et “propriétés”, mais nous avons estimé que les termes “intrinsèque” et “extrinsèque” répondaient plus précisément à cette dichotomie.

<sup>40</sup><http://atilf.atilf.fr>



qui sont variables dans une construction, mais c'est leur instanciation qui dépend des données produites à analyser.

Autrement dit, on fera une distinction au sein des constructions entre leurs caractéristiques intrinsèques, qui sont stables et inhérentes à leur nature, et leurs caractéristiques extrinsèques, qui sont variables et dépendent de la nature de leurs constituants et des relations qu'ils entretiennent.

### 1.1.2 Représentation

Comme l'illustre la fig. 10, chaque construction est représentée sous la forme d'une boîte, qui porte un cartouche contenant son identifiant permettant de la situer de façon unique dans la grammaire, et est organisée en deux parties, à la manière de la représentation traditionnelle de CxG (p.ex. dans [Mathieu, 2003] ou [Kay & Fillmore, 1999]), légèrement adaptée. Le bloc supérieur de la boîte contient les informations intrinsèques, et le bloc inférieur les informations extrinsèques. L'héritage (indiqué dans l'identifiant) fait référence à la (aux) construction(s) dont la présente hérite en mentionnant son (ses) étiquette(s), et s'opère par unification du contenu des constructions.

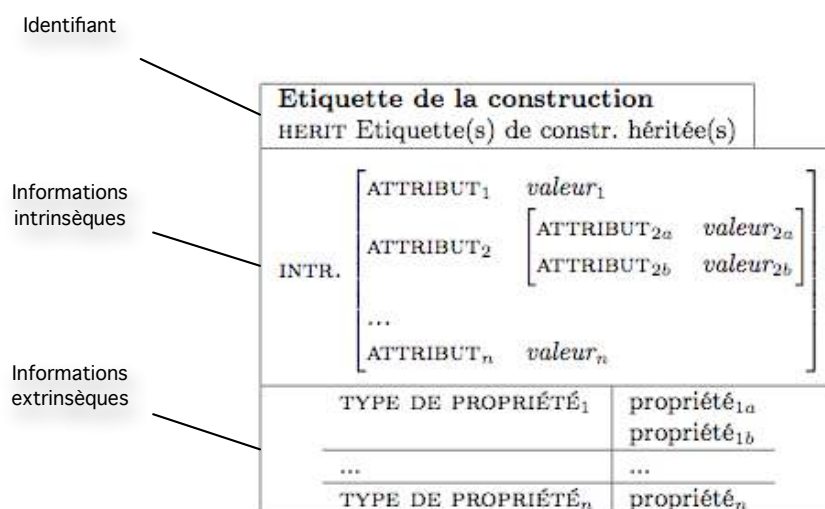


FIG. 10 – Forme de base de la représentation d'une construction dans la grammaire.

### 1.1.3 Exemple

Prenons un exemple de la grammaire présentée en partie 2 : la construction représentant le Syntagme Nominal déterminé (fig. 11).

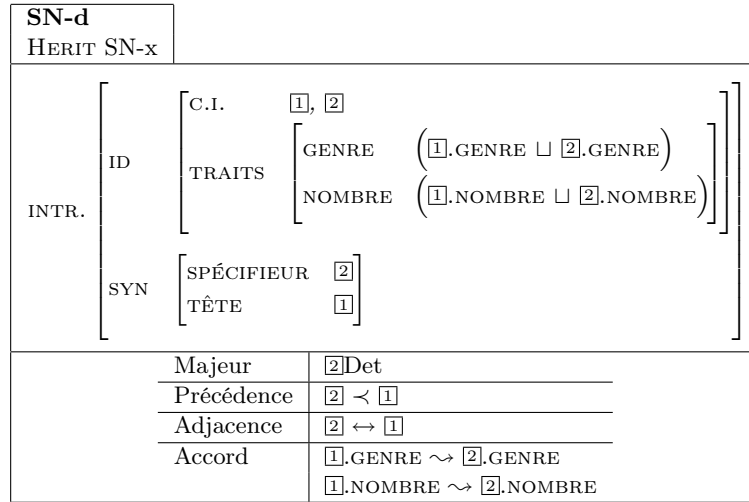


FIG. 11 – Exemple de construction : Le Syntagme Nominal déterminé.

On voit que cette construction est organisée en trois parties : le cartouche en haut à gauche qui permet d'identifier la construction, puis les deux blocs qui reçoivent les informations intrinsèques, et les informations extrinsèques de la construction. Nous laissons de côté pour l'instant ce qui concerne la forme et le contenu précis de ces blocs, nous y reviendrons dans les sections qui suivent.

## 1.2 Identifiant

### 1.2.1 Définition

L'identifiant est l'ensemble d'informations qui permet de désigner d'une manière unique et de placer chaque construction dans la grammaire. Il est constitué de deux informations :

- L'**étiquette** de la construction (son nom en quelque sorte),
- L'**héritage** de la construction dans la hiérarchie de la grammaire.

C'est l'ensemble constitué de ces deux informations qui doit permettre de désigner l'objet sous une forme unique, et non sa seule étiquette. Par exemple, on pourra avoir deux constructions partageant une même étiquette (p.ex. un nom lexical et un nom syntaxiquement construit comme on le verra dans nos éléments de grammaire en partie 2) mais n'ayant pas le même héritage (et *vice versa*, bien entendu).

### 1.2.2 Représentation

Comme on l'a dit plus haut, l'identifiant prend la forme d'un cartouche accolé aux deux blocs de la construction contenant ses informations intrinsèques et extrinsèques. La liste des

constructions héritées est précédée de l'attribut `HERIT`, et les objets de la liste sont liés par des “et logiques” ( $\wedge$ ).

### 1.2.3 Exemple

Si l'on reprend l'exemple précédent (fig. 11), on peut y voir que la construction a reçu l'étiquette `SN-d`, et qu'elle hérite d'une autre construction dont l'étiquette est `SN-x` (ce qui signifie que les informations définissant `SN-x` seront unifiées avec celles présentes ici). On reproduit le détail de l'identifiant en fig. 12.

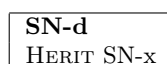


FIG. 12 – Exemple d'identifiant dans le Syntagme Nominal déterminé.

## 1.3 Informations intrinsèques

### 1.3.1 Définition

Les informations intrinsèques sont constituées, comme on vient de le voir, de l'ensemble des caractéristiques inhérentes à la construction définie. Elles ont pour spécificité d'être stables, c'est-à-dire régulières indépendamment de leur formes effectives possibles.

### 1.3.2 Représentation

Les informations intrinsèques sont représentées sous la forme d'une structure de traits, tout à fait classique et telle que présentée par exemple dans [Abeillé, 1993]. L'organisation des informations au sein de cette structure de traits dépend non pas du formalisme, mais de la grammaire <sup>41</sup>.

Les structures peuvent être réentrantes, ce qui sera représenté par des index selon la convention en usage en HPSG notamment, sous la forme  $\boxed{1}$ ,  $\boxed{2}$ , etc. Ces index peuvent également être utilisés dans les informations extrinsèques de la construction.

Chaque attribut peut recevoir une valeur atomique ou complexe. La valeur d'un trait peut en outre être le résultat d'une opération, dans laquelle il est possible d'utiliser l'un des opérateurs suivants :

- L'opérateur de **sélection**, symbolisé par le point “.”, comme par exemple dans

Adj.GENRE

<sup>41</sup>Nous détaillons l'organisation qui concerne la grammaire présentée ici au chapitre 3 §1.

qui fait référence au trait de genre d'un objet de catégorie Adj, ou dans

$$N.[\text{TYPE } \textit{propre}]$$

qui fait référence à un objet de catégorie N et dont le type de valeur "propre".

- La **négation**, symbolisée par " $\neg$ ". Par exemple,

$$\neg \textit{Adj}$$

fait référence à un objet dont la catégorie n'est pas Adj.

- Le **et logique**, symbolisé par " $\wedge$ ". Par exemple,

$$[\text{GENRE } \textit{masculin}] \wedge [\text{NOMBRE } \textit{singulier}]$$

fait référence à un objet dont le genre est "masculin" et le nombre est "singulier".

- Le **ou exclusif** symbolisé par "**w**". Par exemple,

$$[\text{GENRE } \textit{masculin}] \textbf{w} [\text{NOMBRE } \textit{singulier}]$$

fait référence à un objet dont le genre est "masculin" ou bien dont le nombre est "singulier", mais qui ne porte pas ces deux caractéristiques à la fois.

- Le **ou inclusif** symbolisé par " $\vee$ " ou par le point-virgule " $;$ ". Par exemple,

$$[\text{GENRE } \textit{masculin}] \vee [\text{NOMBRE } \textit{singulier}]$$

fait référence à un objet dont le genre est "masculin" ou dont le nombre est "singulier", ou encore qui porte ces deux caractéristiques.

- L'opérateur d'**unification** <sup>42</sup> représenté par " $\sqcup$ ". Par exemple,

$$[\text{GENRE } \textit{masculin} ; \textit{féminin}] \sqcup [\text{GENRE } \textit{masculin}]$$

produira

$$[\text{GENRE } \textit{masculin}]$$

alors que pour

$$[\text{GENRE } \textit{masculin}] \sqcup [\text{GENRE } \textit{féminin}]$$

l'unification échouera.

- Enfin, nous avons ajouté un opérateur binaire que nous appelons d'"**alternative**", symbolisé par " $?:$ ". Il s'agit d'une notation choisie en référence à l'opération ternaire correspondante par exemple en C/C++, qui se note

$$x = a ? b : c$$

---

<sup>42</sup>Telle que défini par [Colmerauer, 1975], [Kay, 1979].

et se lit “si  $a$  est vrai alors  $x = b$ , sinon  $x = c$ ”. Pour ce qui nous intéresse,  $x$  est un attribut, et  $a ? b : c$  est la valeur qui lui est affectée. Nous l’avons réduit à un opérateur binaire parce que dans notre cas,  $a$  équivaut à “ $b$  existe”, et on en a donc fait une opération binaire pour de simples raisons de lisibilité afin d’éviter les répétitions sous la forme  $b ? b : c$ . Dans ce cas donc,  $b$  est à la fois une condition de vérité et une valeur à affecter. L’alternative sert à exprimer, par exemple dans

$$x = ?b : c$$

que l’on affecte à  $x$  la valeur de  $b$  à la condition que  $b$  existe, et dans le cas contraire on affecte à  $x$  la valeur de  $c$ . Si  $c$  non plus n’existe pas, alors  $x$  recevra une valeur indéterminée : il sera sous-spécifié. Par exemple, dans la description du syntagme verbal à attribut de l’objet (chapitre 7 §2.3.1), on voit que la valeur du DÉPENDANT de la relation objet direct (OBJ-DIR) de la construction est

$$?[2] : [1].\text{OBJ-DIR} | \text{DÉPENDANT}$$

Cela signifie que si l’objet indexé  $[2]$  fait référence à un objet présent alors l’attribut DÉPENDANT le prendra pour valeur, mais s’il ne fait pas référence à un objet présent alors il prendra pour valeur une référence au DÉPENDANT de l’OBJ-DIR de l’objet indexé  $[1]$ . Si ce dernier n’existe pas non plus, alors notre trait ne sera pas spécifié.

### 1.3.3 Exemple

Reprenons l’exemple du SN déterminé présenté en fig. 11 ; on en reproduit les informations intrinsèques en fig. 13.

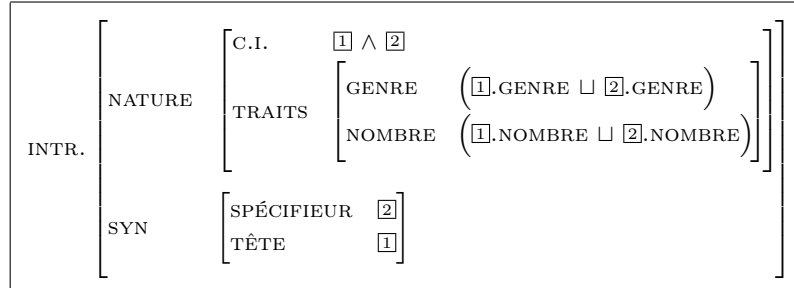


FIG. 13 – Exemple d’informations intrinsèques dans le Syntagme Nominal déterminé.

On y voit que les informations intrinsèques (INTR.) à cette construction sont organisées en deux parties : d’une part sous un attribut ID (pour “identité”), d’autre part sous un attribut SYN (pour “syntaxe”) <sup>43</sup>. La valeur de chacun d’eux est complexe. Dans les traits d’ID, la construction reçoit deux informations, une première dont l’attribut est C.I. (pour “constituants immédiats”) et qui a pour valeur une liste d’index (qui font en l’occurrence référence à des objets indiqués dans les informations extrinsèques), et une seconde dont l’attribut est TRAITS et dont la valeur, complexe, est formée d’un GENRE dont la valeur est le résultat de l’unification des valeurs de GENRE des deux objets indexés  $[1]$  et  $[2]$ , et d’un NOMBRE dont la valeur est le

<sup>43</sup>Les choix concernant l’organisation de la structure de traits sont présentés au chapitre 3 §1 ; ce n’est pas l’objet du présent chapitre de les expliquer.

résultat de l'unification des valeurs de NOMBRE de ces mêmes objets. Enfin, sous l'attribut SYN figure un attribut TÊTE qui a pour valeur l'index [2] et un autre SPÉCIFIEUR qui a pour valeur l'index [2].

## 1.4 Informations extrinsèques

### 1.4.1 Définition

Comme on l'a vu plus haut, les informations extrinsèques sont celles qui correspondent à la mise en relation d'objets permettant d'introduire une construction donnée dont ils sont les constituants. Elles forment le pivot de la grammaire puisque ce sont elles qui vont permettre l'analyse des énoncés. Et c'est la façon dont les objets sont mis en relation via l'expression de ces informations extrinsèques qui constitue l'une des caractéristiques fondamentales, sinon la caractéristique centrale, de GP.

**Propriétés.** — Dans le formalisme, les informations relationnelles ne consistent pas simplement en l'expression de la constituance ou de la dépendance, considérés comme des blocs homogènes <sup>44</sup>. Il s'agit ici de pouvoir isoler au sein de ces deux notions chacun des indices qui permettent de déduire l'une ou l'autre, et de regrouper ces indices possibles en relations-types, que l'on appelle propriétés. C'est ce que [Blache, 2001] appelle l'**encapsulation** de l'information linguistique. Il s'agit du même principe que celui de la dissociation des règles de réécriture globales en "dominance immédiate" d'une part et "précédence linéaire" d'autre part qui a été proposée en GPSG par [Gazdar *et al.*, 1985], mais cette dissociation est ici appliquée à d'autres relations existantes que ces deux-là. On va pouvoir utiliser un ensemble de propriétés de divers types (notamment de constituance, de cooccurrence, de position et de dépendance <sup>45</sup>) afin de décrire avec précision chacun des indices caractérisant une construction.

Les différents types de propriétés ne sont pas hiérarchisés entre eux (et les propriétés ne sont pas non plus hiérarchisées entre elles), ou tout au moins ne le sont pas *a priori* (néanmoins on peut imaginer assez facilement qu'ils gagneraient à pouvoir l'être, et comme le montrent [Blache *et al.*, 2006] il semble que certains types de relations soient plus importants que d'autres pour l'analyse d'une production linguistique par un locuteur). Toutefois, si la possibilité d'une hiérarchisation des types de propriétés (et/ou des propriétés elles-mêmes) relève du formalisme GP, le contenu de la hiérarchie dépendra de la théorie formalisée (à chaque grammaire formelle correspond un jeu de types de propriétés qui s'organisent en un système, et les poids relatifs de chacun des types dépend directement de l'ensemble).

Cette façon d'exprimer les relations entre objets constitue selon nous une approche tout à fait distincte des autres modes de représentation de la structure linguistique que l'on a pu rencon-

---

<sup>44</sup>Quoique l'on pourrait les exprimer ainsi si telle le voulait une théorie donnée.

<sup>45</sup>Initialement, [Bès & Blache, 1999] classaient les propriétés en trois grands types : **existence** ("les pièces qui [...] composent" un énoncé - cela correspond à notre "constituance" et à la "dominance immédiate" de [Gazdar *et al.*, 1985]), **linéarité** ("leur ordre" — notre "position", et la "précédence linéaire" de [Gazdar *et al.*, 1985]) et **fléchage** ("les liens entre les pièces"). Nous considérons que la notion de "lien" est bien trop floue (toute relation est un lien) et préférons distinguer les deux autres sortes que l'on propose : "dépendance" d'une part, "cooccurrence" d'autre part.

trer, et c'est précisément cette particularité qui confère à GP sa grande flexibilité d'expression mais aussi d'analyse, en permettant de prendre en compte tous les indices disponibles, quelle que soit leur quantité.

### 1.4.2 Représentation

A une grammaire donnée correspond un ensemble de propriétés défini au départ. En effet, il appartient à chaque grammaire choisir les relations qu'elle souhaite pouvoir exprimer entre les objets, ainsi que la façon dont ces relations fonctionnent (ce dernier point fait l'objet de la spécification, cf. §2). Elles sont en revanche toujours exprimées sous la même forme : dans un tableau dont la première colonne indique le type de propriété et dont la seconde indique les propriétés correspondant à ce type et impliquant des constituants possibles de la construction.

Une construction donnée ne doit pas nécessairement contenir des propriétés de tous les types possibles de la grammaire, et pour chaque construction il peut figurer plusieurs propriétés du même type, sans limitation de nombre. Les propriétés ne sont pas ordonnées entre elles, de même que les types de propriétés entre eux. Les objets mentionnés dans les propriétés peuvent être indexés ( $\boxed{1}$ ,  $\boxed{2}$ ,...), auquel cas ils font référence aux mêmes objets que dans les propriétés intrinsèques de la construction.

### 1.4.3 Exemple

Intéressons-nous une nouvelle fois à l'exemple du SN déterminé (fig. 11) : nous reproduisons le contenu de ses informations extrinsèques ci-dessous en fig. 14.

Majeur	$\boxed{2}$ Det
Précédence	$\boxed{2} < \boxed{1}$
Adjacence	$\boxed{2} \leftrightarrow \boxed{1}$
Accord	$\boxed{1}.\text{GENRE} \rightsquigarrow \boxed{2}.\text{GENRE}$ $\boxed{1}.\text{NOMBRE} \rightsquigarrow \boxed{2}.\text{NOMBRE}$

FIG. 14 – Exemple d'informations extrinsèques dans le Syntagme Nominal déterminé.

Sans entrer dans le détail de la signification de ces propriétés <sup>46</sup>, on remarque que cette construction présente des propriétés de quatre types : Majeur, Précédence, Adjacence et Accord. Les trois premiers types portent chacun une propriété, et le quatrième en porte deux. En outre, la première propriété ne mentionne qu'un seul objet, alors que les suivantes en mentionnent chacune deux. Toutes ces propriétés font référence à des objets en utilisant leurs index, et l'on voit dans la propriété Majeur que l'objet indexé  $\boxed{2}$  correspond à un Det (pour "déterminant"). On voit également que les deux propriétés d'Accord s'intéressent à une sous-partie des objets auxquels ils font référence : les traits de GENRE pour la première, ceux de NOMBRE pour la seconde.

En plus de la représentation formelle des propriétés au sein des constructions, on a besoin

<sup>46</sup>Elles sont définies dans la présentation du modèle qui en fait usage, chapitre 3 §2.

pour en connaître le fonctionnement de faire référence à leur spécification. C’est ce que l’on présente dans la section qui suit.

## 2 Spécification des propriétés

La mise en place de la spécification des propriétés, initialement présentée dans [Vanrullen *et al.*, 2003], est issue de deux origines. La première est un travail de systématisation de l’expression des propriétés afin de pouvoir répondre à des questions comme “*Est-ce que le jeu de propriétés proposé dans [Blache, 2001] est le seul possible ?*”, “*Qu’est-ce qui justifie que l’on introduise une nouvelle propriété ?*”, “*Sur quoi se base-t-on pour définir une nouvelle propriété ?*”, etc. A travers une série de tests sur les caractéristiques de chacun des types de propriétés (combien de membres mettent-ils en jeu, est-ce que les objets sont ordonnés ou pas, est-ce que la négation de la propriété a un sens, comment évalue-t-on chacun des types et à quoi se réfère-t-on, peut-on exprimer la même chose différemment et alors sous quelle forme, etc.), on a progressivement dégagé un certain nombre d’invariants des différents types de propriétés. A la suite de cela on a pu définir avec plus de précision les critères selon lesquels la définition d’un type de propriété est nécessaire, de remettre en question sur cette base la définition des propriétés initialement utilisées dans les premières applications de GP, et finalement de définir une série de nouvelles propriétés <sup>47</sup>.

La seconde origine est un travail que nous avons mené en coopération avec Tristan Vanrullen, visant à concevoir une implémentation de GP qui exploite réellement les capacités du formalisme et soit utilisable et intéressante aussi bien du point de vue informatique que du point de vue linguistique. En l’occurrence, il s’agissait de rendre réellement possible d’implanter des grammaires issues de modèles linguistiques différents dans un même parseur. A cette fin il a été nécessaire de commencer par isoler les informations qui devaient faire l’objet d’un développement linguistique, et qui devaient donc rester accessibles et modifiables lors de chaque implantation de grammaire, des informations qui relevaient du mécanisme d’analyse général et qui devaient rester stables d’une utilisation à l’autre. Il ne s’agit pas là d’une question triviale, puisque l’enjeu était de permettre une implantation simple et systématisée des différentes grammaires GP et ainsi d’en permettre, notamment, une évaluation et une comparaison simplifiées <sup>48</sup>.

**Architecture générale.** — La spécification des propriétés est une seconde ressource, séparée de la première (la grammaire) mais qui y est directement liée. Elle permet de définir l’identité et le fonctionnement de chacun des types de propriétés utilisés dans la grammaire à laquelle elle correspond.

Cette spécification rend possible de faire varier le fonctionnement (et donc l’interprétation,

---

<sup>47</sup>Elles sont utilisées en Partie 2 et présentées au chapitre 3 §2.

<sup>48</sup>L’objectif a été en grande partie rempli par l’analyseur *SeedParser* (cf. [Vanrullen, 2005]) et de l’environnement de développement de grammaire associé *Accolade* (cf. [Guénot & Vanrullen, 2003]) dans lesquels on a pu implanter un certain nombre de grammaires relevant de théories fort différentes, dont notamment un certain nombre de versions de celle présentée en partie 2, ainsi que la grammaire réalisée pour les besoins de la campagne EASY (cf. annexe A) qui présente des caractéristiques tout à fait différentes du présent travail.



la signification) de chaque type de propriété en fonction des besoins, tout comme de créer de nouveaux types, en utilisant un attirail formel dédié. La ressource se présente sous la forme d'un ensemble de **définitions de type de propriété** (§2.1), qui sont chacune constituées de deux parties : on commence par donner l'**identité** et la **forme** du type de propriété (§2.2), puis on donne ses conditions de **disponibilité** et de **satisfaction** (§2.3). On donne une vue d'ensemble de cette organisation en fig. 15.

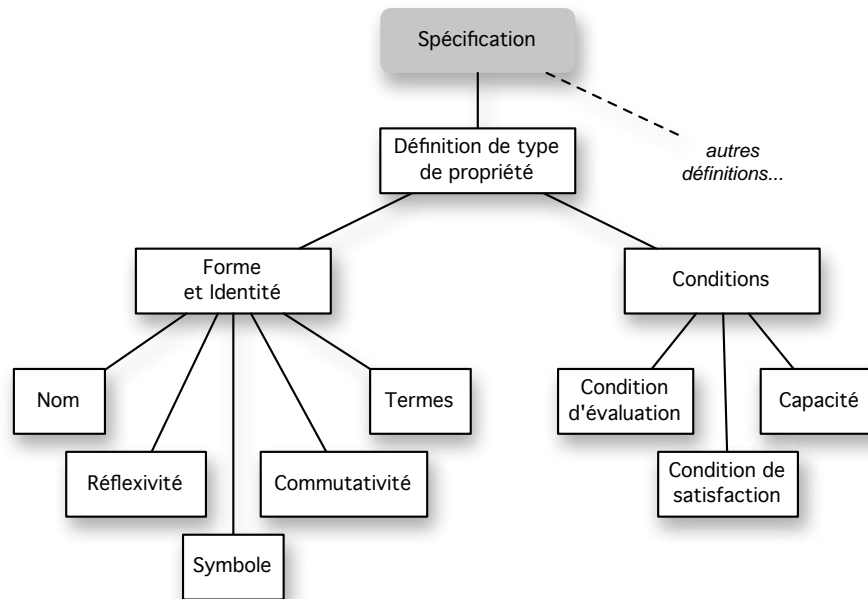


FIG. 15 – Architecture la spécification des propriétés.

## 2.1 Définition de type de propriété

### 2.1.1 Définition

A chaque relation-type que l'on souhaite utiliser dans une grammaire donnée, on fait correspondre une définition de type de propriété. Celle-ci renferme toutes les informations permettant de lire et d'écrire des propriétés dans la grammaire (dans les informations extrinsèques des constructions) et d'utiliser ces propriétés pour effectuer des analyses de productions.

On introduit dans une spécification autant de définitions de type de propriété que l'on utilise de types de propriétés dans la grammaire à laquelle la présente ressource est liée. En fait on peut tout à fait y faire figurer plus de définitions de type de propriété différentes qu'on n'utilise de types de propriétés dans la grammaire, mais ceci n'aurait que peu d'intérêt. À l'inverse on ne peut pas utiliser de type de propriété dans la grammaire qui n'ait pas reçu de définition dans la spécification correspondante.

Les indications fournies dans les définitions servant à lire et écrire des propriétés figurent dans un ensemble d'informations que l'on appelle **forme et identité** (§2.2), celles servant à définir leur comportement figurent dans un autre ensemble d'informations que l'on appelle **conditions** (§2.3).

### 2.1.2 Représentation

Dans la spécification, chaque définition de type de propriété prend la forme d'une boîte dans laquelle on fait figurer l'ensemble des informations nécessaires à sa définition, comme présenté en fig. 16. Toutes ces informations doivent obligatoirement apparaître dans chacune des définitions.

Nom du type de propriété	
Symbole	<i>Symbole utilisé</i>
Termes	<i>Liste des termes</i>
Réflexivité	Obligatoire $\vee$ Autorisée $\vee$ Interdite
Commutativité	Vrai $\vee$ Faux
Disponibilité	<i>Condition de disponibilité des propriétés du type</i>
Satisfaction	<i>Condition de satisfaction des propriétés du type</i>
Capacité	<i>Capacité des propriétés du type</i>

FIG. 16 – Forme de base d'une définition de type de propriété dans la spécification.

Comme on le voit en fig. 16, le nom du type de propriété est mis en exergue dans un cartouche au-dessus du reste des informations (bien qu'il fasse partie du bloc du dessous en réalité). Suit un tableau en deux parties, contenant d'abord les informations de forme et d'identité du type de propriété, ensuite ses conditions. Nous voyons dans ce qui suit le détail de contenu de chacun.

### 2.1.3 Exemple

Pour illustrer notre propos, prenons l'exemple de la propriété de Précédence que l'on définit dans la spécification correspondant à la grammaire présentée en Partie 2 (fig. 17).

Précédence	
Symbole	$<$
Termes	$\boxed{1}$ Construction $\wedge$ $\boxed{2}$ Construction
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Satisfaction	$\boxed{1}.\text{INDEX.DEBUT} < \boxed{2}.\text{INDEX.DEBUT}$
Capacité	$\infty$

FIG. 17 – Exemple de définition de type de propriété : la Précédence.

Sans entrer dans le détail de la signification des informations (qui font l'objet des sections qui suivent), on peut y lire que cet objet est une définition d'un type de propriété dont le nom

est Précédence, et qu'elle est effectivement définie par l'ensemble complet des informations figurant dans la structure de base des définitions de type de propriété (fig. 16).

## 2.2 Forme et identité

### 2.2.1 Définition

Cette première partie de la spécification introduit les informations qui permettent de reconnaître un type de propriété au sein de l'ensemble qu'ils constituent et de le représenter dans la grammaire (c'est-à-dire de pouvoir écrire et lire des propriétés dans la grammaire). On y mentionne les informations suivantes :

- Son **nom**, qui doit être unique pour un jeu de propriétés donné puisque c'est lui qui permet d'identifier le type parmi les autres. P.ex. Précédence ou Exigence.
- Le **symbole** utilisé pour représenter les propriétés de ce type (qui doit lui aussi être unique pour un jeu de propriétés), p.ex.  $\prec$  pour la Précédence ou  $\Rightarrow$  pour l'Exigence.
- La liste des **termes** utilisés dans chaque propriété de type. Cette liste indique à la fois le nombre d'objets que la propriété manipule (à travers le nombre d'objets figurant dans la liste) ainsi que leur identité : les propriétés peuvent faire référence soit à des informations extrinsèques (i.e. des propriétés), soit à des informations intrinsèques (i.e. des structures de traits), soit à des constructions complètes (i.e. des ensembles intrinsèques + extrinsèques). Pour chacune de ces possibilités, les termes peuvent faire référence soit à une partie de l'objet (p.ex. le trait de GENRE d'un N), soit à un objet complet (p.ex. les SN), soit encore à un ensemble d'objets (p.ex. les objets dont le NOMBRE est *pluriel*, ou alors une liste comme (SN  $\vee$  N  $\vee$  Adj)).
- On ajoute à la liste précédente une information qui permet de savoir si les termes que l'on a introduits peuvent faire référence à un même objet ou non, c'est-à-dire si les termes peuvent (voire, doivent) être **réflexifs** ou pas. Par exemple, on peut avoir besoin dans une grammaire d'exprimer des propriétés d'Exclusion qui prendraient la forme  $\boxed{V} \not\Leftarrow \boxed{1}.\text{TEMPS}$ , ce qui signifierait que le Verbe indexé  $\boxed{1}$  ne peut porter de trait de temps. A l'inverse, dans une autre grammaire les propriétés d'Exclusion pourraient interdire la réflexivité des termes, de manière à ce qu'une propriété telle que  $\boxed{1}V \not\Leftarrow \text{TEMPS}$  permette d'exprimer que le Verbe indexé  $\boxed{1}$  exclut la présence d'objets portant un trait de temps à l'exception de lui-même (qui peut en porter un).
- Enfin, on indique si les termes sont ordonnés entre eux ou non, i.e. si la propriété est **commutative** ou pas. P.ex. les termes d'une propriété de Précédence sont ordonnés puisque  $Det \prec N$  ("le Déterminant précède le Nom") et  $N \prec Det$  ("Le Nom précède le Déterminant") n'ont pas la même interprétation, alors que ceux d'une propriété d'Adjacence ne le sont pas puisque  $Det \leftrightarrow N$  ("le Déterminant est adjacent au Nom") signifie, tout comme  $N \leftrightarrow Det$  ("le Nom est adjacent au Déterminant"), que  $Det$  et  $N$  doivent être contigus, quel que soit leur ordre relatif.

La position relative des termes peut avoir une pertinence linguistique sans avoir de signification formelle. Par exemple, l'ordre d'une propriété d'Accord n'est pas important pour son évaluation (il s'agit de vérifier la correspondance de valeurs de traits entre deux objets), mais elle a un sens en linguistique : on dit que c'est l'adjectif qui s'accorde avec le nom et non le contraire. On a convenu d'ordonner donc les termes des propriétés d'Accord, le premier étant celui qui pose la valeur et le second celui qui doit s'accorder (i.e., on note par exemple  $N \rightsquigarrow Adj$  pour signifier que c'est l'adjectif qui s'accorde avec le nom <sup>49</sup> (ou, en d'autres termes, que c'est le nom qui impose ses marques à l'adjectif — d'où le sens de la flèche symbolisant la relation).

## 2.2.2 Représentation

Comme on l'a vu en fig. 16, les informations de forme et d'identité se présentent dans la représentation sous la forme d'un tableau dans lequel figurent en colonne de gauche le nom de l'information (symbole, termes, etc.) et dans celle de droite les valeurs propres à chacune de ces informations pour le type qui nous intéresse. Seule exception à ceci, le nom du type de propriété, qui est mis en exergue dans un cartouche en haut à gauche de la boîte représentant la définition. Voyons plus précisément quelle est la forme de chacune de ces valeurs :

- Le **nom** du type de propriété permet de l'identifier de manière unique dans la spécification. La plupart du temps on s'efforce d'attribuer aux types de propriétés qui noms qui s'efforcent d'indiquer de quelle relation il s'agit.
- Le **symbole** doit être un symbole qui n'est pas utilisé par ailleurs (p.ex. il ne peut être le  $\wedge$  qu'on utilise pour exprimer le “et logique”), et doit être également spécifique à chaque type de propriété <sup>50</sup>.
- Les **termes** se présentent sous la forme d'une liste d'objets liés entre eux par des  $\wedge$ . Chacun de ces termes reçoit deux informations : d'une part un index (de la même forme que les index dans la grammaire, i.e.  $\boxed{1}$ ,  $\boxed{2}$ , etc. — mais ces index-là ne font pas référence aux mêmes index que ceux de la grammaire) et d'autre part la spécification de la nature des objets pouvant être impliqués dans la relation qui nous intéresse. Cette nature peut être, comme on l'a dit précédemment, *Construction* quand les objets manipulés sont des constructions complètes, *Intrinsèque* quand les objets sont exclusivement des informations intrinsèques, ou *Extrinsèque* quand les objets sont exclusivement des informations extrinsèques. La mention de *Construction* indique que l'on peut faire référence aux trois types : soit des constructions entières (en mentionnant leur étiquette), soit des informations intrinsèques, soit des informations extrinsèques.

---

<sup>49</sup>Mais cet ordre n'a pas de valeur formelle, et dans la pratique si l'on écrit  $N \rightsquigarrow Adj$  ou  $Adj \rightsquigarrow N$  l'évaluation sera la même.

<sup>50</sup>Rien ne l'y contraint formellement, mais c'est préférable pour d'évidentes raisons de lisibilité des propriétés. On peut toutefois imaginer “surcharger” un symbole tout comme on “surchargerait” une fonction en programmation, en utilisant un même symbole pour deux propriétés dont la signification linguistique serait proche, mais les définitions différentes (p.ex. entre une propriété d'Exclusion Absolue qui mentionnerait un terme unique, et une autre d'Exclusion Relative qui mentionnerait des restrictions de cooccurrence relatives à certains constituants présents, donc à deux termes.)

- La **réflexivité** peut prendre trois valeurs : *Obligatoire*, *Autorisée* ou *Interdite*, chacun de ces termes explicitant clairement leur signification. Concrètement, une réflexivité obligatoire impose que les valeurs d'INDEX des termes soient égales (c'est-à-dire que  $\boxed{n}$ .INDEX =  $\boxed{n+1}$ .INDEX) ainsi que les valeurs d'étiquette (c'est-à-dire que  $\boxed{n}$ .ETIQ =  $\boxed{n+1}$ .ETIQ) <sup>51</sup>, une réflexivité interdite impose que ces mêmes valeurs soient différentes d'un terme à l'autre (c'est-à-dire que  $\boxed{n}$ .INDEX  $\neq$   $\boxed{n+1}$ .INDEX) ainsi que les valeurs d'étiquette (c'est-à-dire que  $\boxed{n}$ .ETIQ  $\neq$   $\boxed{n+1}$ .ETIQ) <sup>52</sup>, et une réflexivité autorisée n'impose aucune contrainte à ce propos.

Ajoutons à cela que dans une propriété à plusieurs termes, on pose que  $t_n \neq t_{n+1}$ . Cela signifie non pas que les deux termes doivent être de nature différente (ils peuvent tous deux être des SN, p.ex.), mais qu'ils ne doivent pas correspondre à un même objet, ce qui revient à dire que  $(t_n.index \neq t_{n+1}.index) \wedge (t_n.etiquette \neq t_{n+1}.etiquette)$ .

- La **commutativité**, enfin, porte la valeur *Vrai* quand les termes ne sont pas ordonnés entre eux, et *Faux* quand il sont ordonnés.

### 2.2.3 Exemple

Reprenons notre exemple de Précédence présenté en fig. 17 : nous rappelons le détail de ses informations de forme et d'identité en fig. 18.

Précédence	
Symbole	$\prec$
Termes	$\boxed{1}$ Construction $\wedge$ $\boxed{2}$ Construction
Réflexivité	Interdite
Commutativité	Faux

FIG. 18 – Définition de type de propriété : forme et identité d'une propriété de Précédence.

On y voit que le nom de la propriété est Précédence comme on l'avait vu plus haut ; on y voit également que le symbole utilisé pour représenter des propriétés de ce type sera  $\prec$ , que les propriétés compteront deux termes qui seront chacun des références à des constructions distinctes (puisque la réflexivité est interdite), et dont l'ordre sera pertinent (puisque le type n'est pas commutatif).

<sup>51</sup>Si  $\boxed{1}$  et  $\boxed{2}$  ont le même index et deux étiquettes différentes il s'agit de deux objets linguistiques simultanés (dans le cas d'une analyse multimodale p.ex.) ; s'ils ont la même étiquette et des index différents il s'agit de deux objets successifs (comme les deux adjectifs dans (9)) ; la conjonction des deux différences de valeurs rend impossible que les deux membres fassent référence au même objet.

<sup>52</sup>C'est la conjonction de ces deux différences qui importe : on n'interdit pas les termes portant la même étiquette mais des index différents, ou le même index mais des étiquettes différentes.

## 2.3 Conditions

### 2.3.1 Définition

Ce que l'on rassemble dans les “conditions” sont les informations qui permettent non pas d'écrire (et de lire), mais d'interpréter les propriétés, c'est-à-dire ce qui concerne leur fonctionnement. On y indique trois informations :

- La **condition de disponibilité** détermine les conditions dans lesquelles une propriété du type donné peut être évaluée ou non, étant donnée une entrée observée. Par exemple, dans la description du SN d'une grammaire figure une propriété d'Accord entre le nom et l'adjectif (qualificatif) ; or dans une entrée telle que (8),

(8) la maison de mon frère

il figure bel et bien un nom, mais pas d'adjectif : la propriété d'Accord sus-citée n'est donc pas pertinente pour la caractérisation du SN, ce qui se traduit formellement par le fait qu'elle ne remplit pas la condition de disponibilité pour l'entrée observée.

La disponibilité d'une propriété dépend directement de la disponibilité des termes qui la composent : pour qu'une propriété remplisse sa condition de disponibilité, il faut que le (ou les) membre(s) nécessaire(s) à son évaluation soi(en)t lui-même (eux-mêmes) disponible(s), i.e., présent(s). Tous les termes d'une propriété n'ont pas forcément à être disponibles pour que la propriété elle-même soit disponible : pour qu'une contrainte d'Accord soit disponible il faut effectivement que les deux termes soit présents, mais pour qu'une contrainte d'Exigence (par exemple  $N \Rightarrow Det$ ) soit disponible, seul le premier terme de la propriété ( $N$ ) doit être disponible ; de la présence ou l'absence du second ( $Det$ ) dépendra sa satisfaction, mais sa disponibilité est déjà établie par la présence du premier.

- La **condition de satisfaction**, comme son nom l'indique, permet quant à elle de signifier quelles conditions doivent être remplies pour qu'une propriété du type donné soit satisfaite (pour reprendre l'exemple de l'Accord, s'il est vérifié), sinon violée (si les objets ne s'accordent pas). Lors d'une analyse, elle n'est évidemment observée que pour les propriétés qui ont rempli leur condition de disponibilité (l'Accord entre adjectif et nom ne peut être ni satisfait ni violé pour (8) puisqu'il n'y a pas d'adjectif).

Elle peut elle-même porter soit sur la disponibilité d'un terme (p.ex. la satisfaction d'une propriété d'Exigence dépend de la disponibilité de son second terme : s'il est présent alors elle est satisfaite, s'il ne l'est pas alors elle est violée), soit sur tout ou partie des informations correspondant aux objets auxquels les termes font référence (p.ex. les index pour une Précédence).

- Enfin, la **capacité** indique le nombre de fois qu'une même propriété du type concerné peut être évaluée pour une même construction caractérisée. P.ex. dans un SN tel que (9),

(9) un gentil petit chien

la propriété indiquant la Tête de la construction ne pourra être évaluée qu'une seule fois (pour *chien* en l'occurrence) alors que la propriété indiquant la présence possible d'un adjectif sera évaluée deux fois, une pour chaque occurrence de la catégorie (*gentil* d'une

part, *petit* d'autre part).<sup>53</sup>

Il est à noter que la capacité concerne le nombre de fois qu'une propriété peut être évaluée, et non pas qu'un type de propriété peut être évalué pour une construction, ce qui signifie que si la capacité d'un type de propriété est de 1 alors chaque propriété de ce type ne pourra être évaluée qu'une fois pour une construction donnée, mais si cette construction présente plusieurs propriétés de ce même type alors chacune d'elles pourra être évaluée une fois. Par exemple on peut imaginer une grammaire pour laquelle dans le SN déterminé figurent deux propriétés d'Obligation : une première donnant la liste des têtes possibles, une seconde celle des spécifieurs possibles. Chacune de ces propriétés (tête d'une part, spécifieur d'autre part) pourra être évaluée une (et une seule) fois pour une construction.

### 2.3.2 Représentation

Comme on l'a vu en fig. 16 l'expression des conditions prend la même forme de tableau que celle des informations de forme et d'identité du type de propriété. Chacune d'elles prendra la forme suivante :

- La **disponibilité** prend la forme d'une liste dont chaque item correspond à la mention de la disponibilité de l'un des termes du type de propriété, représenté par l'index lui correspondant. Par exemple dans une propriété d'Exigence dont la disponibilité requiert la présence du premier terme (mais pas de celle du second), on indique  $\boxed{1} = disponible$ . Pour une propriété d'Accord dont la disponibilité requiert la présence des deux termes de la propriété, on indique  $\boxed{1} = disponible \wedge \boxed{2} = disponible$ .
- La **satisfaction** prend la forme d'une liste dont chaque item correspond soit à la disponibilité de l'un des termes indexés précédemment (p.ex. dans le cas d'une propriété d'Exigence, la condition de satisfaction prendra la forme  $\boxed{2} = disponible$  puisque pour qu'une telle propriété soit satisfaite il faut que le second terme soit présent), soit à la comparaison de parties des termes précédemment indexés (p.ex. pour une propriété de Précédence la condition de satisfaction pourra prendre la forme  $\boxed{1}.INDEX|DEBUT < \boxed{2}.INDEX|DEBUT$  ce qui signifie que pour qu'une telle propriété soit satisfaite il faut que la valeur de l'index de début de l'objet représenté par le terme indexé  $\boxed{1}$  soit inférieure à la valeur de l'index de début de l'objet représenté par le terme indexé  $\boxed{2}$ ).
- La **capacité**, enfin, reçoit une valeur numérique comprise entre 1 et  $\infty$ <sup>54</sup>.

---

<sup>53</sup>[Vanrullen, 2005] a découpé cette information en deux : une **capacité maximale** qui correspond à ce que l'on vient de définir, et une **capacité minimale** qui sert à imposer (dans une certaine mesure) qu'un type de propriété soit évalué. Par exemple cela permet de ne pas introduire de construction dont aucune propriété Tête n'est évaluée. Nous considérons pour notre part que cette "capacité minimale" constitue typiquement une heuristique dont le formalisme, en soi, n'a pas besoin : ce n'est que dans le cadre de son *implémentation* que l'on peut avoir recours à ceci en tant que moyen de limiter la surgénération.

<sup>54</sup>Une capacité de 0 reviendrait à définir un type dont les propriétés ne peuvent jamais être évaluées, ce qui n'aurait aucun sens.

### 2.3.3 Exemple

Reprenons une nouvelle fois notre exemple de définition de type Précédence ; nous en rappelons la partie qui nous intéresse en fig. 19.

Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Satisfaction	$\boxed{1}.\text{INDEX.DEBUT} < \boxed{2}.\text{INDEX.DEBUT}$
Capacité	$\infty$

FIG. 19 – Définition de type de propriété : Conditions d'une propriété de Précédence.

Notre définition présente bien la mention des trois informations : on y voit d'abord que pour que cette propriété soit disponible il faut que les termes précédemment indexés  $\boxed{1}$  et  $\boxed{2}$  soient tous deux disponibles, puis que la valeur l'index de début de l'objet auquel fait référence le terme indexé  $\boxed{1}$  soit inférieure à la valeur l'index de début de l'objet auquel fait référence le terme indexé  $\boxed{2}$  pour que la propriété soit satisfaite ; enfin on voit que chaque propriété de ce type peut être évaluée autant de fois qu'elle sera disponible (puisque sa capacité est infinie).

## 2.4 Evolution de la spécification

Il est notable que la présentation de la spécification des propriétés que l'on vient de faire est différente de celles proposées dans [Vanrullen *et al.*, 2003] puis dans [Vanrullen, 2005] ; cela provient du fait que cette ressource est en pleine évolution et qu'elle est complétée et affinée à mesure des recherches effectuées, conjointement et/ou individuellement. Sa forme n'est donc pas encore tout à fait stable à l'heure actuelle, et il reste un certain nombre d'informations à spécifier de manière sans doute plus précise que ce que l'on fait ici <sup>55</sup>.

Toutefois, même dans sa version actuelle, pour imparfaite qu'elle puisse être, elle constitue une avancée importante dans la définition de GP en ce qu'elle permet une expression régulière et approfondie du fonctionnement de chacune des propriétés que l'on peut utiliser, et c'est la raison pour laquelle, bien qu'elle nécessite des recherches complémentaires, elle nous est d'ores et déjà utile et plus précise que ce qui était auparavant disponible pour représenter une grammaire formelle.

Jusqu'ici nous avons vu les deux ressources permettant la formalisation d'un modèle théorique. Il est important maintenant de présenter le déroulement d'une analyse en GP afin de bien saisir la façon dont fonctionnent les objets que nous pouvons décrire et la raison pour laquelle les informations sont représentées comme on vient de le voir : c'est ce que nous présentons dans la section qui suit.

---

<sup>55</sup>On revient sur ce point dans les Perspectives, §5.2.



### 3 Analyse

La caractérisation est le moyen d'analyse utilisé en GP. Il consiste à poser qu'une construction donnée est introduite si une certaine proportion <sup>56</sup> des propriétés pertinentes (figurant dans la grammaire) a été évaluée (en vertu de la spécification des propriétés). En d'autres termes, cela signifie que chaque ensemble de propriétés, mettant en relation un certain nombre d'objets, faisant partie de la définition d'une construction donnée, et dont une proportion minimale convenue a été évaluée (comme satisfaite ou non), aura pour conséquence l'introduction de la construction définie. La caractérisation est donc l'ensemble de propriétés satisfaites et non satisfaites correspondant à une construction donnée <sup>57</sup>.

**Architecture générale.** — Plus concrètement, on se représente l'analyse de la façon suivante : dans un premier temps on repère tous les indices susceptibles de mettre en relation les objets présents les uns avec les autres. Dans cet ensemble d'indices, certains signifient que des objets peuvent être en relation (une ou plusieurs propriétés qui les lient sont satisfaites) et d'autres signifient au contraire que des objets ne peuvent pas être en relation (une ou plusieurs propriétés violées entre eux). De ceci émerge des ensembles de relations créant un réseau dense entre certains objets, et correspondant à la définition d'une construction donnée : celle-ci est alors introduite, et l'on pourra par la suite évaluer les propriétés qui la mettent en relation avec les autres objets de l'énoncé en suivant le même processus.

L'analyse en GP se déroule en plusieurs étapes successives, correspondant chacune à un **niveau** d'analyse, et s'effectuant en deux "mouvements" : un premier qui aboutit à l'établissement d'un **réseau de relations** entre les objets (§3.1), un second qui mène à l'**introduction de nouvelles constructions** (§3.2). Une fois le deuxième mouvement effectué on repasse au premier mouvement pour mettre en relation les constructions obtenues entre elles et avec le reste de l'énoncé, et ainsi de suite jusqu'à ce qu'aucune nouvelle relation ne puisse être établie. On donne une illustration de ceci en fig. 20, et on présente chacun de ces deux mouvements dans les sections qui suivent <sup>58</sup>.

#### 3.1 Premier mouvement : Construction du réseau de relations

Le premier mouvement consiste à mettre en relation les constructions alors disponibles à l'aide des propriétés contenues dans la grammaire et obtenir ainsi un **réseau de relations** complexe.

##### 3.1.1 Evaluation de la disponibilité des propriétés

Lors de la construction de ce réseau de relations, toutes les propriétés (i.e., toutes les relations potentielles figurant dans les informations extrinsèques de toutes les constructions de

<sup>56</sup>Cette proportion fait l'objet d'une convention, et est variable suivant les objectifs.

<sup>57</sup>C'est-à-dire ce qu'on appelle l'*état du système de contraintes* d'un point de vue procédural.

<sup>58</sup>On ne présente pas d'exemple ici : le lecteur pourra se référer à la partie 3 pour une illustration de l'analyse en GP, et plus spécialement au chapitre 12 qui présente un exemple d'analyse pas à pas.

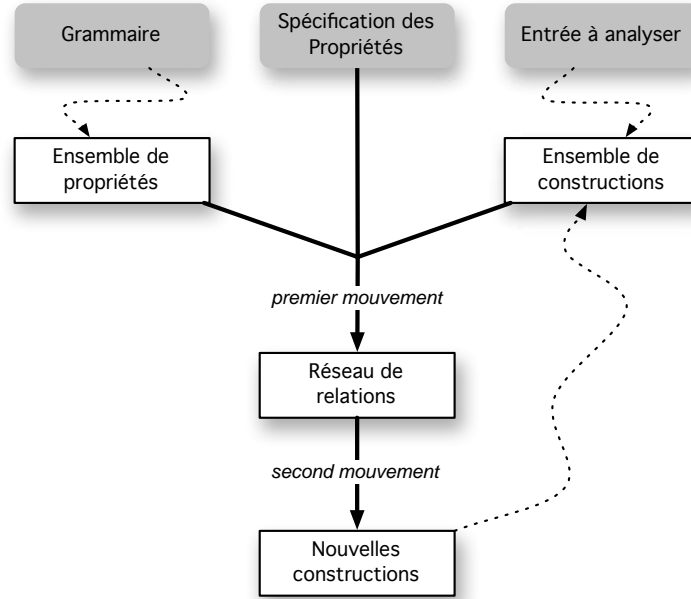


FIG. 20 – Architecture générale du processus d'analyse.

la grammaire) sont observées, sans prise en considération des constructions auxquelles elles appartiennent. Dans un premier temps, cet ensemble général de propriétés (appelons-le  $\mathcal{P}_G$ ,  $G$  étant la référence à la grammaire utilisée) est divisé en deux sous-ensembles :

- Le premier sous-ensemble,  $\mathcal{P}_{G,e,n}^{\neg disp}$  (où  $e$  fait référence à l'énoncé observé et  $n$  au niveau d'analyse de cet énoncé), est constitué des **propriétés non disponibles**, i.e. dont la condition de disponibilité n'a pas été remplie, ce qui signifie qu'elles ne sont pas pertinentes pour l'analyse des objets concernés. Par exemple, si le niveau  $n$  de l'énoncé  $e$ , que l'on est en train d'observer, est une séquence déterminant - nom - verbe, les propriétés concernant les adverbes ne seront pas pertinentes et seront donc placées dans cet ensemble.
- Le second sous-ensemble,  $\mathcal{P}_{G,e,n}^{disp}$ , contient les **propriétés disponibles**.

### 3.1.2 Evaluation de la satisfaction des propriétés

Une fois  $\mathcal{P}_{G,e,n}^{disp}$  établi, chacune des propriétés y figurant est évaluée (c'est-à-dire que l'on vérifie sa condition de satisfaction), et c'est le résultat de cette évaluation qui constitue le réseau de relations du niveau d'analyse auquel on se situe. Ce réseau, appelons-le  $\mathcal{P}_{G,e,n}^{eval}$  se divise lui-même en deux sous-parties : l'ensemble  $\mathcal{P}_{G,e,n}^+$  des **propriétés satisfaites**, et l'ensemble  $\mathcal{P}_{G,e,n}^-$  des **propriétés non satisfaites**.

A chaque nouveau niveau d'analyse, on réutilise l'ensemble  $\mathcal{P}_G$  comme point de départ, i.e. on

remet toutes les propriétés de la grammaire en jeu.

### 3.2 Second mouvement : Introduction des nouvelles constructions

Le second mouvement consiste à observer le système de contraintes issu du premier mouvement (c'est-à-dire  $\mathcal{P}_{\mathcal{G},e,n}^{eval}$ ) et à en déduire les **nouvelles constructions** que l'on peut introduire.

#### 3.2.1 Caractérisation

A cette fin, la première étape consiste à retrouver à quelle(s) construction(s) chaque propriété évaluée correspond. Il est à noter qu'une propriété donnée peut faire partie de la description grammaticale de plusieurs constructions. Par exemple, la propriété d'accord en genre entre nom et adjectif,  $N.GENRE \rightsquigarrow Adj.GENRE$ , figure dans plusieurs constructions dont notamment le N construit, le SN exocentrique, le SN endocentrique, etc.

A l'issue de ceci, les constituants de l'énoncé sont liés entre eux non plus par des propriétés, mais par des ensembles de propriétés <sup>59</sup>, lesquels ensembles sont rapportés aux constructions auxquelles ils correspondent. Chacun de ces ensembles est nommé  $\mathcal{C}_{\mathcal{G},e,n}^s$ , c'est-à-dire que chaque ensemble constitue la **caractérisation**  $\mathcal{C}$  d'une construction donnée  $s$  <sup>60</sup> relativement à la grammaire  $\mathcal{G}$ , à l'énoncé  $e$  et au niveau d'analyse  $n$ . L'ensemble des constructions caractérisées est symbolisé par  $\mathcal{C}_{\mathcal{G},e,n}$  (en quelque sorte, il s'agit de la caractérisation générale du niveau d'analyse  $n$  de l'énoncé  $e$ ).

Puisque cette caractérisation est établie à partir de  $\mathcal{P}_{\mathcal{G},e,n}^{eval}$ , qui contient à la fois les propriétés satisfaites et non satisfaites, cela signifie que de la même façon, chaque  $\mathcal{C}_{\mathcal{G},e,n}^s$  sera constitué de deux sous-ensembles,  $\mathcal{C}_{\mathcal{G},e,n}^{s+}$  des propriétés satisfaites pour cette construction et  $\mathcal{C}_{\mathcal{G},e,n}^{s-}$  des propriétés non satisfaites pour cette construction.

#### 3.2.2 Introduction des constructions caractérisées

Tous les objets ainsi caractérisés ne correspondent pas nécessairement à autant de constructions effectivement introduites en tant que résultat possible. En effet l'étape d'**introduction des nouvelles constructions** qui suit l'étape de caractérisation, consiste à sélectionner parmi  $\mathcal{C}_{\mathcal{G},e,n}$  les caractérisations que l'on estime suffisamment pertinentes pour être considérées comme une analyse possible.

La notion de "pertinence" à cette étape dépend du type d'analyse que l'on souhaite effectuer : si l'on ne veut proposer que des résultats parfaitement conformes à la grammaire on n'introduira que les constructions dont  $\mathcal{C}_{\mathcal{G},e,n}^{s-}$  est vide, i.e. qui n'ont pas de propriétés violées. Si en revanche

<sup>59</sup>Dans lesquels la mention de chaque propriété n'est pas exclusive, puisque l'on vient de voir qu'une même propriété peut être affectée à plusieurs constructions.

<sup>60</sup>On note  $s$  pour "syntagme", de manière à éviter une confusion possible entre le  $\mathcal{C}$  de "caractérisation" et le  $c$  de "construction".

on veut rendre l'analyse plus souple on peut convenir d'introduire les constructions dont  $\mathcal{C}_{\mathcal{G},e,n}^{s-}$  n'est pas vide mais contient moins d'un nombre donné de propriétés, ou alors ne contient pas de propriétés de tel(s) type(s) (constituance, précédence,...), ou encore dont  $\mathcal{C}_{\mathcal{G},e,n}^{s+}$  représente plus de  $x\%$  de  $\mathcal{C}_{\mathcal{G},e,n}^s$ , ou de convenir de tout autre moyen de sélection des constructions à introduire en fonction de leurs caractérisations respectives.

Cette possibilité offerte de tenir compte d'informations précises aussi bien concernant la qualité d'une construction (son *degré de grammaticalité*, cf. notamment [Blache *et al.*, 2006] à ce propos) en référence à une grammaire (proportion de propriétés satisfaites parmi les propriétés évaluées) que la quantité d'informations disponibles la concernant (nombre de propriétés évaluées parmi les propriétés existantes, *densité d'information* existant pour une construction donnée, cf. à ce propos [Blache, 2004b]), permettent de proposer plusieurs types de résultat en fonction des objectifs propres à un besoin particulier, et c'est l'une des spécificités de GP que de permettre un tel choix.

Au terme de ce second mouvement, on obtient un nouvel ensemble d'objets, constitué des constructions que l'on vient d'introduire et des objets n'étant pas des constituants de ces nouvelles constructions le cas échéant, ensemble qui fera l'objet du niveau d'analyse suivant, jusqu'à ce que les objets ne puissent plus être mis en relation les uns avec les autres.

**Contenu des constructions introduites.** — Pendant une analyse, les constructions introduites représentent des instanciations des définitions de la grammaire en fonction d'une production linguistique à analyser, et correspondent à un ensemble de propriétés (leur caractérisation) qui est un sous-ensemble de celui ayant servi à la définition de la construction : toutes les propriétés figurant dans la grammaire pour une telle construction n'ont pas nécessairement été évaluées, et certaines ont pu être instanciées plusieurs fois pour représenter un même type de contrainte s'exerçant sur plusieurs objets ou couples d'objets différents.

De plus, chaque propriété figurant dans la caractérisation d'une construction contient une information supplémentaire par rapport aux propriétés de la grammaire : le fait qu'elle soit satisfaite ou non. Finalement, une instance de construction introduite à la suite de l'évaluation du système de contraintes ressemblera à la figure 21 : Dans les informations extrinsèques on ne mentionne que les propriétés évaluées, en précisant pour chacune si elle est satisfaite ou non.

Construction		
INTR. [...]		
TYPE DE PROPRIÉTÉ <sub>1</sub>	propriété <sub>1a</sub> propriété <sub>1b</sub>	$\pm$ <i>satisfaite</i> $\pm$ <i>satisfaite</i>
...	...	...
TYPE DE PROPRIÉTÉ <sub>n</sub>	propriété <sub>n</sub>	$\pm$ <i>satisfaite</i>

FIG. 21 — Forme de base d'une construction introduite lors d'une analyse.

**Contraintes.** — Pour finir, notons que l'ensemble de ce processus d'analyse est conçu comme un problème de satisfaction de contraintes. En effet, pour pouvoir tenir compte des informa-

tions de natures diverses (de plusieurs domaines) et dans des proportions diverses (de finesse, d'importance et de grammaticalité variables), l'utilisation des contraintes, dont l'évaluation individuelle (par flexibilité qu'elle permet) et la mise en relation (par l'homogénéité qu'elle représente) procurent la souplesse nécessaire à ce type de considération ([Johnson & Lappin, 1999], [Blache, 2000b]), sans être réduites à un système de filtrage.

D'autres systèmes ont recours à la résolution de contraintes pour diverses tâches : pour désambiguïser des informations catégorielles (*Grammaire de Contraintes*, [Karlsson, 1990]), pour contrôler la forme des structures syntaxiques (*Grammaires de Dépendances par Contraintes* [Maruyama, 1990], *eXtensible Dependency Grammars* [Debusmann *et al.*, 2004]), pour contrôler la grammaticalité des objets construits (HPSG), ou pour filtrer des analyses (OT, MG [Villemonte de la Clergerie, 2005], le *Programme Minimaliste*). En revanche, GP est le seul formalisme à faire reposer la totalité de l'analyse sur ce principe, sans nécessiter le recours à un autre mécanisme.

## Résumé

Dans ce chapitre on a présenté l'architecture du formalisme des *Grammaires de Propriétés*. On a vu que la formalisation d'un modèle de grammaire prend la forme d'un couple de ressources : une **grammaire** (§1) d'abord, qui rassemble la totalité des informations proprement descriptives du modèle, et une **spécification des propriétés** (§2) qui permet de décrire la forme et le fonctionnement des propriétés utilisées dans cette grammaire.

On a vu que la grammaire est constituée d'un ensemble de **constructions** qui sont liés entre elles en une hiérarchie d'héritage. Chacune de ces constructions se présente sous la forme d'un ensemble constitué d'un **identifiant** permettant de pointer la construction dans la grammaire, et d'**informations intrinsèques** d'une part, représentées en une structure de traits, et **extrinsèques** d'autre part, représentées sous forme de propriétés.

On a vu également que la spécification des propriétés définit l'interprétation et le fonctionnement des **types de propriétés** de la grammaire. C'est l'application de ceci aux descriptions de la première ressource qui constitue le pivot de GP puisque c'est cela qui permet d'effectuer une analyse.

Enfin au §3 nous avons présenté le déroulement d'une analyse en GP afin d'illustrer le fonctionnement du formalisme en action et de montrer ce que cette technique permet.

Dans le chapitre suivant, nous présentons une application possible de cette architecture, à travers la formalisation des positions théoriques présentées au précédent chapitre. Puis au chapitre 4 nous montrerons quelques formalisations possibles de modèles différents afin de montrer quelles sont les possibilités offertes par GP.

## Chapitre 3

# Notre modèle de grammaire

Dans le premier chapitre on a introduit les positions théoriques qui sont les nôtres ; puis dans le suivant on a présenté le formalisme des *Grammaires de Propriétés*, que l'on a utilisé pour représenter notre proposition. Ici nous présentons maintenant le modèle de grammaire que nous proposons, qui est la mise en oeuvre des idées du chapitre 1 appliquée à la description du français, et représentée selon le formalisme du chapitre 2. Ce modèle n'est qu'une utilisation précise et personnelle de GP, et l'on pourra voir dans le chapitre 4 qu'il est possible de représenter des grammaires issues de théories fort différentes de celle-ci selon ce même formalisme ; c'est par ailleurs ce qui nous a permis de développer une autre grammaire que celle qui fait l'objet de cette thèse, et qui est présentée en annexe A.

Nous présentons dans ce chapitre les choix d'organisation générale des informations qui sont spécifiques à la grammaire dont des éléments constitutifs sont proposés en partie 2, en fonction du formalisme de représentation que l'on a choisi d'utiliser.

### 1 Structuration des informations intrinsèques

Dans nos constructions nous utilisons la structuration de base présentée en figure 22.

Selon cette structuration, l'information est divisée en plusieurs grandes parties. Nous détaillons dans ce qui suit les parties que l'on aura l'occasion d'utiliser dans la Partie 2, mais il est important de garder à l'esprit que les éléments de grammaire qui font l'objet du présent travail ne constituent qu'une partie de la ressource complète ; cela signifie que les traits que l'on présente dans les sections qui suivent ne constituent qu'une partie de la structure de traits complète, et que la liste sera à compléter à chaque nouvelle étape de complétion de la grammaire.

INTR.	ANCRAGE	INDEX	$\begin{bmatrix} \text{DÉBUT} & \text{début} \\ \text{FIN} & \text{fin} \end{bmatrix}$
	FORME	GRAPH	forme graphique
	ID	DOM	domaine d'appartenance
		C.I.	constituants immédiats
		NATURE	TYPE type de construction
			CAT catégorie
			SCAT sous-catégorie
			TRAITS traits
		ACCORD	informations d'accord
SYN	DÉPENDANCE <sub>1</sub>	RECTEUR	index vers recteur
		DÉPENDANT	index vers dépendant
	DÉPENDANCE <sub>2</sub>	RECTEUR	index vers recteur
		DÉPENDANT	index vers dépendant
			etc.

FIG. 22 – Structuration des informations intrinsèques dans la grammaire.

## 1.1 Ancrage

Les premier ensemble de caractéristiques propres à une construction est son ancrage <sup>61</sup>. A ce titre plusieurs types de références et d'index peuvent être utilisés simultanément ou optionnellement pour localiser une même construction. Celle-ci pourra être référencée non seulement par sa **position temporelle** dans le signal, mais aussi par son **index** dans la chaîne linguistique correspondante (cf. [Bird & Liberman, 1999]). On peut également au besoin ajouter un attribut s'appuyant sur l'**univers de discours** (cf. [Kamp & Reyle, 1993]). Enfin, pour toutes les unités on pourra ajouter l'**identification du (ou des) locuteur(s)** qui les a (ont) produites <sup>62</sup>.

Cette technique d'ancrage multiple, telle que proposée dans [Blache, 2003a] puis [Guénot & Bellengier, 2004], permet de localiser la construction non seulement dans la chaîne linguistique produite mais également dans toutes les autres dimensions jugées nécessaires à l'analyse (pragmatique, sémantique, etc.), et dont l'organisation n'est pas nécessairement aussi linéaire que la chaîne de production (comme le fait remarquer [Kracht, 2005b]).

<sup>61</sup>En effet, on pourrait considérer que l'ancrage dans un énoncé dépend d'une occurrence précise de la construction et non pas de la construction elle-même, mais pourtant le fait qu'une construction ait un ancrage est bel et bien une information qui lui est propre (ce sont les valeurs d'ancrage qui sont variables, pas l'ancrage lui-même). C'est la raison pour laquelle on le fait figurer dans les informations intrinsèques de la construction.

<sup>62</sup>Sur la possibilité de considérer plusieurs locuteurs comme producteurs d'une même construction, cf. par exemple [Loufrani, 1981; Loufrani, 1984].



## 1.2 Forme

Cet attribut rassemble les informations concernant la forme de la construction. Elle peut recevoir des informations sur la forme ORTHographique ou PHONique (auquel cas, phonologique et/ou phonétique).

Cet attribut ne sert que peu dans la grammaire, surtout au degré de généralité des éléments qui suivent. On n'est ici pas descendus dans la finesse de description assez profondément pour décrire des constructions dont la forme est contrainte (ce qui serait le cas de formes plus ou moins figées, comme par exemple *il y a* / *il y avait* / etc. ou *y compris*). Par contre on trouve ce trait à la fois dans les analyses (elles font référence aux formes des objets constituant les constructions), et évidemment dans les constructions lexicales (où à chaque construction est associée une forme).

## 1.3 Identité (ID)

Cet attribut rassemble ce qui permet de définir réellement l'identité grammaticale de la construction. Elle contient les informations suivantes :

- Le **domaine** (DOM) d'appartenance de la construction (morphologie, syntaxe, etc.). Les constructions étant par définition des objets pouvant appartenir à plusieurs domaines, les valeurs de cet attribut sont par conséquent non exclusives (un objet peut être de [DOMAINE *morphologie*  $\wedge$  *syntaxe*] par exemple)

Concrètement, puisque les éléments de grammaire que l'on présente ci-après ne sont, comme on l'a dit, que des constructions syntaxiques, on ne précisera pas ce trait, qui sera sous-entendu, mais qui devrait figurer pour que la structure soit réellement complète.

- La liste des **constituants immédiats** (C.I.) de la construction, ces constituants immédiats étant des références à des constructions. Nous utilisons ici cette appellation sans la limiter à la notion de “constituants immédiats” en syntaxe, c'est-à-dire que toutes les constructions, quel que soit leur domaine d'appartenance, reçoivent une liste de constituants immédiats (une unité syntaxique minimale, par exemple, recevra des constructions morphématiques en guise de constituants immédiats, et rien n'interdit que cette liste soit constituée de constructions de domaines différents, dès lors qu'il s'agit bien de constructions).

### 1.3.1 Nature

Il reçoit également un ensemble de traits rassemblés sous l'appellation **nature** <sup>63</sup>, et qui sont les suivants :

- Le **type de construction** (TYPE), dont la valeur est une partie du discours (pour les syntagmes) ou bien une étiquette correspondante (pour les entassements paradigmatiques).

---

<sup>63</sup>On parle de “nature” ici comme quand on l'emploie par opposition à “fonction” dans les grammaires scolaires.

Ceci correspond à la “*classe formelle*” de [Bloomfield, 1961] (cf. discussion à ce propos plus bas).

- La **catégorie** syntagmatique (CAT) de la construction, i.e. N, SN, etc. La catégorie, contrairement au type, correspond à une étiquette syntagmatique, ce qui signifie qu’elle ne peut pas être “entassement paradigmatique” (une telle construction, comme on le verra en détail au chapitre les décrivant, est de [TYPE *EP*] et porte la (ou les) catégorie(s) correspondant à ses constituants, p.ex. [CAT *SN*] pour *Pierre et Marie*).
- Eventuellement, la **sous-catégorie** (SCAT) de la construction, i.e. *propre* ou *commun* pour un Nom, *qualificatif* ou *relationnel* pour un Adjectif, *principal*, *auxiliaire* ou *modal* pour un Verbe, etc. Il s’agit donc littéralement d’une sous-classe de la catégorie, et non pas de l’acception de “*sous-catégorisation*” que l’on trouve dans de nombreuses grammaires formelles (comme en HPSG par exemple).
- La liste des **traits** (TRAITS) de la construction, i.e. les informations spécifiques qui lui sont *intrinsèques*, comme par exemple le GENRE pour les Noms.
- La liste des **accords** (ACC) de la construction, i.e. les informations spécifiques qui lui sont *extrinsèques*, comme par exemple le GENRE pour les Adjectifs ou les Déterminants. Nous appelons ce trait “accord” parce que ces marques sont la manifestation d’un accord avec l’objet dont ils dépendent (déterminant et adjectif s’accordent avec le nom dont ils dépendent, et à ce titre portent la marque du genre du nom).

Il est à noter que l’on effectue, dans la version de la grammaire qui suit, une utilisation abusive des traits TRAITS et ACCORD, puisque l’on y mentionne non seulement des informations qui sont réellement des traits (et des accords) à proprement parler (comme c’est le cas du GENRE), mais on y mentionne également des informations qui ne sont pas des traits de l’objet mais des **morphèmes grammaticaux** <sup>64</sup>, comme notamment le NOMBRE, la PERSONNE, etc. Nous sommes consciente que ces deux informations devront être séparées, mais nous ne l’avons pas fait dans la version actuelle de la grammaire pour des raisons de compatibilité avec les analyseurs automatiques et avec les ressources existantes avec lesquelles nous avons mené nos recherches (notamment un lexique qui ne fait pas la différence entre traits et morphèmes grammaticaux, et qui considère que les unités syntaxiques minimales sont des mots et non des morphèmes, ce qui est tout à fait discutable). De fait, même si la distinction entre les deux est nécessaire et devra être faite, en ce qui concerne notre tâche présente, d’analyse syntaxique, la distinction formelle entre les deux ne s’est pas révélée indispensable (puisque dans l’état actuel on arrive à accéder à ces informations, sans avoir absolument besoin de les distinguer pour faire de la syntaxe <sup>65</sup>); elle le sera lorsqu’il s’agira de développer les constructions de taille inférieure aux constructions syntaxiques, afin de permettre une analyse morphologique notamment.

---

<sup>64</sup>“*Morphèmes grammaticaux*” selon l’appellation standard, que l’on trouve notamment chez [Touratier, 2002], qui correspondent aux diverses “*modalités*” (nominales, verbales, etc.) de [Martinet, 1979].

<sup>65</sup>Notons d’ailleurs qu’il n’est pas rare qu’elles ne soient pas distinguées dans les modèles formels que l’on peut rencontrer.

## 1.4 Syntaxe (SYN)

Cet attribut rassemble les informations syntaxiques internes de la construction.

Dans les constructions syntagmatiques (cf. chapitre 5 §2.2), les constituants Majeurs sont référencés sous les attributs TÊTE, et SPÉCIFIEUR le cas échéant ; les constituants Mineurs sont référencés comme RECTEUR ou DÉPENDANT d’une relation binaire s’établissant avec la tête de la construction. Ces dépendances syntaxiques auront alors une forme ressemblant à :

$$\text{SYN} \left[ \begin{array}{l} \text{TÊTE} \quad \boxed{1}SV \\ \text{SUJET} \quad \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{2}SN \\ \text{DÉPENDANT} \quad \boxed{1}. \text{TÊTE} \end{array} \right] \\ \text{OBJ-DIR} \quad \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{1}. \text{TÊTE} \\ \text{DÉPENDANT} \quad \boxed{3}SN \end{array} \right] \end{array} \right]$$

On peut ajouter à ce cela les relations hypotactiques s’établissant entre deux constituants d’une construction macrosyntaxique, c’est-à-dire définir un NOYAU macrosyntaxique, qui entretiendrait des relations avec des PRÉFIXES, des POSTFIXES et/ou des SUFFIXES ; on n’a pour l’instant pas fait figurer ces relations dans nos descriptions, mais rien n’empêche de le faire. L’une des raisons pour lesquelles on n’a pas encore fait figurer ces informations est ce que la relation syntaxique entre un NOYAU et un SUFFIXE est censée être paratactique, alors même que l’on définit cette relation en posant que le SUFFIXE dépend (informativement) du NOYAU.

pour une Proposition de la forme *Pierre mange une pomme*, ou bien pour un SN de la forme *le petit chien* on aura une structure syntaxique ressemblant à :

$$\text{SYN} \left[ \begin{array}{l} \text{TÊTE} \quad \boxed{1}N \\ \text{SPÉCIFIEUR} \quad \boxed{2}Det \\ \text{ÉPITHÈTE} \quad \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{1} \\ \text{DÉPENDANT} \quad \boxed{3}Adj \end{array} \right] \end{array} \right]$$

Dans les constructions paradigmatisées (cf. chapitre 5 §2.3), les constituants n’entretenant que des relations paratactiques, il n’existe pas de dépendance syntaxique entre eux. On ne spécifie donc rien à leur propos sous l’attribut SYN <sup>66</sup>.

### 1.4.1 La notion de “tête”

Dans les syntagmes nous avons mentionné la présence d’un attribut TÊTE. Il est utile de préciser que nous utilisons ici ce terme dans une acception légèrement différente de celle

<sup>66</sup>Il serait néanmoins intéressant de réfléchir au type d’information que l’on pourrait indiquer sous cet attribut pour ces constructions. [Mithun, 2005], p.ex., postule que les relations paradigmatisées sont (essentiellement) pragmatiques ; [Deulofeu, 2005] y voit pour sa part des relations syntaxiques “primitives” qui, à la différence des relations syntagmatiques, ne sont pas basées sur les catégories.

utilisée par la syntaxe X-barre ([Harris, 1968]) et jusqu'à HPSG, et basée sur la notion d'endo-/exocentricité définie par [Bloomfield, 1961] :

*Chaque construction syntaxique nous présente deux formes libres (et quelquefois plus) combinées en un syntagme, que l'on peut appeler syntagme résultant. Le syntagme résultant peut appartenir à une classe formelle autre que celle d'un constituant quelconque. Par exemple, John ran n'est pas une expression nominale (comme John) ni une expression verbale conjuguée (comme ran). C'est pourquoi nous disons que la construction anglaise acteur-action est exocentrique : le syntagme résultant appartient à la classe formelle de constituant non-immédiat. Par ailleurs le syntagme résultant peut appartenir à la même classe formelle que l'un (ou davantage) de ses constituants. Par exemple, poor John est une expression de nom propre, de même que le constituant John ; les formes John et poor John ont, dans l'ensemble, les mêmes fonctions. Par conséquent, nous dirons que la construction anglaise qualité-substance (comme dans poor John, fresh milk, etc.) est une construction endocentrique.*

Dans cette définition, l'auteur pose qu'un syntagme endocentrique est un objet qui est de la même "classe formelle" que l'un de ses constituants immédiats. Au sein des constituants immédiats, il définit ce qu'il appelle "centre" (le N *milk* dans le SN *very fresh milk*), c'est-à-dire la "tête". Il décrit le SN comme un syntagme endocentrique, tout en précisant bien qu'il considère les SN anglais ; or pour notre part, nous ne considérons pas que l'endocentricité soit aussi évidente pour les SN français <sup>67</sup>, mais nous souhaitons toutefois conserver cette notion de "tête", que l'on pense pertinente aussi bien pour les syntagmes endocentriques que pour les syntagmes exocentriques. D'une certaine manière, on peut interpréter l'idée de similitude de "classe formelle" de Bloomfield comme étant traduite par le "type de construction" dans notre approche <sup>68</sup>, où un N construit tel que *lait frais* et un SN tel que *du lait frais* sont de [TYPE N]. Dans les deux cas, *lait* sera la tête de la construction ; mais le N construit *lait frais*, commutant <sup>69</sup> avec sa tête, sera considéré comme un syntagme endocentrique, alors que le SN *du lait frais*, ne commutant pas avec sa tête, sera considéré comme un syntagme exocentrique. La différence entre les deux est signifiée par leurs catégories distinctes, N pour l'un, SN pour l'autre, et leur point commun est signifié par leur type identique, N dans les deux cas.

On peut s'étonner que cette notion que l'on convient d'appeler "tête" pour les besoins du présent exercice et bien que cet emploi du terme ne corresponde pas à sa définition habituelle, ne figure pas dans la littérature linguistique. On peut rencontrer la notion de "**noyau**" (p.ex. [Riegel *et al.*, 1999]), qui ne répond pas à nos besoins puisque il ne fait pas nécessairement référence à une unité minimale : ainsi le noyau de *du lait frais* serait *lait frais* et non simplement *lait*. Même chose pour le "**nucleus**" de [Tesnière, 1959] (dont le noyau est une transposition en grammaires syntagmatiques) qui est constitué de l'ensemble des objets régissant directement ou non un noeud donné. Dans les grammaires de dépendance on trouve aussi la notion de "**recteur / gouverneur**", qui porte bien sur une unité minimale, mais dont la définition ne correspond pas à ce que l'on cherche puisqu'il s'agit d'une relation binaire s'établissant entre deux objets lexicaux, et qu'il est impossible d'extrapoler cette idée à une rection qui se ferait, dans une grammaire syntagmatique, entre une unité minimale et un syntagme dont elle est constituant. On peut également rencontrer la notion d'"**ancree**" dans les grammaires lexicalisées, qui diffère elle aussi de notre besoin puisqu'il s'agit ici certes d'un élément lexical, mais qui est associé à une structure élémentaire (p.ex. à un arbre élémentaire en TAG, cf. à

<sup>67</sup>Sur ce point nous rejoignons par exemple [Touratier, 2005].

<sup>68</sup>Cf. Partie 2, Introduction.

<sup>69</sup>Dans la plupart des cas.

ce propos [Joshi & Schabès, 1992]). La différence entre l’ancre des grammaires lexicalisées et notre tête est que cette dernière a une acception plus large : alors que l’ancre est toujours une forme lexicale précise (le nom commun masculin singulier *cheval*, le verbe au présent de l’indicatif 1<sup>ère</sup> personne du singulier *vais*, etc.), notre tête peut être nettement moins spécifiée que cela : lui affecter une forme lexicale aussi fine que l’ancre des grammaires lexicalisées représente le degré extrême de précision de la tête, qui le plus souvent, en fait, correspondra à une catégorie syntaxique. Toutefois, selon un certain point de vue leur “ancre” et notre “tête” ont en commun le fait qu’elles désignent un objet qui revêt cette même importance toute particulière au regard des constructions dont il est le constituant.

On peut ajouter deux dernières différences entre notre présente acception de “tête” et sa définition habituelle. La première est que notre grammaire n’est littéralement pas “*guidée par les têtes*” : certaines de nos constructions n’ont pas de tête ; c’est le cas des constructions Paradigmatiques (cf. chapitre 5 §2.3). La seconde différence est que nous considérons que les constructions à tête ne peuvent en porter qu’une seule.

## 1.5 Autres traits

Il est bien entendu possible d’ajouter à cet ensemble d’autres attributs, tels que MORPHologie, SEMantique, PROsodie ou PRAGmatique, mais nous ne les introduisons pas ici puisque nous n’en faisons pas usage dans l’état actuel de la grammaire.

## 2 Propriétés et spécification

La formulation des informations extrinsèques prend, comme on l’a vu précédemment, la forme de propriétés. Ce qui varie d’une grammaire à l’autre est d’une part l’ensemble de propriétés choisi, d’autre part leur spécification.

Le jeu de propriétés que nous utilisons dans notre modèle est le suivant : nous utilisons deux propriétés de constituance que l’on nomme **Majeur** (§2.1.1) et **Mineur** (§2.1.2) ; trois propriétés de cooccurrence que l’on nomme **Exigence** (§2.2.1), **Exclusion** (§2.2.2) et **Unicité** (§2.2.3) ; deux propriétés de position que l’on nomme **Précédence** (§2.3.1) et **Adjacence** (§2.3.2) ; et enfin, une propriété de dépendance que l’on nomme **Accord** (§2.4.1).

Le choix et la spécification de ces propriétés a été justifié par les phénomènes que l’on a décrits ; du fait que l’on s’est concentrée sur le domaine syntaxique, il reste tout à fait envisageable d’ajouter de nouvelles propriétés à cet ensemble afin de traiter des types de relations qui émergeraient dans d’autres domaines (ou d’interactions entre domaines) et qui ne pourraient pas être décrites à l’aide des présentes. A titre d’exemple nous présentons au §2.3.3 une propriété de **Simultanéité** que l’on a eue l’occasion de définir pour l’analyse multimodale, et qui n’est (pour l’instant) pas utilisée dans la Partie 2.

Dans chaque section qui suit, nous présentons les propriétés que nous utilisons et leur spécification, et les comparons le cas échéant avec les propriétés qui avaient été définies auparavant

en GP afin d'en faire ressortir les différences et les raisons de nos éventuelles modifications.

## 2.1 Propriétés de constituance : Majeur et Mineur

Les propriétés de constituance rassemblent les types de relations qui s'établissent non pas entre les divers constituants d'une construction, mais entre cette construction et ses constituants. C'est la raison pour laquelle ce sont ici typiquement des relations unaires, c'est-à-dire que l'on n'y mentionne que les constituants intéressés et non la construction (puisque ce sont des propriétés qui feront partie de ses informations extrinsèques donc y seront associées par définition).

Les deux propriétés **Majeur** (§2.1.1) et **Mineur** (§2.1.2) remplacent respectivement celles d'**Obligation** et **Facultativité** employées dans les anciennes grammaires GP, que l'on évoque au §2.1.3.

### 2.1.1 Majeur

**Définition.** — On appelle Majeure la relation établie entre une construction et l'un de ses constituants qui joue un rôle spécifique et notable dans le choix de la NATURE (cf. §1.3.1) de la construction.

Typiquement et pour les relations, syntaxiques, qui nous intéressent, cette propriété permet de pointer deux constituants entretenant une relation majeure avec la construction : sa **tête** d'une part (au sens défini au §1.4.1), qui influe directement sur le TYPE de la construction (une construction à tête nominale aura tendance à être de type nominal, une construction à tête verbale aura tendance à être de type verbal, etc.), et son **spécifieur** d'autre part, qui influe directement sur la CATÉGORIE de la construction (une construction spécifiée par un déterminant sera un syntagme nominal, une construction spécifiée par une préposition sera un syntagme prépositionnel, etc.).

Nous posons dans notre grammaire que pour une construction donnée il ne peut y avoir qu'une seule tête et un seul spécifieur, sans que la présence de ceux-ci ne soit obligatoire pour autant (c'est-à-dire qu'il peut exister des constructions sans spécifieur et/ou sans tête, mais pas de construction à plusieurs têtes et/ou à plusieurs spécifieurs).

**Représentation.** — La spécification du type Majeur est représentée en fig. 23.

On y voit que les propriétés Majeur sont constituées d'un unique terme (elles sont donc nécessairement non commutatives et non réflexives), qui fait référence à une Construction. Pour qu'une telle propriété soit disponible, il faut que le terme mentionné soit lui-même disponible (c'est-à-dire que l'objet auquel il fait référence soit présent dans l'entrée à analyser), auquel cas elle est automatiquement satisfaite. Enfin, sa capacité est de 1, ce qui signifie qu'elle ne peut être évaluée qu'une seule fois pour une construction : ceci est l'application de l'idée selon laquelle il ne peut y avoir qu'une seule tête et un seul spécifieur par construction.

Majeur	
Symbole	M
Termes	$\boxed{1}$ Construction
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1}$ = disponible
Satisfaction	Vrai
Capacité	1

FIG. 23 – spécification du type Majeur.

### 2.1.2 Mineur

**Définition.** — Le type Mineur est la contrepartie du type Majeur que l'on vient de définir : il s'agit d'une propriété représentant une relation entre la construction dont elle fait partie des informations extrinsèques et ses constituants dont la présence n'a pas d'incidence sur la NATURE de la construction.

Typiquement il s'agit des constituants étant des compléments (au sens large) soit des constituants majeurs, soit de la construction elle-même. Nous posons dans notre grammaire qu'il n'y a pas de limitation *a priori* du nombre de constituants Mineurs dans une construction (p.ex. il peut y avoir plusieurs Adjectifs épithètes dans un Nom construit (10)).

(10) vieille bonne grande écharpe toute usée rayée bleue et verte

**Représentation.** — La spécification du type Mineur est représentée en fig. 24.

Mineur	
Symbole	m
Termes	$\boxed{1}$ Construction
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1}$ = disponible
Satisfaction	Vrai
Capacité	$\infty$

FIG. 24 – spécification du type Mineur.

On constate qu'elle a la même spécification que le type Majeur (ce qui signifie qu'elle a le même fonctionnement) à deux différences près : son symbole d'une part (*m* pour Mineur au lieu de *M* pour Majeur), et sa capacité d'autre part : puisque l'on a posé qu'il n'y a pas de limitation dans les Mineurs d'une construction, cela signifie qu'une propriété Mineur peut être évaluée autant de fois que c'est pertinent pour une même construction (p.ex. pour (10) la propriété  $m(Adj)$  sera évaluée une fois pour chaque adjectif présent).

### 2.1.3 Remarques sur les propriétés d'Obligation et de Facultativité

Après une première version de grammaire GP où la constituance était constituée d'un unique type de propriété du même nom (p.ex. [Bès & Blache, 1999]), celle-ci a rapidement été scindée en deux types, Obligation et Facultativité, dont les définitions respectives étaient “*liste des têtes possibles*” et “*liste des constituants non-tête*” ([Blache, 2001]). Or ces deux propriétés nous ont posé deux problèmes dans le développement de notre grammaire : d'une part, comme nous l'avons dit pour le type Majeur, nous ne considérons pas que ce que les auteurs appelaient Obligation ne renfermait que les têtes possibles, d'autre par les notions même d'obligation et de facultativité, relativement aux principes de GP, nous ont semblé inappropriées.

**Obligation vs. Majeur.** — Comme on vient de le dire, la propriété d'Obligation telle que définie dans les multiples publications de Philippe Blache est définie comme étant la “*liste des têtes possibles*”. Notons avant toute chose que l'utilisation de “tête” par l'auteur n'ayant pas été précisément définie, on en déduit qu'il s'agit de la définition classique de la tête (et non pas, notamment, de celle dont on a parlé plus haut). Pourtant, l'auteur indique bien que selon l'utilisation de cette propriété, il ne peut y avoir qu'une seule tête par construction, ce qui pose d'emblée un problème par rapport aux autres théories utilisant cette notion puisque nombre d'entre elles postulent que certaines constructions syntaxiques peuvent porter plusieurs têtes (typiquement la coordination dans un certain nombre d'analyses) : on ne sait pas clairement comment l'auteur se justifie de cette différence de taille et comment il la met en pratique.

Ce que l'on ignore également, c'est pourquoi est-ce que l'auteur choisit de différencier ce type précis de constituant alors qu'il ne distingue pas les autres : on a dit dans notre définition du type Majeur que l'on justifie la distinction entre celui-ci et le type Mineur du fait que les premiers ont une incidence notable sur la nature de la construction dont ils sont constituants, au contraire des seconds. On ne sait pas, pour Philippe Blache, pourquoi est-ce que les spécifieurs sont rassemblés au sein du même type de propriété que tous les autres constituants qui n'ont véritablement pas le même statut dans la construction.

**Les notions d'obligation et de facultativité.** — Du fait que l'on souhaite rassembler de préférence pour notre part les têtes et spécifieurs possibles des constructions d'un côté, et le reste des constituants possibles de l'autre, rien de nous empêcherait (ni formellement, ni théoriquement) d'utiliser simplement l'Obligation et la Facultativité en ce sens précis pour notre grammaire, si ce n'est les notions mêmes véhiculées par les noms choisis à ces deux types de propriétés.

En effet, l'une des caractéristiques de GP est que les types de propriétés ne sont pas hiérarchisés *a priori* dans le formalisme. Il est envisageable de leur associer une importance relative dans le cadre d'une théorie particulière, mais ceci dépend de la grammaire et non du formalisme, dans lequel aucune propriété n'est censée prévaloir sur les autres, ni être indispensable à l'introduction d'une construction en dehors des choix à ce propos qui relèvent de la théorie. Or, en utilisant les termes d’“obligation” et de “facultativité”, on en vient à contredire cela en affirmant que les objets mentionnés dans la première propriété sont indispensables, alors que ceux mentionnés dans la seconde (ainsi que dans les autres types, nécessairement) ne le sont



pas<sup>70</sup>.

De plus, même si l'on souhaite ajouter à la spécification une notion d'importance des types de propriétés relativement les uns aux autres (en leur affectant un poids par exemple), cette information devra alors être exprimée pour tous les types, et non simplement pour les propriétés de constituance, et ne s'exprimera sans doute pas sous une forme binaire (obligation *vs.* facultativité), mais à l'aide d'une échelle où l'on pourra p.ex. indiquer que la Linéarité est plus importante que l'Accord, mais qu'elle est moins importante que l'Obligation (ce qui fait l'objet de [Blache *et al.*, 2006]).

Pour ces raisons, nous avons finalement décidé de ne pas faire usage des types Obligation et Facultativité définis dans les premières versions de GP et d'introduire ceux de Majeur et Mineur, dont la définition linguistique (et formelle) est certes proche, mais nous l'espérons plus en phase avec d'une part notre théorie, d'autre part les caractéristiques du formalisme.

## 2.2 Propriétés de cooccurrence : Exigence, Exclusion et Unicité

Les types de propriétés suivants rassemblent les réclamations et restrictions de cooccurrence entre constituants d'une construction. Il s'agit des trois propriétés définies dans [Blache, 2001] p.ex. : **Exigence** (§2.2.1), **Exclusion** (§2.2.2) et **Unicité** (§2.2.3).

### 2.2.1 Exigence

**Définition.** — Dans une construction A, on dit qu'une construction B exige une construction C lorsque, si B fait partie des constituants de A alors B requiert la présence de C également parmi les constituants de A. L'inverse n'est pas vraie : si B exige C (dans la construction A), alors C peut être présent dans A sans que B ne compte parmi ses constituants. Par exemple, dans un syntagme nominal déterminé, le nom exige un déterminant <sup>71</sup> (11).

(11) le petit chat de ma grand-mère

**Représentation.** — La spécification du type Exigence est représentée en fig. 25.

On y voit que les propriétés d'Exigence seront représentées sous la forme  $\boxed{1} \Rightarrow \boxed{2}$ , les deux index faisant référence à des constructions différentes (non réflexive), et étant ordonnés (pas de commutativité). La propriété est évaluée si l'objet auquel le premier terme fait référence est présent, et satisfaite si l'objet auquel réfère le second est présent. Enfin, chacune des propriétés d'Exigence est évaluable autant de fois que cela est pertinent pour la caractérisation d'une construction (capacité infinie).

<sup>70</sup>Et il s'agit bien ici de notions qui étaient adossées au formalisme lui-même puisque la possibilité de hiérarchisation des (types de) propriétés n'était pas encore envisagée au moment de la définition de ces deux types.

<sup>71</sup>Pour les besoins de l'exemple on a simplifié l'expression de cette propriété, qui est en fait plus générale dans la description précise donnée au chapitre 6 §2.2.1.

Exigence	
Symbole	$\Rightarrow$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1}$ = disponible
Satisfaction	$\boxed{2}$ = disponible
Capacité	$\infty$

FIG. 25 – spécification du type Exigence.

## 2.2.2 Exclusion

**Définition.** — Dans une construction A, on dit qu’une construction B exclut une construction C lorsque, si B fait partie des constituants de A alors C ne doit pas en faire partie. Cette relation est réciproque : si C est présent alors B ne peut l’être (puisque B exclut C). Par exemple, dans une construction verbale contenant des pronoms clitiques, un clitique de forme *me* interdit celle d’un autre de la forme *lui* (12) (et par conséquent, si *lui* est présent alors *me* ne peut l’être) (12).

(12) \* il me lui donne

**Représentation.** — La spécification du type Exclusion est représentée en fig. 26.

Exclusion	
Symbole	$\nrightarrow$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Interdite
Commutativité	Vrai
Disponibilité	$\boxed{1}$ = disponible ; $\boxed{2}$ = disponible
Satisfaction	$\boxed{1} \neq \text{disponible} \vee \boxed{2} \neq \text{disponible}$
Capacité	$\infty$

FIG. 26 – spécification du type Exclusion.

On y voit que les propriétés d’Exclusion seront représentées sous la forme  $\boxed{1} \nrightarrow \boxed{2}$ , les deux index faisant référence à des constructions différentes (non réflexive), et n’étant pas ordonnés (commutativité autorisée). La propriété est évaluée si au moins l’un des deux objets auxquels les termes font référence est présent (p.ex. dans (12) les deux termes sont disponibles), et elle est satisfaite si l’un des deux ne l’est pas (dans (12) les deux termes étant disponibles, la propriété est violée). Enfin, chacune des propriétés d’Exclusion est évaluable autant de fois que cela est pertinent pour la caractérisation d’une construction (capacité infinie).

### 2.2.3 Unicité

**Définition.** — Dans une construction A, on dit qu’une construction B est unique lorsque parmi les constituants de A il ne peut figurer qu’un seul objet dont la catégorie est celle de B. Par exemple, dans un syntagme verbal, le SN (objet direct) est unique (13).

(13) \* il offre son cadeau une surprise le paquet à Marie

**Représentation.** — La spécification du type Unicité est représentée en fig. 27.

Unicité	
Symbole	$u$
Termes	$\boxed{1}$ Construction
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1}$ = disponible
Satisfaction	$\neg((\boxed{2}$ Construction = disponible) $\wedge$ ( $\boxed{1}$ .CAT = $\boxed{2}$ .CAT))
Capacité	1

FIG. 27 – spécification du type Unicité.

On y voit que les propriétés d’Unicité seront représentées sous la forme  $u(\boxed{1})$ , le terme indexé faisant référence à une construction. Une telle propriété sera disponible si le terme indexé  $\boxed{1}$  fait référence à un objet présent dans l’entrée à analyser, et elle sera satisfaite s’il n’existe pas dans cette même entrée une construction (indexée  $\boxed{2}$  et non coréférente à  $\boxed{1}$  puisque la propriété l’interdit) qui soit disponible, et dont la CATégorie soit identique à celle de l’objet auquel  $\boxed{1}$  fait référence. Cette propriété a une capacité de 1 puisque pour chaque construction, l’unicité d’un objet dépend du fait qu’elle ne soit évaluée qu’une fois (sinon, dans (13) l’unicité serait satisfaite une fois pour chaque occurrence du SN (*son cadeau* puis *une surprise* puis *le paquet*), ce qui n’aurait pas de sens).

## 2.3 Propriétés de position : Précédence, Adjacence et Simultanéité

Les propriétés de position indiquent les relations d’ordre linéaire relatif entre les constituants d’une construction. Dans la présente version de la grammaire on en utilise deux types : **Précédence** (§2.3.1) d’une part et **Adjacence** (§2.3.2) d’autre part. On ajoute à celles-ci la spécification d’un type supplémentaire, celui de **Simultanéité** (§2.3.3) qui nous a servi pour la description de constructions multimodales dans [Guénot & Bellengier, 2004].

### 2.3.1 Précédence

**Définition.** — On appelle Précédence la relation établie entre deux constituants d’une construction dont le premier doit être situé dans l’énoncé avant le second. Par exemple, dans un syntagme nominal déterminé, le déterminant précède le nom (14).

- (14) a. le petit chat de ma grand-mère  
 b. \* petit chat de ma grand-mère le  
 c. \* petit chat le de ma grand-mère

**Représentation.** — La spécification du type Précédence est représentée en fig. 28.

Précédence	
Symbole	$\prec$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Satisfaction	$\boxed{1}.\text{INDEX} \text{DEBUT} < \boxed{2}.\text{INDEX} \text{DEBUT}$
Capacité	$\infty$

FIG. 28 – spécification du type Précédence.

On y voit que les propriétés de Précédence seront représentées sous la forme  $\boxed{1} \prec \boxed{2}$ , les deux termes faisant référence à des constructions différentes (non réflexive), et étant ordonnés (commutativité interdite). La propriété est évaluée si les deux objets auxquels les termes font référence sont présents, et elle est satisfaite si l'index de début de l'objet auquel le premier terme est inférieur à celui de l'objet auquel le second terme fait référence. Enfin, chacune des propriétés de Précédence est évaluable autant de fois que cela est pertinent pour la caractérisation d'une construction (capacité infinie).

**Précédence vs. Linéarité.** — Dans les versions précédentes de grammaires GP le type équivalent à notre Précédence est celui nommé Linéarité. La différence entre les deux types (différence qui nous a conduite à affecter à la nôtre un nom différent de façon à éviter les confusions possibles), vient du fait que pour un énoncé tel que (15),

- (15) Pierre relit soigneusement sa copie

l'adverbe *soigneusement*, placé entre les deux constituants du SV *relit sa copie*, ne satisfaisait pas la propriété de Linéarité telle qu'initialement définie puisque celle-ci stipulait que pour être satisfaite il était nécessaire que le second terme se situe après le premier, autrement dit que l'index de fin du SV soit inférieur ou égal à l'index de début de l'Adverbe. Or ceci n'est pas le cas ici : le SV se termine après l'Adv, bien que ce dernier commence après le SV. Nous avons donc reformulé la condition de satisfaction afin qu'elle exprime que l'index de *début* du premier terme doit être *strictement* inférieur à l'index de début du second terme ; cette condition-là est remplie par l'adverbe dans (15).

### 2.3.2 Adjacence

Il s'agit ici d'un type de relation nouveau, qui n'est pas une reformulation d'autres propriétés mais un ajout à celles existantes.

**Définition.** — On appelle Adjacence la relation établie entre deux constituants d’une construction qui doivent être contigus quel que soit leur ordre relatif. Par exemple, dans un nom construit, le nom et l’adjectif doivent être adjacents (16) <sup>72</sup>.

- (16) a. gros fumeur  
b. fumeur gros

**Représentation.** — La spécification du type Adjacence est représentée en fig. 29.

Adjacence	
Symbole	$\leftrightarrow$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Interdite
Commutativité	Vrai
Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Satisfaction	$(\boxed{1}.\text{INDEX} \text{FIN} = \boxed{2}.\text{INDEX} \text{DEBUT}) \vee (\boxed{2}.\text{INDEX} \text{FIN} = \boxed{1}.\text{INDEX} \text{DEBUT})$
Capacité	$\infty$

FIG. 29 – spécification du type Adjacence.

Les propriétés d’Adjacence seront représentées sous la forme  $\boxed{1} \leftrightarrow \boxed{2}$ , les deux termes faisant référence à des constructions différentes (non réflexive), et n’étant pas ordonnés (commutativité autorisée). La propriété est évaluée si les deux objets auxquels les termes font référence sont présents, et elle est satisfaite si l’index de fin de l’un des deux objets auxquels les termes font référence est égal à l’index de début de l’autre. Chacune des propriétés d’Adjacence est évaluable autant de fois que cela est pertinent pour la caractérisation d’une construction (capacité infinie).

### 2.3.3 Simultanéité

Il s’agit ici également d’un type de relation nouveau, qui a été introduit dans [Guénot & Bellen-gier, 2004] afin de traiter des entrées multimodales et de pouvoir introduire des constructions dont les constituants (un geste et un son, par exemple) peuvent être produits dans le même temps.

On peut très bien imaginer que cette propriété soit utilisée dans des développements ultérieurs de la grammaire, même en dehors du cadre de l’analyse multimodale. Par exemple, on peut imaginer une construction syntaxique ou syntaxico-sémantique donnée, qui répond à un ensemble de caractéristiques propres, et une autre construction, prosodique cette fois, qui répond à un autre ensemble de caractéristiques. Chacune de ces constructions a son autonomie et peut être introduite indépendamment de l’autre. Mais dans un cas spécifique donné, on peut décrire une troisième construction dont les constituants seraient les deux premières (ou alors qui hériterait des deux premières) et dont l’une des caractéristiques propres serait leur simultanéité

<sup>72</sup>On donne cet exemple précisément pour souligner le fait que l’on ne s’intéressera pas, dans les descriptions qui suivent en Partie 2, à l’évidente nuance de sens existant entre les deux exemples : il s’agit là d’une nuance sémantique, mais la relation syntaxique existant entre le nom et l’adjectif est dans les deux cas la même, et c’est ce dont on tient compte ici pour l’instant.

(une autre caractéristique pourrait être par exemple une interprétation sémantique ou un effet pragmatique précis).

**Définition.** — On appelle Simultanéité la relation établie entre plusieurs constituants d’une construction, dont au moins une partie des réalisations se déroulent en même temps. Par exemple, si au cours d’une production verbale est commencé un geste qui, lui, se termine après la fin de la production verbale, alors ceux-ci seront considérés comme simultanés (compte tenu du fait que la “simultanéité” des productions n’est qu’accidentellement parfaitement exacte dans les faits).

**Représentation.** — La spécification du type Simultanéité est représentée en fig. 30.

Simultanéité	
Symbole	$\updownarrow$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction} \wedge \dots \wedge \boxed{n}\text{Construction}$
Réflexivité	Interdite
Commutativité	Vrai
Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible} \wedge \dots \wedge \boxed{n} = \text{disponible}$
Satisfaction	$\max(\boxed{1}.\text{INDEX} \text{DEBUT}, \boxed{2}.\text{INDEX} \text{DEBUT}, \dots, \boxed{n}.\text{INDEX} \text{DEBUT}) < \min(\boxed{1}.\text{INDEX} \text{FIN}, \boxed{2}.\text{INDEX} \text{FIN}, \dots, \boxed{n}.\text{INDEX} \text{FIN})$
Capacité	$\infty$

FIG. 30 — spécification du type Simultanéité.

Les propriétés de Simultanéité seront représentées sous la forme  $\updownarrow(\boxed{1}, \boxed{2}, \dots, \boxed{n})$ , les termes faisant référence à des constructions différentes (non réflexive), et n’étant pas ordonnés (commutativité autorisée). La propriété est évaluée si tous les objets auxquels les termes font référence sont présents, et elle est satisfaite si le plus petit des index de fin des objets est supérieur au plus grand des index de début. Sa capacité étant infinie, elle peut être évaluée autant de fois que cela sera pertinent pour une même construction.

## 2.4 Propriété de dépendance : Accord

Ces propriétés rassemblent les relations entre constituants d’une construction qui sont la manifestation d’une dépendance (pouvant être d’un ordre quelconque). Au niveau de développement actuel de notre grammaire nous n’en utilisons qu’une seule, que l’on appelle **Accord** (§2.4.1), et qui nous sert à représenter les relations de dépendance morphologique comme par exemple l’accord en genre entre le déterminant et le nom.

A dire vrai, elle nous sert même à marquer, sous une forme artificielle et inexacte qu’il conviendra de corriger dans les futurs développements (mais toutefois suffisante en l’état actuel parce que les deux phénomènes se manifestent globalement de la même façon), les relations entre les différentes parties de morphèmes à signifiant discontinu comme la personne (pronom clitique nominatif + désinence verbale) ou le pluriel (quand il est marqué sur plusieurs constituants de la construction dont il est lui-même constituant). A terme, cette utilisation abusive de l’Accord devrait être remplacée par la définition de constructions permettant d’introduire ces

morphèmes à signifiant discontinu de la même façon que l'on pourra introduire les morphèmes à signifiant continu, et ils seront rétablis en tant qu'unités d'analyse plutôt qu'en tant que marques de relation comme c'est uniquement le cas actuellement.

### 2.4.1 Accord

**Définition.** — On appelle Accord la relation établie entre deux constituants d'une construction, pour lesquels deux traits précisés portent des valeurs équivalentes.

**Représentation.** — La spécification du type Accord est représentée en fig. 31.

Accord	
Symbole	$\rightsquigarrow$
Termes	$\boxed{1}$ Intrinsèque $\wedge$ $\boxed{2}$ Intrinsèque
Réflexivité	Interdite
Commutativité	Vrai
Disponibilité	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Satisfaction	$\boxed{1} = \boxed{2}$
Capacité	$\infty$

FIG. 31 — spécification du type Accord.

Les propriétés d'Accord seront représentées sous la forme  $\boxed{1} \rightsquigarrow \boxed{2}$ , les termes faisant référence aux informations intrinsèques de constructions différentes (non réflexive), et n'étant pas ordonnés (commutativité autorisée). La propriété est évaluée si tous les objets auxquels les termes font référence sont présents. Notons que dans cette propriété, les termes font référence à un trait d'une construction (puisque les termes font référence à des informations intrinsèques) et non à une construction complète : pour qu'elle soit satisfaite donc, il faut que les valeurs des traits mentionnés soient égales, i.e. que les termes soient égaux. Sa capacité étant infinie, elle peut être évaluée autant de fois que cela sera pertinent pour une même construction.

Rappelons que la propriété n'est pas ordonnée dans la description de son fonctionnement, puisque le résultat de l'évaluation de la compatibilité des genres du nom et de l'adjectif (par exemple) sera le même quel que soit l'ordre des deux objets. Cependant l'ordre est pertinent du point de vue linguistique puisque c'est l'adjectif qui s'accorde avec le nom et non le contraire, et donc on posera conventionnellement que dans ce cas on notera  $N.GENRE \rightsquigarrow Adj.GENRE$  plutôt que  $Adj.GENRE \rightsquigarrow N.GENRE$  (la signification en termes de dépendance morphologique est donc que le premier objet est le recteur et le second le dépendant) <sup>73</sup>.

**Accord vs. Dépendance.** — Cette propriété remplace celle de Dépendance qui avait été utilisée dans les version précédentes de GP. La différence entre les deux types de propriétés (qui nous ont menée, comme pour Linéarité vs. Précédence, à proposer un nom différent afin

<sup>73</sup>Par contre pour ce qui concerne l'utilisation de cette propriété pour mettre en relation les parties de signifiants discontinus d'un même morphème, l'ordre n'a aucune importance puisque l'on ne saurait dire qu'une telle partie de la forme en régit une autre.

d'éviter la possible confusion) est que notre propriété d'Accord fait référence à un trait (plus ou moins précis) d'une structure de traits alors que la propriété de Dépendance de [Blache, 2001] faisait référence à la totalité de la structure de traits des constructions.

Par exemple, une propriété de Dépendance qui avait la forme  $N \rightsquigarrow Det$  signifiait que les traits de genre et de nombre du déterminant devaient correspondre à ceux du nom, mais sans préciser pourquoi il s'agissait de ces deux traits-là précisément d'une part (et pourquoi pas les autres, et dans quelles limites), ni prévoir de différence de résultat pour les cas où par exemple un déterminant était accordé en genre mais pas en nombre, de cas où un déterminant n'était accordé ni en genre ni en nombre : dans les deux cas l'unique propriété était violée.

Avec l'Accord, l'équivalent de cette propriété de Dépendance prendra la forme de deux propriétés,  $N.GENRE \rightsquigarrow Det.GENRE$  d'une part, et  $N.NOMBRE \rightsquigarrow Det.NOMBRE$  d'autre part, chacune étant évaluée séparément et précisant clairement sur quel trait (que sa valeur soit simple ou complexe) la comparaison porte.

En fait, cette différence est en grande partie due au fait que la Dépendance était une propriété qui représentait à la fois les dépendances morphologiques (qui se manifestent par un Accord, mais dont l'expression était alors trop large) et les dépendances syntaxiques (qui ne se manifestent pas par une marque particulière mais sont en réalité le résultat d'un ensemble de relations de plusieurs types, c'est-à-dire qu'il ne s'agit typiquement pas d'une propriété mais d'une information intrinsèque étant le résultat d'une caractérisation). Pour notre part, nous indiquons les accords par la propriété du même nom, et nous indiquons les dépendances syntaxiques dans les informations intrinsèques de la construction concernée.



## Résumé

Dans ce chapitre on a présenté notre **modèle de grammaire**, c'est-à-dire la mise en oeuvre des positions théoriques exprimées au chapitre 1 à travers le formalisme GP dont l'architecture générale était présentée au chapitre 2. Une formalisation en GP étant constituée de deux ressources complémentaires, on a d'abord présenté les caractéristiques de notre **grammaire** (§1), qui consistent en l'explicitation des choix de contenu de la structure de traits contenant les informations intrinsèques de nos constructions, puis on a présenté le **jeu de propriétés** que l'on utilisera ainsi que leur **spécification** (§2). L'ensemble de ces informations constitue le modèle dans lequel on a effectué nos descriptions grammaticales figurant en deuxième partie.

Il est important de souligner que les descriptions qui constituent la partie 2 de cet ouvrage ne représentent qu'une partie, volontairement limitée, de la grammaire dont le présent chapitre veut présenter le modèle : celle-ci est par nature vouée à être complétée, et cette complétion pourra évidemment engendrer des ajouts d'informations à ce qui vient d'être présenté (aussi bien dans la structure des constructions que dans le jeu de propriétés). Toutefois, le contenu de ce chapitre permet dès à présent de représenter un certain nombre de descriptions de constructions syntaxiques, et c'est ce qui fait l'objet de ce travail.

Dans le chapitre suivant nous présentons quelques autres formalisations possibles à l'aide de GP, représentant des théories aux positions nettement différentes des nôtres présentes, ceci afin de permettre une comparaison des modèles basée sur une même représentation formelle, et également d'illustrer l'éventail de possibilités de formalisation en GP.



## Chapitre 4

# D’autres modèles en GP

Ce que l’on a donné au chapitre précédent constitue l’application formelle de la théorie que l’on a présentée au premier chapitre, mais le formalisme de GP permet de représenter des grammaires basées sur des théories tout à fait différentes de ce que l’on vient de voir. Les premières versions de grammaires GP en sont le premier exemple, puisque ni leur jeu de propriétés (qui à l’époque était constitué de Obligation + Facultativité<sup>74</sup> + Unicité + Exigence + Exclusion + Linéarité + Dépendance) ni le contenu des informations intrinsèques (dites alors “locales” ou “statiques”) n’étaient les mêmes. Le second exemple est celui de la grammaire développée pour les besoins de la campagne EASY qui figure en annexe A. D’autres exemples sont présentés dans ce chapitre, volontairement choisis pour leur éloignement d’avec notre modèle, afin de pouvoir comparer l’usage que l’on fait ici de GP et ce qu’il est possible d’en faire par ailleurs.

Il coexiste actuellement une considérable variété de modèles formels et il nous serait impossible de les observer tous ici ; nous nous concentrons dans ce chapitre sur trois modèles, que nous considérons comme représentatifs de trois grands types d’approches formelles : une **Grammaire d’Unification** à travers le cas de la *Grammaire d’Unification Sens-Texte* (GUST) de [Kahane, 2004] (§1), une **Grammaire Catégorielle à représentation procédurale** à travers le cas de *Dynamic Syntax* ([Kempson *et al.*, 2001], §2), et enfin une **Métagrammaire** à travers l’exemple de la *Métagrammaire TAG* de [Crabbé, 2005b] (§3).

Les trois grands types d’approche retenus ne s’opposent pas les uns aux autres. Les Grammaires d’Unification ont pour point commun un **mécanisme d’analyse**, celui de l’unification ; les Grammaires Catégorielles, quant à elles, on en commun un mode de **représentation des objets** ; les Métagrammaires enfin, partagent un même mode de **factorisation des informations**. [Uszkoreit, 1986], [Tugwell, 2004] ont p.ex. proposé des modèles combinant mécanisme d’unification et grammaires catégorielles. Quant aux Métagrammaires, elles sont la plupart du temps basées sur des grammaires d’unification, notamment TAG (p.ex. [Crabbé, 2005b]), LFG (p.ex. [Clément & Kinyon, 2003a]), ou les deux à la fois (p.ex. [Clément & Kinyon, 2003b]) ; on trouve même un avatar de Métagrammaire Catégorielle chez [Morrill, 1987]. Ce qui nous a

---

<sup>74</sup>Voir même uniquement Constituanse dans les toutes premières versions, à la place du couple Obligation + Facultativité.

poussée à choisir ces trois grands types d'approches est qu'ils sont tous trois fortement différents de la nôtre, chacun à leur façon. Nous avons donc cherché à faire émerger, à travers ces différentes comparaisons, différentes facettes de notre modèle.

## 1 Une grammaire d'unification modulaire en GP

Les Grammaires d'Unification sont un ensemble constitué de grammaires qui utilisent le **mécanisme d'unification** ([Colmerauer, 1975], [Kay, 1979]) en tant que procédure d'analyse des énoncés. D'après [Abeillé, 1993, p. 33],

*(...) l'unification vérifie la compatibilité entre deux structures de traits et produit une structure résultante qui est la plus petite structure qui contient toute l'information contenue dans la première structure et toute l'information contenue dans la deuxième structure.*

Pour guider notre comparaison de GP et des grammaires d'unification nous nous sommes basée sur l'exemple de la **Grammaire d'Unification Sens-Texte** (GUST, [Kahane, 2002]) formalisée sous forme de **Grammaire d'Unification Polarisée** (GUP, [Kahane, 2004]). La spécificité des GUP au sein des grammaires d'unification est la suivante :

*Une grammaire d'unification polarisée (GUP) est définie par une famille finie  $F$  de types d'objets (avec des fonctions attachées aux différents types d'objets), un système  $(P, \cdot)$  de polarités, un sous-ensemble  $N$  de  $P$  de polarités neutres, et un ensemble fini de structures élémentaires polarisées, dont les objets sont décrits par  $F$  et dont une est marquée comme la structure initiale. Les structures générées par la grammaire sont les structures neutres obtenues par combinaison de la structure initiale et d'un nombre fini de structures élémentaires. Les polarités sont uniquement utilisées pour contrôler la saturation au cours du calcul et ne sont pas considérées lorsqu'on évalue la capacité générative de la grammaire. Bien que données sous une forme déclarative, les polarités jouent en fait un rôle essentiellement procédural. [Kahane, 2004]*

D'après l'auteur, GUP constitue un mode de représentation générique des grammaires d'unification, et en particulier des grammaires d'arbres de [Nasr, 1995], des grammaires de réécriture (d'après ce qui a été montré par [Burroni, 1993]), des TAG telles que présentées par [Vijay-Shanker, 1992], de HPSG (en considérant les structures de traits comme des graphes acycliques orientés comme l'ont fait p.ex. [Kesper & Mönnich, 2003]), et de LFG. Il permet bien sûr également de représenter les *Grammaires d'Interaction* de [Perrier, 2002], qui sont elles aussi basées sur une représentation polarisée. Par conséquent la comparaison de GP et GUP (à travers la comparaison des deux formalisations de GUST) devrait nous permettre de tirer des conclusions sur les relations qui existent entre GP et les grammaires d'unification d'une manière générale.

Pour comparer GP et GUST nous avons choisi de nous baser sur la présentation qui en est faite dans [Kahane & Lareau, 2005]. Rappelons que GUST est un modèle basé sur la *Théorie Sens-Texte* ([Mel'čuk, 1997]), et à ce titre "une" grammaire d'unification sens-texte est constituée en fait d'autant de grammaires qu'il y est défini de modules d'analyse, en l'occurrence sémantique, syntaxique, morphotopologique et phonologique (que les auteurs appellent "*grammaires de bonne formation*"), auxquelles s'ajoutent une grammaire par interface entre chaque niveau (qu'ils appellent "*grammaires de correspondance*"). Dans l'article qui nous intéresse, les auteurs

présentent les “grammaires de bonne formation” correspondant aux niveaux syntaxique et sémantique, et les “grammaires de correspondance” permettant l’interface sémantique-syntaxe et l’interface syntaxe-morphotopologie. Pour illustrer leur propos ils se basent sur les extraits de ces grammaires qui permettent d’analyser l’énoncé (17).

(17) Pierre mange deux pommes

Pour notre comparaison, nous tâchons ici de transcrire les grammaires GUST exprimées en GUP en des grammaires GUST exprimées en GP pour observer s’il émerge de cela des rapports remarquables entre les deux formalismes. On suit donc le développement de l’article, en montrant point par point comment il est possible de représenter cela en GP. Notons toutefois avant de commencer la comparaison, que les extraits de grammaire présentés dans l’article ayant été très simplifiés pour les besoins de l’exercice, il en sera de même pour leur transcription en GP.

## 1.1 Grammaire sémantique

### 1.1.1 Version originale ( $\mathcal{GUP}_{sem}$ )

L’extrait de grammaire sémantique proposée dans [Kahane & Lareau, 2005] est constitué de quatre entrées (partie encadrée de la fig. 32), chacune correspondant à une unité sémantique lexicale de (17). Ces quatre entrées, via le jeu des polarités, peuvent être combinées de façon à produire le graphe donné à droite de la figure 32.

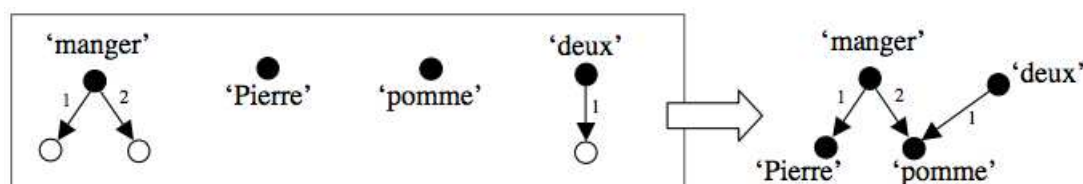


FIG. 32 – Un extrait de  $\mathcal{GUP}_{sem}$  générant le graphe sémantique de (17).

### 1.1.2 Transcription ( $\mathcal{GP}_{sem}$ )

**Graphes.** — Chaque **graphe** de GUP représente une entrée de la grammaire, et correspond donc en GP à une **construction**.

Il y a une différence à noter entre les graphes de  $\mathcal{GUP}_{sem}$  et les constructions de  $\mathcal{GP}_{sem}$  : les premiers représentent des entrées lexicales (bien qu’elles constituent une grammaire), et les secondes des entrées grammaticales. Cela vient du fait que GUST est une grammaire lexicalisée. Les entrées lexicales qu’elle contient font donc mention d’objets autres que l’entrée elle-même, ce qui correspond en GP à des informations extrinsèques, et puisque l’on a dit précédemment que les objets lexicaux de GP ne contenaient pas d’informations extrinsèques,

cela signifie qu'ici l'on aura affaire à des constructions grammaticales, et non lexicales. On voit à travers cet exemple que la frontière entre grammaires lexicalisées et non-lexicalisées est essentiellement une distinction notationnelle et conventionnelle, et que l'on peut tout à fait représenter une grammaire lexicalisée en GP (contrairement à ce qui est le cas de notre modèle).

**nœuds et arcs.** — Les graphes de  $\mathcal{GUP}_{sem}$  sont constitués de nœuds et d'arcs, les **nœuds** représentant des objets portant ou réclamant un sémantème (signifié de l'unité minimale représentée par le nœud) et éventuellement des relations prédicat-argument établies avec d'autres nœuds, et figurées par les **arcs**.

Dans  $\mathcal{GP}_{sem}$ , chaque nœud de GUP correspond à un trait SEM (pour "nœud sémantique"). Chaque SEM peut recevoir comme valeur deux traits, son SÉMANTÈME d'une part, et ses ARGUMENTS d'autre part. Les arguments pouvant être plusieurs, chacun sera numéroté comme le sont les arcs de GUP. Par exemple, les informations intrinsèques de 'manger' seront représentées de la façon suivante :

$$\boxed{1}SEM \left[ \begin{array}{l} \text{SÉMANTÈME} \\ \text{ARGUMENTS} \end{array} \left[ \begin{array}{l} 'manger' \\ \begin{array}{l} 1 \mid \boxed{2}SEM \mid \text{SÉMANTÈME} \mid \boxed{3} \\ 2 \mid \boxed{4}SEM \mid \text{SÉMANTÈME} \mid \boxed{5} \end{array} \end{array} \right] \right]$$

où  $\boxed{1}$ ,  $\boxed{2}$  et  $\boxed{4}$  représentent les nœuds du graphe, et  $\boxed{3}$  et  $\boxed{5}$  représentent les sémantèmes des deux objets réclamés par  $\boxed{1}$ .

**Polarisation.** — Les nœuds et les arcs de GUP sont polarisés : dans chaque entrée de l'extrait, le nœud correspondant au prédicat est saturé  $\bullet$ , ainsi que les arcs des relations prédicats-arguments. Les nœuds représentant les sémantèmes des arguments, eux, ne sont pas saturés  $\circ$ .

Dans  $\mathcal{GP}_{sem}$ , on considère que les informations qui sont saturées sont considérées comme déjà connues, et sont donc placées dans les informations intrinsèques de la construction. De plus, puisqu'elles sont non seulement indiquées dans le graphe, mais aussi saturées, cela signifie que la construction refuse de recevoir des objets les contenant : on fait figurer cette information sous forme de propriétés d'Exclusion. Les informations non saturées sont considérées comme attendues, et on les fait figurer sous la forme d'une propriété d'Exigence.

Ces deux propriétés que l'on vient d'évoquer, Exigence et Exclusion, n'ont pas la même spécification que celles portant le même nom dans notre modèle, et qu'on a définies dans le chapitre précédent. Elles présentent une différence de forme : elles sont réflexives (les deux termes peuvent faire référence à un même objet). Les deux propriétés sont toujours disponibles (c'est ce qui fait la spécificité de la polarisation) ; l'Exigence est satisfaite quand son terme est disponible, et l'Exclusion au contraire quand il ne l'est pas. Puisque la première représente les objets non saturés et la seconde les objets saturés on les représente respectivement  $\Rightarrow^\circ$  et  $\nRightarrow^\bullet$ . On en donne les représentations en fig. 33.

Les propriétés extrinsèques de 'manger', par conséquent, seront donc les suivantes :

Exigence		Exclusion	
Symbole	$\Rightarrow^\circ$	Symbole	$\nRightarrow^\bullet$
Termes	$\boxed{1}\text{Constr.} \wedge \boxed{2}\text{Constr.}$	Termes	$\boxed{1}\text{Constr.} \wedge \boxed{2}\text{Constr.}$
Réflexivité	Autorisée	Réflexivité	Autorisée
Commutativité	Faux	Commutativité	Faux
Disponibilité	Vrai	Disponibilité	Vrai
Satisfaction	$\boxed{2} = \text{disponible}$	Satisfaction	$\boxed{2} \neq \text{disponible}$
Capacité	1	Capacité	$\infty$

FIG. 33 – Spécification des types Exigence et Exclusion pour GUST.

Exigence	$\boxed{2} \Rightarrow^\circ \boxed{3}\text{SÉMANTEME}$
	$\boxed{4} \Rightarrow^\circ \boxed{5}\text{SÉMANTEME}$
Exclusion	$\boxed{1} \nRightarrow^\bullet \boxed{1}.\text{SÉMANTEME}$
	$\boxed{1} \nRightarrow^\bullet \boxed{1}.\text{ARGUMENTS} \mid 1$
	$\boxed{1} \nRightarrow^\bullet \boxed{1}.\text{ARGUMENTS} \mid 2$

c'est-à-dire que les objets indexés  $\boxed{2}$  et  $\boxed{4}$  (les index correspondent à ceux que l'on a donnés plus haut) exigent chacun de recevoir leur sémantème, et l'objet  $\boxed{1}$  interdit de recevoir un sémantème ainsi que ses arguments, puisque leur polarisation (donc les informations intrinsèques en GP) indique qu'ils ont déjà été reçus.

**Grammaire.** — Selon les conventions que l'on vient de poser, l'extrait de  $\mathcal{GUP}_{sem}$  de la figure 32 prend la forme en GP de la figure 34.

Un structure neutralisée en GUP correspond alors à une structure entièrement satisfaite en GP : toutes les propriétés doivent être évaluées puisque toutes sont disponibles par définition. En effet, pour que les structures soient “neutralisées” cela signifie que tous les objets cités dans les propriétés sont présents, et qu'ils satisfont les propriétés citées dans les informations extrinsèques.

**Analyse.** — Le graphe résultat de l'analyse sémantique prendra en GP la forme de la figure 35.

## 1.2 Grammaire syntaxique

### 1.2.1 Version originale ( $\mathcal{GUP}_{synt}$ )

La figure 36 représente l'extrait de  $\mathcal{GUP}_{synt}$  (partie encadrée) nécessaire à l'analyse syntaxique (partie de droite) de (17). Il s'agit d'arbres de dépendances où chaque nœud figure un lexème. A chaque lexème peut être lié un certain nombre d'informations de sous-catégorisation que les auteurs appellent “fonctions grammaticales” (p.ex. *pdd* : *V* pour “partie du discours : Verbe”, *nbre* : *sg* pour “nombre singulier”, etc.).

Les arcs qui relient les lexèmes entre eux figurent les fonctions syntaxiques et portent les noms correspondants (p.ex. “sujet”, “détermination”).

'manger'	
INTR.   [1]SEM	$\left[ \begin{array}{l} \text{SÉMANTEME} \quad \text{'manger'} \\ \text{ARGUMENTS} \quad \left[ \begin{array}{l} 1 \mid [2]SEM \left[ \begin{array}{l} \text{SÉMANTEME} \quad [3] \end{array} \right] \\ 2 \mid [4]SEM \left[ \begin{array}{l} \text{SÉMANTEME} \quad [5] \end{array} \right] \end{array} \right] \end{array} \right]$
Exigence	$\begin{array}{l} [2] \Rightarrow {}^\circ [3] \text{SÉMANTEME} \\ [4] \Rightarrow {}^\circ [5] \text{SÉMANTEME} \end{array}$
Exclusion	$\begin{array}{l} [1] \not\Rightarrow [1].\text{SÉMANTEME} \\ [1] \not\Rightarrow [1].\text{ARGUMENTS} \mid 1 \\ [1] \not\Rightarrow [1].\text{ARGUMENTS} \mid 2 \end{array}$
'Pierre'	
INTR.   [1]SEM	SÉMANTEME 'Pierre'
Exclusion	[1] $\not\Rightarrow$ [1].SÉMANTEME
'pomme'	
INTR.   [1]SEM	SÉMANTEME 'pomme'
Exclusion	[1] $\not\Rightarrow$ [1].SÉMANTEME
'deux'	
INTR.   [1]SEM	$\left[ \begin{array}{l} \text{SÉMANTEME} \quad \text{'deux'} \\ \text{ARGUMENTS} \quad \left[ 1 \mid [2]SEM \left[ \begin{array}{l} \text{SÉMANTEME} \quad [3] \end{array} \right] \right] \end{array} \right]$
Exigence	$[2] \Rightarrow {}^\circ [3] \text{SÉMANTEME}$
Exclusion	$\begin{array}{l} [1] \not\Rightarrow [1].\text{SÉMANTEME} \\ [1] \not\Rightarrow [1].\text{ARGUMENTS} \mid 1 \end{array}$

 FIG. 34 – Extrait de  $\mathcal{GP}_{sem}$  correspondant à l'extrait de  $\mathcal{GUP}_{sem}$ .

Enoncé	
INTR.	$\left[ \begin{array}{l} \text{SEM} \left[ \begin{array}{l} \text{SÉMANTEME} \quad \text{'manger'} \\ \text{ARGUMENTS} \quad \left[ \begin{array}{l} 1 \mid \text{SEM} \mid \text{SÉMANTEME} \quad \text{'Pierre'} \\ 2 \mid [1]SEM \mid \text{SÉMANTEME} \quad \text{'pomme'} \end{array} \right] \end{array} \right] \\ \text{SEM} \left[ \begin{array}{l} \text{SÉMANTEME} \quad \text{'deux'} \\ \text{ARGUMENTS} \quad \left[ 1 \mid [1] \end{array} \right] \end{array} \right] \end{array} \right]$

 FIG. 35 – Représentation GP du graphe sémantique généré par  $\mathcal{GP}_{sem}$ .

### 1.2.2 Transcription ( $\mathcal{GP}_{synt}$ )

Pour transcrire  $\mathcal{GUP}_{synt}$  en  $\mathcal{GP}_{synt}$ , on conserve les conventions que l'on avait utilisées précédemment pour  $\mathcal{GUP}_{sem}$ , y compris la spécification des propriétés. Puisqu'on est ici dans un autre module, cela n'était pas une obligation puisqu'on a vu au chapitre 2 qu'à chaque module correspondait un couple grammaire-spécification ; le fait que la spécification reste la même ici n'est en fait qu'une coïncidence et non un besoin. En revanche, la grammaire est modifiée de la façon suivante :



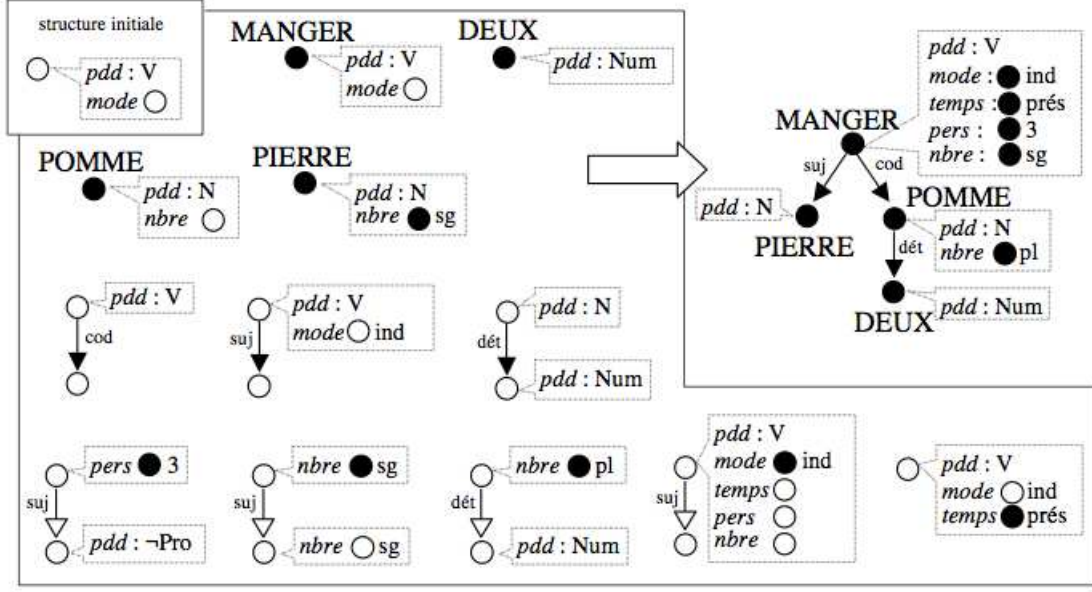


FIG. 36 – Un extrait de  $\mathcal{GUP}_{synt}$  générant le graphe syntaxique de (17).

**nœuds et arcs.** — Les nœuds ici figurent des unités syntaxiques et non plus sémantiques. Ceux-ci portent l'attribut SYNT (au lieu de SEM précédemment), qui peut contenir un LEXÈME (à la place du SÉMANTEME), des fonctions grammaticales rassemblées sous l'attribut F-GRAM et des fonctions syntaxiques rassemblées sous l'attribut F-SYNT. L'entrée MANGER par exemple portera les informations intrinsèques suivantes :

$$\boxed{1}_{\text{SYNT}} \begin{bmatrix} \text{LEXÈME} & \text{MANGER} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & V \\ \text{MODE} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

où  $\boxed{1}$  figure le nœud, et  $\boxed{2}$  la valeur de mode qui est attendue.

**Structure initiale.** — Il s'agit d'une construction à part dans la grammaire syntaxique, en ceci qu'elle constitue le sommet de l'arbre, i.e. le seul nœud qui ne peut être gouverné. C'est la raison pour laquelle on lui donne un attribut spécial, qu'on nomme INIT à la place de SYNT<sup>75</sup>.

On sait qu'elle est constituée d'un SYNT qui est de  $[\text{PDD } V]$ , et qu'elle réclame la présence dans ce même SYNT d'un lexème et d'un mode : on représente cela comme dans la figure 37.

On introduit ici un nouveau type de propriété, que par analogie avec notre modèle on a appelé Majeur. Sa spécification est donnée en figure 38. Il sert à mentionner les nœuds exigés (et non

<sup>75</sup>D'ailleurs cette structure initiale ne figure pas un objet syntaxique, donc le fait de ne pas lui donner l'attribut SYNT est cohérent.

Enoncé	
INIT	$\left[ \begin{array}{c} \boxed{1} \text{SYNT} \left[ \begin{array}{c} \text{LEXÈME } \boxed{2} \\ \text{F-GRAM} \left[ \begin{array}{c} \text{PDD } V \\ \text{MODE } \boxed{3} \end{array} \right] \end{array} \right] \end{array} \right]$
Majeur	$M^\circ(\boxed{1}[\text{F-GRAM} \text{PDD } V])$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2} \text{LEXÈME}$ $\boxed{1} \Rightarrow^\circ \boxed{3} \text{MODE}$

FIG. 37 – La structure initiale en  $\mathcal{GP}_{synt}$ .

pas les parties de nœuds, comme dans l'Exigence), et il puisqu'il figure des "besoins" on accole à son symbole le signe  $\circ$ .

Majeur	
Symbole	$M^\circ$
Termes	$\boxed{1} \text{Construction}$
Réflexivité	Autorisée
Commutativité	Faux
Disponibilité	Vrai
Satisfaction	$\boxed{1} = \text{disponible}$
Capacité	1

FIG. 38 – Spécification du type Majeur pour GUST.

**Règles lexicales.** — Les quatre règles qui suivent dans la grammaire syntaxique présentent des objets portant un lexème et une partie du discours, et pouvant en outre soit porter une fonction grammaticale (le nombre de PIERRE), soit la réclamer (le nombre de POMME, le mode de MANGER). Le nombre de PIERRE fait l'objet d'une propriété d'Exclusion, tandis que le nombre de POMME et le mode de MANGER sont Exigés par le recteur de la construction. Les valeurs de lexème, toutes saturées, font l'objets d'Exclusions. Ces règles lexicales prennent la forme présentée en figure 39.

**Règles sagittales.** — Viennent ensuite trois règles qui décrivent les relations syntaxiques possibles en fonction des parties du discours, tout en leur affectant une fonction syntaxique. Dans ces règles, les arcs sont saturés, mais les nœuds ne le sont pas : les premiers font donc l'objet de propriétés d'Exclusion, les seconds d'Exigence. Leur représentation dans  $\mathcal{GP}_{synt}$  est donnée en fig. 40.

**Règles grammaticales.** — Enfin, les cinq règles qui complètent l'extrait représentent les accords nécessaires entre les divers dépendants, que l'on a représenté en GP dans la fig. 41.

Dans les quatre premières, les arcs ne sont pas saturés : ceci dans  $\mathcal{GP}_{synt}$  ne correspond pas simplement à des propriétés d'Exigence, mais à un ensemble constitué d'une propriété Mineur mentionnant l'objet attendu par la structure, une propriété spécifiant son Unicité dans la construction, et une propriété d'Exigence précisant son attente. La spécification des types

MANGER	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEXÈME} & \text{MANGER} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & V \\ \text{MODE} & \boxed{2} \end{bmatrix} \end{bmatrix}$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2}\text{MODE}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{1}.\text{LEXÈME}$
POMME	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEXÈME} & \text{POMME} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & N \\ \text{NBRE} & \boxed{2} \end{bmatrix} \end{bmatrix}$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2}\text{NBRE}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{1}.\text{LEXÈME}$
DEUX	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEXÈME} & \text{DEUX} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & \text{Num} \end{bmatrix} \end{bmatrix}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{1}.\text{LEXÈME}$
PIERRE	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEXÈME} & \text{PIERRE} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & N \\ \text{NBRE} & \boxed{2}sg \end{bmatrix} \end{bmatrix}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{1}.\text{LEXÈME}$ $\boxed{1} \not\Leftarrow^\bullet \boxed{2}\text{NBRE}$

FIG. 39 – Les règles lexicales dans  $\mathcal{GP}_{\text{synt}}$ .

COD	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEX} & \boxed{2} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & V \end{bmatrix} \\ \text{F-SYNT} & \begin{bmatrix} \text{COD} \mid \boxed{3}\text{SYNT} \mid \text{LEX} & \boxed{4} \end{bmatrix} \end{bmatrix}$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2}\text{LEXÈME}$ $\boxed{3} \Rightarrow^\circ \boxed{4}\text{LEXÈME}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{3}$
Sujet	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEX} & \boxed{2} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & V \\ \text{MODE} & \boxed{3} \end{bmatrix} \\ \text{F-SYNT} & \begin{bmatrix} \text{SUJ} \mid \boxed{4}\text{SYNT} \mid \text{LEX} & \boxed{5} \end{bmatrix} \end{bmatrix}$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2}\text{LEXÈME}$ $\boxed{1} \Rightarrow^\circ \boxed{3}[\text{MODE } ind]$ $\boxed{4} \Rightarrow^\circ \boxed{5}\text{LEXÈME}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{4}$
Déterminant	
$\boxed{1}\text{SYNT}$	$\begin{bmatrix} \text{LEXÈME} & \boxed{2} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & N \end{bmatrix} \\ \text{F-SYNT} & \begin{bmatrix} \text{DÉT} \mid \boxed{3}\text{SYNT} & \begin{bmatrix} \text{LEXÈME} & \boxed{4} \\ \text{F-GRAM} & \begin{bmatrix} \text{PDD} & \text{Num} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$
Exigence	$\boxed{1} \Rightarrow^\circ \boxed{2}\text{LEXÈME}$ $\boxed{3} \Rightarrow^\circ \boxed{4}\text{LEXÈME}$
Exclusion	$\boxed{1} \not\Leftarrow^\bullet \boxed{3}$

FIG. 40 – Les règles sagittales dans  $\mathcal{GP}_{\text{synt}}$ .

Mineur et Unicité dans GUST sont donnés en fig. 42.

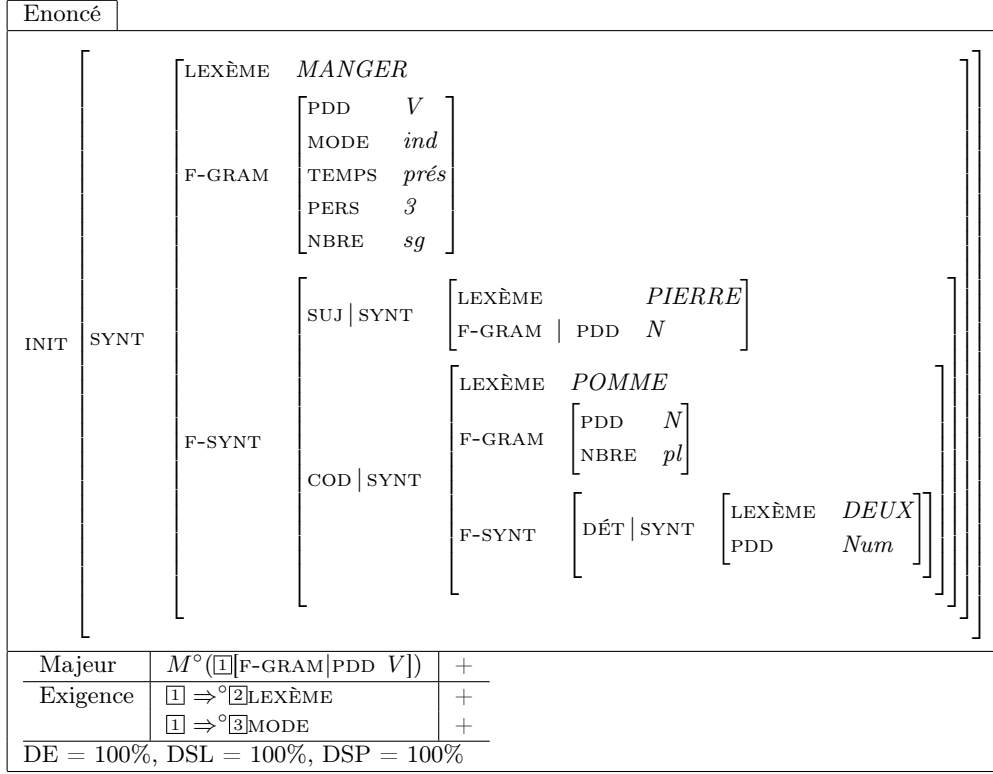
Suj : accord à la 3e personne	
$[1]_{\text{SYNT}}$	$\left[ \begin{array}{l} \text{LEXEME } [2] \\ \text{F-GRAM } \left[ \begin{array}{l} \text{PERS } [3] \end{array} \right] \\ \text{F-SYNT } \left[ \begin{array}{l} \text{SUIJ} \mid [3]_{\text{SYNT}} \left[ \begin{array}{l} \text{LEXEME} \\ \text{F-GRAM} \mid \text{PDD } \neg \text{Pro} \end{array} \right] \end{array} \right] \end{array} \right]$
Mineur	$m^\circ([3] \mid \text{F-GRAM} \mid \text{PDD } \neg \text{Pro}))$
Unicité	$U([3])$
Exclusion	$[1] \not\Rightarrow [1] \bullet [1] \text{F-GRAM} \mid \text{PERS}$
Exigence	$[1] \Rightarrow [3]$ $[1] \Rightarrow [2] \text{LEXÈME}$ $[3] \Rightarrow [4] \text{LEXÈME}$
Suj : accord au singulier	
$[1]_{\text{SYNT}}$	$\left[ \begin{array}{l} \text{LEXEME } [2] \\ \text{F-GRAM } \left[ \begin{array}{l} \text{NBRE } sg \end{array} \right] \\ \text{F-SYNT } \left[ \begin{array}{l} \text{SUIJ} \mid [3]_{\text{SYNT}} \left[ \begin{array}{l} \text{LEXEME} \\ \text{F-GRAM} \mid \text{NBRE } sg \end{array} \right] \end{array} \right] \end{array} \right]$
Mineur	$m^\circ([3] \mid \text{F-GRAM} \mid \text{NBRE } sg))$
Unicité	$U([3])$
Exclusion	$[1] \not\Rightarrow [1] \bullet [1] \text{F-GRAM} \mid \text{NBRE}$
Exigence	$[1] \Rightarrow [3]$ $[1] \Rightarrow [2] \text{LEXÈME}$ $[3] \Rightarrow [4] \text{LEXÈME}$
Dét : accord au pluriel	
$[1]_{\text{SYNT}}$	$\left[ \begin{array}{l} \text{LEXEME } [2] \\ \text{F-GRAM } \left[ \begin{array}{l} \text{NBRE } pl \end{array} \right] \\ \text{F-SYNT } \left[ \begin{array}{l} \text{DÉT} \mid [3]_{\text{SYNT}} \left[ \begin{array}{l} \text{LEXEME} \\ \text{F-GRAM} \mid \text{PDD } Num \end{array} \right] \end{array} \right] \end{array} \right]$
Mineur	$m^\circ([3] \mid \text{F-GRAM} \mid \text{PDD } Num))$
Unicité	$U([3])$
Exclusion	$[1] \not\Rightarrow [1] \bullet [1] \text{F-GRAM} \mid \text{NBRE}$
Exigence	$[1] \Rightarrow [3]$ $[1] \Rightarrow [2] \text{LEXÈME}$ $[3] \Rightarrow [4] \text{LEXÈME}$
Suj : indicatif	
$[1]_{\text{SYNT}}$	$\left[ \begin{array}{l} \text{LEXEME } [2] \\ \text{F-GRAM } \left[ \begin{array}{l} \text{PDD } V \\ \text{MODE } ind \\ \text{TEMPS } [3] \\ \text{PERS } [4] \\ \text{NBRE } [5] \end{array} \right] \\ \text{F-SYNT } \left[ \begin{array}{l} \text{DÉT} \mid [6]_{\text{SYNT}} \left[ \begin{array}{l} \text{LEXEME } [7] \end{array} \right] \end{array} \right] \end{array} \right]$
Mineur	$m^\circ([6]_{\text{SYNT}})$
Unicité	$U([6])$
Exclusion	$[1] \not\Rightarrow [1] \bullet [1] \text{F-GRAM} \mid \text{MODE}$
Exigence	$[1] \Rightarrow [6]$ $[1] \Rightarrow [2] \text{LEXÈME}$ $[1] \Rightarrow [3] \text{TEMPS}$ $[1] \Rightarrow [4] \text{PERS}$ $[1] \Rightarrow [5] \text{NBRE}$ $[6] \Rightarrow [7] \text{LEXÈME}$
Verbe à l'indicatif présent	
$[1]_{\text{SYNT}}$	$\left[ \begin{array}{l} \text{LEXEME } [2] \\ \text{F-GRAM } \left[ \begin{array}{l} \text{PDD } V \\ \text{MODE } [3] \\ \text{TEMPS } prés \end{array} \right] \end{array} \right]$
Exclusion	$[1] \not\Rightarrow [1] \bullet [1] \text{F-GRAM} \mid \text{TEMPS}$
Exigence	$[1] \Rightarrow [3] \text{MODE}$ $[1] \Rightarrow [2] \text{LEXÈME}$

 FIG. 41 – Les règles grammaticales dans  $\mathcal{G}^p_{\text{synt}}$ .

<b>Mineur</b>		<b>Unicité</b>	
Symbole	$m^\circ$	Symbole	$U$
Termes	$[1]_{\text{Construction}}$	Termes	$[1]_{\text{Construction}}$
Réflexivité	Autorisée	Réflexivité	Autorisée
Commutativité	Faux	Commutativité	Faux
Disponibilité	Vrai	Disponibilité	$[1] = \text{disponible}$
Satisfaction	$[1] = \text{disponible}$	Satisfaction	$\neg((\text{[2] Constr.} = \text{disponible}) \wedge (\text{[1].CAT} = \text{[2].CAT}))$
Capacité	1	Capacité	1

FIG. 42 – Spécification des types Mineur et Unicité pour GUST.

**Représentation syntaxique générée.** — Le résultat de l'analyse syntaxique prendra finalement la forme de la figure 43.


 FIG. 43 – Représentation syntaxique générée par  $\mathcal{G}_{\text{synt}}^{\mathcal{P}}$ .

### 1.3 Grammaire de correspondance syntaxe-sémantique

#### 1.3.1 Version originale ( $\mathcal{G}_{\text{synt-sem}}^{\mathcal{U}\mathcal{P}}$ )

La grammaire de correspondance entre structures sémantique et syntaxique en GUP prend la forme de l'extrait présenté en figure 44.

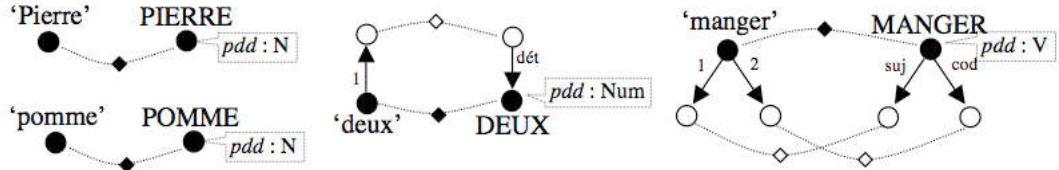


FIG. 44 – Extrait de grammaire de correspondance syntaxe-sémantique en GUP.

Cette grammaire a pour objet de mettre en correspondance les deux structures produites par la grammaire sémantique et par la grammaire syntaxique. Dans l'exemple de l'article, elle est utilisée en fonctionnement transductif (c'est-à-dire qu'on lui fournit en entrée une structure sur les deux) en synthèse (c'est-à-dire que l'on va, selon la TST, du *sens* vers le *texte*, et par conséquent qu'on lui fournit la structure sémantique en entrée).

On lui fournit donc en entrée le graphe sémantique de la figure 32 à ceci près que chaque objet du graphe, au lieu de ne porter qu'une unique polarité, en portera deux, l'une sémantique, l'autre d'interface syntaxe-sémantique. C'est ce que l'on voit dans la première structure de la figure 45. La deuxième structure montre le résultat après l'application  $\mathcal{GUP}_{\text{synt-sem}}$ , et la troisième celle après l'application de  $\mathcal{GUP}_{\text{synt}}$ .

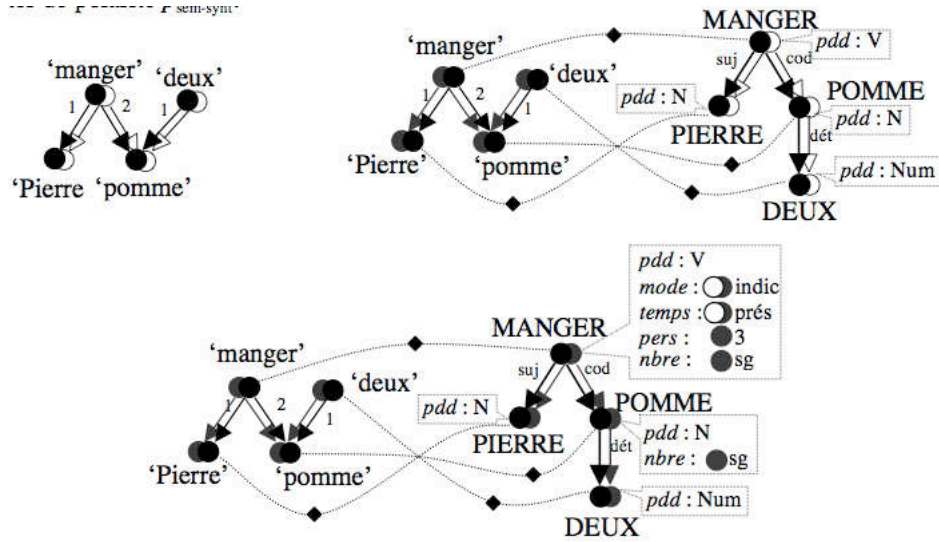


FIG. 45 — Le graphe sémantique de (17) avant, pendant et après l'interfaçage syntaxe-sémantique.

### 1.3.2 Transcription ( $\mathcal{GP}_{\text{synt-sem}}$ )

En suivant les mêmes conventions de transcription que précédemment et en réutilisant encore une fois la spécification que l'on a donnée pour les deux premières grammaires, on peut proposer les éléments suivants :

**Graphe sémantique.** — Commençons par décrire le graphe sémantique puisque c'est sa forme qui expliquera la forme des graphes suivants. Comme on l'a vu dans la figure 45 le graphe sémantique fourni en entrée doit être doublement polarisé, chaque élément recevant une polarité sémantique et une polarité de correspondance syntaxe-sémantique (fig. 46).

On double donc le graphe sémantique (G-SEM) en deux sous-graphes, comme dans la figure GUP d'origine : un premier sous-graphe correspond au graphe sémantique que l'on avait déjà donné (SG-SEM) et un second correspond à la polarisation pour l'interface sémantique-syntaxe (SG-SYNT-SEM). Ce qui permet de faire le lien entre les objets des deux sous-graphes est le partage des sémantèmes ([5], [6], [7] et [8]).

Les quatre premières Exigences correspondent aux nœuds non saturés, les trois dernières aux arcs non saturés.



$$\boxed{1} \text{G-SYNT-SEM} \left[ \begin{array}{l} \text{G-SEM} \left[ \text{SG-SYNT-SEM} \left[ \boxed{2} \text{SEM} \left[ \text{SÉMANTEME} \quad 'Pierre' \right] \right] \right] \\ \text{G-SYNT} \left[ \text{SG-SYNT-SEM} \left[ \boxed{3} \text{SYNT} \left[ \begin{array}{l} \text{LEXÈME} \quad PIERRE \\ \text{F-GRAM} \mid \text{PDD} \quad V \end{array} \right] \right] \right] \\ \boxed{4} \text{SYNT-SEM} \left[ \begin{array}{l} \text{SYNT} \quad \boxed{3} \\ \text{SEM} \quad \boxed{2} \end{array} \right] \end{array} \right]$$

La grammaire  $\mathcal{GP}_{\text{synt-sem}}$  aura donc la forme présentée en figures 47 à 49.

Pierre	
$\boxed{1} \text{G-SYNT-SEM}$	$  \left[ \begin{array}{l} \text{G-SEM} \mid \text{SG-SYNT-SEM} \mid \boxed{2} \text{SEM} \left[ \text{SÉMANTEME} \quad 'Pierre' \right] \\ \text{G-SYNT} \mid \text{SG-SYNT-SEM} \mid \boxed{3} \text{SYNT} \left[ \begin{array}{l} \text{LEXÈME} \quad PIERRE \\ \text{F-GRAM} \mid \text{PDD} \quad V \end{array} \right] \\ \boxed{4} \text{SYNT-SEM} \left[ \begin{array}{l} \text{SYNT} \quad \boxed{3} \\ \text{SEM} \quad \boxed{2} \end{array} \right] \end{array} \right]  $
Exclusion	$  \begin{array}{l} \boxed{1} \not\vdash \bullet \boxed{2} \\ \boxed{1} \not\vdash \bullet \boxed{3} \\ \boxed{1} \not\vdash \bullet \boxed{4} \end{array}  $
pomme	
$\boxed{1} \text{G-SYNT-SEM}$	$  \left[ \begin{array}{l} \text{G-SEM} \mid \text{SG-SYNT-SEM} \mid \boxed{2} \text{SEM} \left[ \text{SÉMANTEME} \quad 'pomme' \right] \\ \text{G-SYNT} \mid \text{SG-SYNT-SEM} \mid \boxed{3} \text{SYNT} \left[ \begin{array}{l} \text{LEXÈME} \quad POMME \\ \text{F-GRAM} \mid \text{PDD} \quad V \end{array} \right] \\ \boxed{4} \text{SYNT-SEM} \left[ \begin{array}{l} \text{SYNT} \quad \boxed{3} \\ \text{SEM} \quad \boxed{2} \end{array} \right] \end{array} \right]  $
Exclusion	$  \begin{array}{l} \boxed{1} \not\vdash \bullet \boxed{2} \\ \boxed{1} \not\vdash \bullet \boxed{3} \\ \boxed{1} \not\vdash \bullet \boxed{4} \end{array}  $

FIG. 47 – La grammaire de correspondance  $\mathcal{GP}_{\text{synt-sem}}$  : les entrées Pierre et pomme.



deux	
[1]G-SYNT-SEM	$\left[ \begin{array}{l} \text{G-SEM} \mid \text{G-SYNT-SEM} \mid [2]\text{SEM} \\ \text{G-SYNT} \mid \text{G-SYNT-SEM} \mid [5]\text{SYNT} \\ [8]\text{SYNT-SEM} \\ [9]\text{SYNT-SEM} \end{array} \right]$
	$\left[ \begin{array}{l} \left[ \begin{array}{l} \text{SÉMANTEME} \\ \text{ARG} \mid 1 \mid [3]\text{SEM} \mid \text{SÉMANTEME} \end{array} \right] \begin{array}{l} 'deux' \\ [4] \end{array} \\ \left[ \begin{array}{l} \text{LEXÈME} \\ \text{F-SYNT} \mid \text{DÉT} \mid [7]\text{SYNT} \end{array} \right] \begin{array}{l} [6] \\ \left[ \begin{array}{l} \text{LEXÈME} \quad DEUX \\ \text{PDD} \quad Num \end{array} \right] \end{array} \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYNT} \quad [6] \\ \text{SEM} \quad [2] \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYNT} \quad [4] \\ \text{SEM} \quad [3] \end{array} \right]$
Exigence	$\begin{array}{l} [3] \Rightarrow [4]\text{SÉMANTEME} \\ [5] \Rightarrow [6]\text{LEXÈME} \\ [1] \Rightarrow [9] \end{array}$
Exclusion	$\begin{array}{l} [1] \not\Rightarrow [2] \\ [1] \not\Rightarrow [7] \\ [1] \not\Rightarrow [8] \\ [2] \not\Rightarrow [3] \\ [7] \not\Rightarrow [5] \end{array}$

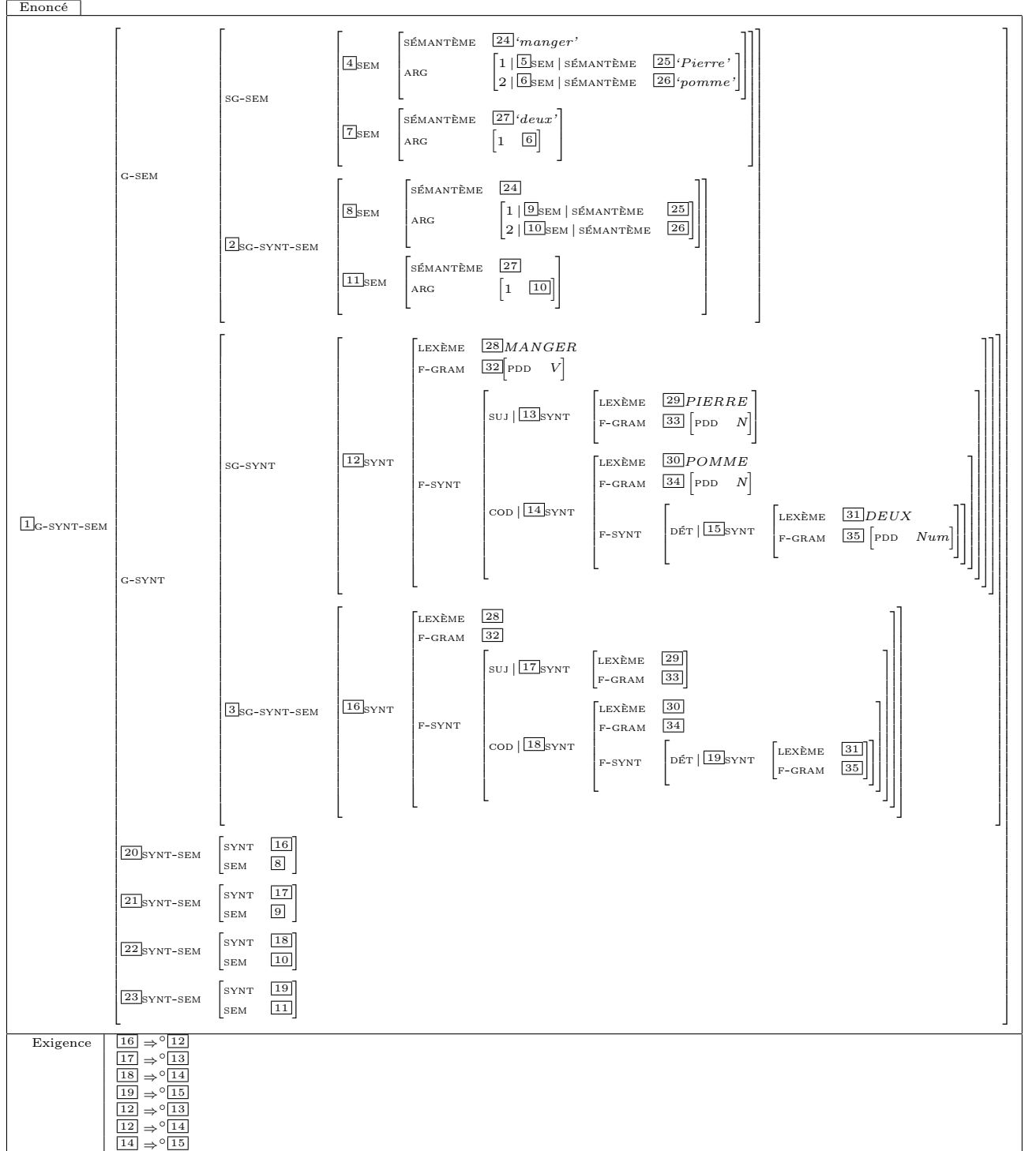
FIG. 48 – La grammaire de correspondance  $\mathcal{GP}_{\text{synt-sem}}$  : l'entrée deux.

**Application de  $\mathcal{GP}_{\text{synt-sem}}$  sur le graphe sémantique.** — A partir de ce que l'on a transcrit jusqu'à présent, l'application de la grammaire de correspondance telle que décrite dans les figures 47 à 49 sur le graphe sémantique doublement polarisé, tel que décrit en figure 46 mènera à la structure présentée en figure 50.

mange	
[1]G-SYNT-SEM	$\left[ \begin{array}{l} \text{G-SEM} \mid \text{G-SYNT-SEM} \mid [2]\text{SEM} \\ \\ \text{G-SYNT} \mid \text{G-SYNT-SEM} \mid [7]\text{SYNT} \\ \\ [12]\text{SYNT-SEM} \\ \\ [13]\text{SYNT-SEM} \\ \\ [14]\text{SYNT-SEM} \end{array} \right]$
	$\left[ \begin{array}{l} \text{SÉMANTEME} \quad 'manger' \\ \text{ARG} \quad \left[ \begin{array}{l} 1 \mid [3]\text{SEM} \mid \text{SÉMANTEME} \quad [4] \\ 2 \mid [5]\text{SEM} \mid \text{SÉMANTEME} \quad [6] \end{array} \right] \end{array} \right]$
	$\left[ \begin{array}{l} \text{LEXÈME} \quad MANGER \\ \text{F-GRAM} \mid \text{PDD} \quad V \\ \text{F-SYNT} \quad \left[ \begin{array}{l} \text{SUJ} \mid [8]\text{SYNT} \quad \left[ \text{LEXÈME} \quad [9] \right] \\ \text{COD} \mid [10]\text{SYNT} \quad \left[ \text{LEXÈME} \quad [11] \right] \end{array} \right] \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYNT} \quad [7] \\ \text{SEM} \quad [2] \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYNT} \quad [8] \\ \text{SEM} \quad [3] \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYNT} \quad [10] \\ \text{SEM} \quad [5] \end{array} \right]$
Exigence	$\begin{array}{l} [3] \Rightarrow {}^\circ [4]\text{SÉMANTEME} \\ [5] \Rightarrow {}^\circ [6]\text{SÉMANTEME} \\ [8] \Rightarrow {}^\circ [9]\text{LEXÈME} \\ [10] \Rightarrow {}^\circ [11]\text{LEXÈME} \\ [1] \Rightarrow {}^\circ [13] \\ [1] \Rightarrow {}^\circ [14] \end{array}$
Exclusion	$\begin{array}{l} [1] \not\Rightarrow [2] \\ [1] \not\Rightarrow [7] \\ [1] \not\Rightarrow [12] \\ [2] \not\Rightarrow [3] \\ [2] \not\Rightarrow [5] \\ [7] \not\Rightarrow [8] \\ [7] \not\Rightarrow [10] \end{array}$

 FIG. 49 – La grammaire de correspondance  $\mathcal{G}_{\text{synt-sem}}^{\mathcal{P}}$  : l'entrée mange.

Les Exigences sont du même ordre que celles du graphe sémantique doublement polarisé, si ce n'est que cette fois elles concernent, dans le graphe syntaxique (G-SYNT) les éléments du sous-graphe syntaxique (SG-SYNT), réclamés par leurs équivalents du sous-graphe d'interface syntaxe-sémantique (SG-SYNTS-SEM) qui viennent d'être introduits par l'analyse.


 FIG. 50 – La sortie de l'analyse du graphe sémantique par  $\mathcal{G}^{L_{synt-sem}}$ .

**Application de  $\mathcal{G}^{P_{synt}}$  sur le graphe d'interface syntaxe-sémantique.** — Quand enfin l'on applique la grammaire syntaxique à la structure obtenue précédemment, on obtient la

même structure à ceci près que plus aucune exigence n'est présente et que les objets syntaxiques ont reçu toutes les informations qu'ils réclamaient (ce qui était donné dans la grammaire syntaxique).

## 1.4 Synthèse

### 1.4.1 Résumé des équivalences

Voici en figure 51 un tableau qui synthétise les équivalences entre représentations que l'on a utilisées au cours de cette section.

GUP	GP
Graphe	Construction
nœud	Information intrinsèque : attribut SEM ou SYNT
Arc	Information intrinsèque : attribut ARG, F-SYNT ou SYNT-SEM
Fonction grammaticale	Information intrinsèque : attribut F-GRAM
nœud saturé •	Information extrinsèque : Exclusion d'un SEM ou d'un SYNT
Arc saturé ▼	Information extrinsèque : Exclusion d'un ARG, d'un F-SYNT ou d'un SYNT-SEM
nœud non saturé ○	Information extrinsèque : Exigence d'un SEM ou d'un SYNT
Arc non saturé ▽	Informations extrinsèques : Propriétés d'Exigence, Unicité et Mineur d'un SEM ou d'un SYNT

FIG. 51 – Résumé des équivalences entre représentations GUP et GP.

### 1.4.2 Remarques sur la transcription

Plusieurs remarques peuvent être faites concernant cette transcription.

Tout d'abord, les extraits de grammaire cités dans [Kahane & Lareau, 2005] et leur fonctionnement sont tellement simplifiés (p.ex. l'accord en nombre avec le numéral) qu'il est difficile de savoir clairement si la transcription que l'on en a proposée ici est effectivement exacte ou non. Pour essayer de s'en rapprocher le plus possible on a fait en sorte de traduire le plus régulièrement possible chaque partie d'information, et le plus explicitement possible également, ce qui nous a menée à une représentation qui n'est pas la plus simple qu'on puisse imaginer sur la base des exemples GUP (cf. la complexité de la figure 50 par rapport à son équivalent graphique en GUP, fig. 44). On aurait pu se permettre une interprétation un peu plus libre de la structuration des informations indiquées par les auteurs et ainsi transcrire en GP en utilisant une notation plus légère, plus courte; nous avons refusé de le faire parce que nous avons souhaité représenter ce dont on disposait aussi objectivement qu'il nous était possible de le faire.

La deuxième remarque concerne les propriétés utilisées, i.e. essentiellement Exigence et Exclusion, quelques rares fois Mineur et Unicité, une unique fois Majeur (dans la structure initiale de la grammaire syntaxique). Ceci vient de deux raisons bien distinctes. La première est que les extraits de grammaire fournis ne donnent aucune indication sur l'ordre relatif des objets, ce qui est normal puisque cette information, selon la théorie utilisée dans GUST, prend sa place dans le module morphotopologique qui n'est pas développé ici ; le résultat de cela est l'absence de propriétés d'ordre dans les constructions GP correspondantes. La seconde raison est que les informations d'accord donnent lieu à des entrées spécifiques de la grammaire, donc en GP à des constructions dont l'accord se traduit non pas par la propriété du même nom, mais par un jeu d'Exigences/Exclusions entre informations d'accord des objets (p.ex., en fig. 41 on a une construction qui indique que le recteur d'une relation sujet-verbe dont le nombre est singulier exige un dépendant dont le nombre est lui aussi singulier). Cette seconde raison nous fait penser que même si l'on avait disposé d'un extrait de grammaire morphotopologique ici, et donc d'informations sur l'ordre relatif des objets, ceux-ci auraient probablement revêtu la même forme que les informations d'accord au niveau syntaxique, i.e. celle d'un ensemble d'Exigences et d'Exclusions.

Ce que l'on peut conclure de cela est que ce modèle, par rapport au nôtre, a une façon bien distincte d'organiser, et par là d'exprimer les types d'information : là où GUST utilise un **type de relation régulier** en termes de "ressources" et de "besoins" <sup>76</sup>, qui permet la polarisation des structures, et qui implique parallèlement de décliner des graphes exprimant des **informations de toutes natures** (règles lexicales, sagittales, grammaticales,...), notre modèle utilise un **type de représentation régulier** en termes de constructions, qui permet d'envisager l'analyse dans les termes uniques d'un processus de satisfaction de contraintes, et qui implique parallèlement de décliner des propriétés exprimant des **relations de toutes natures** (constituance, ordre, accord,...).

Il émerge également de cette comparaison des différences notables entre les deux formalismes GUP et GP : GUP permet de représenter des grammaires lexicalisées alors que le nôtre est envisagé de manière à ne pas imposer cela. Toutefois comme on l'a vu plus haut, la frontière entre lexicalisation et non lexicalisation n'est que conventionnelle et l'on peut passer outre. Tandis que celle que l'on vient d'explicitier semble plus difficile à traverser, à tel point qu'il ne nous paraît pas possible, en l'état de nos connaissances de GUP <sup>77</sup>, de polariser une GP. De là, si l'on en croit [Kahane, 2004] démontrant que la polarisation consiste en la mise en évidence de processus communs aux Grammaires d'Unification, on peut déduire que GP ne fait pas partie de ce groupe.

En effet, comme on a pu le dire précédemment l'objectif de GP est de proposer des structures analytiques sur la base de l'ensemble des informations disponibles, sans que ces structures doivent nécessairement aboutir à une forme complète, finie, générale. Le fait d'y parvenir est une possibilité, certes heureuse, néanmoins contingente. A l'opposé le principe même de la polarisation d'une grammaire revient à considérer que le but d'une analyse est de produire une structure englobante, neutralisée, stable. Ces deux objectifs semblent, à la lumière de ce que l'on vient de voir, incompatibles : polariser une GP reviendrait soit à en modifier les

---

<sup>76</sup>Que l'on exprime en GP avec des propriétés de constituance et cooccurrence.

<sup>77</sup>Bien entendu cette affirmation est directement liée au fait que GUP n'est pas l'objet de nos recherches et que l'on n'en a donc qu'une connaissance limitée.

buts, soit à reconsidérer la métaphore de la polarisation en posant que la neutralisation des structures ne serait plus qu'une anecdote, ce qui semble s'opposer à la raison même de son utilisation.

## 2 Un modèle catégoriel procédural en GP

Dans cette section on s'intéresse à la représentation d'une grammaire catégorielle procédurale, c'est-à-dire d'une ressource dont les objets prennent la forme de procédures, et non pas directement d'informations linguistiques. Comme on l'a vu au chapitre 1, l'un des modèles typiques de ce type d'approche est *Dynamic Syntax* (ci-après DS). Plus précisément, on s'intéresse ici à la présentation du modèle proposée dans [Kempson *et al.*, 2001].

On cherche ici à voir si l'on peut représenter des informations procédurales en GP, sans tenir compte de la question de l'implémentation de ces représentations, qui relèvent exactement du même mécanisme que celui utilisé par DS. Autrement dit, on ne s'intéresse qu'à représenter la grammaire, et non le mécanisme d'analyse.

### 2.1 Présentation de DS

DS essaie de représenter la structure d'une langue selon les termes d'un processus qui consiste à construire progressivement une interprétation à partir d'une séquence de mots, en fonction de leur ordre d'apparition et de ce que chacun implique sur la structuration de l'énoncé. Dans l'architecture dynamique, les propriétés des phrases expriment l'établissement progressif de leur interprétation.

#### 2.1.1 DS en tant que grammaire procédurale

L'idée de départ est de proposer un processus qui reflète celui que les interlocuteurs suivent pour interpréter un énoncé, quand vraisemblablement ils n'utilisent qu'assez peu d'informations issues des mots et les enrichissent en une sorte d'interprétation sélectionnée. Voulant représenter prioritairement non pas ce qui est donné comme information mais plutôt *comment* elle en vient à être comprise, Kempson *et al.* considèrent l'interprétation d'une occurrence comme un **processus incrémental inférentiel**.

L'analyse, donc l'interprétation d'un énoncé en DS, consiste en l'élaboration d'un arbre (ou plus précisément, d'une description d'arbre telle que définie dans [Blackburn & Meyer-Viol, 1994], [Blackburn *et al.*, 1996]). Cela commence par l'introduction d'un nœud initial (le formalisme étant *goal-oriented*, i.e. on pose *a priori* que ce que l'on essaie de reconstituer est un énoncé), et se termine (dans un cas où l'analyse réussit) par un arbre dont les nœuds sont des formes logiques associées aux objets linguistiques. Chaque étape intermédiaire, correspondant à l'analyse d'un élément de l'énoncé, consiste à ajouter de l'information à l'arbre, c'est-à-dire à le construire progressivement.

Toute l'information linguistique de DS est située dans son lexique. Dans celui-ci, à chaque entrée correspond non pas un ensemble de traits linguistiques, mais une suite d'opérations à effectuer dans l'arbre. Ces opérations peuvent consister à ajouter de l'information, à compléter les informations existantes, ou à se déplacer d'un nœud à l'autre. À chaque étape on sait ce qui existe déjà dans la structure, et ce qui reste à compléter. En effet chaque nœud porte deux types d'informations (*decorations*) : celles qui se réfèrent à des opérations effectuées (*annotations*), et celles qui se réfèrent à des opérations attendues (*requirements*). Chaque nœud est introduit en fonction des attentes des autres, c'est-à-dire qu'il prend place à un certain endroit de l'arbre parce qu'il remplit un certain nombre d'exigences exprimées par les nœuds préexistants.

À ce lexique procédural est ajouté un ensemble de règles générales que l'on peut appliquer à l'arbre. Le choix dans l'application d'une règle plutôt qu'une autre est opéré selon la méthode du *labelled deductive system* de [Gabbay, 1996]. Ces règles sont considérées comme universelles, au sens où elles sont censées être valables pour toutes les langues (contrairement au contenu du lexique qui lui, diffère suivant les langues).

### 2.1.2 DS en tant que Grammaire Catégorielle

Bien qu'il ne s'agisse pas d'un stéréotype de grammaire catégorielle (ci-après CG), on peut toutefois considérer que DS s'inscrit dans cette lignée. En effet elle est basée sur les quatre idées-clés des grammaires catégorielles (telles que définies dans [Moortgat, 1990]) :

- Le **lexicalisme**. — Les approches catégorielles sont celles qui sont les plus marquées parmi les grammaires lexicalisées (ce que confirment, p.ex., [Baldrige & Kruijff, 2003]). En CG, la composante syntagmatique est intégralement supprimée au profit des informations lexicales, et c'est typiquement le cas de DS, qui ne dispose d'aucune ressource en dehors du lexique et de quelques règles de combinaison (qui, elles, forment une combinatoire libre, qui n'est réduite que par le lexique, propre à une langue donnée).
- La **structure fonction-argument**. — Il est fait, en CG, une opposition de base entre deux types d'objets : les arguments d'une part, et les foncteurs d'autre part. La façon dont les objets se combinent est gérée par l'expression même des objets, c'est-à-dire leur nature de foncteur (qui introduit une relation) ou d'argument (qui prend place dans une relation). On retrouve cela en DS dans l'utilisation des types sémantiques, où chaque objet est défini dans la structure soit comme une entité (*e* ou *tt*), soit comme un objet réclamant une entité ( $e \rightarrow t$  p.ex.).
- La **flexibilité de la structure de constituants** : À une même expression non ambiguë est associé un ensemble de dérivations équivalentes (et non pas une unique analyse), à l'aide d'un système combinatoire plus souple que celui des grammaires syntagmatiques. En DS, cela correspond au choix dans l'enchaînement de l'application des règles.
- La **compositionnalité** : La relation entre syntaxe et sémantique est considérée comme homomorphique (ce qui est souligné aussi p.ex. par [Retoré & Moot, 2006]), c'est-à-dire que l'on y postule une correspondance régulière entre les règles d'analyse syntaxique (i.e., la distinction entre foncteur et argument) et la structure sémantique (c'est la "*rule-by-rule hy-*

*pothesis*" de [Montague, 1974]). C'est également le cas de DS, où la grammaire ne mentionne pas explicitement des informations syntaxiques vs. sémantiques, et où la structure construite est une structure dite d'"*interprétation*", c'est-à-dire qu'elle mêle structure syntaxique et informations sémantiques comme deux visions d'une même organisation linguistique.

## 2.2 Exemple d'analyse

Afin de voir un peu plus précisément à quoi ressemble le modèle DS et comment il fonctionne, [Kempson *et al.*, 2001] illustrent leurs propos (p. 56-76) par l'analyse pas à pas de (18a) ; pour nos besoins, nous nous inspirons de cette présentation en l'adaptant au traitement de (18b)<sup>78</sup>.

- (18) a. John admires him  
b. Jean admire Pierre

Dans ce qui suit nous nous basons sur cet exemple pour présenter ce que l'on peut formaliser en GP, et de quelle façon, en suivant la progression de l'analyse de l'exemple. Notons que puisqu'il s'agit d'un exemple d'analyse, nous présenterons ici à la fois des constructions prenant leur place dans la grammaire, et des constructions représentant le résultat de l'analyse produite.

### 2.2.1 Création du nœud racine

La toute première étape d'analyse consiste en la création du nœud initial de l'arbre. Toutes les analyses DS, quelle que soit l'entrée, commencent par cette même étape. Le nœud introduit porte les informations contenues dans la fig. 52.

$$\boxed{\{Tn(a), ?Ty(t), \diamond\}}$$

FIG. 52 – Le nœud racine en DS.

Dans la description d'arbre, chaque nœud de l'arbre est présenté entre accolades (l'arbre ne porte ici que le nœud initial donc une seule paire d'accolades), et est constitué de plusieurs informations qui sont séparées par des virgules. La première information que ce nœud contient,  $Tn(a)$  indique que ce nœud est l'ancre<sup>79</sup> de la description d'arbre ; cette ancre porte pour valeur, pour l'instant, la variable  $a$  (il arrive dans certaines analyses que plusieurs nœuds initiaux soient introduits, auquel cas ils portent chacun une valeur différente).  $?Ty(t)$  est une exigence (signalée par le point d'interrogation) indiquant que le nœud pointé exige d'être un

<sup>78</sup>Le modèle DS a été illustré par des exemples de diverses langues, principalement l'anglais mais aussi l'arabe et le japonais, mais n'a pas encore été développé pour le français. L'analyse que nous en proposons dans ce qui suit (ou plus précisément, le contenu du lexique) ne vient donc pas d'une proposition en DS mais de notre propre interprétation de DS, adaptée au français. On remplace l'objet pronominal de l'exemple original par l'emploi d'un nom propre parce que cela simplifiera l'exemple sans en modifier le fonctionnement.

<sup>79</sup>Il ne s'agit pas ici de l'ancre au sens des TAG p.ex., mais de l'ancre par rapport à laquelle se situeront les autres nœuds de l'arbre.



élément de type  $t$ <sup>80</sup>. Enfin, le  $\diamond$  signale un pointeur permettant d'indiquer sur quel nœud de l'arbre on se trouve chaque étape de l'analyse (il pourra être déplacé dans l'arbre en fonction des indications fournies par les objets analysés), et ainsi de préciser où les règles de construction s'appliquent.

**Transcription de la description d'arbre en GP.** — La description d'arbre à la suite de cette première étape prendra la forme de la fig. 53.

Arbre		
DESC	$\diamond$	$\boxed{\boxed{1}\text{NOEUD} \left[ \begin{array}{ll} \text{POSITION} & a \\ \text{2TYPE} & t \end{array} \right]}$
Exigence	$\boxed{1} \rightarrow ?\boxed{2}$	—

FIG. 53 — La description d'arbre à l'issue de la première étape d'analyse, en GP.

Obéissant aux principes de GP, elle fait l'objet d'une construction, dont les informations sont les suivantes :

- *Identifiant.* — On affecte à la construction l'étiquette “arbre”.
- *Informations intrinsèques.* — On sait que cet arbre est constitué, pour l'instant, d'un NOEUD. On connaît le TYPE de ce noeud :  $t$ . On connaît aussi sa POSITION dans la structure arborescente : il en est l'ancree,  $a$ . Enfin, on sait que le pointeur est placé sur ce noeud : on le fait donc figurer devant lui.
- *Informations extrinsèques.* — On sait que le  $[\text{TYPE } t]$  du noeud est exigé ; on peut exprimer cela sous la forme d'une propriété de type Exigence, à laquelle on attribue le même symbole qu'en DS (le ?), qui prend deux termes (le premier étant celui qui exige — c'est-à-dire le noeud dans lequel est indiquée l'Exigence —, et le second celui qui est exigé), qui est toujours évaluée une et une seule fois (compte-tenu de la générativité du modèle) et est satisfaite si le second terme est disponible (fig. 54).

Exigence	
Symbole	$\rightarrow ?$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Autorisée
Commutativité	Faux
Disponibilité	Vrai
Satisfaction	$\boxed{2}$ = disponible
Capacité	1

FIG. 54 — Spécification du type Exigence pour DS.

<sup>80</sup>Il s'agit du type sémantique de la *logique des types* (*typed logical language*, cf. p.ex. [Cann, 1993]), où l'on pose deux types de base,  $t$  pour l'énoncé (qui dénote une valeur de vérité) et  $e$  pour le SN (qui dénote une entité), et où les autres objets sont exprimés en fonction de leur relation à ces deux entités, à la manière des grammaires catégorielles (p.ex.  $e \rightarrow t$  pour un verbe intransitif, *i.e.* un objet qui réclame un  $e$  pour former un  $t$ ).

A ce stade de l'analyse, la propriété d'Exigence n'est pas satisfaite (puisque l'exigence est mentionnée dans la description d'arbre en DS).

**Mécanisme d'introduction.** — Pour que cette construction “Arbre” soit introduite dès le début de l'analyse, cela demande d'utiliser les mécanismes d'analyse qui sont propres à DS. Le mécanisme d'analyse de GP n'est en effet pas étudié *a priori* pour effectuer des analyses descendantes (*top-down*) mais pour des ascendantes (*bottom-up*), et il en résulte que cette construction ne pourrait être introduite sans qu'une partie de ses propriétés ne soit, *a minima*, évaluées (or ce ne serait pas le cas ici, en tout début d'analyse). Il s'agit là d'une différence de taille entre GP et ce type d'analyse, et qui signifie que mis à part la question de la représentation des informations, GP nécessiterait une extension de ses processus pour effectuer ce type d'analyse. Cette remarque étant exprimée, cela n'empêche pas de proposer une représentation GP de DS, et de voir comment l'on utilise ce qui existe déjà, afin d'étudier ce qui serait à ajouter dans une telle perspective.

### 2.2.2 Développement du nœud

On applique ici la règle d'introduction (*Introduction rule*), qui consiste à développer le contenu du nœud sur lequel se trouve le pointeur en vertu d'une abduction : pour obtenir une formule de type  $t$ , l'une des possibilités est d'exiger deux autres formules, l'une de type  $e$  et l'autre de type  $(e \rightarrow t)$ . La première fait office d'argument ( $\langle \downarrow_0 \rangle$ ), la seconde de foncteur ( $\langle \downarrow_1 \rangle$ ); dans le modèle, les positions des branches sont directement interprétées comme argument pour la branche de gauche, foncteur pour celle de droite. On obtient alors la description d'arbre de la fig. 55.

$$\boxed{\{Tn(a), ?Ty(t), ?\langle \downarrow_0 \rangle Ty(e), ?\langle \downarrow_1 \rangle Ty(e \rightarrow t), \diamond\}}$$

FIG. 55 – Description d'arbre à l'issue de la deuxième étape, en DS.

Ces deux objets sont bien entendu réclamés (puisque'ils n'ont pas été analysés, mais sont le résultat de l'application de la règle d'introduction). A l'issue de ceci, on se trouve toujours avec un arbre à nœud unique, mais dont le contenu a été développé en vertu d'une règle qui indique un principe général de structuration.

**Transcription de la description d'arbre en GP.** — La description d'arbre correspondant à cette étape de l'analyse prendra la forme de la fig. 56.

On y a ajouté, par rapport à la description précédente, la mention des deux BRANCHES que l'on a introduites, en y mentionnant les informations que l'on en connaît : leurs types et leurs positions. On les ajoute également à la liste des Exigences.

En termes de mécanismes d'analyse, ce nouvel état de la description d'arbre ne correspond pas à l'introduction d'une nouvelle construction dont la description précédente serait un constituant, mais à l'unification de la description dans son état précédent avec une construction lui ajoutant les nouvelles informations (pour preuve, notamment le fait que la propriété précé-

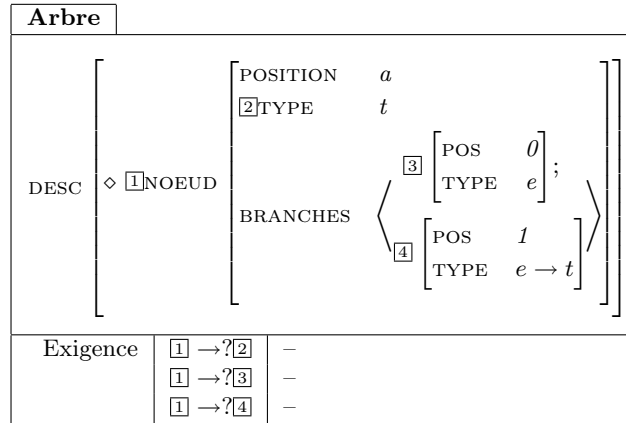


FIG. 56 – La description d'arbre à l'issue de la deuxième étape d'analyse, en GP.

dente est reproduite ici, et ajoutée aux nouvelles contraintes). Cette construction avec laquelle la description a été unifiée correspond à ce qui permet d'effectuer une règle d'introduction, i.e. une construction qui contient les deux branches (dont les types sont exprimés non pas absolument, mais relativement au type de l'objet avec lequel elle s'unifie) et les deux propriétés les concernant. Le mécanisme d'analyse ne fonctionne donc pas par caractérisation, mais par unification, ce qui est différent de GP.

On pourrait proposer une construction "Introduction" qui correspondrait au contenu à unifier avec celui de l'arbre lors de l'application de la règle d'introduction (fig. 57).

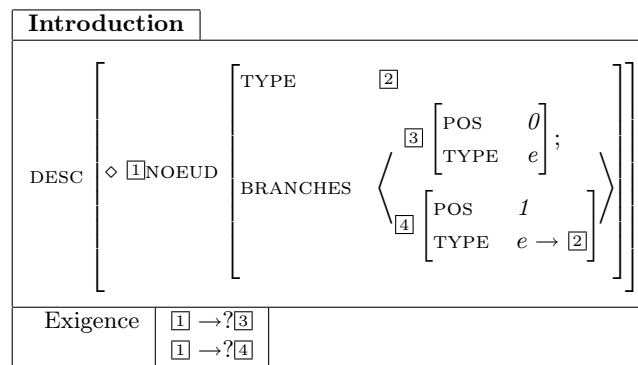


FIG. 57 – Une construction figurant la règle d'introduction, en GP.

Dans cette construction, on précise l'emplacement du pointeur, puisqu'il indique à quel noeud l'on ajoute les informations (ceci apparaîtra comme nécessaire quand la description d'arbre comportera plus d'un seul noeud). La première branche correspond à un argument de type  $e$ , et la seconde à un foncteur de type  $e \rightarrow [2]$ , c'est-à-dire qu'il est exprimé relativement au type du noeud auquel on applique la règle.

### 2.2.3 Introduction de nouveaux nœuds

Une fois l'existence des deux nœuds-fils annoncée par la règle précédente au niveau du nœud-tête, on introduit les descriptions des nœuds argument et foncteur dans l'arbre en appliquant deux fois la règle de prédiction (*prediction rule*). On ajoute alors à la description d'arbre les informations correspondant, successivement, aux deux nœuds : celui de gauche (l'argument) est de type  $e$ , et celui de droite (le foncteur) est de type  $(e \rightarrow t)$ , cf. fig. 58 (on a également représenté l'arbre correspondant pour en faciliter la lecture).

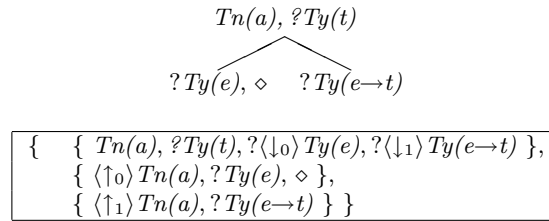


FIG. 58 – Introduction de nouveaux noeuds en DS.

Pour chacun des nœuds fils, on indique également leur position par rapport à l'ancree de l'arbre ( $Tn(a)$ ), en précisant le parcours à suivre dans l'arbre pour y accéder (en remontant une branche gauche pour l'argument,  $\langle \uparrow_0 \rangle$ , ou une branche droite pour le foncteur,  $\langle \uparrow_1 \rangle$ ).

A noter aussi qu'à l'issue des applications de cette règle, une décision a été prise concernant le déplacement du pointeur vers l'argument ; ceci, d'après les auteurs, peut différer en fonction de l'ordre des mots dans les langues. La position en sortie de l'application règle de prédiction est essentielle puisqu'elle conditionne à quoi va s'appliquer la règle suivante (mais les auteurs ne précisent pas comment cette décision est prise, ni si cela est dépendant de la règle de prédiction, auquel cas on sortirait de l'universalité des règles qui s'oppose à la spécificité des lexiques dans ce type de grammaire). On peut supposer que le déplacement du pointeur lors de l'application d'une règle relève d'une "propriété" au même titre que dans certaines grammaires syntagmatiques, *i.e.* une régularité propre à une langue, mais exprimée au même niveau que les "principes".

**Transcription de la description d'arbre en GP.** — La description d'arbre prendra à l'issue de cette étape la forme de la fig. 59.

On y voit que la description contient maintenant trois nœuds, le second portant le pointeur. Comme précédemment, toutes les propriétés correspondant aux noeuds sont ajoutées aux informations extrinsèques, et aucune exigence n'est pour l'instant satisfaite.

**Transcription de la prédiction en GP.** — Comme pour la règle d'introduction, la description d'arbre obtenue à l'issue de cette étape correspond à l'unification de la description précédente et d'une autre construction, que l'on peut appeler "Prédiction" ; cette unification s'est effectuée deux fois, une première pour le foncteur, une seconde pour l'argument (et lors de la seconde, le pointeur a été déplacé). Ces deux actions ne donnent pas le même résultat sur la description d'arbre, donc on peut imaginer deux constructions héritant de la construc-

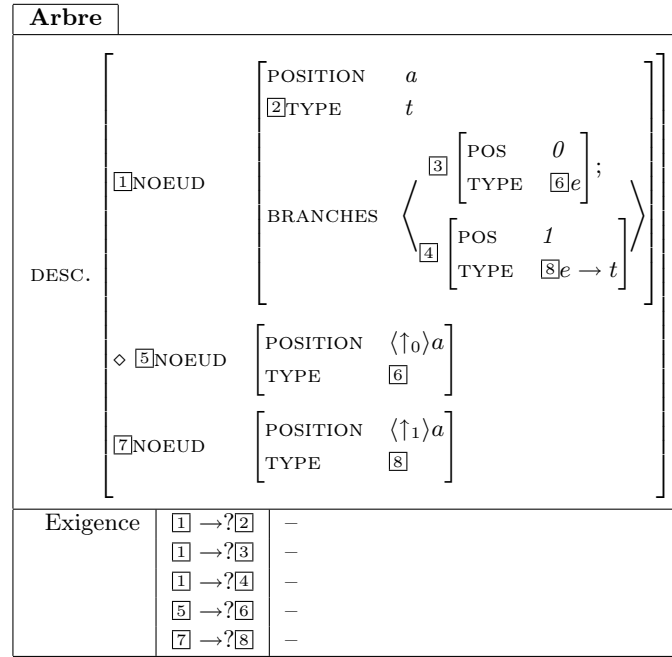


FIG. 59 – Description de l'arbre à l'issue de la troisième étape, en GP.

tion générale “Prédiction”, une première concernant la prédiction d’un foncteur (on l’appellera “Pred-f”) et une seconde concernant la prédiction d’un argument (et l’appellera “Pred-a”).

La construction correspondant à la règle générale de prédiction est donnée en fig. 60.

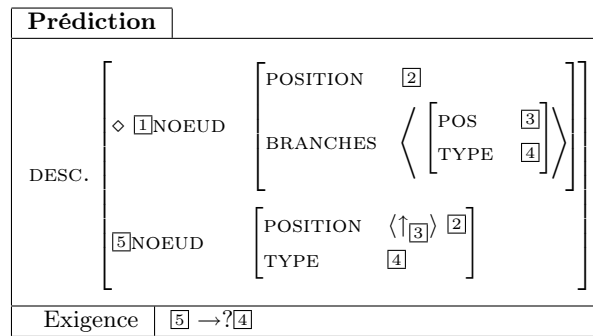


FIG. 60 – La construction figurant la règle générale de prédiction, en GP.

Cette construction exprime la partie régulière de la règle de prédiction : on y indique la création d’un nouveau NOEUD (ici indexé [5]), dont la POSITION correspond à celle d’argument ou foncteur (suivant la valeur de [3]) de l’objet qui lui-même est en position [2], et qui porte le TYPE mentionné dans la BRANCHE considérée, i.e. celle correspondant à [4]. On ajoute à cela la propriété d’Exigence du TYPE par le nouveau NOEUD.

La construction correspondant à la prédiction d’un foncteur est donnée en fig. 61.

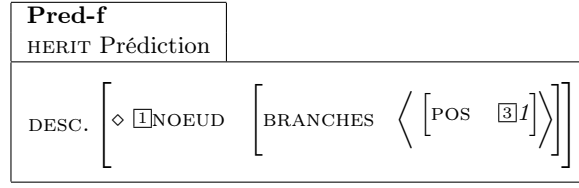


FIG. 61 – La construction figurant la prédiction d'un foncteur, en GP.

Celle-ci hérite de la précédente, et y ajoute uniquement une précision concernant la POSITION de la BRANCHE considérée, qui doit être de valeur 1.

Enfin, la construction correspondant à la prédiction d'un argument est donnée en fig. 62.

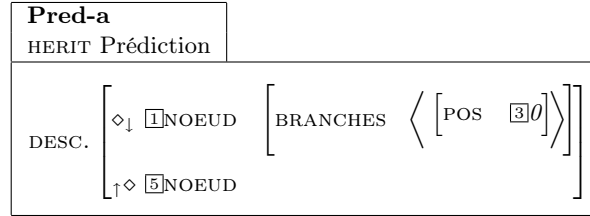


FIG. 62 – La construction figurant la prédiction d'un argument, en GP.

Cette dernière construction correspondant à la règle de prédiction hérite de la construction générale Prédiction, et y ajoute deux nouvelles informations : tout d'abord une précision concernant la valeur de la POSITION de la BRANCHE considérée, qui doit cette fois être 0, et ensuite elle indique le déplacement du pointeur. Pour ce faire, nous proposons de représenter par  $\diamond_{\downarrow}$  la position de départ, et par  $\uparrow \diamond$  la position d'arrivée du pointeur.

#### 2.2.4 Incorporation d'information lexicale

C'est à ce stade que l'on en vient à observer le premier mot de l'entrée, *Jean*. L'analyse d'un mot de l'énoncé consiste à appliquer à la description d'arbre la séquence d'instructions correspondant à ce mot, dans le lexique. Dans le modèle, les entrées lexicales prennent la forme d'une suite IF - THEN - ELSE qui peuvent contenir des instructions :

- d'ajout d'information au nœud sur lequel se trouve le pointeur au moment de l'application de l'instruction,
- de création de nouveaux nœuds dépendant de celui sur lequel se trouve le pointeur,
- de déplacement du pointeur.

Dans notre exemple, l'entrée lexicale correspondant à *Jean* est donnée en fig. 63.

Cette entrée indique que SI, quand on rencontre cette entrée, le pointeur se trouve sur un objet contenant l'exigence d'un type *e*, ALORS on ajoute à ce nœud l'information disant qu'il s'agit de la forme *John*, l'indication qu'il s'agit bien d'un objet de type *e* et celle indiquant que c'est un nœud terminal (i.e. pas de descendance possible,  $[\downarrow] \perp$ ). SINON, si le pointeur se

<i>Jean</i>	IF	$?Ty(e)$ ,
	THEN	add $Fo(Jean)$ , $Ty(e)$ and $[\downarrow]\perp$ to the description,
	ELSE	ABORT.

 FIG. 63 – L'entrée lexicale *Jean* en DS.

situé sur un objet qui n'exige pas de type  $e$ , alors le processus d'analyse est arrêté.

Quand on applique cela à notre description d'arbre, on obtient le résultat donné en fig. 64.

$$\boxed{\begin{array}{l} \{ \{ Tn(a), ?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_1\rangle Ty(e \rightarrow t) \}, \\ \{ \langle\uparrow_0\rangle Tn(a), Fo(Jean), Ty(e), [\downarrow]\perp, \diamond \}, \\ \{ \langle\uparrow_1\rangle Tn(a), ?Ty(e \rightarrow t) \} \} \end{array}}$$

 FIG. 64 – Application des informations de *Jean* à l'arbre en DS.

C'est-à-dire que l'on a ajouté les informations contenues dans l'entrée lexicale à celles que l'on avait déjà. On a, pour ce faire, appliqué une règle implicite de réduction (*thinning*) indiquant que si l'on a dans un même nœud les informations  $\phi$  et  $? \phi$ , alors on retire l'exigence du contenu du nœud (puisque'elle est alors remplie) : c'était le cas pour  $Ty(e)$  et  $?Ty(e)$ , et on a donc retiré  $?Ty(e)$ .

**Transcription de la description d'arbre en GP.** — A l'issue de cette étape, la description d'arbre prendra en GP la forme de la fig. 65.

On a ajouté dans les informations intrinsèques la FORME *Jean* et l'indication que ce nœud est terminal (i.e., qu'il ne peut porter de BRANCHES). Dans les propriétés, l'ajout de l'information selon laquelle il s'agit d'un objet de type  $e$  permet de satisfaire l'Exigence  $\boxed{5} \rightarrow ?\boxed{6}$ .

**Transcription de l'entrée lexicale *Jean* en GP.** — Ici encore, cette description d'arbre a été obtenue par l'unification de la structure précédente, avec celle correspondant à l'objet *Jean* dans le lexique. On peut considérer que ce dernier, en GP, aura la forme donnée en fig. 66.

Cette construction contient le contenu des informations à ajouter à la structure dans la suite d'instruction THEN, mais la simple mention de ceci est suffisant à l'expression de la forme IF - THEN - ELSE de l'entrée en DS. En effet, si la condition initiale (celle du IF) n'est pas satisfaite alors l'unification échouera (ce qui équivaut à l'instruction "*abort*" du ELSE). Si elle est satisfaite, alors l'unification fonctionnera.

On doit maintenant propager à l'arbre (en l'occurrence, au nœud supérieur) les informations que l'on vient d'ajouter ici. C'est l'objet de la règle suivante.

**Satisfaction descendante des propriétés.** — On remarque qu'ici, ce qui provoque l'évaluation (et la satisfaction le cas échéant) de la propriété d'Exigence est ce qui est contenu dans les propriétés intrinsèques de la construction que l'on est en train de construire. Il ne

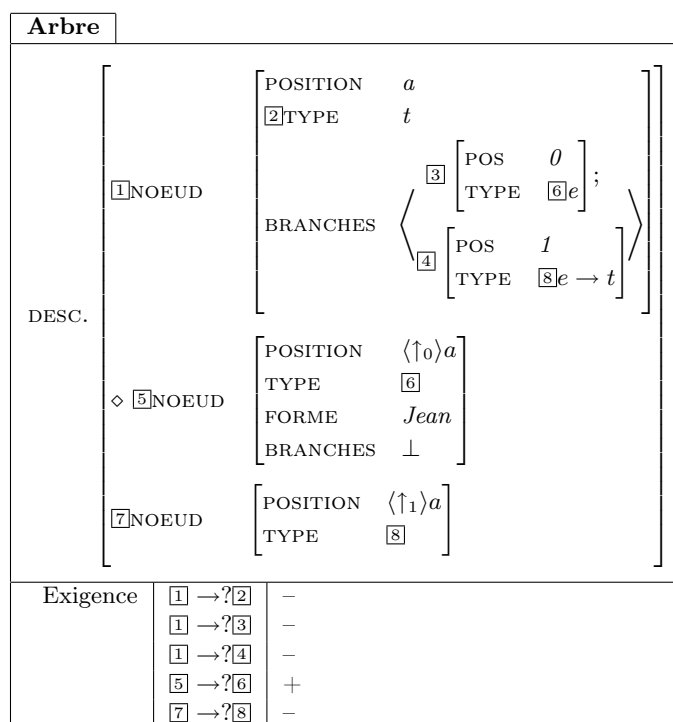
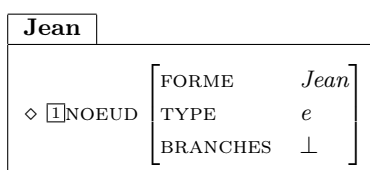


FIG. 65 – Description de l'arbre à l'issue de la quatrième étape, en GP.


 FIG. 66 – L'entrée lexicale *Jean* en GP.

s'agit donc pas là de la même façon d'évaluer les propriétés que ce que l'on fait normalement dans GP. En fait, on peut considérer que cela est dû au fait qu'on est ici dans une analyse descendante comme on l'avait constaté précédemment : dans une analyse standard GP, qui s'effectue de façon montante, les informations proviennent des constituants de la construction et *remontent* dans la structure, des informations extrinsèques vers les intrinsèques. A l'inverse, dans une analyse descendante, on constate que les informations proviennent de l'unification des structures (et non de la caractérisation), et c'est le contenu des informations intrinsèques qui *descend* vers les extrinsèques.

Il s'agit là, bien entendu, de quelque chose qui n'a pas été pris en considération dans le formalisme GP dans son état actuel ; nous ne faisons que constater les différences qu'il existe entre ce qui est déjà disponible, et ce qui devrait (ou tout au moins, pourrait) être fait pour envisager une analyse descendante. Nous avons toutefois trouvé intéressant de souligner le parallélisme que l'on voit apparaître ici entre le sens de l'analyse et le sens de la propagation des informations dans la construction.



### 2.2.5 Complétion de l'arbre

Cette étape consiste à ajouter au nœud père, puisque son exigence d'un argument de  $Ty(e)$  a été satisfaite, l'information qui va exprimer qui est son argument, finalement. On lui ajoute donc cette information, sous la forme de deux éléments,  $\langle \downarrow_0 \rangle Fo(Jean)$  qui indique sa forme, et  $\langle \downarrow_0 \rangle Ty(e)$  qui indique son type<sup>81</sup>. Cette dernière étape consiste également à placer le curseur dans la position requise pour la suite, i.e. sur le dernier élément en l'occurrence. On obtient le résultat présenté en fig. 67.

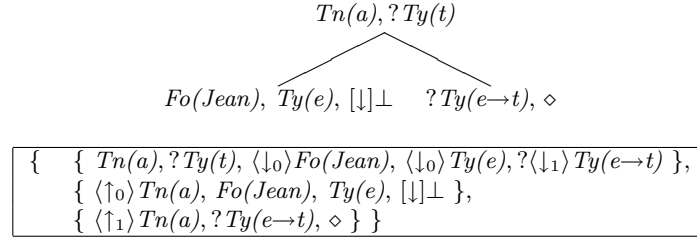


FIG. 67 – Complétion de l'arbre en DS.

**Transcription de la description d'arbre en GP.** — Cette description d'arbre prend la forme de la fig. 68 en GP.

On y indique que l'Exigence du nœud  $\boxed{3}$  est satisfaite, et on ajoute à l'argument de  $\boxed{1}$  la forme correspondante. A l'issue de cette étape, le pointeur s'est déplacé sur le dernier nœud.

**Transcription de la règle de complétion en GP.** — La règle de complétion, selon les mêmes principes que les règles précédentes d'introduction et de prédiction, peut faire l'objet en GP d'une construction, qui prendra la forme de la fig. 69.

On y indique l'ajout de la FORME au niveau de la BRANCHE correspondante du NOEUD supérieur, ainsi que le déplacement du pointeur du nœud argument vers le foncteur.

### 2.2.6 Observation du mot suivant

Le mot suivant dans l'énoncé est *admire* ; l'entrée lexicale lui correspondant est donnée en fig. 70.

Cette entrée demande d'ajouter à l'emplacement du pointeur une Exigence pour  $? \langle \downarrow_0 \rangle Ty(e)$  (c'est-à-dire pour un argument de type  $e$ ), puis d'introduire un foncteur (make  $\langle \downarrow_1 \rangle$ ), de s'y déplacer (go  $\langle \downarrow_1 \rangle$ ) et d'y ajouter les informations sur la forme correspondante ( $Fo(admire)$ ), le type ( $Ty(e \rightarrow (e \rightarrow t))$ ) et l'indication qu'il s'agit d'un nœud terminal ( $[\downarrow]\perp$ ) ; enfin, de remonter au nœud supérieur (go  $\langle \uparrow \rangle$ ).

<sup>81</sup>Pour des raisons de lisibilité, on peut alléger cette notation en omettant l'un des deux. L'information est alors incomplète, mais plus lisible.

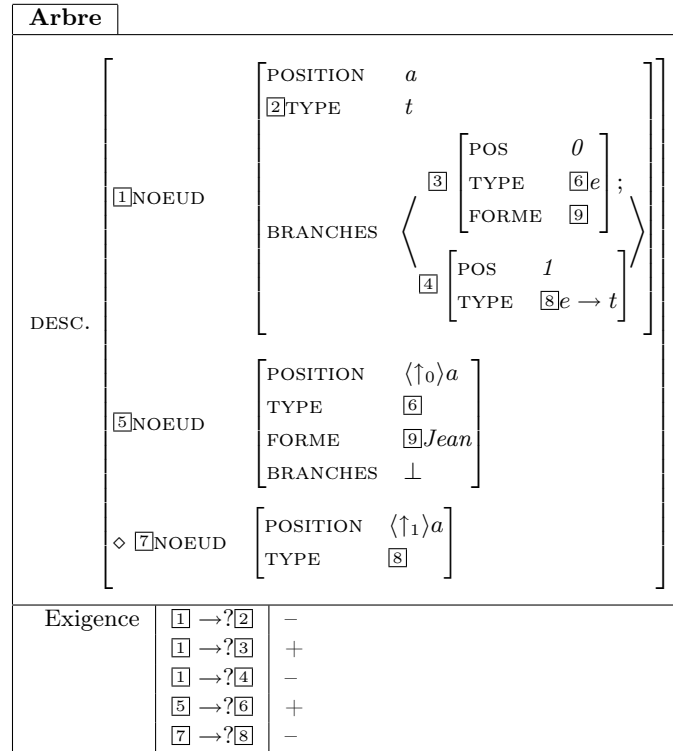


FIG. 68 – Description de l'arbre à l'issue de la cinquième étape, en GP.

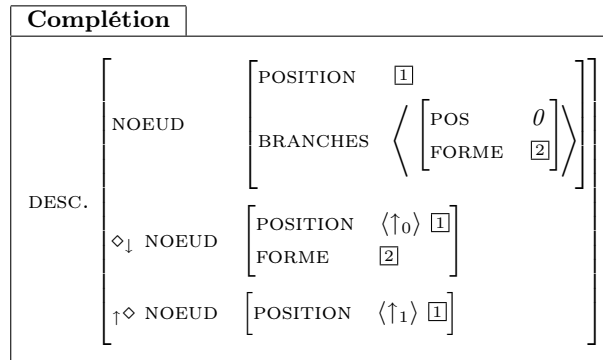


FIG. 69 – La construction figurant la règle de complétion, en GP.

<i>admire</i>		IF $?Ty(e \rightarrow t)$ , THEN   put $? \langle \downarrow_0 \rangle Ty(e)$ ; make $\langle \downarrow_1 \rangle$ ; go $\langle \downarrow_1 \rangle$ ; add $Fo(admire)$ , $Ty(e \rightarrow (e \rightarrow t))$ , $[\downarrow] \perp$ ; go $\langle \uparrow \rangle$ , ELSE    ABORT
---------------	--	---

FIG. 70 – L'entrée lexicale *admire* en DS.

A l'issue de cela, la description d'arbre prend la forme de la fig. 71.

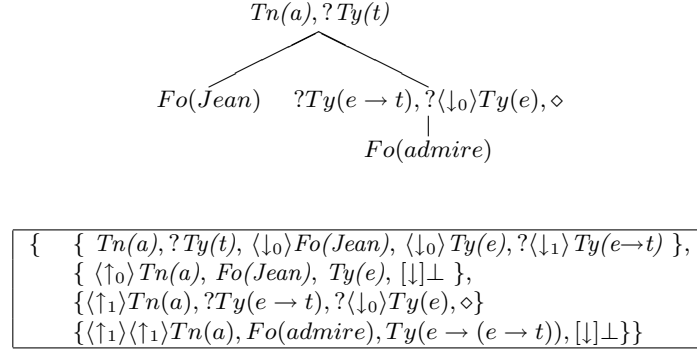
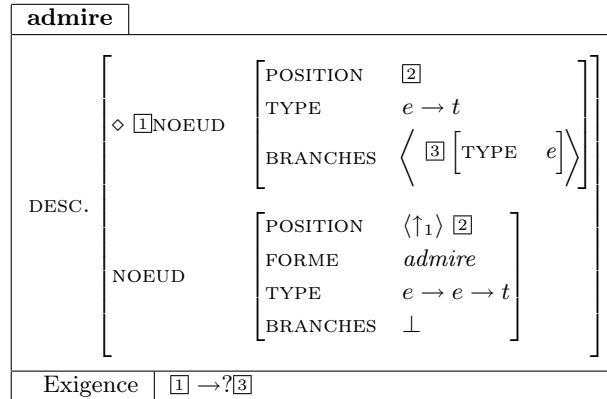


FIG. 71 – Description d'arbre à l'issue de l'étape 6, en DS.

Notons que l'ordre des objets dans l'arbre ne correspond pas, d'après ce qu'on peut voir ici, à l'ordre syntaxique des mots dans la phrase : il dépend du statut des noeuds (argument ou foncteur) et non pas de leur linéarisation : *Pierre* correspondra à la branche de gauche et *admire* à la branche de droite du nœud qui les régit.

**Transcription de l'entrée *admire* en GP.** — Selon les mêmes conventions que précédemment, l'entrée lexicale correspondant à *admire* prendra en GP la forme de la fig. 72.


 FIG. 72 – L'entrée lexicale *admire* en GP.

Elle indique la création d'un nouveau noeud (indexé [4] ici) qui porte les indications données dans l'instruction THEN de l'entrée correspondante en DS (concernant sa FORME, son TYPE et son absence de BRANCHES). Elle ajoute également au noeud sur lequel porte le pointeur la BRANCHE correspondant au nouveau NOEUD créé.

**Transcription de la description d'arbre en GP.** — L'équivalent de la description d'arbre en GP est donné en fig. 73.

On a y ajouté le nouveau nœud correspondant à *admire* que l'on a indexé [11], et on a également ajouté (dans les informations intrinsèques et dans les extrinsèques) l'argument exigé par le nœud indexé [7], qui doit être de type  $e$ .

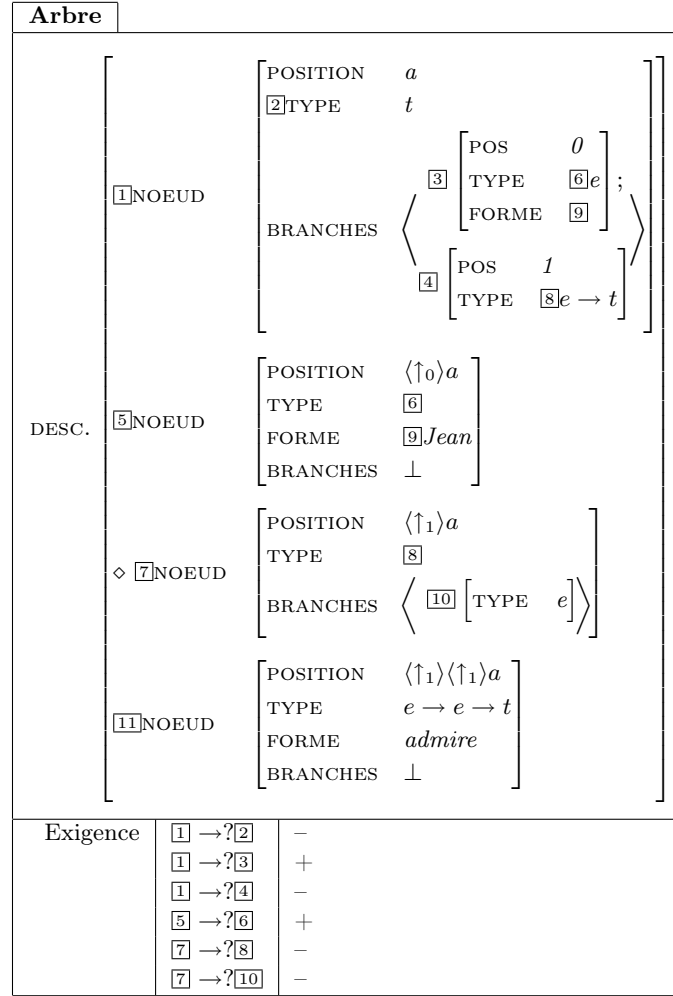


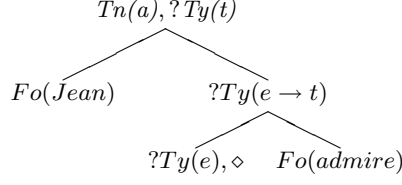
FIG. 73 – Description de l'arbre à l'issue de la sixième étape, en GP.

### 2.2.7 Prédiction

L'application à ce stade de la règle de prédiction permet de construire le noeud correspondant à l'argument que l'on a introduit sous l'index [10] dans l'étape précédente,  $?Ty(e)$ . Le résultat de cette étape correspond à la fig. 74. Le pointeur est déplacé sur ce noeud nouvellement introduit.

Notons que le résultat produit par l'application de cette règle et de la précédente aurait pu être obtenu en effectuant une introduction, puis la prédiction des deux noeuds correspondants, et enfin l'observation de l'entrée *admire* (à laquelle aurait alors correspondu un autre ensemble d'instructions). D'après les auteurs, cette variante de ce que l'on avait vu lors des premières étapes de l'analyse provient de l'organisation linguistique, et varie d'une langue à l'autre, ce qui demande de faire des choix quant à l'introduction d'informations dans le lexique ou dans les principes généraux <sup>82</sup>.

<sup>82</sup>Et dans notre cas, comme nous suivons un exemple anglais, nous nous sommes basée sur ce qui était fait



$$\boxed{
 \begin{array}{l}
 \{ \{ Tn(a), ?Ty(t), \langle \downarrow_0 \rangle Fo(Jean), \langle \downarrow_0 \rangle Ty(e), ?\langle \downarrow_1 \rangle Ty(e \rightarrow t) \}, \\
 \{ \langle \uparrow_0 \rangle Tn(a), Fo(Jean), Ty(e), [\downarrow] \perp \}, \\
 \{ \langle \uparrow_1 \rangle Tn(a), ?Ty(e \rightarrow t), ?\langle \downarrow_0 \rangle Ty(e) \} \\
 \{ \langle \uparrow_0 \rangle \langle \uparrow_1 \rangle Tn(a), ?Ty(e), \diamond \} \\
 \{ \langle \uparrow_1 \rangle \langle \uparrow_1 \rangle Tn(a), Fo(admire), Ty(e \rightarrow (e \rightarrow t)), [\downarrow] \perp \} \}
 \end{array}
 }$$

FIG. 74 – Description d'arbre à l'issue de la septième étape, en DS.

**Transcription de la description d'arbre en GP.** — L'équivalent de la figure précédente en GP est donné en fig. 75.

Cette construction est le résultat de l'unification de l'arbre de obtenu lors de l'étape précédente et de la construction “Pred-a” (prédiction d'un argument) que l'on avait proposée en fig. 62, et qui cette fois a été appliquée à la position du pointeur sur le noeud indexé  $\boxed{7}$  et non sur le  $\boxed{1}$  comme la première fois : le noeud  $\boxed{11}$  a été créé et porte les informations de TYPE et de POSITION relativement aux informations qui figuraient sur le noeud d'où est parti le pointeur, lequel s'est finalement déplacé sur le nouveau noeud, et une nouvelle Exigence est ajoutée (celle de  $\boxed{12} \rightarrow ?\boxed{13}$ ).

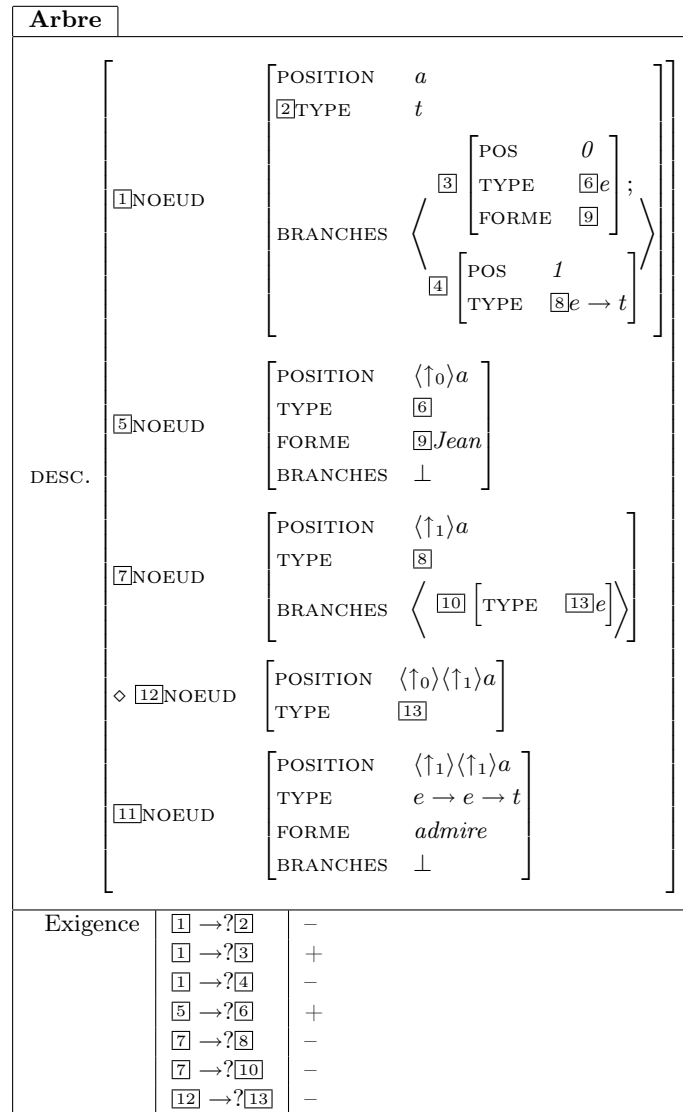


FIG. 75 – Description de l'arbre à l'issue de la septième étape, en GP.

### 2.2.8 Observation du mot suivant

**Description.** — On observe alors le mot suivant, *Pierre*. Celui-ci correspond à l'entrée suivante dans le lexique (fig. 76) :

<i>Pierre</i>	IF	$?Ty(e),$
	THEN	$add\ Fo(Pierre), Ty(e), [\downarrow]\perp,$
	ELSE	ABORT

FIG. 76 – L'entrée lexicale *Pierre* en DS.

Elle porte les mêmes informations que l'entrée *Jean*, à la valeur de la forme près.

A l'issue de cette étape, la description d'arbre est de la forme de la fig. 77.

$$\{ \{ Tn(a), ?Ty(t), \langle \downarrow_0 \rangle Fo(Jean), \langle \downarrow_0 \rangle Ty(e), ?\langle \downarrow_1 \rangle Ty(e \rightarrow t) \}, \\ \{ \langle \uparrow_0 \rangle Tn(a), Fo(Jean), Ty(e), [\downarrow] \perp \}, \\ \{ \langle \uparrow_1 \rangle Tn(a), ?Ty(e \rightarrow t), ?\langle \downarrow_0 \rangle Ty(e) \} \\ \{ \langle \uparrow_0 \rangle \langle \uparrow_1 \rangle Tn(a), Fo(Pierre), Ty(e), [\downarrow] \perp, \diamond \} \\ \{ \langle \uparrow_1 \rangle \langle \uparrow_1 \rangle Tn(a), Fo(admire), Ty(e \rightarrow (e \rightarrow t)), [\downarrow] \perp \} \}$$

FIG. 77 – Description d'arbre à l'issue de la huitième étape, en DS.

**Transcription de la description d'arbre en GP.** — L'entrée lexicale de *Pierre* en GP est la même que celle de *Jean* présentée en figure 66, à la valeur de la FORME près : on ne la reproduit donc pas ici. Quand elle fait l'objet d'une unification avec la description d'arbre obtenue lors de l'étape précédente, on obtient la construction présentée en fig. 78.

L'Exigence du type  $e$  pour l'index  $\boxed{13}$  est satisfaite par l'entrée *Pierre*, et les informations de cette entrée sont ajoutées à la description d'arbre.

### 2.2.9 Evaluation de l'arbre

Maintenant que les noeuds terminaux de l'arbre ont été annotés, on en vient à évaluer l'arbre. Cela consiste à le reparcourir afin de voir si toutes les exigences sont satisfaites, ou non. Du point de vue du noeud représentant le prédicat, en nous basant sur les informations des deux noeuds qui en dépendent, on peut conclure que son argument est  $\langle \downarrow_0 \rangle Fo(Pierre)$  et son foncteur  $\langle \downarrow_1 \rangle Fo(admire)$ .

A partir de là, ces informations peuvent être mises en relation en vertu d'une règle d'élimination (*elimination rule*) qui est une sorte de  $\beta$ -réduction. Cette règle permet, en l'occurrence, d'effectuer les implications suivantes :

$$[\langle \downarrow_0 \rangle Fo(Pierre) \wedge \langle \downarrow_1 \rangle Fo(admire)] \rightarrow Fo(admire(Pierre)) \\ [\langle \downarrow_0 \rangle Ty(e) \wedge \langle \downarrow_1 \rangle Ty(e \rightarrow (e \rightarrow t))] \rightarrow Ty(e \rightarrow t)$$

Le même procédé de complétion et d'élimination permet d'évaluer le noeud racine de l'arbre comme  $Fo(admire(Pierre))(Jean), Ty(t)$ .

Une fois ceci effectué, alors il est possible d'évaluer le noeud ancre de l'arbre comme en étant la racine, en remplaçant  $Tn(a)$  par  $Tn(0)$ , ce qui équivaut à  $[\uparrow] \perp$ . Cette modification peut être propagée aux branches de l'arbre, où p.ex.  $\langle \uparrow_0 \rangle Tn(a)$  exprimé relativement à  $a$  devient  $Tn(00)$  en notation "absolue", etc. La description de l'arbre a alors la forme de la fig. 79.

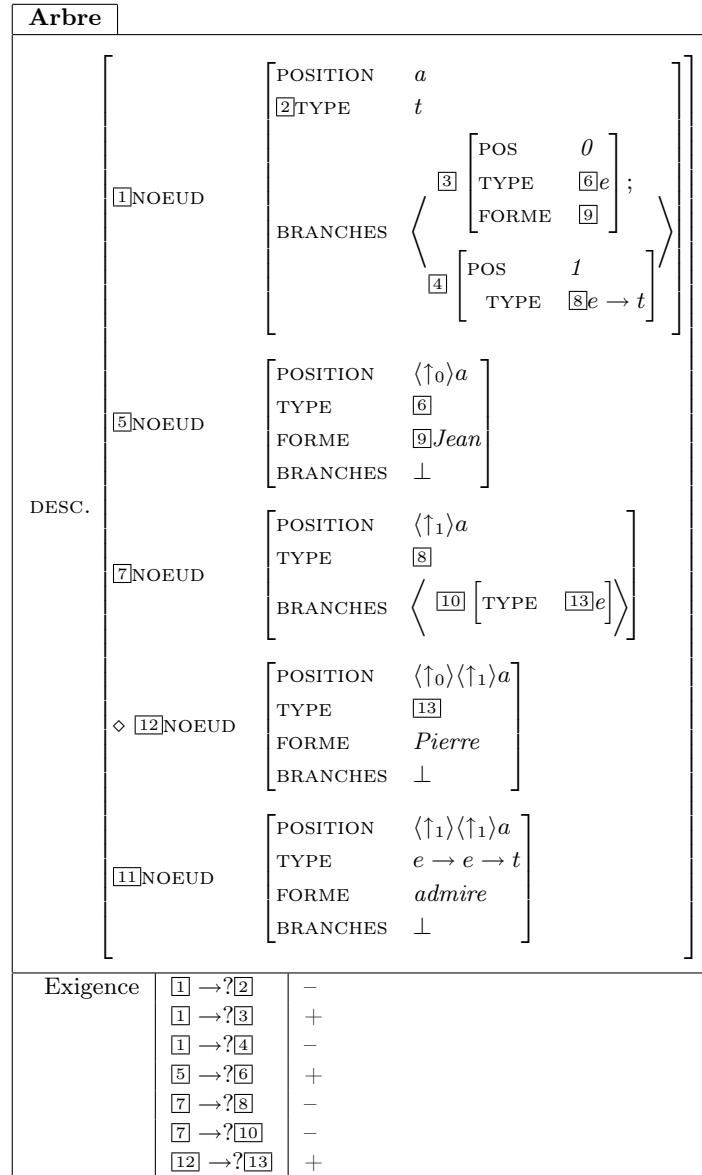


FIG. 78 – Description de l'arbre à l'issue de la huitième étape, en GP.

$$\begin{aligned}
 \{ & \{Tn(0), Ty(t), Fo(admire(Pierre)(Jean)), \diamond\}, \\
 & \{Tn(00), Ty(e), Fo(Jean), [\downarrow] \perp\}, \\
 & \{Tn(01), Ty(e \rightarrow t), Fo(admire(Pierre))\} \\
 & \{Tn(001), Fo(Pierre), Ty(e), [\downarrow] \perp\} \\
 & \{Tn(011), Fo(admire), Ty(e \rightarrow (e \rightarrow t)), [\downarrow] \perp\}
 \end{aligned}$$

FIG. 79 – Description d'arbre à l'issue de la neuvième étape, en DS.

**Transcription de la description d'arbre en GP.** — L'équivalent de la figure précédente en GP est donné en fig. 80.



Arbre		
DESC.	◇ [1]NOEUD	$\left[ \begin{array}{ll} \text{POSITION} & 0 \\ \text{[2]TYPE} & t \\ \text{FORME} & \textit{admire}(\textit{Pierre})(\textit{Jean}) \end{array} \right]$
	[5]NOEUD	$\left[ \begin{array}{ll} \text{POSITION} & 00 \\ \text{TYPE} & \text{[6]}e \\ \text{FORME} & \text{[9]}Jean \\ \text{BRANCHES} & \perp \end{array} \right]$
	[7]NOEUD	$\left[ \begin{array}{ll} \text{POSITION} & 01 \\ \text{TYPE} & \text{[8]}e \rightarrow t \\ \text{FORME} & \textit{admire}(\textit{Pierre}) \end{array} \right]$
	[12]NOEUD	$\left[ \begin{array}{ll} \text{POSITION} & 001 \\ \text{TYPE} & \text{[13]}e \\ \text{FORME} & \textit{Pierre} \\ \text{BRANCHES} & \perp \end{array} \right]$
	[10]NOEUD	$\left[ \begin{array}{ll} \text{POSITION} & 011 \\ \text{TYPE} & e \rightarrow (e \rightarrow t) \\ \text{FORME} & \textit{admire} \\ \text{BRANCHES} & \perp \end{array} \right]$
Exigence		
	[1] → ?[2]	+
	[1] → ?[5]	+
	[1] → ?[7]	+
	[5] → ?[6]	+
	[7] → ?[8]	+
	[7] → ?[12]	+
	[12] → ?[13]	+

FIG. 80 – Description de l'arbre à l'issue de la neuvième étape, en GP.

La réduction des types du noeud représentant la prédicat, i.e.

$$[\langle \downarrow_0 \rangle Ty(e) \wedge \langle \downarrow_1 \rangle Ty(e \rightarrow (e \rightarrow t))] \rightarrow Ty(e \rightarrow t)$$

conduit à satisfaire l'Exigence [7] → ?[8]. De même, la réduction des types du noeud ancre, i.e.

$$[\langle \downarrow_0 \rangle Ty(e) \wedge \langle \downarrow_1 \rangle Ty(e \rightarrow t)] \rightarrow Ty(t)$$

satisfait l'Exigence [1] → ?[2].

Les réductions de formes, quant à elles, permettent de satisfaire les Exigences [1] → ?[3], [1] → ?[4] et [7] → ?[10]. Plus exactement, lors de cette étape les index des arguments en foncteurs des noeuds non-terminaux sont remplacés par ceux des noeuds auxquels ils font référence : [3] est coindexé avec [5], [4] avec [7], et [10] avec [12].

Par la suite, la valeur relative de POSITION du noeud ancre est transformée en valeur absolue, et cette modification est répercutée sur les valeurs relatives de position des autres noeuds. Ceci n'est pas opéré par l'unification, mais par une modification de la valeur initiale  $a$ .

**Transcription de l'évaluation.** — Cette évaluation ne relève pas d'une unification comme cela a été le cas jusqu'ici pendant l'analyse, mais d'un mécanisme différent. Elle ne fait donc pas l'objet d'une construction comme avaient pu le faire les règles d'introduction, de prédiction et de complétion. Il s'agit plutôt ici d'un parcours des informations intrinsèques de la construction afin de mettre à jour les informations extrinsèques.

En outre, dans les informations intrinsèques les valeurs de POSITION sont modifiées dans leur expression, l'attribut BRANCHES disparaît suite à la coindexation de ses éléments constitutifs avec des noeuds de l'arbre, et celui de FORME apparaît pour les noeuds non-terminaux (et il fait référence au contenu des FORMES des noeuds dépendants). La règle d'élimination et donc constituée d'un ensemble d'instructions qui lui est très spécifique et n'a que peu de rapport avec le reste de l'analyse. D'ailleurs, il s'agit de la seule règle que l'on ne peut appliquer qu'une fois lors d'une analyse : en fin d'énoncé, ce qui souligne bien son caractère spécial.

Elle représente typiquement un ensemble d'actions qui n'ont pas été prévues dans le formalisme GP. Ceci n'empêche pas la formalisation de son résultat en GP, mais les mécanismes permettant de passer de la figure 78 à la figure 80 relèvent de possibilités d'action qui restent à définir intégralement.

## 2.3 Synthèse et conclusions

Dans cette deuxième section, on a comparé GP avec un formalisme catégoriel et procédural, celui de *Dynamic Syntax*, en nous basant sur un exemple d'analyse pas à pas en DS et en tentant de transcrire au fur et à mesure les informations en GP. Ceci nous a permis de mettre en évidence certaines des différences essentielles entre les deux formalismes, et qui constituent autant de limitations des échanges entre l'un et l'autre.

On a vu tout d'abord que la représentation des informations issues de l'analyse de DS sont formalisables en GP : on peut y représenter des descriptions d'arbres, où dans une construction la description elle-même est fournie dans les informations intrinsèques et les éléments exigés dans les informations extrinsèques.

On a vu ensuite que la majorité des étapes d'analyse en DS (mis à part la dernière) peuvent relever, en ce qui concerne GP, non pas d'une caractérisation comme c'est le cas normalement pour une analyse, mais d'une unification entre la structure constituant le résultat progressivement complété, et une construction de la grammaire. On a vu que les règles de DS (introduction, prédiction, complétion) peuvent faire l'objet de constructions grammaticales et donc être utilisées comme telles pour l'analyse.

On a vu également que les objets lexicaux de DS peuvent être eux aussi représentés en GP sous forme de constructions. Celles-ci, lors de l'analyse, sont elles aussi combinées par unification à la construction résultante, tout comme le sont les constructions représentant les règles. On a vu, à travers cela, qu'il est possible de représenter l'équivalent d'informations procédurales dans des constructions GP : en fait, ce ne sont pas les procédures que l'on a représentées, mais leur résultat ; la procédure est représentée par la conjonction entre les informations que l'on fournit dans les entrées lexicales d'une part, et le résultat de l'unification d'autre part.

A partir de ces deux dernières constatations, on peut imaginer qu’une grammaire GP représentant une DS porterait deux racines (elle aurait donc deux hiérarchies parallèles en quelque sorte), la première étant l’origine des constructions correspondant aux règles, et la seconde celle correspondant aux entrées lexicales. Puisqu’elles sont utilisées de la même façon et font l’objet de la même représentation, elles peuvent donc être toutefois représentées dans une même ressource.

On a également vu que l’une des différences essentielles entre GP et DS ne vient pas du contenu des objets, mais du déroulement de l’analyse : GP est un formalisme prévu pour représenter des analyses ascendantes, alors que DS effectue des analyses descendantes. Sans que GP dans son état actuel ne permette donc réellement de formaliser une DS, on peut toutefois constater que dans le cadre d’une analyse descendante, le flux des informations entre les deux blocs d’informations de la construction s’effectuent dans le sens inverse de celui d’une analyse ascendante : dans le dernier cas les propriétés font remonter les informations dans le bloc intrinsèque, alors que dans une analyse descendante c’est le bloc intrinsèque qui fournit les éléments permettant l’évaluation des propriétés extrinsèques.

Au delà du fait que ce parallélisme d’orientation entre le déroulement de l’analyse et la diffusion des informations au sein des constructions puisse être assez intéressant pour être noté, cela montre une dans ce cas, les notions d’informations “intrinsèques” et “extrinsèques” peuvent être remises en question. Non pas du simple fait de la différence dans l’orientation du flux d’information, mais parce que dans le cas de l’utilisation de GP que l’on vient de faire, le bloc supérieur ne fait plus réellement référence à des informations proprement intrinsèques (mais à une description d’arbre), ni le bloc inférieur à des informations proprement extrinsèques (mais à des exigences de la part de la description d’arbre). On peut en conclure que la différence entre GP et DS ne se situe pas uniquement en termes de déroulement d’analyse, mais plus exactement en termes d’organisation générale des informations.

### 3 Une Métagrammaire en GP

Le dernière section de ce chapitre s’intéresse à la comparaison en GP et les Métagrammaires. Une *Métagrammaire* <sup>83</sup> (ci-après MG) est un formalisme représentant l’information grammaticale sous la forme factorisée d’une hiérarchie d’héritage de descriptions arborescentes (des fragments d’arbres) représentant des généralisations linguistiques. Elle ne constitue pas en soi une grammaire à proprement parler, mais un ensemble de fragments que l’on combine pour construire une grammaire.

Cette approche est utilisée pour représenter principalement des grammaires TAG et LFG, afin d’en permettre une implémentation efficace comme par exemple la Métagrammaire de [Candito, 1999], LexOrg de [Xia, 2001], le compilateur MCG de [Gaiffe *et al.*, 2001], le système DyaLOG de [Thomasset & de la Clergerie, 2005] ou le compilateur XMG de [Debusmann *et al.*, 2004].

---

<sup>83</sup>Le terme est introduit par [Candito, 1999], qui s’est basée sur les travaux de [Vijay-Shanker & Schabès, 1992], [Rogers & Vijay-Shanker, 1994].

Ce qui nous intéresse ici n'est pas de comparer les différents systèmes de compilation de métagrammaires avec GP, mais d'essayer de mettre en évidence quelles sont les contraintes théoriques et formelles imposées par l'utilisation d'une MG à la description grammaticale d'une part, et quels sont les points communs et différences entre GP et MG d'autre part. La question ici n'est pas de savoir s'il est possible de représenter une TAG en GP, ni s'il est possible de factoriser une GP en MG ; elle est de savoir si l'on pourrait, par exemple, représenter une MG en une sorte de méta-GP, c'est-à-dire utiliser GP comme une métagrammaire.

Pour réaliser cette première tentative de comparaison <sup>84</sup> entre GP et MG nous nous sommes basée sur la proposition de [Crabbé, 2005b] concernant TAG. Quand on fera référence aux MG dans ce qui suit, on fera donc implicitement référence aux propositions de cet auteur précis, bien que d'autres approches existent (telles que celles que l'on a citées plus haut).

### 3.1 Caractéristiques des MG

Les MG ont pour objectif principal de permettre une **généralisation des informations**, afin d'éviter les redites ; en d'autres termes, elles visent à **factoriser** le plus possible l'information linguistique et de permettre la **réutilisation** de ces éléments factorisés en plusieurs lieux de la grammaire. Dans ce cadre et d'après l'approche à laquelle on s'intéresse ici, on peut factoriser les informations selon deux axes :

- Un axe de **structure** où l'on réutilise des *descriptions grammaticales partielles* (en d'autres termes, des fragments de grammaire) que l'on organise en une *hiérarchie d'héritage*.
- Un axe d'**alternatives** <sup>85</sup>, au même titre que les *transformations* de Chomsky et les *règles lexicales* des grammaires d'unification, afin de lier les représentations alternatives d'une même notion (e.g. alternance actif / passif). Ces alternatives sont représentées à l'aide de *conjonctions* et de *disjonctions*.

On représente donc dans une MG des fragments de descriptions grammaticales, fragments que l'on organise les uns par rapport aux autres au sein de deux systèmes parallèles (celui des structures et celui des alternatives).

On sait déjà qu'il serait impossible d'exprimer GP en une métagrammaire telle que celles qui ont été actuellement définies, parce que celles-ci sont étudiées pour représenter des grammaires lexicalisées <sup>86</sup> ; or les capacités de GP lui rendent possible de représenter des modèles non lexicalisés (comme c'est le cas de notre grammaire). Ce que l'on va chercher à voir, à l'inverse, c'est s'il est possible d'utiliser une GP pour représenter de telles généralisations, autrement

---

<sup>84</sup>Il est bien évident que les quelques points que l'on abordera ici ne sont que des pistes à étudier avec plus d'approfondissement, et ne sont pas les seules possibles. On ne fait ici que montrer quelques caractéristiques comparées des deux formalismes.

<sup>85</sup>Toutes les MG n'offrent pas ce second axe de factorisation. Et c'est parce qu'elle le propose que l'on s'est intéressée à cette approche en particulier, afin de voir comment GP pourrait représenter ce type de relation, que l'on n'utilise pas dans notre modèle.

<sup>86</sup>Cela étant, ceci vient du fait que les MG déjà définies n'ont travaillé qu'à représenter des modèles lexicalisés ; il serait intéressant de voir s'il serait en réalité possible d'étendre un formalisme de MG à des modèles non-lexicalisés.

dit d'utiliser GP comme formalisme méta-grammatical (et donc ainsi d'élargir les capacités de formalisation des métagrammaires à des modèles non-lexicalisés, notamment). A cette fin l'on commence par observer comment représenter des fragments de descriptions grammaticales en GP (§3.2), puis on voit comment ceux-ci peuvent être organisés d'abord selon l'axe des structures (§3.3) puis selon l'axe des alternatives (§3.4).

### 3.2 Représentation de fragments d'arbres

**Principe.** — Il s'agit de factoriser l'expression des constructions récurrentes, en représentant indépendamment des fragments d'arbres que l'on combine par la suite, et en permettant à ces fragments d'être organisés selon une hiérarchie d'héritage (telle que définie dans [Flickinger, 1987]). Par exemple, le "sujet en position canonique" fait l'objet d'une généralisation en MG (fig. 81).

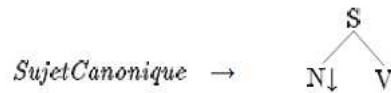


FIG. 81 — Représentation du sujet en position canonique selon [Crabbé, 2005].

**Représentation en GP.** — Les informations factorisées correspondent à un ensemble variable d'informations intrinsèques et extrinsèques d'une construction en GP. D'après l'exemple de la fig. 81 on peut penser que l'ensemble d'informations factorisé en une construction "sujet en position canonique" est le suivant :

- *Identifiant.* La construction se voit affecter une étiquette, qui est ici *SujetCanonique*.
- *Informations intrinsèques.* La construction représente un fragment d'arbre syntagmatique, qui représente un SYNTAGME dont on a la CATÉGORIE de la RACINE (S) et celles des CONSTITUANTS IMMÉDIATS (N et V). On sait du premier constituant qu'il est lui-même un SYNTAGME puisqu'il porte une ↓ ce qui signifie qu'il est un "nœud de substitution".
- *Informations extrinsèques.* La construction fait référence à trois objets : un premier de catégorie S, et deux autres qui en sont les constituants immédiats et qui sont de catégorie N et V. On introduit donc un type de propriété DI (pour "dominance immédiate"), qui est binaire et mentionne la règle DI correspondant à ce fragment d'arbre, à savoir  $S \triangleleft N, V$ . Une propriété DI est toujours évaluée, et satisfaite si les deux termes sont disponibles. Elle n'est ni réflexive ni commutative, et sa capacité est de 1 puisqu'elle ne peut être évaluée qu'une fois pour une construction donnée (cf. fig 82).

On sait aussi grâce à cet arbre que les deux objets N et V sont liés par une contrainte de précedence immédiate que l'on introduit sous la forme d'un type PL (pour "précedence linéaire") et qui prendra la forme  $N \prec V$ . Celle-ci est toujours évaluée, et satisfaite si (les deux termes sont disponibles et si) l'index de fin de l'objet auquel le premier terme fait

<b>DI</b>	
Symbole	$\triangleleft$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Autorisée
Commutativité	Faux
Disponibilité	Vrai
Satisfaction	$\boxed{1} = \text{disponible} \wedge \boxed{2} = \text{disponible}$
Capacité	1

FIG. 82 – Spécification du type Dominance Immédiate pour MG.

référence est égal à l'index de début de l'objet auquel le second terme fait référence. Elle n'est elle non plus ni commutative ni réflexive, et a une capacité de 1 (cf. fig. 83).

<b>PL</b>	
Symbole	$\prec$
Termes	$\boxed{1}\text{Construction} \wedge \boxed{2}\text{Construction}$
Réflexivité	Interdite
Commutativité	Faux
Disponibilité	Vrai
Satisfaction	$(\boxed{1} = \text{disponible}) \wedge (\boxed{2} = \text{disponible}) \wedge (\boxed{1}.\text{INDEX} \text{FIN} = \boxed{1}.\text{INDEX} \text{DÉBUT})$
Capacité	1

FIG. 83 – Spécification du type Précédence Linéaire pour MG.

Rien n'est précisé à propos de l'accord entre sujet et verbe dans le fragment d'arbre : on n'en dit donc rien non plus en GP.

Deux caractéristiques de ces types viennent des propriétés de base des arbres syntagmatiques : la première est leur capacité de 1 (chaque propriété d'un type donné ne peut être évaluée qu'une fois parce que ne peut porter que sur un objet ou une relation entre objets), la seconde leur condition de disponibilité toujours vraie (puisque par définition, dans une grammaire syntagmatique générative, toutes les conditions demandées dans la grammaire sont requises pour qu'un objet soit introduit).

Pour représenter le "sujet en position canonique", on peut donc proposer une sorte de "méta-construction" GP qui prendrait la forme de la fig. 84.

<b>SujetCanonique</b>	
SYNT	$\left[ \begin{array}{l} \text{RACINE} \mid \text{CAT} \quad \boxed{1} \\ \text{C.I.} \quad \left\langle \begin{array}{l} \text{SYNT} \left[ \text{RACINE} \mid \text{CAT} \quad \boxed{2} \right], \\ \left[ \text{CAT} \quad \boxed{3} \right] \end{array} \right\rangle \end{array} \right]$
DI	$\boxed{1}\text{S} \triangleleft (\boxed{2}\text{N} \wedge \boxed{3}\text{V})$
PL	$\boxed{2} \prec \boxed{3}$

FIG. 84 – Représentation de l'équivalent de la figure 81 en GP.

S'agissant d'une "méta-construction" elle est par définition fragmentaire, et c'est la raison pour

laquelle elle manque d'information ; il en est de même pour le *SujetCanonique* en MG. Ce sont ces fragments qui sont organisés selon les deux axes que l'on a mentionnés plus haut, afin de proposer des généralisations de structure (§3.3) et d'alternatives (§3.4).

### 3.3 Généralisations de structure

#### 3.3.1 Hiérarchie d'héritage

On a vu comment représenter un fragment de grammaire en GP ; voyons maintenant comment l'on peut représenter l'héritage des fragments entre eux.

**Principe.** — On vient de voir le fragment correspondant au *SujetCanonique* ; dans l'axe de structure, celui-ci hérite de la classe *ArgumentVerbal* dont deux autres représentants sont le *SujetRelativisé* et l'*ObjetCanonique* représentés en fig. 85.

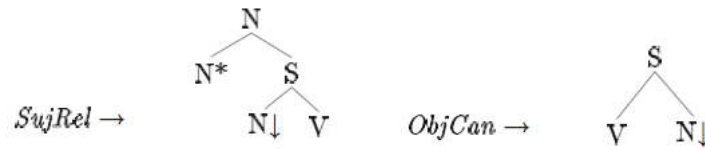


FIG. 85 — Représentation du Sujet Relativisé et de l'Objet Canonique dans la MG de [Crabbé, 2005].

La hiérarchie d'héritage (incomplète ici, pour les besoins de l'exemple) se présente sous la forme de la fig. 86.

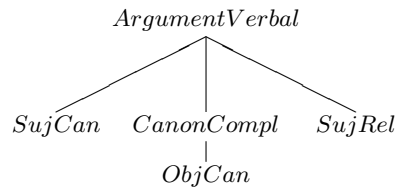


FIG. 86 — Extrait de la hiérarchie des Arguments Verbaux dans la MG de [Crabbé, 2005].

**Représentation en GP.** — On peut tout à fait, pour représenter ceci, utiliser la hiérarchie d'héritage existant en GP. On aurait alors pour les deux objets que l'on vient de présenter, en nous basant sur ce qu'on la posé précédemment et en y ajoutant l'information d'héritage, les “méta-constructions” de la figure 87.

On a ici, dans le fragment correspondant au *SujRel*, représenté l'information disant que le premier constituant immédiat du *N* était un “nœud pied” (*N\**), en posant qu'il s'agit d'un SYNTAGME ; en effet, selon la définition que l'on peut en trouver dans [Abeillé, 2002, p. 25],

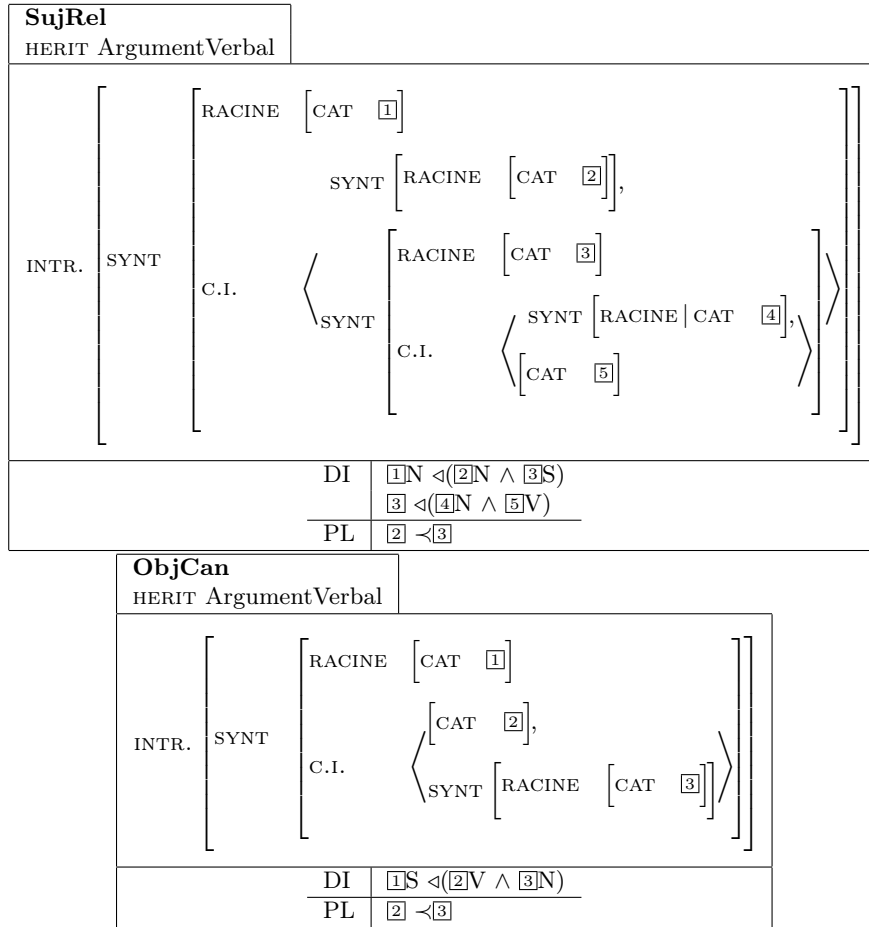


FIG. 87 – Représentation de l'équivalent de la figure 85 en GP.

*Les arbres auxiliaires ont en outre un nœud feuille (appelé “nœud pied”) étiqueté par un non terminal de même catégorie que celle de leur nœud racine.*

On peut remarquer qu'en l'espèce, l'héritage d'*ArgumentVerbal* n'est pas d'une grande utilité en matière de factorisation d'information puisque rien n'est généralisable à son niveau : on se contente de rassembler les fragments, sans chercher à généraliser les informations qu'ils contiennent. La conséquence de cela est que seuls les nœuds terminaux de la hiérarchie portent des informations.

### 3.3.2 Construction d'arbres à partir des fragments

On a vu comment représenter des fragments d'arbres et comment les organiser en une hiérarchie de structure ; voyons à présent comment ceux-ci se combinent pour former des arbres, c'est-à-dire des objets grammaticaux.

**Principe.** — Pour construire des arbres grammaticaux à partir des fragments d'arbres, on



combine ces derniers avec des conjonctions (fig. 88).

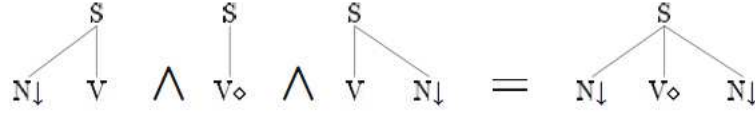


FIG. 88 – Combinaison de fragments de descriptions selon l’approche de [Crabbé, 2005].

**Représentation en GP.** — On peut penser que la conjonction utilisée par [Crabbé, 2005b] pour combiner les fragments n’est en fait pas une conjonction à proprement parler mais plutôt une unification : en effet, la structure résultante rassemble les informations des structures combinées exactement de la même façon qu’une unification l’aurait fait (notamment, les trois V des trois fragments d’arbres en constituent un seul dans l’arbre résultant, qui porte l’ensemble des informations des trois). En ce cas, on peut tout à fait envisager le même mécanisme pour la combinaison des “méta-constructions” GP (fig. 89).

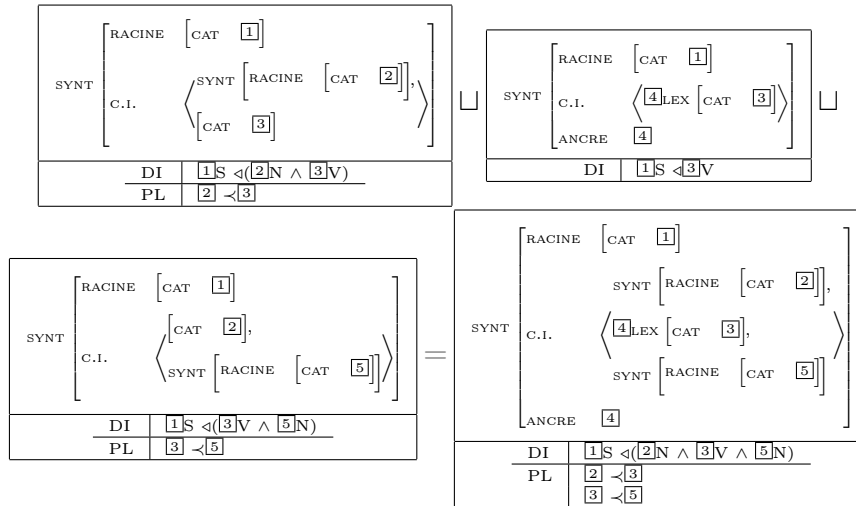


FIG. 89 – Représentation de l’équivalent de la figure 88 en GP.

Notons que dans cette figure, on a représenté l’information  $V\Diamond$  du deuxième fragment d’arbre, représentant un “nœud tête”, comme indiquant que l’objet ainsi pointé correspond à l’ANCRE de l’arbre (i.e. à un objet LEXICAL et non à un SYNTAGME).

### 3.4 Généralisations d’alternatives

#### 3.4.1 Organisation des fragments en alternatives

**Principe.** — Le second axe de factorisation des informations de la MG de [Crabbé, 2005b] permet une généralisation des alternatives de constitution syntagmatique. Par exemple il per-

met de lier les deux fragments *SujCan* et *SujRel* comme étant deux réalisations possibles (combinées par des disjonctions) de la fonction *Sujet* :

$$Sujet \rightarrow SujCan \vee SujRel$$

**Représentation.** — Pour représenter cela en GP on peut également utiliser la représentation hiérarchique, qui posera que les deux fragments *SujCan* et *SujRel* héritent de la classe *Sujet* ; le principe même de la hiérarchie d'héritage aura pour conséquence qu'un *Sujet* pourra être soit un *SujRel*, soit un *SujCan*, mais pas les deux à la fois. On aurait donc une seconde hiérarchie d'héritage, parallèle à la première, de la forme de la fig. 90.

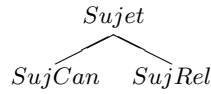


FIG. 90 – Interprétation des alternatives comme une hiérarchie d'héritage en GP.

Dans les “méta-constructions” correspondant aux fragments d'arbres, il faut alors mentionner non seulement l'héritage de la première hiérarchie, et celui de la seconde hiérarchie (fig 91). Pour de simples raisons de lisibilité on a appelé le premier héritage STRUC. et le second ALTERN., mais il ne s'agit que d'une convention de notation qui n'a pas d'incidence sur le mécanisme.

### 3.4.2 Construction de familles d'arbres

**Principe.** — La généralisation d'alternatives peut se faire en mettant en relation non seulement des fragments d'arbres comme on vient de le voir, mais également des classes d'alternatives elles-mêmes. Par exemple, la classe d'alternatives *Sujet* que l'on vient de présenter peut entrer dans l'expression d'une classe plus complexe, comme celle de *DiathèseTransitive*, où l'on indique le lien s'opérant entre la succession d'un *Sujet*, d'une *FormeActive* (verbale) et d'un *Objet* (qui sont tous trois des classes d'alternatives), et la succession d'un *Sujet*, d'une *FormePassive* (verbale) et d'un *ParObjet* (i.e., un objet indirect introduit par la préposition *par*), ainsi que celle d'un *Sujet* et d'une *FormePassive* seuls dans le cas d'un passif sans agent :

$$DiatheseTransitive \rightarrow (Sujet \wedge FormeActive \wedge Objet) \vee (Sujet \wedge FormePassive \wedge ParObjet) \vee (Sujet \wedge FormePassive)$$

**Représentation.** — Dans ce cas, et d'après ce que l'on a vu précédemment, la hiérarchie d'héritage représentant les alternatives se représente sous la forme de la fig. 92.

Dans cette figure on voit que ce qui fait l'objet d'une structuration hiérarchique sont les disjonctions, mais pas les conjonctions (c'est-à-dire ce qui lie  $(Sujet \wedge FormeActive \wedge Objet)$  à  $(Sujet \wedge FormePassive \wedge ParObjet)$ , mais pas ce qui lie *Sujet* à *FormeActive*). Les conjonctions, qu'elles rassemblent des classes d'alternatives comme ici, ou des fragments d'arbres

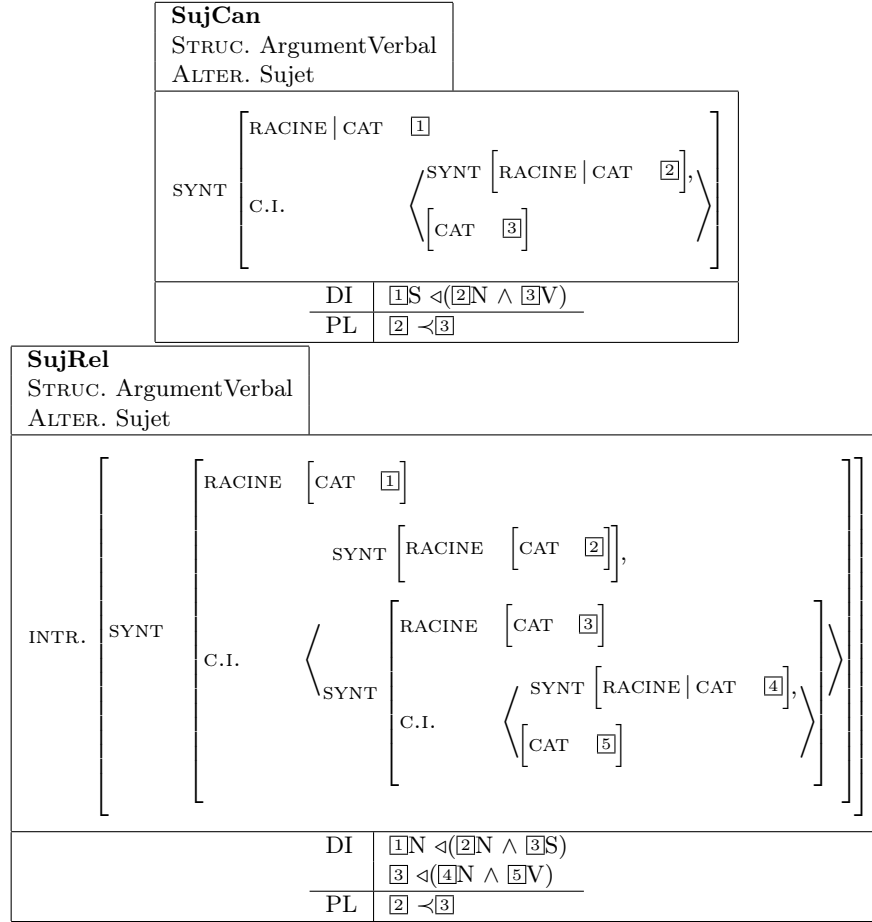


FIG. 91 – Représentation des “méta-constructions” avec double héritage.

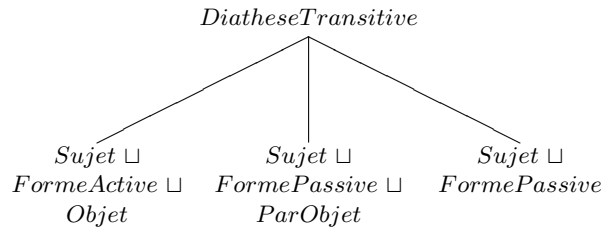


FIG. 92 – Interprétation des alternatives de combinaisons d’alternatives en GP.

comme on l’avait vu au §3.3.2, constituent des arbres qui sont le résultat de l’unification des structures correspondantes (d’où le remplacement des  $\wedge$  par des  $\sqcup$ ). En l’occurrence, puisqu’il s’agit ici de classes d’alternatives et non simplement de fragments d’arbres, le résultat de l’unification peut varier en fonction du fragment utilisé pour chaque alternative (par exemple, la classe *Sujet* pourra faire référence soit au fragment *SujCan* soit à *SujRel*).

Le même principe, et donc la même représentation, préside à l’expression des “familles” (c’est-à-dire des ensembles d’arbres qui représentent diverses réalisations possibles d’une même struc-

ture prédicat-argument). [Crabbé, 2005b] propose à titre d'exemple la famille :

$$FamilleDitransitive \rightarrow DiatheseTransitive \wedge ObjetIndirect$$

Dans ce cas,  $DiatheseTransitive \wedge ObjetIndirect$  correspondra en GP à l'unification de la classe d'alternatives représentée par  $DiatheseTransitive$  (que l'on vient de voir) et à celle correspondant à  $ObjetIndirect$  (pouvant, d'après l'auteur, prendre la forme d'un groupe prépositionnel, en position canonique  $-IndObjCan-$  ou extraite  $-IndObjWh-$ ).

### 3.5 Principes et Propriétés

On a fait remarquer plus haut que rien dans les fragments d'arbres n'était précisé quant aux contraintes d'accord entre le sujet et son verbe. Ceci vient du fait que la grammaire TAG qui est ici formalisée fait appel à une série de *Principes* et de *Propriétés* qui sont situées en-dehors de la grammaire elle-même, et où sont exprimées ces contraintes.

Or, on a dit précédemment que les Principes (et il en est de même pour les Propriétés), en GP, n'étaient pas exprimés en-dehors de la grammaire, mais plutôt dans les constructions de base de celle-ci, dont toutes les autres constructions de la grammaire héritent (ce qui a pour conséquence que ces contraintes s'appliquent à toutes les constructions, tout comme c'est le cas des principes, mais sans avoir à les stipuler en-dehors de la grammaire).

On peut exprimer l'équivalent des Principes, dans une "méta-GP", de deux façons alternatives :

- Soit, en plus des constructions produites par la combinaison des fragments de descriptions, on considère que la grammaire contient des constructions de base qui ne font pas l'objet d'une "méta-formalisation". Dans ce cas, on doit poser que les constructions issues de la combinaison des fragments d'arbres viendront s'y ajouter, en en héritant. Cette première possibilité serait plus proche de l'approche théorique du modèle TAG, mais s'éloignerait quelque peu 1/ de l'intérêt de la factorisation des informations grammaticales (puisque le but reste de mettre en commun le plus d'informations possible, or celles-ci sont très simples à mettre en commun puisqu'elles sont censées être générales à toutes les constructions de la grammaire), et 2/ des possibilités d'utilisation de GP en tant que métagrammaire, puisque cela demanderait d'ajouter un mécanisme supplémentaire à ceux déjà existants, pour créer une ressource en intégrant dans un héritage pré-existant des informations supplémentaires (et non pas de générer une grammaire complète à partir d'une structuration factorisée). Notons toutefois que c'est exactement ce que font les MG puisque c'est de cette façon que fonctionne l'application de Principes à une grammaire (cf. par exemple [Duchier *et al.*, 2005], qui mentionnent 3 principes : un principe de couleurs, un principe d'unicité et un principe de rang). Dans ce cas même, les principes font partie du méta-formalisme lui-même, ce qui constitue une contrainte théorique de taille pour un système de représentation.
- Soit l'on ne souhaite pas faire exister dans la grammaire d'autres constructions que celles faisant l'objet d'une factorisation, et qu'on ne souhaite pas introduire de nouvel élément dans GP, alors on peut exprimer les principes sous forme, eux aussi, de fragments de descriptions. C'est dans ce cas que l'on exploite pleinement les possibilités de GP, puisqu'il est alors tout à fait possible de mentionner des propriétés d'Accord, p.ex., dans des fragments d'arbres

consacrés aux principes (alors que dans une grammaire TAG — de même que dans une méta-TAG — on ne peut pas formuler les accords directement dans la grammaire parce que le formalisme n’est pas étudié pour rendre cela possible : dans la grammaire elle-même, on ne peut que donner des informations de dominance immédiate et de précédence linéaire <sup>87</sup>). Ceci ne demanderait pas de modifier les bases théoriques du modèle, mais uniquement la forme sous laquelle celles-ci sont exprimées.

D’ailleurs, on peut constater qu’une partie des Principes est déjà utilisée dans la spécification des deux types de propriétés MG que l’on a définies plus haut (DI et PL) : en effet, la capacité de 1 de ces propriétés relève en réalité du Principe d’unicité. On pourrait tout à fait imaginer ne pas contraindre la capacité des propriétés, mais ajouter un type Unicité qui ferait l’objet d’un fragment de description représentant un Principe, et qui constituerait la racine de la hiérarchie d’héritage des structures (donc, dont tous les fragments hériteraient). De cette façon, s’il advient que ce Principe doive ne pas être appliqué à un cas particulier (l’auteur donne à titre d’exemple le doublement du syntagme prépositionnel dans *Jean parle de Paul à Marie*) alors on pourrait imaginer une partie de la hiérarchie de structure qui n’hériterait pas de ce principe, et éviterait le recours qu’a l’auteur à un système coloration des nœuds qui simule leur polarisation (à la manière de [Perrier, 2003]).

### 3.6 Synthèse et conclusions

On a vu dans cette section qu’il est possible d’utiliser GP comme une “méta-grammaire”, c’est-à-dire comme un formalisme de représentation de la factorisation des descriptions grammaticales. Ceci ne demande le recours à aucun mécanisme ni mode de représentation autre que ce qui a été défini dans l’architecture du formalisme.

En revanche, ce qui est modifié, et c’est bien évident puisqu’on change de niveau de représentation (on représente ici une méta-grammaire et non plus une grammaire), c’est la façon dont les représentations et mécanismes s’organisent :

- Tout comme dans une GP, dans une “méta-GP” les objets (ici, des fragments d’arbres et des classes de fragments d’arbres) sont représentés sous forme de constructions. Ces constructions obéissent exactement à la même définition que celle que l’on a donnée au chapitre 2, et ici aussi héritent les unes des autres. Ce qui change ici (dans cet exemple précis mais pas nécessairement dans toutes les MG), c’est que l’héritage n’est pas simple, mais double : les constructions sont reliées entre elles d’une part selon un axe de structures, d’autre part selon un axe d’alternatives. Cela ne pose pas de problème au formalisme puisqu’il est tout à fait possible d’y représenter un héritage multiple.
- Une chose qui est modifiée <sup>88</sup>, c’est l’utilisation qui est faite de l’héritage : on a vu que dans la proposition de [Crabbé, 2005b] seuls les terminaux de la hiérarchie, les fragments d’arbres en l’occurrence (par opposition aux classes qui sont les non-terminaux), contiennent de l’in-

<sup>87</sup>Bien entendu ici on ne parle que des propriétés syntaxiques ; on pourrait y ajouter les informations sémantiques qui sont liées aux fragments d’arbres mais cela ne réglerait pas la question de la formulation des accords directement dans la grammaire et non via des principes qui y sont ajoutés.

<sup>88</sup>Ou tout au moins, différente de notre modèle.

formation linguistique. Les classes, nœuds non-terminaux, ne portent que des informations sur l'héritage lui-même. Cela ne modifie pas réellement l'utilisation de l'héritage, mais cela a pour conséquence que l'on n'utilise pas l'unification pour rassembler les informations de plusieurs constructions qui héritent les unes des autres. Plus exactement, on peut considérer qu'on l'utilise, mais en unifiant une structure vide avec une structure contenant de l'information, donc le résultat est identique à la seconde structure.

- On utilise l'unification dans un cas qui n'existe pas pour une grammaire : celui de la combinaison de fragments. En effet, on a vu que ce que [Crabbé, 2005b] représentait sous forme d'une conjonction revenait, en termes de combinaison de structures, à une unification, et on a donc fait usage de cela en "méta-GP". Cela ne nous demande pas d'introduire un nouveau mécanisme, mais d'en utiliser un préexistant à des fins qui n'existaient pas dans une vision non "méta-grammaticale" (sans que cela s'oppose à l'usage qui en fait en grammaire).

On considère que les contraintes s'appliquant à la combinaison des arbres, et qui sont d'ailleurs mentionnées en dehors de la MG de [Crabbé, 2005b], ne font pas partie de la métagrammaire elle-même mais d'un ensemble d'heuristiques ayant pour but de limiter la surgénération des constructions grammaticales à partir de la grammaire. On peut mettre cela en parallèle avec les heuristiques que l'on devrait mettre en œuvre dans notre modèle pour limiter la surgénération lors de l'analyse, et qui relèvent à notre sens non pas du modèle lui-même mais plutôt des implémentations qui peuvent en être faites. On constate que le point de vue d'un développeur de métagrammaires est en accord avec nous sur cette question.

L'une des conséquences de l'utilisation de GP comme une métagrammaire est qu'une fois la grammaire construite à partir de la métagrammaire, celle-ci ne portera pas de hiérarchie d'héritage. En effet, l'héritage (éventuellement double) qui est mentionné dans la métagrammaire sert à construire la grammaire, mais ne subsiste pas dans cette dernière. En d'autres termes, les constructions d'une grammaire construite à partir d'une "méta-GP" ne constitueront pas un réseau de descriptions, mais uniquement un ensemble d'objets non liés entre eux. Il ne s'agira donc pas, à proprement parler, d'une Grammaire de Propriétés.

Concernant le modèle (TAG) utilisé dans cet exemple maintenant, on a vu qu'il est possible de représenter un modèle DI/PL en GP. Peu importe que celle-ci soit une "méta-GP" en l'occurrence ou une GP standard (qui aurait été une transcription d'arbres et non de fragments d'arbres), les contraintes sont les mêmes dans les deux cas. On a défini pour la formalisation que l'on a présentée ici deux types de propriétés, Dominance Immédiate et Précédence Linéaire, qui permettent d'exprimer les deux types de relations des grammaires syntagmatiques classiques. On a vu également comment l'on pouvait exprimer l'équivalent des Principes de ces grammaires, soit en les ajoutant à la métagrammaire, soit en les ajoutant à la grammaire <sup>89</sup>.

On a vu également que la spécification des types de propriétés peut permettre de rendre une grammaire "générationnelle" : on la contraint à cela d'une part en ignorant la condition de disponibilité des propriétés, c'est-à-dire en rendant toutes les propriétés toujours disponibles, donc toujours à évaluer ; d'autre part en leur affectant une capacité de 1 ce qui implique que chaque propriété ne peut être évaluée qu'une seule fois pour une construction donnée. Il est

---

<sup>89</sup>Dans ce dernier cas on rétablit la hiérarchie au sein de la grammaire : toutes les constructions dérivées de la métagrammaire hériteront — à un même niveau — des constructions de base représentant les principes.

donc possible de contraindre un modèle à la générativité en GP.

Avec l'utilisation de GP comme une métagrammaire, on peut envisager de factoriser des grammaires qui ne sont nécessairement ni lexicalisées, ni génératives, et par exemple d'envisager de représenter en une MG la grammaire issue de notre modèle (chose qu'on ne pourrait pas faire avec une MG classique, puisque les formalismes utilisés pour les descriptions de fragments d'arbres n'offrent pas suffisamment de variété dans les relations possibles, et donc contraignent trop la théorie).

Un dernier point qui semble intéressant mais que l'on ne peut développer ici pour des raisons de temps, est que dans le cas de l'emploi de GP en tant que métagrammaire, l'utilisation qui en est faite revient à de la génération (de grammaire) et non plus de l'analyse (d'énoncés). Il serait sans aucun doute très intéressant d'explorer cette piste pour évaluer comment il serait possible d'implémenter cela, et ainsi, éventuellement, de commencer à envisager d'utiliser GP, en tant que grammaire cette fois, comme une ressource de génération d'énoncés.

### Résumé

Dans ce chapitre on a tenté de représenter divers modèles de grammaire, éloignés du nôtre selon divers aspects, en GP, afin d'en faire émerger les possibilités de représentation.

On a d'abord vu qu'il était possible de représenter un modèle modulaire et lexicalisé, avec le cas de la GUST. A travers cet exemple, on a montré également qu'une GP peut formaliser une grammaire d'unification, mais que notre modèle n'entre pas dans le groupe des grammaires d'unification puisqu'elle ne répond pas à certains de ces points de définition.

On a vu ensuite qu'il était possible de représenter l'équivalent d'informations procédurales et catégorielles en GP, mais qu'en revanche il était impossible, faute d'y ajouter un certain nombre de mécanismes, de représenter une analyse ascendante.

On a enfin vu qu'il était possible d'employer GP comme une métagrammaire, et que cela permettait d'utiliser la représentation métagrammaticale sous une forme qui contraigne théoriquement moins les modèles (en l'occurrence, qui ne l'oblige pas à être lexicalisée). Par cet exemple, on a aussi montré qu'il est possible de contraindre une GP à la générativité.



Deuxième partie

Éléments de grammaire du français



# Introduction

Une première grammaire GP très simple avait été proposée dans [Blache, 2001], qui avait plus pour objectif d’illustrer le fonctionnement du formalisme, que d’être couvrante et fine. Puis à l’occasion de la campagne EASY (*Evaluation des Analyseurs SYntaxiques*, cf. p.ex. [Paroubek, 2005]) nous avons réalisé une deuxième grammaire<sup>90</sup> ; celle-ci devait répondre à un cahier des charges très strict (le *Protocole d’Evaluation des Analyseurs Syntaxiques*, [Gendner & Vilnat, 2004]), et notamment fournir des analyses dont la forme plate n’est pas notre objectif ici, et dont la finesse syntaxique est largement discutable (cf. les discussions à ce propos *a priori* dans [Blache & Morin, 2003] et *a posteriori* dans [Balfourier *et al.*, 2005]). Nous avons également participé à l’élaboration d’une grammaire-jouet permettant l’utilisation d’un configurateur pour l’analyse en GP (cf. [Estratat, 2006]). Les éléments de grammaire que nous présentons ici ont un objectif différent de ces deux premières grammaires : ils visent à fournir une **base couvrante et stable**, théoriquement pertinente du point de vue développé dans la première partie, permettant tous les développements ultérieurs envisagés.

## 1 Objectifs

Il ne s’agit bien évidemment pas d’une grammaire qui se prétend complète : le temps nécessaire à cela est très nettement supérieur à celui dévoué au développement de la présente ressource. Néanmoins, elle veut proposer une représentation de la plupart des **phénomènes syntaxiques basiques du français**, que ce soit sous sa forme écrite ou orale, et autoriser par là le traitement d’entrées tout venant.

Cette grammaire n’a pas non plus pour objectif la finesse de traitement : nous avons préféré proposer une **ressource la plus générale possible** plutôt que de nous concentrer sur un (ou sur quelques) phénomène(s) précis. Ce choix s’est effectué pour trois raisons. La première est que comme on l’a dit, il n’existait pas encore de grammaire GP qui soit suffisamment développée et qui exploite réellement les positions théoriques du modèle : nous avons donc estimé qu’il était nécessaire de commencer par là. La deuxième est que, comme on a eu l’occasion de voir, il n’existe pas à notre connaissance de grammaire formelle qui soit basée sur les mêmes positions théoriques que celles qui sont les nôtres ; ils nous aurait donc été impossible d’intégrer une partie de nos descriptions à une ressource pré-existante, sauf à abandonner une part plus ou moins importante de nos choix théoriques. Enfin, la troisième est que nous

---

<sup>90</sup>Cette grammaire est présentée en annexe A.

avons décidé de nous concentrer sur un problème à notre sens trop rarement posé de manière explicite en linguistique formelle malgré son omniprésence : celui de la **cohérence globale des descriptions**. En effet, dans les modèles les plus populaires on trouve en abondance des descriptions très fines des phénomènes les plus divers, mais ces descriptions sont trop peu souvent intégrées à une ressource grammaticale globale <sup>91</sup>. Quand on travaille sur le développement d'une grammaire, on se rend compte que l'intégration d'un phénomène à un ensemble requiert de prendre en considération la totalité de l'objet pour ne pas utiliser de représentations ni de mécanismes *ad hoc*, pour ne pas générer de conflits avec le reste des descriptions, et pour tenir compte de ces autres descriptions (et donc des informations auxquelles on a accès : les objets construits, leur forme, leur contenu, les choix théoriques). Ce point, dans les décisions concernant et la description elle-même, et sa représentation, est crucial pour pouvoir utiliser les travaux à des fins de traitement automatique notamment, mais pas uniquement (cela vaut aussi simplement pour la conception de grammaires complètes, même si elles ne sont pas implémentées). Nous nous sommes donc plus concentrée sur la cohérence générale des descriptions que sur l'approfondissement de certains d'entre eux <sup>92</sup>.

Une conséquence de ceci est que cette grammaire n'a pas non plus pour but de proposer une analyse résolument novatrice de tel ou tel phénomène. Mis à part notre volonté de prise en compte de la multi-dimensionnalité des analyses qui a une incidence assez remarquable sur certaines descriptions, dans la plupart des cas nous avons utilisé plus ou moins fidèlement des descriptions pré-existantes, parce qu'elles répondaient à nos besoins. Néanmoins, les théories dont nous nous sommes inspirée n'avaient pour la plupart pas été formalisées ; nous pensons notamment aux descriptions de l'Approche Pronominale. Celles-ci ont souvent retenu notre attention et nous en avons donc utilisé une partie pour l'intégrer à cette grammaire formelle générale. Cette intégration pourra, nous l'espérons, une fois la grammaire implantée dans un parseur <sup>93</sup>, permettre la vérification à grande échelle de leurs propositions et fournir un outil de travail supplémentaire à cette approche théorique. Dès lors, elle permet déjà d'intégrer ces travaux à un ensemble et de vérifier leur homogénéité et leur interdépendance.

Enfin, les extraits qui suivent n'ont pas pour objectif la prise en compte de tous les domaines d'analyse : bien que le modèle permette, encourage même, l'interaction de ces domaines, nous nous sommes ici et dans un premier temps concentrée exclusivement sur les **phénomènes syntaxiques**. Notre grammaire en son état actuel ne contient donc que des constructions syntaxiques, décrites à l'aide d'informations de cette même nature <sup>94</sup>. Néanmoins il est tout à fait possible d'adjoindre aux constructions existantes des informations, ainsi que d'ajouter de nouvelles constructions de tous autres domaines, puisque le modèle a été conçu dans cette perspective ; il suffit de les insérer dans la hiérarchie, quitte à la modifier en partie si nécessaire.

---

<sup>91</sup>Mis à part dans des travaux tels que la *French Lexicalized TAG* (FTAG, [Abeillé & Candito, 2000], [Abeillé et al., 2000]), l'*English Resource Grammar* basée sur HPSG (ERG, [Copestake & Flickinger, 2000]), ou encore la *Grammaire de référence du français contemporain* en préparation (projet ILF sous la direction d'Anne Delaveau).

<sup>92</sup>Nous avons tout de même décrit quelques phénomènes avec plus de précision que d'autres : c'est le cas du traitement des entassements paradigmatiques (coordinations et disfluences, cf. chapitre 11), qui sont la manifestation la plus remarquable des spécificités de notre modèle, ainsi que de certains types de verbes (à montée, causatifs, cf. chapitre 7).

<sup>93</sup>Cette grammaire a été implantée dans des parseurs à plusieurs étapes de son développement, mais la version présentée ici ne l'est pas encore. Rien ne s'y oppose toutefois, si ce n'est le temps imparti à la tâche.

<sup>94</sup>Et bien évidemment des informations portées par les constructions syntaxiques minimales, c'est-à-dire des informations des domaines lexical et morphologique.

Cette limitation à la prise en compte d'informations syntaxiques illustrera le fait que même pour ne décrire que des constructions strictement d'un domaine donné, on ne peut accéder à suffisamment d'information en se bornant à ne considérer que ce seul domaine : concrètement, un nombre important des constructions décrites dans ce qui suit manque de précision, ceci étant dû au fait que les informations syntaxiques ne sont souvent pas suffisantes pour permettre une description d'une finesse satisfaisante. On en revient ainsi à démontrer à quel point l'information linguistique est **éparse et inconstante**, tel qu'on le notait au début de ce travail.

## 2 Notations simplifiées

Dans la rédaction des descriptions grammaticales, pour des raisons de facilité d'écriture (et par là même de lecture), il nous est arrivé d'avoir recours à quelques simplifications dans la notation. Nous les expliquons dans cette section.

### 2.1 Référence aux catégories et étiquettes de constructions

Quand, dans une propriété, on veut faire référence à la catégorie d'une construction, on note simplement la valeur du trait correspondant. Par exemple pour faire référence à tous les objets de [CAT V] on note simplement V, tout comme pour faire référence à tous les objets qui sont des verbes modaux on note uniquement V-m sans préciser qu'il s'agit de l'étiquette.

Nous avons adopté la même notation pour les deux cas puisqu'il est facile de les distinguer : les étiquettes, lorsqu'on leur fait référence sous cette forme simplifiée, portent toujours un trait d'union (V-m, N-n, Pro-pcn, etc.)<sup>95</sup> alors que ce n'est jamais le cas des catégories (V, N, Pro, etc.).

### 2.2 Référence aux formes orthographiques de constructions

Pour les mêmes raisons il nous est arrivé de simplifier la notation de la forme orthographique d'un objet, en remplaçant p.ex. [FORME | ORTH *cheval*] par simplement *cheval*.

### 2.3 Portée des index

Ensuite, pour simplifier les descriptions de constructions<sup>96</sup>, nous avons choisi de faire porter les index non seulement localement pour la construction au sein de laquelle ils sont définis,

---

<sup>95</sup>Une étiquette qui ne porte pas de trait d'union correspond à une construction dont la catégorie est identique à l'étiquette, et pointe donc tous les objets de cette catégorie, c'est-à-dire que dans ce cas que l'on fasse référence à l'une ou l'autre revient au même.

<sup>96</sup>Notons bien que ceci concerne les constructions, c'est-à-dire ce qui se trouve dans la grammaire et non pas dans la spécification des types de propriétés.

mais également pour les constructions qui en héritent. Par exemple, dans les verbes construits (V) nous avons défini une Tête portant l'index  $\boxed{1}$ , puis dans les constructions qui héritent de V nous avons parfois fait référence à cette Tête sans la rappeler intégralement mais simplement en mentionnant le même index,  $\boxed{1}$ . Dans toutes les constructions qui héritent de V, la mention de cet index  $\boxed{1}$  fait référence à la Tête définie dans V. Par contre, dans la description des Syntagmes Verbaux (SV) qui suivent, l'index  $\boxed{1}$  fait référence à autre chose puisque SV n'hérite pas de V (et dans ce cas l'index est défini localement pour la construction SV et celles qui en héritent).

D'une manière générale, aucun index que nous utilisons n'est libre ; tous sont liés à des variables, qui peuvent être définies en local (i.e. directement dans la construction qui les utilise) ou alors dans une construction située plus haut dans la hiérarchie et dont la construction qui l'emploie hérite.

## 2.4 Héritage au choix

Enfin, pour simplifier la représentation graphique des hiérarchies d'héritage, nous avons fait usage d'une convention de notation (un *syntactic sugar*) qui nous permet de ne pas démultiplier les constructions, en permettant à un objet d'hériter de plusieurs autres *au choix*.

Par exemple dans les verbes construits, on fait hériter la construction “verbe à négation” (p.ex. *ne mange pas*) de la construction “verbe” (p.ex. *mange*) ; par la suite on introduit une construction “verbe modal”, qui peut hériter soit du “verbe” (*peux*), soit du “verbe à négation” (*ne peux pas*) ; puis on introduit un “verbe à auxiliaire *avoir*” qui peut hériter soit du “verbe” (*ai mangé*), soit du “verbe à négation” (*n'ai pas mangé*), soit du “verbe modal” qui lui-même avait déjà hérité soit du “verbe” (*ai pu*) soit du “verbe à négation” (*n'ai pas pu*).

Dans la notation hiérarchique standard, pour représenter les deux formes possibles du “verbe modal” (négatif ou positif) on devrait créer deux constructions différentes pour celui-ci, l'une héritant du “verbe” et l'autre du “verbe à négation” ; ensuite pour représenter les quatre héritages possibles du “verbe à auxiliaire *avoir*” (positif non modal, négatif non modal, positif modal, négatif modal) on devrait créer autant de constructions différentes pour celui-ci, une première héritant du “verbe”, une deuxième du “verbe à négation”, une troisième du “verbe positif modal” et une quatrième du “verbe négatif modal”. Et ainsi de suite pour le reste de la hiérarchie, ce qui a pour inconvénient de multiplier à chaque fois le nombre de constructions. On donne un exemple de cette représentation en figure 1.

Notre convention de représentation de l'héritage *au choix* nous permet de créer un seul “verbe modal” qui peut être négatif ou non selon qu'il hérite immédiatement du “verbe” ou du “verbe à négation”, puis un seul “verbe à auxiliaire *avoir*” qui peut être positif ou négatif, modal ou non, en fonction des constructions précédentes dont il hérite.

Graphiquement, on représente cela en faisant se rejoindre les deux arcs d'héritage (p.ex. un partant du “verbe” et un autre partant du “verbe à négation”) en un point matérialisé par un  $\bullet$ , dont ne part qu'un unique arc qui pointe vers le “verbe modal” (p.ex.), et qui illustre le fait que le “verbe modal” n'hérite pas à la fois du “verbe” et du “verbe à négation”, mais soit de

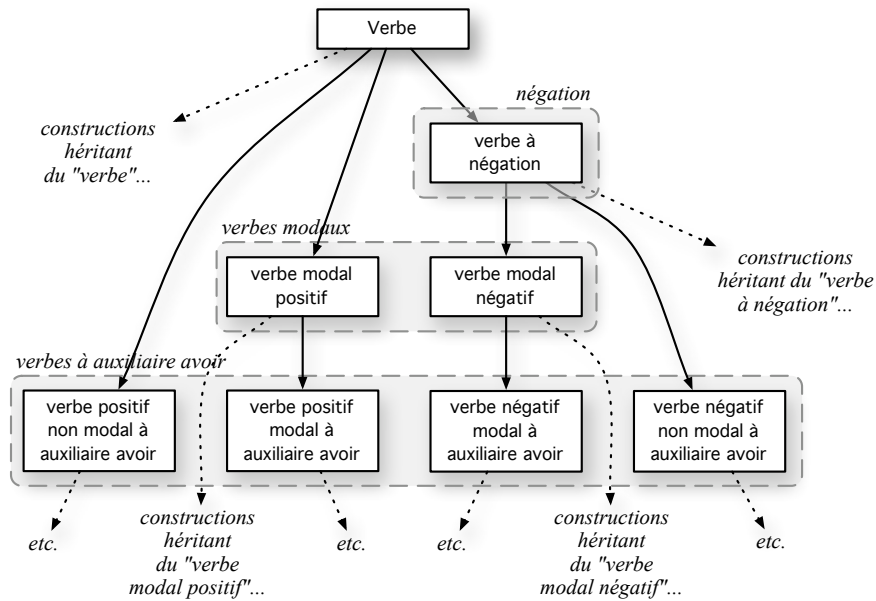


FIG. 1 – Représentation dans la hiérarchie de l'héritage sans la notation "au choix".

l'un, soit de l'autre. La représentation dans la hiérarchie prend alors la forme de la figure 2.

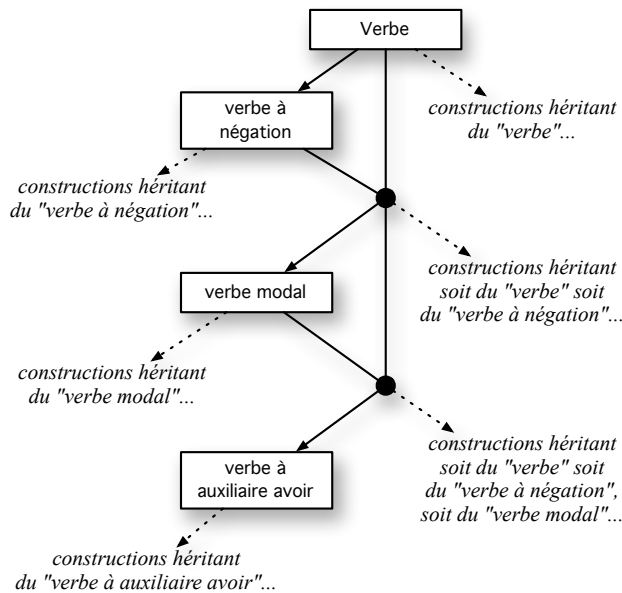


FIG. 2 – Représentation dans la hiérarchie de l'héritage avec la notation "au choix".

Le "verbe modal", s'il hérite du "verbe", ne récupérera que ses caractéristiques ; si par contre il hérite du "verbe à négation", il récupérera les caractéristiques de ce dernier (et accessoirement

celles du “verbe” également puisque le “verbe à négation” hérite lui-même du “verbe” - mais ceci n’est pas systématique).

Dans les représentations formelles, les constructions héritées *au choix* sont représentées sous un unique attribut HERIT, et séparées par des  $\vee$  (i.e., des “ou inclusifs” suivant nos conventions) :



FIG. 3 – Représentation dans une construction de l’héritage *au choix*.

L’exemple de la fig. 3 signifie que la construction d’étiquette A hérite soit de la construction d’étiquette B, soit de la construction d’étiquette C (sachant qu’il est possible, p.ex. que C elle-même ait hérité de B).

A titre de comparaison, une construction A’ qui hérite *conjointement* des deux constructions B et C comportera autant de traits HERIT que de constructions héritées (fig. 4) :

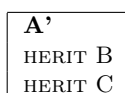


FIG. 4 – Représentation dans une construction de l’héritage conjoint.

### 3 Organisation de la partie

Dans les chapitres qui suivent nous présentons une série de descriptions de constructions syntaxiques. Nous commençons par présenter la hiérarchie générale de la grammaire et les **constructions de base** (chapitre 5), puis dans chacun des chapitres qui suivent nous détaillons une partie de la hiérarchie syntagmatique, classés par **type de construction** (nominales, verbales, etc.) : nous proposons d’abord la description des constructions nominales (chapitre 6) qui nous permet d’illustrer assez précisément le fonctionnement du modèle puisqu’elle fait usage de tout ce que l’on a pu montrer, tout en possédant une hiérarchie relativement simple ; nous verrons dans les chapitres qui suivent que cette hiérarchie peut être plus complexe (chapitre 7 sur les verbes) ou plus simple (chapitres 10 sur les adjectifs, etc.), tout en conservant un même squelette de base et une certaine homogénéité dans l’organisation. Nous finissons notre ensemble de descriptions par le chapitre décrivant les **entassements paradigmatiques** (chapitre 11), qui constitue l’un des points d’intérêt majeur de cette grammaire puisqu’il donne une place originale à l’étude des coordinations et des disfluences dans la grammaire. Nous y verrons de quelle façon nous justifions notre position à ce propos, et ce que cela apporte à la description grammaticale.

Cette partie a donc pour objectif de montrer comment nous représentons les informations dans la grammaire, quelle cohérence cela nous permet de conserver tout au long de ces descriptions de phénomènes, et de mettre en avant au fil des chapitres certaines différences (et certains rapprochements) de choix théoriques avec d’autres grammaires. Puis pour compléter ceci, dans



la partie suivante nous présenterons le fonctionnement de quelques-unes de ces descriptions de constructions à travers des exemples d'analyses d'énoncés, afin de montrer quels sont les apports de notre proposition en termes d'analyse également.

Deux remarques pour finir. Tout d'abord, présenter une grammaire sous la forme linéaire d'un manuscrit est un exercice ardu ; le lecteur saura donc d'avance nous pardonner les fréquents renvois d'un chapitre à l'autre, d'une section à l'autre, malheureusement inévitables compte tenu de la nature de l'objet décrit.

D'autre part, le développement d'une grammaire requiert de faire usage de nombreuses références à tout moment. Aussi n'avons-nous pas systématiquement listé au fil des descriptions la totalité des sources de toutes nos informations ; nous ne les avons précisées que dans certains cas où cela nous a semblé indispensable. Parmi les sources que nous avons régulièrement utilisées pour rédiger cette ressource, nous pouvons toutefois citer (même si la liste n'est pas exhaustive) , [Abeillé, 1993], [Abeillé, 2002], [Arnauld & Lancelot, 1830], [Baccus, 2002], [Berrendonner *et al.*, 1983], [Blache, 2001], [Chevalier *et al.*, 1997], [Damourette & Pichon, 1968], [Damourette & Pichon, 1987], [Dubois & Lagane, 1973], [Le Goffic, 1993], [Gross, 1986], [Guénou, 2002], [Morin, 1985], [Morin, 1989], [Riegel *et al.*, 1999], et [Wilmet, 1998].

### Résumé

Dans cette introduction nous avons présenté les objectifs de cette deuxième partie. Il s'agit de fournir une représentation de la plupart des phénomènes syntaxiques de base du français, sous sa forme écrite et orale. Nous avons eu dans le développement de cette grammaire un souci constant de maintien de la cohérence globale des descriptions, afin que la ressource soit la plus générale possible.

Nous avons ensuite présenté quelques simplifications de notations que l'on a utilisées dans les descriptions qui suivent : la notation des références aux catégories et aux formes orthographiques des objets, la portée des index, la représentation de l'héritage *au choix*.

Les chapitres qui suivent présentent chacun une partie de la grammaire : le premier décrit les constructions de base de la hiérarchie, puis les suivants portent chacun sur un type de construction. On commence par présenter les constructions syntagmatiques, pour finir par les constructions paradigmatiques qui illustrent le point le plus spécifique à notre proposition.

# Chapitre 5

## Constructions de base

Afin d'introduire une structure uniforme des informations, et afin de rassembler tout ce qui est commun à la totalité des constructions de la grammaire, toute la hiérarchie est construite à partir d'une construction de base, que nous appelons Construction Grammaticale (CG), qui constitue la racine de la grammaire et dont toutes celles qui suivent héritent.

Nous commençons ici par donner une vue générale de la **hiérarchie de la grammaire** (§1). A la suite de cela nous présentons une descriptions des constructions de base de cette hiérarchie (§2), parmi lesquelles la **Construction Grammaticale** (§2.1), suivie des constructions qui en héritent directement : **Syntagme** (§2.2) et **Paradigme** (§2.3), ainsi que les deux constructions qui héritent du Syntagme : **X construits** (§2.2.2) et **Syntagmes X** (§2.2.3).

### 1 Hiérarchie générale des constructions

Comme on vient de le dire, toutes les constructions de la grammaire héritent, directement ou indirectement, de la Construction Grammaticale générale. Toute la grammaire s'articule donc autour d'elle.

Notre ensemble de descriptions n'a prétention à être ni complet ni fin puisqu'il n'est qu'une première étape de développement, néanmoins il serait d'ores et déjà difficile d'embrasser la totalité de la hiérarchie en une seule figure. Nous donnons donc en fig. 5 l'articulation initiale de la grammaire, i.e. la Construction Grammaticale et ses premiers niveaux d'héritage. Les parties qui sont développées dans ce chapitre sont encadrées ; celles qui ne sont pas développées ici font chacune l'objet d'un chapitre ou d'une section, et leurs hiérarchies locales seront introduites dans les chapitres les concernant.

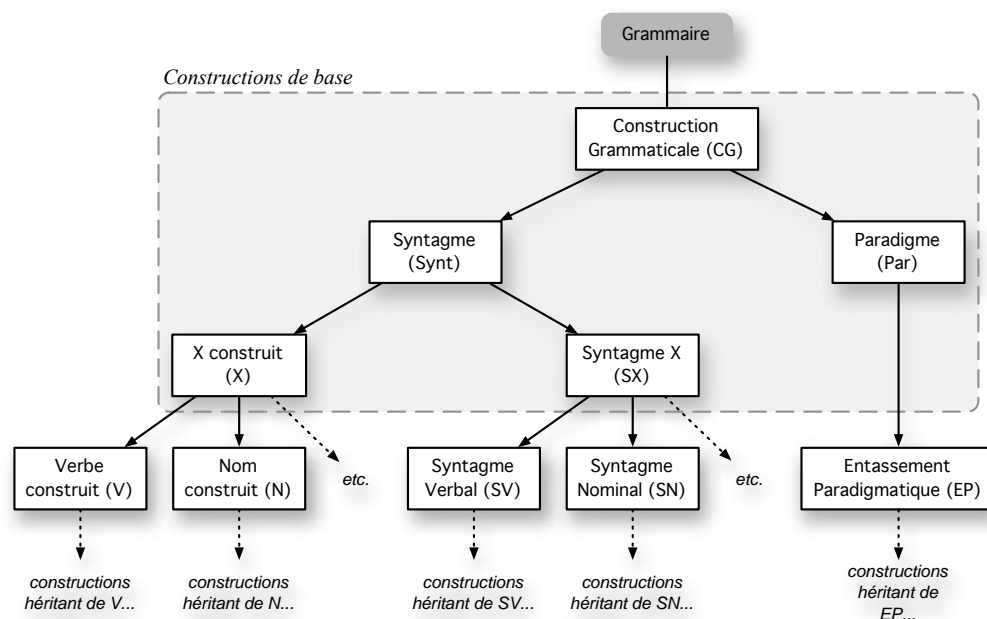


FIG. 5 – Hiérarchie générale de la grammaire : racine et premiers niveaux.

## 2 Constructions de Base

On présente ici la description de la Construction Grammaticale (§2.1), des constructions Syntagme (§2.2) et Paradigme (§2.3), et des constructions X construit (§2.2.2) et Syntagme X (§2.2.3). Toutes ces constructions n'ont pas réellement vocation à l'introduction de quelconques objets lors d'une analyse <sup>97</sup>, mais uniquement de permettre l'organisation générale des objets de la grammaire par la mise en commun des caractéristiques partagées.

### 2.1 Construction Grammaticale (CG)

La Construction Grammaticale (CG, fig. 6) est la racine de la grammaire : elle est donc la seule à n'hériter d'aucune construction. Elle est constituée d'au moins un objet Majeur, indexé  $\boxed{1}$  et de type quelconque (d'où sa mention sous forme de structure de traits vide,  $[]$ ) <sup>98</sup>, dont elle récupère (dans ses informations intrinsèques) les informations syntaxiques. En outre, CG est définie comme faisant partie du domaine syntaxique.

Ce dernier point, ainsi que le précédent dans une moindre mesure, vient du fait que les construc-

<sup>97</sup>Concrètement, dans une implémentation de cette grammaire, il faut clairement interdire que les constructions de cette section puissent être à l'origine de l'introduction d'objets, sous peine de se retrouver avec une explosion combinatoire évidente.

<sup>98</sup>La spécification de cet objet pourra par la suite être précisée plus localement, mais son index restera toujours  $\boxed{1}$ .

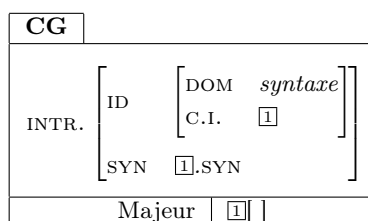


FIG. 6 – La Construction Grammaticale (CG), racine de la grammaire.

tions que l'on décrit par la suite sont uniquement des constructions syntaxiques, et c'est la raison pour laquelle on a fait remonter cette information jusqu'à la racine de la grammaire ; mais on peut supposer que quand cette grammaire sera étoffée cette information devra redescendre dans la hiérarchie jusqu'au point au-dessus duquel les domaines pourront être variables.

Le fait que notre grammaire ait une unique racine est également un effet secondaire de l'ensemble des constructions qu'elle contient dans son état actuel : on peut tout à fait concevoir qu'elle puisse nécessiter une (ou plusieurs) autre(s) construction(s) racine, présentant des caractéristiques différentes de celle-ci, et dont les constructions héritantes n'auront rien en commun avec celles que l'on présente dans ce qui suit <sup>99</sup>. Le fait de n'avoir qu'une unique racine à notre grammaire est un cas particulier et non pas une obligation.

De cette construction héritent immédiatement deux autres :

- D'une part les **Syntagmes**, qui sont des constructions *syntaxiques* (en l'occurrence, puisque l'on a défini notre domaine comme syntaxique uniquement) dont les constituants entretiennent des relations *hypotactiques*.
- D'autre part les **Paradigmes**, qui sont des constructions *syntaxiques* dont les constituants entretiennent des relations *paratactiques*.

## 2.2 Syntagme (Synt)

Les Syntagmes (Synt) sont, comme on vient de le dire, des constructions syntaxiques dont les relations entre les constituants sont hypotactiques (de même que les relations entre les constituants et la construction elle-même). Cela signifie qu'ils entretiennent des relations de dépendance syntaxique au sein desquelles l'un des objets est recteur, l'autre dépendant.

La construction (fig. 7) hérite de CG. Elle ajoute aux caractéristiques héritées principalement des informations intrinsèques : on sait dans ce cas que la construction précédemment indexée [ ] fournira à la construction son TYPE (sous le trait ID|NATURE), et occupera la fonction de TÊTE de la construction (visible sous le trait SYN) s'il s'agit d'une construction lexicale ; sinon c'est la tête de [ ] qui en fera office (qui, elle, sera une construction lexicale puisque la sélection entre les deux se sera faite également au niveau d'analyse précédent, et ainsi de suite depuis le début de l'analyse).

<sup>99</sup>Dans ce cas, modifier l'étiquette de celle-ci ne serait pas inutile.

A cela on ajoute une information extrinsèque, qui indique que les constituants disfluents (étiquetées EP-d) précèdent les constituants non-disfluents. Ceci fait écho aux études menées à ce propos (p.ex. [Blanche-Benveniste, 1987], [Pallaud, 2006]) qui indiquent une certaine régularité dans ce fait : les disfluences ont tendance à se produire en début de syntagme.

Synt	
HERIT CG	
INTR.	$\left[ \begin{array}{l} \text{ID} \left[ \text{NATURE} \left[ \text{TYPE} \left[ \underline{1}.\text{TYPE} \right] \right] \right] \\ \text{SYN} \left[ \text{TÊTE} \left[ \underline{1}.\text{TÊTE} ? : \underline{1} \right] \right] \end{array} \right]$
Précédence	
EP-d $\prec$ $\neg$ EP-d	

FIG. 7 – La construction syntagmatique générale, Syntagme (Synt).

### 2.2.1 La Tête comme spécificité des Syntagmes

Revenons sur ce que l’on vient de dire : *tous les syntagmes ont une tête*. Sur ce point, nous nous accordons parfaitement avec l’immense majorité des grammaires. Mais il y a pourtant une différence de taille dans les implications de cette affirmation : pour ce qui nous concerne, *toutes les constructions syntaxiques ne sont pas des syntagmes*. En effet d’après ce que l’on vient de voir, la construction Syntagme, celle qui a pour caractéristique de posséder une tête, n’est pas la seule construction syntaxique de notre grammaire : une autre hérite de CG, la construction Paradigme (sur laquelle nous revenons au §2.3 qui lui est consacré). Et le fait que nous ayons choisi de n’introduire cette tête qu’au niveau du Syntagme et pas plus haut dans la hiérarchie signifie, comme on peut s’en douter, que les Paradigmes ne portent pas de tête.

Cette différence n’est pas en soi une différence de conception de l’analyse *syntagmatique* ; il s’agit d’une différence de conception de l’analyse *syntactique*. Là où l’immense majorité des grammaires, fort judicieusement appelées syntagmatiques, réduisent l’analyse syntaxique à la constitution de syntagmes et *in fine* ont tendance à confondre les deux, pour notre part nous mettons un point d’honneur à considérer les deux dimensions, complémentaires et distinctes, de l’analyse syntaxique : syntagmatique certes, mais également paradigmatique. De ce fait, l’assertion selon laquelle *tous les syntagmes ont une tête* devient une caractéristique locale à la dimension syntagmatique, mais n’est pas généralisée en une constante syntaxique.

**Pas de principes.** — De cela résulte une autre différence avec de nombreuses théories concurrentes : comme on a eu l’occasion de le dire précédemment, notre grammaire n’est pas “*guidée par les têtes*” ; ceci est normal puisque toutes les constructions n’en portent pas nécessairement. Mais ceci a une conséquence directe, qui est que l’on n’utilise pas de “*principes*” pour gérer le passage des traits d’un objet à un autre : en l’occurrence, on n’a pas recours à un “*principe des traits de tête*” pour faire remonter la totalité des traits de la tête *a priori* à la construction : on a la possibilité de préciser une sélection des informations intrinsèques qui remontent de l’une à l’autre (et l’on peut généraliser ceci au point de le rendre équivalent à l’utilisation de principes), et c’est ce que l’on fait lorsqu’on indique dans une construction aussi

générale que celle de Syntagme le fait que la construction récupère les traits syntaxiques de sa tête p.ex., mais ce mécanisme n'est pas défini comme une donnée qui se placerait au-dessus ou au-delà de la grammaire, en tant que constante générale, i.e. en tant que “*principe*” tel qu'on l'entend dans ces modèles.

On a montré en première partie que pour formaliser une théorie en GP on avait besoin de définir deux ressources en l'espèce d'une grammaire et de la spécification des types de propriétés qui y sont utilisés. On vient de montrer ici l'une des conséquences de ceci : on n'a pas besoin d'ajouter à cela un ensemble de principes qui expliciteraient toute une partie du fonctionnement de la grammaire. Si l'on postule des caractéristiques qui ont une portée générale, et ce quel que soit le degré de généralité (i.e. même si l'on postule un certain degré d'universalité des caractéristiques linguistiques), alors on les indique dans les constructions les plus générales de la grammaire ; mais on n'a nullement besoin de les en séparer.

Nous avons dit plus haut que nous n'utilisons pas de principes, et avons donné l'exemple du “*principe des traits de tête*”, selon lequel les informations de la tête remontent progressivement le long de la structure. A ce propos nous pouvons noter que dans notre modèle, le passage des informations intrinsèques se fait toujours du constituant vers le constitué (du “bas” vers le “haut”) et jamais en sens inverse (du “haut” vers le “bas”). Cela vient du fait que notre grammaire n'est pas lexicalisée, et plus généralement du fait que l'on ne présume pas de la structure dans laquelle les constructions sont impliquées au sein même de leurs informations intrinsèques. On n'a donc jamais besoin de faire “descendre” une information d'une construction à l'un de ses constituants.

**Héritage du Syntagme.** — De cette construction héritent deux types de Syntagmes :

- D'une part les syntagmes (canoniquement) *endocentriques*, i.e. dont la catégorie est la même que celle de leur tête. On leur donne l'appellation générale de **X construits**, X représentant la catégorie de leur tête. On aura par exemple des Adjectifs construits (d'étiquette Adj) comme *très grand*<sup>100</sup>, des Noms construits (N) comme *gentil chien*<sup>101</sup>, des Verbes construits (V) comme *l'ai mangé*<sup>102</sup>, etc.
- D'autre part les syntagmes (canoniquement) *exocentriques*, i.e. dont la catégorie est différente de celle de leur tête. On leur attribue l'appellation générale de **Syntagmes X**, X représentant ici également la catégorie de leur tête. On aura par exemple des Syntagmes Nominaux (SN) comme *le gentil chien*, des Syntagmes Verbaux (SV) comme *ai mangé un fruit*, mais pas de Syntagmes Adjectivaux.

On décrit ces deux constructions dans les sections qui suivent.

<sup>100</sup>Souvent appelé — à tort, selon nos définitions — “syntagme adjectival”.

<sup>101</sup>Ce que certains appellent “membre nominal”, ou même parfois “groupe nominal” par opposition au “syntagme nominal”.

<sup>102</sup>Il s'agit de l'“amas verbal” de [Gerdes & Kahane, 2004].

### 2.2.2 X construit (X)

Les syntagmes que l'on appelle "X construits" sont des constructions syntaxiques canoniquement endocentriques, c'est-à-dire qu'elles portent la même catégorie que leur tête.

<b>X</b>					
HERIT Synt					
INTR.	$\left[ \begin{array}{l} \text{ID} \left[ \begin{array}{l} \text{C.I.} \quad [1], [2] \\ \text{NATURE} \quad \left[ \text{CAT} \quad [1.\text{CAT}] \right] \end{array} \right] \end{array} \right]$				
<table> <tr> <td>Mineur</td><td>[2] [ ]</td></tr> <tr> <td>Exigence</td><td>[1] <math>\Rightarrow</math> [2]</td></tr> </table>		Mineur	[2] [ ]	Exigence	[1] $\Rightarrow$ [2]
Mineur	[2] [ ]				
Exigence	[1] $\Rightarrow$ [2]				

FIG. 8 – Le X construit (X).

Comme on le voit en fig. 8, les X construits héritent de la construction Syntagme (Synt), et portent comme on vient de le dire la catégorie de leur tête (précédemment indexée [1] plus haut dans la hiérarchie). Puisqu'il s'agit d'objets complexes, par définition <sup>103</sup>, le X construit porte également un autre constituant, Exigé par le premier, qui cette fois est mentionné (de manière sous-spécifiée comme pour la tête précédemment) dans les objets Mineurs, et indexé [2] (ce qui signifie que toutes les constructions qui héritent de celle-ci, quand elles feront référence à l'index [2], feront référence à cet objet Facultatif ou alors au même qui aura été précisé plus localement à l'aide du même index). Ce second constituant est mineur puisque sa présence implique l'introduction d'un syntagme de la même catégorie que l'autre constituant immédiat ; cela signifie que canoniquement cet objet indexé [2] sera une expansion de la tête et en tout cas n'a pas d'incidence sur la nature de la construction. Enfin, on le mentionne dans la liste des constituants immédiats.

### 2.2.3 Syntagme X (SX)

Le Syntagme X est l'autre construction héritant de notre construction générale Syntagme. Elle rassemble les syntagmes canoniquement exocentriques, c'est-à-dire dont la catégorie est distincte de celle de leur tête.

<b>SX</b>
HERIT Synt

FIG. 9 – Le Syntagme X (SX).

La fig. 9 nous montre que mis à part son héritage, cette construction ne porte aucune information localement définie. Notons toutefois que cela signifie que le Syntagme X hérite du Syntagme le fait qu'il porte, non pas la catégorie, mais tout au moins le type de sa Tête.

<sup>103</sup>Un "X construit" qui serait constitué uniquement d'un constituant X n'aurait aucun sens en termes d'analyse.



Ici l'on ne spécifie pas d'objet Facultatif, parce que concrètement les Syntagmes X pourront être constitués soit effectivement d'une tête (Majeur) et d'un (ou plusieurs) objet(s) Mineur(s) (par exemple dans un SV, un Verbe *mange* et un SN *un fruit*), soit d'une tête et d'un autre constituant non pas Mineur mais Majeur (par exemple dans un SN, un Déterminant *le* et un N *petit chien*), soit encore par des objets seuls qui changent de paradigme (par exemple une Proposition uniquement constituée d'un SV *je la mange*).

### 2.3 Paradigme (Par)

Le Paradigme est la seconde construction qui hérite de la Construction Générale de la grammaire. Comme on l'a dit, elle rassemble les informations concernant les constructions syntaxiques dont les constituants entretiennent des relations paratactiques. Les constructions répondant à cette définition dans notre grammaire sont pour l'instant limitées aux Entassements Paradigmatiques, i.e. coordinations et disfluences (cf. à ce propos le chapitre 11 qui leur est consacré) <sup>104</sup>.

Par	
HERIT CG	
INTR.	<div> <div>ID</div> <div> <div>C.I.</div> <div>NATURE</div> <div> <div>[1], [2], ..., [n]</div> <div>[TYPE Par]</div> </div> </div> </div>
Majeur	<div> <div>[1]¬Par</div> <div>[2]¬Par</div> <div>...</div> <div>[n]¬Par</div> </div>
Exigence	<div>[x] ⇒ [y]</div>

FIG. 10 – La construction paradigmatique générale, Paradigme (Par).

En fig. 10 on voit que cette construction, en plus d'hériter de CG, y ajoute une information intrinsèque selon laquelle elle est de type *Par* (à la différence des Syntagmes, qui récupéraient le type de leur Tête). Contrairement aux constructions syntagmatiques et comme on l'a dit plus haut, celle-ci n'a effectivement pas de tête. Elle est constituée uniquement d'objets Majeurs (dont le nombre est indéterminé) qui sont tous des objets soit lexicaux soit syntagmatiques, i.e. ne sont pas des Paradigmes, et qui doivent être au minimum deux pour constituer une construction (d'où la mention de la propriété d'Exigence).

**Informations syntaxiques des constituants.** — Comme on le voit dans cette figure, on ne précise pas de relation intrinsèque syntaxique spécifique, ni entre la construction et ses constituants, ni entre les constituants eux-mêmes (pas de tête, pas de spécifieur, ni de dépendances). Il serait toutefois — du moins nous semble-t-il — intéressant d'en faire figurer, tout en gardant à l'esprit que ces relations ne seront pas les mêmes que les dépendances

<sup>104</sup>Mais on peut en envisager d'autres. Dans le cas contraire, on pourra fusionner les constructions "entassement paradigmatique" et "paradigme" en une seule construction "paradigme" dont hériteront directement les constructions qui héritaient de l'"entassement paradigmatique".

syntaxiques des Syntagmes. Néanmoins nous n'avons pas, en l'état actuel de nos recherches, mis en évidence de relations que l'on pourrait mentionner ici, si ce n'est éventuellement une relation d'équivalence paradigmatique, qui serait à notre sens assez redondante avec le type même de la construction. Nous n'en faisons donc pas figurer pour l'instant.

Il serait également envisageable, et intéressant, d'ajouter à cela des informations d'ordre pragmatique, ou sémantique à constructions, parce que celles-ci sont assez typiquement le lieu d'effets de sens et d'effets pragmatiques notables, qu'il s'agisse de coordinations ou de disfluences. Encore une fois, faute de temps nous n'avons pas poussé notre recherche plus loin à ce propos mais ceci serait probablement l'un des compléments les plus urgents à apporter, afin de donner une dimension plus concrète à ces Paradigmes et de les intégrer avec plus de précision dans la grammaire et plus généralement d'en apporter une pleine justification dans la théorie sous-jacente.

### Résumé

On a présenté ici la hiérarchie générale de la grammaire, et l'on a donné les descriptions des constructions de base qui la constituent, c'est-à-dire celles qui forment son squelette et qui rassemblent les informations générales de chacune des grandes classes de constructions syntaxiques.

On a mis en évidence, à travers la présentation de ces constructions de base, l'une des différences essentielles entre notre grammaire et la majorité des autres : toutes nos constructions syntaxiques n'ont pas nécessairement de tête. La raison de cela est que l'on considère les deux axes d'analyse syntaxique, i.e. non seulement syntagmatique (constructions dont les constituants entretiennent des relations hypotactiques, et qui portent une tête) mais aussi paradigmatiques (constructions dont les constituants entretiennent des relations paratactiques, et qui ne portent pas de tête). La conséquence de cela est que notre grammaire n'est pas "guidée par les têtes".

Ceci a été également l'occasion de montrer que la mention d'éventuels "principes" qui sont des généralités linguistiques ne prend pas sa place en GP en-dehors de la grammaire, mais directement au sein de la grammaire. On a vu, dans la même idée, que ces constructions de base sont le lieu des indications très générales sur l'organisation des constructions, et que p.ex. c'est ici que l'on indique la contrainte selon laquelle les disfluences se produisent en début de syntagme (et non dans la description des disfluences elles-mêmes).



## Chapitre 6

# Constructions nominales

On appelle “constructions nominales” celles qui portent dans leurs informations intrinsèques le trait [TYPE *N*]. En font partie d’une manière générale les objets syntaxiquement construits<sup>105</sup> dont la tête est de catégorie nominale (ou se plaçant sur le ce paradigme). Cela implique dès le départ qu’il ne s’agira que de constructions syntagmatiques, puisque l’on a vu dans le chapitre précédent que les Paradigmes portaient le type *Par* (ils seront traités dans le dernier chapitre de cette partie) : en conséquence de quoi, on ne traite pas ici les coordinations de noms ou de SN, par exemple.

On y rassemble à la fois les **Noms construits** (N, §1) qui sont des constructions dont la tête est de catégorie nominale et est accompagnée de ses compléments, et les **Syntagmes Nominaux** (SN, §2) qui sont des constructions dont la tête est soit un nom (construit ou non, ou objet sur ce paradigme) accompagné de son spécifieur, soit un SN accompagné de ses expansions.

La hiérarchie d’héritage des constructions nominales est présentée en figure 11 : on y rappelle un extrait de la hiérarchie générale et on a développé la partie qui nous intéresse ici que l’on a encadrée en trois parties : le premier cadre rassemble les Noms construits (§1), les deux autres rassemblent les SN dont nous avons séparé la description en deux parties, d’une part les SN de base (§2.1) d’autre part ce qui concerne la détermination dans les SN exocentriques (§2.2).

Dans la suite de la grammaire, on verra que pour ce qui concerne les cas des constructions syntagmatiques, très souvent une construction de [TYPE *X*] fera comme ici l’objet d’une part d’une série de constructions héritant du *X* construit (ici, c’est le cas du Nom construit), d’autre part d’une autre série héritant du Syntagme *X* (ici, le SN). C’est l’une des raisons pour lesquelles on a organisé les constructions de base sous cette forme : afin de faire ressortir le parallélisme pouvant se faire entre les différents types de constructions.

Pour les descriptions qui suivent, on reste à un niveau de description très global, comme on s’en est expliquée en introduction. La grammaire étant en cours de développement, elle ne se prétend pas complète. Néanmoins elle veut offrir un squelette permettant par la suite d’affiner

---

<sup>105</sup>En fait, on y fera même, à terme, figurer les constructions morphologiquement construites, notamment, dont le type sera nominal. Mais pour l’instant on n’a pas traité la partie morphologique de la grammaire.

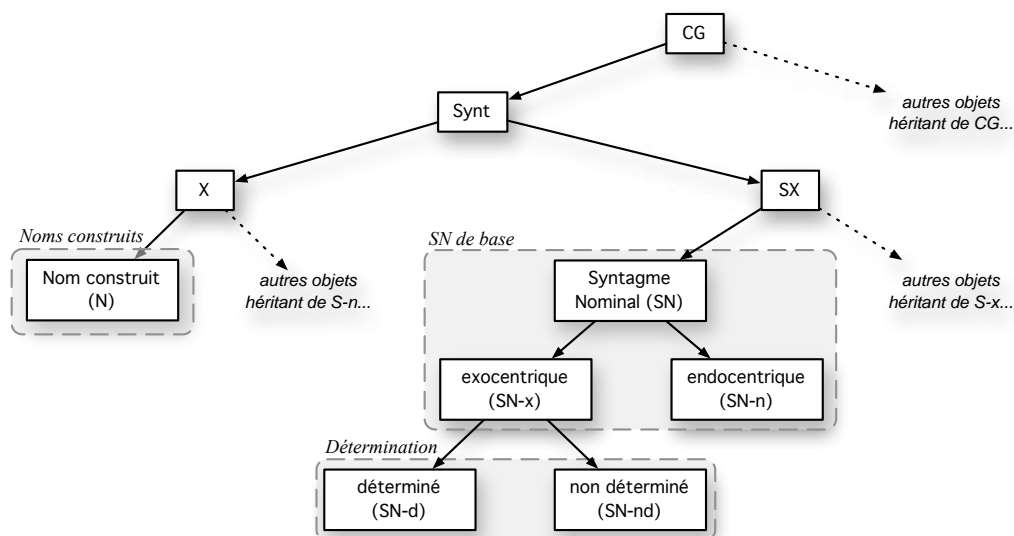


FIG. 11 – Hiérarchie des Noms construits dans la grammaire.

les descriptions en ajoutant des informations, et les constructions ont été pensées spécialement dans ce but-là, i.e. sans ignorer qu'elles manquent actuellement d'informations et ce que l'on pourra ajouter par la suite pour l'améliorer rapidement.

## 1 Noms construits (N)

Les Noms construits sont des constructions dont la tête est un objet de catégorie nominale, et qui restent sur le paradigme du nom (syntaxiquement simple) (19).

- (19) a. une **tarte maison** <sup>106</sup>  
 b. un **livre ennuyeux**, un **tout petit chagrin**, une **voiture facile à réparer**  
 c. un **type bien**  
 d. un **documentaire sur les médecines douces**, votre **réponse à ma demande**  
 e. **l'espion qui venait du froid**, **l'homme qui en savait trop**  
 f. **l'idée que Paul démissionne**  
 g. **l'idée de démissionner**  
 h. **L'endroit où l'on emprunte des livres** se dit bibliothèque en français et library en anglais et **l'endroit où l'on achète des livres** se dit librairie en français, mais bookstore en anglais. <sup>107</sup>  
 i. Il m'a été répondu que j'avais à demander ce dépôt à la **personne venue me remplacer**, ce que j'ai fait, (...) <sup>108</sup>  
 j. le **principe en** est simple <sup>109</sup>

<sup>106</sup>Sauf mention contraire, les exemples suivants sont tirés de [Riegel *et al.*, 1999], pp. 148-149.

<sup>107</sup>fr.comp.algorithmes, 20/03/2004.

<sup>108</sup>fr.misc.droit.immobilier, 15/11/2002.

<sup>109</sup>Celui-ci et suivants : [Le Goffic, 1993].

- k. j'**en** connais les **raisons**  
 l. il m'**en** reste le **souvenir**

Les contraintes qui s'appliquent à leur structure syntaxique ne sont pas très diversifiées, et ne demandent donc pas de hiérarchie d'héritage complexe, comme le montre la figure 11 puisqu'ils sont tous rassemblés en une seule construction. Celle-ci (fig. 12) hérite du X construit (X) : conformément à son héritage elle porte la catégorie et le type de l'objet Majeur indexé  $\boxed{1}$ , à qui l'on attribue la fonction syntaxique de tête, et qui fournit également à la construction ses informations syntaxiques.

N	
HERIT X	
INTR.	ID $\left[ \begin{array}{c} \text{NATURE} \mid \text{SCAT} \mid \boxed{1}.\text{SCAT} \end{array} \right]$
	SYN $\mid$ CPLT $\left[ \begin{array}{c} \text{RECTEUR} \mid \boxed{1} \\ \text{DÉPENDANT} \mid \boxed{2} \end{array} \right]$
Majeur	$\boxed{1}N$
Mineur	$\boxed{2}(N \vee \text{Adj} \vee \text{Adv} \vee \text{SPrep} \vee \text{Conj-s} \vee \text{Pro-pc}.\text{[GRAPH en]})$
Unicité	$\boxed{2}$
Précédence	$\boxed{1} \prec \neg(\text{Adj-q} \vee \text{en})$
Adjacence	$\boxed{1} \leftrightarrow \boxed{2}.\neg \text{en}$
Accord	$\boxed{1}.\text{GENRE} \rightsquigarrow \boxed{2}.\text{ACCORD} \mid \text{GENRE}$
	$\boxed{1}.\text{NOMBRE} \rightsquigarrow \boxed{2}.\text{ACCORD} \mid \text{NOMBRE}$

FIG. 12 – Le Nom construit (N).

A ces informations héritées, on ajoute ici qu'elle récupère également le trait de sous-catégorie (SCAT) de ce même objet. Enfin, on précise la nature de cet objet Majeur dans la propriété correspondante, en indiquant qu'il s'agit d'un objet de catégorie N.

Dans les propriétés extrinsèques, on mentionne également la présence du second constituant, Mineur cette fois, qui récupère l'index  $\boxed{2}$ . Cet objet pouvant être de natures diverses, on liste ici les différentes catégories possibles de ce constituant : N (19a), Adjectif (19b), Adverbe (19c), SPrep (19d) Subordonnée relative (19e), complétive (19f), infinitive (19g), conjonctive (19h) ou participiale (19i), ou encore pronom clitique de la forme *en* (19j, 19k, 19l) <sup>110</sup>. La mention de l'Unicité de ce constituant, ainsi que son Exigence par l'objet Majeur (héritée du X construit) impliquent qu'il devra figurer précisément un Mineur dans une telle construction (pas moins sous peine de violer l'Exigence, pas plus sous peine de violer l'Unicité), si bien que dans un N du type *vieil homme qui habite ici* on aura une structure hiérarchisée telle qu'illustrée en figure 13.

Concernant la position relative de ces deux objets maintenant, on indique que le Majeur précède les objets qui ne sont pas d'étiquette Adj-q (adjectif qualificatif) ni de la forme *en* (19j, 19k, 19l) : cela signifie qu'un objet qui sera soit d'étiquette Adj-q, soit de forme *en* ne

<sup>110</sup>Ce pronom ne fera donc pas partie des constituants de constructions verbales décrites au chapitre 7. Cela se justifie par le fait qu'il s'agit bien là d'un pronom qui occupe une fonction syntaxique de complément du Nom dans le SN, et n'a pas de fonction dans la construction verbale sur laquelle il "monte".

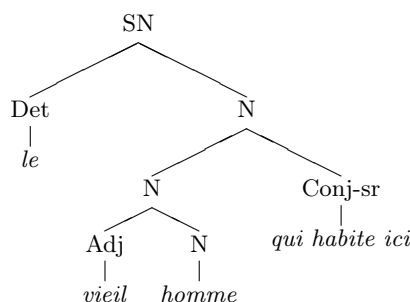


FIG. 13 – Exemple de hiérarchisation des compléments du Nom.

sera pas concerné par cette contrainte (à juste titre puisque dans ces deux cas les constituants ne sont pas ordonnés <sup>111</sup>). On précise aussi que les deux constituants doivent être adjacents, sauf dans le cas de *en* qui pourra figurer à distance du nom.

Enfin, on donne les deux contraintes d'accord s'exerçant entre les deux objets, et selon lesquelles l'expansion s'accorde en genre et en nombre avec le nom (si toutefois elle porte les traits d'accord correspondants, évidemment). A ce propos, le fait de formuler les informations de genre et de nombre en distinguant ceux qui font partie des TRAITS et ceux qui font partie des ACCORDS, comme on l'a présenté au chapitre 3, permet de ne contraindre l'accord que pour les objets qui portent les informations de genre et de nombre en tant que second et non que premier : ainsi, dans un nom construit tel que (20), où l'expansion du Nom, qui est elle-même un Nom (*maison*), porte ces informations dans ses TRAITS et non dans les ACCORDS, la contrainte ne s'appliquera pas.

- (20) a. un **soufflé maison**  
 b. des **soufflés maison**

## 2 Syntagmes Nominaux (SN)

La seconde sorte de constructions nominales est constituée des divers Syntagmes Nominaux. Elle n'a pas d'autre relation avec les Noms construits que de porter le même TYPE, puisqu'on a pu voir en fig. 11 que les SN héritent des Syntagmes X et non des X construits, ce qui implique une différence de forme dès la base.

Les SN sont de constitution plus variable que les Noms construits que l'on vient de voir. Tout d'abord, ils peuvent être **endocentriques** c'est-à-dire constitués d'un SN et d'un complément (§2.1.3), ou bien **exocentriques** c'est-à-dire constitués d'une tête qui n'est pas un SN

<sup>111</sup>On pourra dans un développement ultérieur préciser les contraintes d'ordre qui s'exercent sur les adjectifs qualificatifs. Il serait par exemple tout à fait envisageable d'attribuer un information du même ordre que le trait de "poids" proposé par [Abeillé & Godard, 1999].



(§2.1.2). Dans ce dernier cas, sont considérés comme des SN d’une part les constructions qui sont spécifiées par un Déterminant (on les appelle SN déterminés, §2.2.1), d’autre part les constructions dont la tête est une construction nominale et qui change de paradigme (on les appelle SN non déterminés, §2.2.2).

Nous commençons cette section par une présentation de la distinction entre SN exo- et endocentriques (§2.1), puis nous décrivons les constructions correspondant aux SN déterminés et non déterminés (§2.2).

## 2.1 SN de base

L’ensemble des Syntagmes Nominaux fait tout d’abord l’objet d’une construction SN générale, qui hérite directement du Syntagme X. De celle-ci, qui ne porte qu’une information minimale pour l’instant héritent les deux constructions SN exocentrique (SN-x) et endocentrique (SN-n).

### 2.1.1 Syntagme Nominal (SN)

Cette construction (fig. 14) regroupe les caractéristiques générales de tous les SN, quelle que soit leur forme, y compris les “SN unaires” (canoniquement, les pronoms disjoints ou les noms propres), même si ces derniers ne sont pas à proprement parler des “syntagmes” comme pourraient le faire remarquer [Bloomfield, 1961] ou [Creissels, 2004] puisqu’ils ne sont constitués que d’un unique objet.

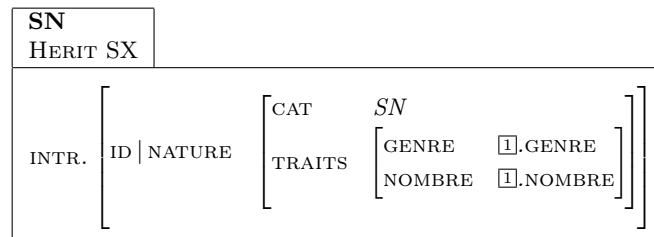


FIG. 14 – Le Syntagme Nominal (SN).

En l’état, leurs caractéristiques générales (outre leur héritage) se limitent aux seules mentions que ces objets portent la catégorie SN, et récupèrent le genre et le nombre de leur tête (indexée ⌈1⌋ plus haut dans la hiérarchie). L’une des informations que l’on pourrait envisager d’ajouter dans cette construction, et qui justifierait l’existence de cette construction générale bien mieux que la seule mention de sa catégorie, serait un ensemble d’informations syntaxiques propres aux syntagmes nominaux. On pense par exemple au fait qu’ils sont le support de l’introduction dans une production d’entités auxquelles on peut faire référence par la suite (entités nommées notamment).

### 2.1.2 Syntagme Nominal exocentrique (SN-x)

Le SN exocentrique (21) hérite du SN (fig. 15).

- (21) a. **Pierre** va au cinéma  
 b. **cette jolie maison** me plaît beaucoup  
 c. **Lui** il agit comme **un bourin** <sup>112</sup>

<b>SN-x</b>
HERIT SN

FIG. 15 – Le Syntagme Nominal exocentrique (SN-x).

Il ne reçoit aucune spécificité : tout ce qui lui est caractéristique est précisé dans les deux constructions qui en héritent : le SN déterminé d'une part, le non-déterminé d'autre part.

### 2.1.3 Syntagme Nominal endocentrique (SN-n)

Le SN endocentrique (22) hérite lui aussi du SN général.

- (22) a. elle (...) ne demande rien à personne et certainement pas à **lui, cet autre dont il se trouve qu'elle partage le lit.** <sup>113</sup>  
 b. (...) les services de renseignements français sont bien au courant de **cet homme, dont j'ai dit dans mon livre que "si la police française voulait l'arrêter elle possédait toutes les preuves pour le retrouver".** <sup>114</sup>  
 c. pas folle la guêpe  
 d. il y avait **environ une vingtaine de personnes**

<b>SN-n</b>	
HERIT SN	
Majeur	[1]SN
Mineur	[2](SN ∨ Adj ∨ Adv-d ∨ SPrep ∨ Conj-s)
Unicité	[2]
Adjacence	[1] ↔ [2]
Accord	[1].GENRE ∼ [2].ACCORD   GENRE [1].NOMBRE ∼ [2].ACCORD   NOMBRE

FIG. 16 – Le Syntagme Nominal endocentrique (SN-n).

Il s'agit (fig. 16) d'une construction constituée d'un SN (tête) et d'une expansion. Celle-ci, qui est généralement du même ordre que les compléments possibles du N, peut être placée indifféremment avant ou après le SN (donc on ne précise pas de contrainte de précédence), mais y est Adjacente. Elle fait également l'objet des mêmes contraintes d'Accord en genre et nombre que le Mineur des Noms construits.

<sup>112</sup>fr.soc.politique, 1er juillet 2006.

<sup>113</sup>fr.rec.arts.litterature, 3 novembre 2002.

<sup>114</sup>fr.soc.complots, 12 décembre 2003.

## 2.2 Détermination

Les SN exocentriques peuvent être déterminés (§2.2.1) ou non (§2.2.2) ; et suivant s'ils le sont ou pas ils portent des contraintes différentes.

### 2.2.1 Syntagme Nominal déterminé (SN-d)

Le SN déterminé (23) hérite du SN exocentrique.

- (23) a. le petit chien de ma grand-mère  
 b. le vrai, le rouge, une blonde, un rapide, le prêt à porter <sup>115</sup>  
 c. le moi, le ça, un rien  
 d. On peut être en désaccord avec **les ceux qui gèrent l'Usenet français** sans user d'un tel vocabulaire qui n'apporte rigoureusement rien au débat. <sup>116</sup>  
 e. le boire, le manger, un militant, les assiégés  
 f. les avants, un contre  
 g. un sans faute  
 h. le pourquoi et le comment, le dessus, le dessous  
 i. des si et des mais  
 j. un je ne sais quoi, un m'as-tu-vu  
 k. Et si on arrêta avec les révélations spectaculaires ? **les qui est qui ? les qui est avec qui ?** <sup>117</sup>  
 l. je crois plus **les qui ont des diplômes que les qui en ont pas (...)** <sup>118</sup>  
 m. **Des où elle ne parle QUE de cul** (et plus gras, plus vulgaire, plus lamentable, tu ne trouveras jamais !), **des où elle s'attaque à ses interlocuteurs ...** <sup>119</sup>

Il rassemble les SN portant un spécifieur, que l'on fait figurer dans les Majeurs et que l'on indexe [2] <sup>120</sup>. Ce dernier se place en tout début de syntagme, est adjacent à la construction qui contient la tête et s'accorde en genre et nombre avec elle (fig. 17).

La tête d'un SN-d peut être bien sûr un nom (23a) mais aussi un adjectif (23b), un pronom (23c, 23d), un verbe (23e), une préposition (23f), un syntagme prépositionnel (23g), un adverbe (23h), une conjonction (23i), une proposition (23j, 23k), une subordonnée (23l, 23m)... en fait, n'importe quel objet peut être tête de SN-d dès lors qu'il est spécifié par un déterminant : c'est la raison pour laquelle on n'avait pas précisé la tête du SN jusqu'ici, et que l'on ne la précise toujours pas.

**Genre et nombre.** — Canoniquement, le SN porte les valeurs de genre et nombre de sa tête (c'est d'ailleurs ce que l'on avait spécifié plus haut dans la hiérarchie). Mais puisque

<sup>115</sup>Sauf mention contraire, ces exemples sont tirés de [Riegel *et al.*, 1999], pp.148 et 546.

<sup>116</sup>fr.usenet.abus.d, 28 sept 2000.

<sup>117</sup>fr.soc.alcoolisme, 1 avr 2004.

<sup>118</sup>fr.soc.religion, 11 mai 2001.

<sup>119</sup>fr.soc.politique, 24 oct 2003.

<sup>120</sup>A ne pas confondre avec le [2] du SN endocentrique indiqué précédemment, qui n'a pas de relation avec celui-ci puisque notre SN déterminé n'hérite pas du SN endocentrique, et aucun objet n'avait été indexé [2] dans les descriptions des constructions dont il hérite.

SN-d	
HERIT SN-x	
INTR.	$\left[ \begin{array}{l} \text{ID} \left[ \begin{array}{l} \text{NATURE} \left[ \begin{array}{l} \text{TRAITS} \left[ \begin{array}{l} \text{GENRE} \quad \boxed{2}.\text{GENRE} \\ \text{NOMBRE} \quad \boxed{2}.\text{NOMBRE} \end{array} \right] \right] \end{array} \right] \end{array} \right]$
	$\left[ \begin{array}{l} \text{SYN} \left[ \begin{array}{l} \text{SPÉCIFIEUR} \quad \boxed{2} \end{array} \right] \end{array} \right]$
Majeur	
Précédence	
Adjacence	
Accord	
$\boxed{2} \text{Det}$ $\boxed{2} \prec \boxed{1}$ $\boxed{2} \leftrightarrow \boxed{1}$ $\boxed{1}.\text{GENRE} \rightsquigarrow \boxed{2}.\text{ACCORD}   \text{GENRE}$ $\boxed{1}.\text{NOMBRE} \rightsquigarrow \boxed{2}.\text{ACCORD}   \text{NOMBRE}$	

FIG. 17 – Le Syntagme Nominal déterminé (SN-d).

concrètement dans un SN-d ces valeurs peuvent être moins précises que celles de son spécifieur (24),

- (24) a. un élève, une élève  
b. le choix, les choix

alors on pose ici qu'il portera la valeur de genre (resp. de nombre) issue de l'unification des valeurs de genre (resp. de nombre) de sa tête et de son spécifieur (formellement, on se contente d'ajouter une référence à ce dernier dans la structure, puisque l'on a dit que l'héritage s'opérait par unification des structures de traits ; or le SN, dont celui-ci hérite, porte déjà la référence à la tête dans ses traits de genre et de nombre). Cela permet de déterminer également le genre et le nombre, réels, des SN dont la tête ne porte pourtant pas de genre (25) ni/ou de nombre (26).

- (25) a. (...) mais pas de la part d'un **m'as-tu-vu** grand format qui se prend pour un Pic de la Mirandole. <sup>121</sup>  
b. Très pénible si c'est aussi **une m'as-tu-vu**. <sup>122</sup>  
c. Nous ne sommes pas tous **des m'as-tu-vu** :-). <sup>123</sup>
- (26) a. J'ai expliqué en DETAIL **le pourquoi** et **le comment**, et Mr. Nightcat et vous-même en avez pris positivement acte, (...). <sup>124</sup>  
b. (...) des rapports de police sur **les pourquoi** et **les comment** d'affaires qui ne sont a l'arrivée que des bagatelles. <sup>125</sup>

**Les formes *N de N*.** — Les SN dont le constituant Majeur indexé  $\boxed{1}$  est d'une forme *N de N* du type (27),

<sup>121</sup>soc.culture.belgium, 3 nov 2002.

<sup>122</sup>fr.lettres.langue.francaise, 28 fév 2003.

<sup>123</sup>soc.culture.belgium, 8 juil 2003.

<sup>124</sup>soc.culture.belgium, 20 jan 2001.

<sup>125</sup>fr.soc.homosexualite, 23 jan 2001.

- (27) a. une bouteille de bière  
 b. un kilo de ces pommes  
 c. ton imbécile de frère  
 d. cet imbécile de Franck

ont une structure syntaxique identique à celles du type (28).

- (28) le chapeau de papa

En effet, dans (27) c'est bien le premier N (*bouteille, kilo, imbécile*) qui est la tête syntaxique du SN : c'est avec lui que s'accordent les autres éléments de l'énoncé le cas échéant (p.ex. le verbe si le SN en question est sujet). Leur différence est sémantique et non pas syntaxique, puisque par contre le gouverneur sémantique de la construction est alors *bière, pommes, frère* et *Franck* (alors que cela restait *chapeau* dans (28)) : ce sont eux qui gèrent la sélection lexicale du reste de la construction notamment ([Flaux & Van de Velde, 2000]). Elles pourraient donc faire l'objet de l'introduction d'une construction spécifique qui hériterait du N construit et dont les spécificités seraient sémantiques.

### 2.2.2 Syntagme Nominal non déterminé (SN-nd)

Cette construction (29) hérite du SN exocentrique (SN-x).

- (29) a. Après une ou deux heures de **mauvais sommeil**, **Pierre** se réveille et se lève. <sup>126</sup>  
 b. **Maison** dans **village calme avec école, commerces, proche de Vittel** <sup>127</sup>  
 c. Les pages me semblent d'un intérêt très limité (**raison pour laquelle je n'ai pas voté**). <sup>128</sup>  
 d. c'est quand même appréciable, comme **petit plus** <sup>129</sup>  
 e. Les autres le voient (**accès aux ressource, connexion lecteur, tout** fonctionne) mais **lui** ne se voit pas et ne voit pas les autres <sup>130</sup>  
 f. **Debian** est pratique sur ce point car il permet d'ajuster très finement ton installation en n'installant que **ce dont tu as besoin**. <sup>131</sup>  
 g. Ils pensent à **ce que eux ont vécu dans leur pays** mais nous, **ce que nous avons vécu**, ils ne veulent pas le savoir, **eux**, ils sont mieux maintenant. <sup>132</sup>

SN-nd	
HERIT SN-x	
Majeur	$\boxed{1}(N \vee \text{Pro.}[\text{CLITIQUE } \textit{non}] \vee \text{Adj} \vee \text{Adv-d})$

FIG. 18 – Le syntagme nominal non déterminé (SN-nd).

<sup>126</sup>fr.education.entraide, 3 juin 2001.

<sup>127</sup>fr.petites-annonces.immobiler, 26 février 2003.

<sup>128</sup>fr.comp.infosystemes, 21 décembre 1995.

<sup>129</sup>fr.rec.anime, 14 juin 2000.

<sup>130</sup>fr.comp.reseaux.internet, 13 juillet 2000.

<sup>131</sup>linux.debian.user.french, 27 mai 2000.

<sup>132</sup>qc.politique, 2 août 2000.

Au contraire de la précédente, elle ne reçoit pas de spécifieur (fig. 18). La catégorie de la tête est alors plus restreinte que celle du SN-d ; il s'agit typiquement de noms propres ou de pronoms non clitiques dans leurs emplois standard, mais cela peut également concerner des noms communs, des adjectifs ou des adverbes non clitiques comme on le voit dans les exemples (29). La contrainte syntaxique s'exerçant sur un SN et exigeant qu'il prenne une forme déterminée ou non ne dépend pas de ses constituants mais lui est externe et provient des constructions dont il est le constituant (p.ex., dans un syntagme prépositionnel certaines prépositions imposent un SN-nd, d'autres un SN-d) : on ne les précise donc pas ici. La tête étant le seul constituant de cette construction, on n'ajoute pas de propriétés supplémentaires.

### Résumé

Dans ce chapitre on a décrit les constructions dites “nominales”, i.e. portant le trait [TYPE N]. Parmi elles figurent les Noms construits qui héritent des X construits, et les Syntagmes Nominaux qui héritent des Syntagmes X.

Les **Noms construits** (N) sont des constructions endocentriques dont la tête est un objet de catégorie N, et qui reçoit un (et un seul) complément pouvant prendre des formes variées (typiquement, *petit chien*).

Les **Syntagmes Nominaux** (SN) sont des constructions pouvant être endo- ou exocentriques. S'ils sont endocentriques ils sont constitués d'un SN et d'une expansion (typiquement *lui, ce grand homme*) et s'ils sont exocentriques ils peuvent être spécifiés par un déterminant (*un chien*) ou non (*Pierre*).

A travers la description de ces constructions, on a mis en avant le parallélisme recherché dans notre grammaire, selon lequel on trouvera dans un nombre assez important de types de constructions, une partie qui héritera de X, une autre de SX, et dont les caractéristiques générales seront assez stables : ceux héritant de X seront endocentriques et porteront la même catégorie que leur type, ceux héritant de SX seront soit endocentriques, soit exocentriques et dans ce cas soit spécifiés, soit non.





## Chapitre 7

# Constructions verbales

Ce chapitre traite de toutes les constructions dont la tête est un verbe ; chacune d’elles comporte donc le trait correspondant [TYPE V], tout en pouvant être de catégorie variable : nous faisons une distinction entre d’une part les **Verbes construits** (§1), constructions endocentriques qui rassemblent en un syntagme les verbes (formes simples et composées) et leurs dépendants clitiques (pronoms et adverbes de négation), et d’autre part les **Syntagmes Verbaux** (§2), constructions exocentriques constituées d’un verbe (construit le cas échéant) et de ses éventuels compléments non clitiques, puis de ses circonstants le cas échéant.

### 1 Verbes construits (V)

Les Verbes syntaxiquement construits correspondent, d’une manière générale, aux objets constitués d’un verbe (non auxiliaire) et de ses dépendants directs : auxiliaires, modaux et clitiques (i.e., pronoms et adverbes de négation). On en donne la hiérarchie (locale) en figure 19. Ils héritent du X construit.

**Equivalents dans d’autres théories.** — Ce que l’on appelle “Verbe construit” ici correspond à un ensemble d’objets que l’on retrouve dans d’autres grammaires : il s’agit à peu de choses près de la même unité que l’**amas verbal** de l’approche topologique de [Gerdes & Kahane, 2004], à ceci près que dans leur proposition chaque verbe compose un amas verbal, qui peut lui-même être constituant d’un amas verbal supérieur (ce qui n’est pas notre cas pour les auxiliaires dans notre proposition).

Notre V construit se rapproche aussi de la définition du **groupe verbal** de la TAG d’[Abeillé, 2002, p. 119], qui dit :

*Nous concluons à la non-pertinence pour le français d’une syntagme verbal classique regroupant verbe et compléments. Nous lui préférons un groupe verbal plus restreint (et sans doute plus traditionnel), comprenant modaux ou aspectuels, pronoms clitiques (plus ne) et formes verbales (simples ou composées), essentiellement sur la base de critères phonétiques.*

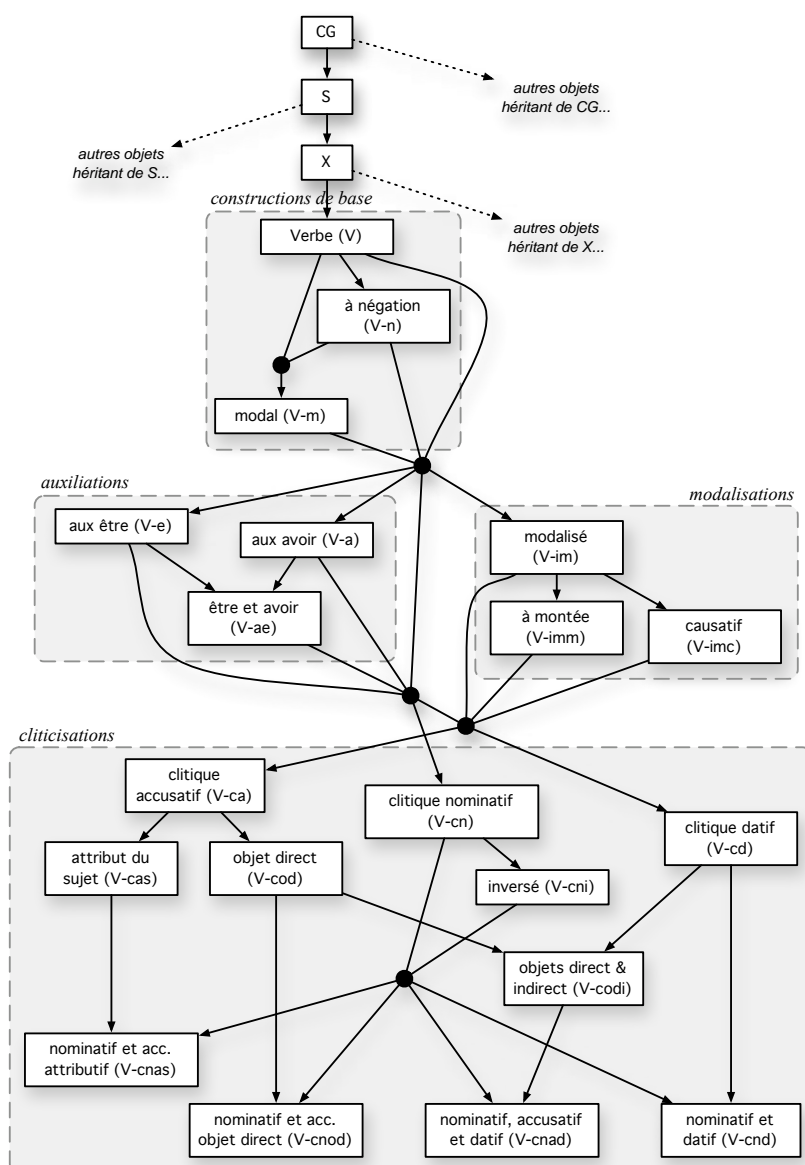


FIG. 19 – Hiérarchie des Verbes construits (V) dans la grammaire.

La différence principale avec notre approche cette fois est que nous maintenons malgré cela les SV, qui sont décrits dans la section suivante, et dont les V sont des constituants.

En HPSG, [Miller & Sag, 1995] posent que les ensembles V + clitiques sont "*traités comme des réalisations à valence réduite de lexèmes verbaux*" <sup>133</sup>, ce qui est proche de notre proposition, à la différence qu’eux, suivant leur approche lexicaliste, les considèrent comme des entrées lexicales là où nous les considérons comme des entrées grammaticales.

<sup>133</sup>"*treated as valence-reduced realizations of verbal lexemes*"

Enfin, p.ex. [Creissels, 2004] considère également les pronoms clitiques comme des **affixes verbaux** : ils peuvent être préfixes ou suffixes, ont un ordre rigoureusement régi, sont attachés à la forme tensée (c’est-à-dire au support des marques de construction, temps, aspect, personne). [Blanche-Benveniste, 1975] appelle **mot syntaxique** cet ensemble verbe+affixes. Cela correspond aussi à la **phrase-mot** de [Guillaume, 1990].

**La construction grammaticale Verbe (V).** — La construction générale correspondant aux verbes syntaxiquement construits (fig. 20) est DOMAINE syntaxique (héritage de CG), et porte deux constituants immédiats, le premier étant un Majeur indexé  $\boxed{1}$  (héritage de CG) et le second un Mineur indexé  $\boxed{2}$  (héritage de X) et Exigé par  $\boxed{1}$  (héritage de X). L’élément Majeur (ou la TÊTE de celui-ci le cas échéant) est la TÊTE syntaxique de la construction (héritage de Synt), et lui fournit sa CATÉGORIE (héritage de X), son TYPE (héritage de Synt) et ses informations syntaxiques (héritage de CG).

A ces informations héritées, on ajoute dans les informations extrinsèques une précision sur la valeur de l’objet Majeur, en indiquant qu’il doit être de CATÉGORIE *V* et de SCAT non *auxiliaire* (i.e. *principal* ou *modal*) ; dans les propriétés intrinsèques, on ajoute qu’il fournit sa sous-catégorie (SCAT) à la construction.

<b>V</b>	
HERIT X	
INTR. $\left[ \text{ID} \mid \text{NATURE} \quad \left[ \text{SCAT} \quad \boxed{1}.\text{SCAT} \right] \right]$	
Majeur	$\boxed{1} \left[ \begin{array}{cc} \text{CAT} & V \\ \text{SCAT} & \neg(\text{aux-être} \vee \text{aux-avoir}) \end{array} \right]$

FIG. 20 – Le Verbe construit (V).

**Valence vs. transitivité.** — Dans notre grammaire, nous faisons une distinction entre “valence” et “transitivité” en nous basant sur les descriptions suivantes ([Creissels, 2004, Chap. 12, p. 1]) :

*Le terme de valence verbale caractérise de la manière la plus générale les propriétés combinatoires des verbes, des deux points de vue sémantique et syntaxique :*

- *sémantiquement, la définition du signifié de chaque verbe implique la mention de la participation d’un nombre déterminé d’arguments, entités qui jouent chacune un rôle précis dans le type d’événement ou de situation auquel se réfère le verbe ; en outre, le signifié du verbe conditionne plus ou moins la possibilité d’introduire dans la construction de l’unité phrastique des mots ou constituants ajoutant tel ou tel type de précision, notamment la mention d’autres entités qui peuvent participer à l’événement, ou des circonstances de l’événement ; par exemple, couper dans son acception la plus courante implique sémantiquement deux arguments, un agent et un patient, mais le sens de ce verbe suggère aussi l’intervention possible d’un instrument, ce qui n’est pas le cas par exemple pour un verbe comme oublier ;*
- *syntactiquement, chaque constituant nominal qui peut entrer dans la construction d’un verbe avec un rôle sémantique déterminé doit présenter des caractéristiques morphosyntaxiques précises, qui généralement le distinguent des autres constituants nominaux susceptibles d’entrer dans la construction du même verbe avec des rôles sémantiques différents.*

Nous posons donc que la **valence** est une propriété **lexicale** des objets linguistiques, qui indique simplement quel est le nombre d'actants que ceux-ci attendent. On la représente donc comme un attribut sémantique dont la valeur peut être nulle (pour les verbes impersonnels : *il pleut*, ou pour les verbes à montée : *Pierre semble dormir*), ou attendre une référence à un actant (*pierre dort*), à deux actants (*je mange une pomme*, *je vais à la plage*), à trois actants (*je donne une pomme à Pierre*), ou à quatre actants (pour les constructions causatives : *Pierre fait donner une pomme à Marie par Paul*).

*En toute rigueur, transitif / intransitif qualifie des constructions, et les verbes ne peuvent être qualifiés de transitifs ou intransitifs que par un raccourci de langage dont il faut être conscient qu'il peut générer des confusions. [Creissels, 2004, Chap. 12, p. 2]*

La **transitivité**, par contre, est une propriété des **constructions syntaxiques** : il s'agit du fait, pour une construction de type verbal (i.e., un verbe construit ou un SV), de recevoir un ou plusieurs compléments de verbe. On ne tient pas compte du sujet dans le cas de la transitivité, alors qu'on le fait figurer dans les actants (donc dans la valence).

Cette distinction entre les deux fait qu'il est non seulement possible qu'un même verbe ait plusieurs valences (30),

- (30) a. L'homme a **pris** une cigarette  
b. L'infirmière lui a **pris** sa température  
c. L'hortensia que j'avais planté le mois dernier n'a pas **pris**

mais aussi qu'un verbe portant une valence donnée puisse être construit selon plusieurs transitivités (31).

- (31) a. Je **mange** de la soupe / Il **mange** mais il ne grossit pas  
b. La compagnie de transport a **baissé** le prix des tickets de bus / Le prix des tickets de bus a **baissé**

Plus précisément, nous suivons [Touratier, 2004, chap. 4, pp. 41-42] dans sa distinction entre ces deux notions :

*Nous préférons donc distinguer le sémantique du syntaxique, et voir dans la transitivité d'un verbe une propriété exclusivement syntaxique, en vertu de laquelle un verbe est construit avec un ou plusieurs compléments de verbe, que ceux-ci soient des compléments directs ou indirects d'objet. Par contre, en ne suivant peut-être pas la lettre, mais du moins l'esprit de Tesnière, nous appellerons la valence du verbe le fait d'avoir sémantiquement besoin, pour constituer un énoncé doué de sens, d'un nombre donné de participants au procès exprimé par le verbe, que Tesnière appelait les actants du verbe, et que les logiciens appellent les arguments (ou les variables) du prédicat.*

*[...] Mais non [sic] préciserons que si aimer est, syntaxiquement, transitif dans j'aime Rose et intransitif dans la maison où j'aime, il est, dans les deux cas, sémantiquement le même, à savoir un verbe bivalent. Le fait que son second actant ne soit pas exprimé dans la maison où j'aime n'a nullement pour conséquence sémantique que je n'aime pas quelqu'un de précis. Ceci permet de rejoindre l'intuition de la grammaire traditionnelle, qui parlait alors d'un verbe transitif employé intransitivement, sans la difficulté terminologique que cette façon de parler traditionnelle revient à dire curieusement qu'un verbe construit avec un complément d'objet (c'est-à-dire transitif) n'est pas construit avec un complément d'objet, ce qui peut paraître contradictoire.*

*Il faut ajouter que, contrairement à ce que semblait admettre Tesnière, il n'y a pas parallélisme entre la structure ou le schéma actanciel du verbe et sa structure syntaxique, même s'il y a entre ces deux types d'organisation un certain rapport. Un verbe monovalent est normalement, au point de vue syntaxique, un verbe intransitif, sauf si ce verbe est morphologiquement impersonnel, auquel cas le premier et seul actant du verbe devient nécessairement un complément d'objet du verbe, qui, alors, est donc transitif :*

- (32) a. Il souffle un vent terrible (Romain).  
 b. Dans le désert où il ne pousse rien, par manque d'eau (Maupassant) (d'après Grevisse & Goosse, 1993).

*Un verbe au moins bivalent est normalement une fois transitif, mais il peut être deux fois transitif, s'il est morphologiquement impersonnel, comme dans :*

- (33) Il est arrivé à Pierre une aventure extraordinaire.

*Mais la valence sémantique peut ne pas déterminer même partiellement la transitivité d'un verbe, dans la mesure où il est possible d'employer transitivement un verbe monovalent personnel, qui, normalement, devrait être toujours intransitif :*

- (34) a. N'ai-je pas sué la sueur de tes nuits ? (Verlaine).  
 b. Dormir son dernier sommeil.  
 c. Vivre sa vie.

Ceci explique pourquoi l'on note la valence comme information intrinsèque sémantique des objets lexicaux, et que l'on pourra parler à présent, dans les verbes construits, de l'information syntaxique de transitivité. Les deux entretiennent des rapports que nous expliciteront au fil de ce chapitre, mais restent à distinguer malgré tout.

## 1.1 Verbe à négation (V-n)

Le Verbe construit négatif (fig. 21) est constitué, en plus de sa Tête verbale, d'une négation clitique sous la forme de l'adverbe *ne* (étiqueté Adv-ng) et/ou d'un adverbe *pas* ou assimilé (*guère, plus, etc.*, étiquetés Adv-np). Pour qu'une construction V-n soit introduite, il faut qu'au moins l'un des deux adverbes de négation soit présent (autrement dit, il faut qu'un "adverbe négatif", Adv-n, soit présent) (35).

- (35) a. ne vais pas / ne vais / vais pas  
 b. ne pas aller / n'aller / pas aller  
 c. ne mange pas / ne mange / mange pas

Les deux adverbes de négation, Adv-ng et Adv-np, sont tous deux signalés comme Mineur sous l'étiquette générale Adv-n. D'une manière générale, *ne* est Unique et se place avant le verbe principal et l'Adv-np. Ses contraintes d'Adjacence dépendent de la présence d'éventuels autres constituants du Verbe construit (pronoms clitiques) et seront donc précisées dans les constructions correspondantes plus loin. L'Adv-np, également unique, est placé avant la Tête si celle-ci est infinitive, après si elle est à un autre mode.

<b>V-n</b> HERIT V	
INTR. $\left[ \text{SYN} \left[ \text{NÉGATION} \left[ \begin{array}{cc} \text{RECTEUR} & \boxed{1} \\ \text{DÉPENDANT} & \text{Adv-n} \end{array} \right] \right] \right]$	
Mineur	Adv-n
Unicité	Adv-ng Adv-np
Exigence	$\boxed{1} \Rightarrow \text{Adv-n}$
Précédence	Adv-ng $\prec \boxed{1}$ Adv-ng $\prec$ Adv-np Adv-np $\prec \boxed{1}.$ [MODE <i>infinitif</i> ] $\boxed{1}.$ [MODE $\neg$ <i>infinitif</i> ] $\prec$ Adv-np

FIG. 21 – Le Verbe à négation (V-n).

## 1.2 Verbes modaux (V-m)

On ne décrit dans cette construction que les verbes modaux qui sont syntaxiquement construits (cela signifie que ceux-ci portent la même étiquette que les verbes modaux qui ne sont pas syntaxiquement construits) ; on n’y évoque pas ceux qui ne sont pas construits, ou alors qui le sont du point de vue morphologique (36).

- (36) a. (...) et maintenir une collection de CD de patch (que tu fais toi-même) que tu dois installer avant de **pouvoir** aller sur le web, (...) <sup>134</sup>  
b. (...) la communication sous toutes ses formes me **paraît** faire partie d’un tel processus. <sup>135</sup>

Concrètement, le V-m ici décrit ne concerne que les modaux constitués d’un verbe modal et d’une préposition exigée par le verbe (37).

- (37) a. Et bien peut-être que c’est qu’Apple **ne vient pas de** faire l’erreur de sa vie, justement... <sup>136</sup>  
b. Nous **sommes sur le point de** gagner. <sup>137</sup>  
c. À quel âge faut-il **commencer à** brosser les dents ? <sup>138</sup>

Tel qu’il est décrit ici, le V-m ne réclame pas de verbe principal auquel être relié ; c’est le verbe principal qui pourra l’exiger mais pas le contraire. Toutefois, l’introduction d’un Verbe modal n’aura de pertinence que si la construction entre elle-même dans la constitution d’une construction de niveau supérieur, or il n’y a que les Verbes modalisés (V-im, §1.5.1) qui mentionnent le V-m, donc il ne pourra être constituant que de telles constructions.

Il est bien évidemment possible de raffiner cette construction afin de préciser quelle préposition est autorisée/exigée par quel modal, mais nous ne rentrons pour l’instant pas plus dans le détail ici.

<sup>134</sup>fr.comp.os.linux.debats, 18 février 2005.

<sup>135</sup>fr.soc.religion, 20 avril 2002.

<sup>136</sup>fr.comp.sys.mac, 6 septembre 1997.

<sup>137</sup>fr.bio.medecine, 17 mai 2006.

<sup>138</sup>fr.bio.medecine, 5 mai 2000.

V-m	
HERIT V ; V-n	
INTR.	$\left[ \text{SYN} \left[ \text{INTRODUCTEUR} \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{1} \\ \text{DÉPENDANT} \quad \text{Prep} \end{array} \right] \right] \right]$
Majeur	$\boxed{1}$ V-m
Mineur	Prep
Unicité	Prep
Exigence	$\boxed{1} \Rightarrow \text{Prep}$
Précédence	$\boxed{1} \prec \text{Prep}$
Adjacence	$\boxed{1} \leftrightarrow \text{Prep}$

FIG. 22 – Le Verbe modal (V-m).

### 1.3 Verbes à contrôle (V-c)

Dans notre proposition, nous ne considérons pas les verbes à contrôle (38) comme des modaux.

- (38) a. Pierre **essaie** de vendre sa voiture  
 b. Marie le **force** à vendre sa voiture  
 c. \* il **essaie** de pleuvoir

Il s'agit de verbes de SCAT *principale*, dont la caractéristique d'être un verbe "à contrôle" est d'origine lexicale, et non syntaxique. Cette caractéristique est donc notée dans l'entrée lexicale V-c sous l'attribut TYPE-PRINCIPAL (qui peut prendre pour valeur *contrôle* ou *non-contrôle*), et le V-c ne fait pas l'objet d'une construction grammaticale.

Nous considérons que les **verbes à contrôle du sujet** sont des verbes bivalents<sup>139</sup> et dans ce cas le premier actant de la tête du complément sera le sujet du verbe à contrôle. Autrement dit, dans l'exemple (38a) *Pierre* est, syntaxiquement, le sujet de *essaie*, et sémantiquement il est le premier actant à la fois de *essaie* et de *vendre*. Dans (38c) *il* satisfait (à juste titre) toutes les propriétés syntaxiques pour être le sujet de *essaie*, de même que les propriétés sémantiques pour être premier actant de *pleuvoir*, mais il ne satisfait pas les propriétés sémantiques pour être le premier actant de *essaie*.

Les **verbes à contrôle de l'objet** sont des verbes trivalents, et dans ce cas le premier actant de la tête de leur complément indirect est le complément direct du verbe à contrôle. Dans l'exemple (38b), *le* est, du point de vue syntaxique, l'objet direct de *force*, et du point de vue sémantique il est à la fois le deuxième actant de *force* et le premier actant de *vendre*.

Cette analyse se distingue de celle de celle de GB (p.ex. [Chomsky, 1981]) en ce qu'on ne fait pas coïncider les structures syntaxiques : là où GB analyserait le sujet (syntaxique) d'un verbe à contrôle comme également sujet (syntaxique) du verbe infinitif, nous n'affectons pas de sujet à ce dernier ; nous nous contentons de lui affecter un actant (sémantique), parce que nous considérons que l'infinitif, syntaxiquement parlant, ne porte pas plus de sujet dans les

<sup>139</sup>On verra les contraintes précises s'exerçant sur leurs compléments dans la section sur le SV, §2.

constructions à contrôle que dans les autres types de constructions.

Elle se distingue aussi des approches telles que celles de LFG (dans [Bresnan & Kaplan, 1982]) ou de HPSG (dans [Miller, 1991] ou [Baschung, 1998]) qui analysent l'ensemble constitué du verbe infinitif + ses compléments comme un SV, lui-même complément du verbe à contrôle (qui se voit affecter, comme dans GB, un sujet syntaxique) : nous analysons effectivement cet ensemble comme un SV, mais comme un SV ayant un statut de subordonnée.

Elle se rapproche en fait de l'analyse TAG proposée dans [Abeillé, 2002] où le complément dont le verbe infinitif est la tête est considéré comme une proposition infinitive (et non un SV), à la différence terminologique près que nous n'appelons pas ces constructions des "propositions".

## 1.4 Auxiliations

Les verbes au participe (V-p) peuvent soit passer sur le paradigme des Adjectifs (cf. alors chapitre 10), soit recevoir un ou plusieurs auxiliaires. Les Verbes à auxiliaire(s) héritent de l'une des constructions verbales que l'on vient de décrire (V, V-n ou V-m). Dans leur structure dépendancielle syntaxique, on pose que c'est la tête, donc le verbe principal, qui régit le ou les auxiliaire(s), en dépit du fait que c'est un auxiliaire qui porte les marques de mode, de personne et de nombre : en effet, du point de vue syntaxique c'est bien le verbe au participe qui régit la construction verbale, i.e. c'est lui qui contraint plus ou moins la sélection des compléments et du sujet (l'auxiliaire, en soi, n'a pas de valence ni de restrictions sémantiques lexicales, il ne fait que porter des marques morphologiques et syntaxiques de la construction verbale).

On a ici l'illustration de l'un des cas où notre approche se distingue des grammaires utilisant des *Principes* pour exprimer la remontée des traits : si l'on donnait à la Tête une valeur primordiale systématique dans nos constructions, alors on devrait soit poser que la Tête de l'ensemble auxiliaire + verbe principal est l'auxiliaire et alors pouvoir automatiquement faire remonter ses valeurs de mode, personne etc., tout en étant syntaxiquement incorrect puisque la transitivité de l'ensemble (puis éventuellement du SV dont il sera constituant) dépend du verbe principal et non de l'auxiliaire ; soit à l'inverse l'on poserait (comme on le fait) que la Tête est le verbe principal, mais on devrait alors trouver un moyen de contourner le *Principe des traits de tête* (ou assimilé) afin de faire remonter *une certaine partie* des traits de l'auxiliaire à la place de ceux de la tête. Ici, puisque l'on ne fait pas usage de Principes, la "règle générale" de remontée des traits est donnée par les constructions les plus hautes dans la hiérarchie d'héritage (et c'est équivalent au principe des traits de tête puisqu'on y indique — pour ce qui est des Syntagmes — que la construction récupère les traits syntaxiques et sémantiques de sa tête), et peut être complétée ou modifiée plus localement (i.e., plus bas dans la hiérarchie). On fait remonter ceux des traits de chacun des constituants qui sont pertinents pour la construction, sans pour autant perdre en généralité des informations grâce à l'héritage.

### 1.4.1 Verbe au participe passé avec auxiliaire *avoir* (V-a)

Le Verbe avec auxiliaire *avoir* (fig. 23) porte les traits et propriétés suivants en plus des



V-a HERIT V ∨ V-n ∨ V-m											
INTR.	<table> <tr> <td>ID   NATURE</td><td> <table> <tr> <td>TRAITS</td><td>MODE [2].MODE</td></tr> <tr> <td>ACCORD</td><td>PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE</td></tr> </table> </td></tr> <tr> <td>SYN</td><td> <table> <tr> <td>AUXILIATION</td><td>RECTEUR [1] DÉPENDANT [2]</td></tr> </table> </td></tr> </table>	ID   NATURE	<table> <tr> <td>TRAITS</td><td>MODE [2].MODE</td></tr> <tr> <td>ACCORD</td><td>PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE</td></tr> </table>	TRAITS	MODE [2].MODE	ACCORD	PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE	SYN	<table> <tr> <td>AUXILIATION</td><td>RECTEUR [1] DÉPENDANT [2]</td></tr> </table>	AUXILIATION	RECTEUR [1] DÉPENDANT [2]
ID   NATURE	<table> <tr> <td>TRAITS</td><td>MODE [2].MODE</td></tr> <tr> <td>ACCORD</td><td>PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE</td></tr> </table>	TRAITS	MODE [2].MODE	ACCORD	PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE						
TRAITS	MODE [2].MODE										
ACCORD	PERSONNE [2].PERSONNE NOMBRE [2].NOMBRE										
SYN	<table> <tr> <td>AUXILIATION</td><td>RECTEUR [1] DÉPENDANT [2]</td></tr> </table>	AUXILIATION	RECTEUR [1] DÉPENDANT [2]								
AUXILIATION	RECTEUR [1] DÉPENDANT [2]										
Majeur	[1]V-ps										
Mineur	[2][SCAT <i>aux-avoir</i> ]										
Unicité	[2]										
Précédence	[2] < [1] [2] < Adv-np Adv-ng < [2]										
Adjacence	[2] ↔ Adv-np ∨ [1] Adv-ng ↔ [2]										

FIG. 23 – Le Verbe au participe passé avec auxiliaire *avoir* (V-a).

caractéristiques dont il hérite :

- La TÊTE de la construction est un Verbe au participe passé (i.e., un V-ps).
- Elle exige un et un seul <sup>140</sup> verbe de SCAT *aux-avoir* (la propriété d'Exigence est héritée du X, on n'ajoute ici que l'Unicité et la précision du Mineur).
- A la différence des cas précédents, c'est ici l'auxiliaire qui fournit à la construction ses valeurs de PERSONNE, de NOMBRE et de MODE. Les autres valeurs (y compris TEMPS et GENRE) restent celles de la Tête comme décrit précédemment (39).

(39) Faut-il faire attention aux poissons qui **auraient mangé** des asticots (...) <sup>141</sup>

- L'auxiliaire Précède la tête. Si des négations sont présentes <sup>142</sup> (40), l'Adv-ng (*ne*) précède l'auxiliaire, l'Adv-np (*pas* etc.) le suit. L'auxiliaire est Adjacent à la tête ou à l'Adv-np d'un côté, et à l'Adv-ng de l'autre.

(40) Lequel **n'avait guère requis** de service, n'avait il pas sa disposition ce qu'il voulait... <sup>143</sup>

#### 1.4.2 Verbe au participe passé avec auxiliaire *être* (V-e)

<sup>140</sup> Ceci reste vrai même dans le cas d'une coordination ou d'une disfluence puisque celles-ci sont analysées sous forme d'Entassements Paradigmatiques, qui eux occupent la même place que des constituants non accumulés, cf. chapitre 11. Cette remarque est valable pour tous les cas similaires suivants.

<sup>141</sup> fr.bio.general, 31 décembre 2000.

<sup>142</sup> I.e., si le V-a hérite soit directement d'un V-n, soit d'un V-m qui hérite d'un V-n.

<sup>143</sup> fr.soc.histoire, 7 janvier 1999.

V-e	
HERIT V ∨ V-n ∨ V-m	
INTR.	<div> <div>ID   NATURE</div> <div> <div>TRAITS   MODE [3].MODE</div> <div>ACCORD   PERSONNE [3].PERSONNE</div> </div> </div>
	<div> <div>SYN</div> <div> <div>AUXILIATION</div> <div> <div>RECTEUR [1]</div> <div>DÉPENDANT [3]</div> </div> </div> </div>
Majeur	[1]V-ps
Mineur	[3][SCAT <i>aux-être</i> ]
Unicité	[3]
Exigence	[1] ⇒ [3]
Exclusion	<div>[3] <math>\not\leftrightarrow</math> [1].[LEXEME <i>avoir</i> ∨ <i>être</i>]</div> <div>[3] <math>\not\leftrightarrow</math> [1].[PERSONNE 0]</div>
Précédence	<div>[3] &lt; [1]</div> <div>[3] &lt; Adv-np</div> <div>Adv-ng &lt; [3]</div>
Adjacence	<div>[3] ↔ Adv-np ∨ [1]</div> <div>Adv-ng ↔ [3]</div>
Accord	[3].NOMBRE ∼ [1].NOMBRE

 FIG. 24 – Le Verbe au participe passé avec auxiliaire *être* (V-e).

Le verbe à auxiliaire *être* (fig. 24) présente les caractéristiques suivantes :

- La Tête (qui est également un V-ps, comme pour le verbe à auxiliaire *avoir*) réclame un et un seul verbe de [SCAT *aux-être*]. Celui-ci précède la Tête, et est adjacent d'une part à l'Adv-ng s'il est présent, d'autre part à la Tête ou à l'Adv-np s'il est présent. Comme l'auxiliaire *avoir*, l'auxiliaire *être* fournit à la construction ses valeurs de mode et de personne ; le nombre de l'auxiliaire, cette fois, s'accorde avec celui du verbe principal.
- Les verbes (principaux) *avoir* (41a) et *être* (41b) interdisent l'auxiliaire *être*. On ajoute donc la propriété d'exclusion correspondante.

- (41) a. il a eu vs. \* il est eu  
b. il a été vs. \* il est été

- Les verbes impersonnels (i.e., de [PERSONNE 0]) interdisent également l'auxiliaire *être* (42).

- (42) a. il a plu vs. \* il est plu  
b. il a fallu que... vs. \* il est fallu que...

### 1.4.3 Verbe au participe passé avec deux auxiliaires (V-ae)

Cette construction (fig. 25) correspond à la fois aux formes dites *surcomposées* (43a) et aux formes passives composées (43b).

- (43) a. Quand Panturle **a eu labouré** son champ, il a déjeuné <sup>144</sup>  
 b. et si c'était Michael Moore qui **avait été assassiné** ? <sup>145</sup>

<b>V-ae</b> HERIT V-a HERIT V-e	
INTR.	$\left[ \begin{array}{c} \text{ID} \left[ \begin{array}{c} \text{MODE} \quad [2].\text{MODE} \\ \text{ACCORD} \left[ \begin{array}{c} \text{PERSONNE} \quad [2].\text{PERSONNE} \\ \text{NOMBRE} \quad [2].\text{NOMBRE} \end{array} \right] \end{array} \right] \end{array} \right]$
Exigence	$[2] \Rightarrow [3]. \left[ \begin{array}{c} \text{MODE} \quad \textit{participe} \\ \text{TEMPS} \quad \textit{passé} \end{array} \right]$
Précédence	$[2] < [3]$ $\text{Adv-np} < [3]$

FIG. 25 – Le Verbe participe passé avec auxiliaires *avoir* et *être* (V-ae).

Elle reçoit les caractéristiques des deux constructions dont elle hérite (V-a et V-e), à ceci près que l'une des contraintes de précédence est modifiée : quand les deux auxiliaires sont présents, l'éventuel Adv-np se place entre les deux, et non plus à la suite de *être* comme c'était le cas dans V-e. En plus de cela, il faut préciser que l'auxiliaire *avoir* (qui avait été indexé [2] dans V-a) précède *être* (indexé [3] dans V-e). La construction reçoit les valeurs de personne, de nombre et de mode de l'auxiliaire *avoir*. Enfin, dans cette construction l'auxiliaire *être* doit être au participe passé : on ajoute la propriété d'exigence correspondante.

## 1.5 Modalisations

Les verbes construits, qu'ils soient négatifs (V-n), modaux (V-m), auxiliés (V-a, V-e ou V-ae) ou rien de cela (V), peuvent recevoir un modal s'ils sont à une forme infinitive.

### 1.5.1 Verbe infinitif modalisé (V-im)

Cette construction (fig. 26) reçoit un Verbe modal (V-m, décrit au §1.2) en plus de sa Tête (44).

- (44) a. une forte allergie **peut elle donner** de la fièvre ? <sup>146</sup>  
 b. La façon dont s'appuie l'un deux sur un cube de pierre **paraît répondre** au pied du Silène appuyé sur la borne, (...) <sup>147</sup>  
 c. Google **ne laisse pas passer** l'UTF-8 en sortie ? <sup>148</sup>

<sup>144</sup>[Riegel *et al.*, 1999]

<sup>145</sup>fr.soc.politique, 3 novembre 2004.

<sup>146</sup>fr.bio.medecine, 29 septembre 2004.

<sup>147</sup>fr.soc.histoire.antique, 19 octobre 2004.

<sup>148</sup>fr.usenet.8bits, 25 mai 2006.

- d. (...) **il me semblait ne pas avoir été** à la hauteur (...) <sup>149</sup>
- e. Bon ben je te trouve d'autres rails, je resoude les garde-boue plus haut, **je la fais repeindre** par mon frère, et je te préviens quand elle est prête; (...) <sup>150</sup>
- f. Les messages **semblent laisser entendre** qu'une cellule suffit et en même temps ils ne donnent aucun détail ("les messages ne sont pas des livres scientifiques (...) <sup>151</sup>

<b>V-im</b>	
HERIT V ∨ V-n ∨ V-m ∨ V-a ∨ V-e ∨ V-ae	
INTR.	$\left[ \begin{array}{l} \text{ID   NATURE} \\ \text{SYN} \end{array} \left[ \begin{array}{l} \text{TRAITS} \left[ \begin{array}{l} \text{MODE } \boxed{4}.\text{MODE} \\ \text{TEMPS } \boxed{4}.\text{TEMPS} \end{array} \right] \\ \text{ACCORD} \left[ \begin{array}{l} \text{PERSONNE } \boxed{4}.\text{PERSONNE} \\ \text{NOMBRE } \boxed{4}.\text{NOMBRE} \end{array} \right] \\ \text{MODALISATION} \left[ \begin{array}{l} \text{RECTEUR } \boxed{1} \\ \text{DÉPENDANT } \boxed{4} \end{array} \right] \end{array} \right]$
Mineur	$\boxed{4}V\text{-m}$
Unicité	$\boxed{4}$
Exigence	$\boxed{1} \Rightarrow \boxed{4}$ $\boxed{1}.\text{[MODE } \neg \textit{infinitif}] \Rightarrow \boxed{2}.\text{[MODE } \textit{infinitif}] \vee \boxed{3}.\text{[MODE } \textit{infinitif}]$
Précédence	$\boxed{4} < \boxed{1}$ $\boxed{4} < \text{Adv-n}$ $\boxed{4} < \boxed{2}$ $\boxed{4} < \boxed{3}$
Adjacence	$\boxed{4} \leftrightarrow \text{Adv-ng} \vee \boxed{2} \vee \boxed{3} \vee \text{Adv-np} \vee \boxed{1}$

FIG. 26 – Le Verbe infinitif à modal (V-im).

Comme on le voit en (44) ce n'est pas nécessairement la Tête qui porte le mode infinitif : ce peut également être l'auxiliaire, mais dans tous les cas, l'un d'eux et uniquement l'un d'eux doit être à l'infinitif. On ne précise donc pas le mode dans la propriété Majeur, mais on pose plutôt une propriété d'Exigence qui indique qu'une Tête non infinitive réclame un (et un seul) auxiliaire infinitif; le fait qu'un seul de ces verbes doive être à l'infinitif est assuré par les propriétés dont la construction hérite : si le V-im hérite de V-a, V-e ou V-ae cela implique que le Verbe principal (et l'auxiliaire *être* pour le V-ae) est (sont) au participe (et donc, pas à l'infinitif).

Le V-m, exigé et unique, est placé en début de construction (il est indexé  $\boxed{4}$  et le restera dans les constructions qui hériteront de celle-ci). S'il a lui-même reçu une négation (45a), une auxiliation (45b), un clitique (45c) <sup>152</sup> et/ou un modal (44f) alors ceux-ci ne sont pas concernés par les présentes propriétés : les relations internes au V-m construit auront été l'objet de l'introduction de la construction V-m utilisée comme constituant ici.

<sup>149</sup>fr.rec.animaux, 12 janvier 2003. Pour ce qui concerne l'intégration des pronoms clitiques dans la construction, cf. plus bas.

<sup>150</sup>fr.rec.moto, 9 septembre 2003.

<sup>151</sup>alt.religion.raelian, 29 avril 2003. Dans ce cas, le verbe principal est *entendre*, et il est construit avec un verbe modal (V-m) lui-même modalisé, *semble laisser*.

<sup>152</sup>On parle ici d'un clitique qui lui est propre, et non d'un clitique qui dépend du verbe principal.

- (45) a. Il **ne semble pas** avoir de succes <sup>153</sup>  
 b. Le problème bien sûr c'est quand il n'ya pas de facture de révision, typiquement quand le vendeur a fait lui-même l'entretien, ou l'**a fait** faire au noir... <sup>154</sup>  
 c. si ta mission consiste comme **il le paraît** à faire vendre des toiles, commence par supprimer les cadres à vomir qui vont avec (...) <sup>155</sup>

Il est adjacent suivant les configurations, soit à l'Adv-ng, soit à un auxiliaire (*avoir*, précédemment indexé [2] ou *être*, précédemment indexé [3]), soit à l'Adv-np, soit à la Tête (toujours indexée [1]), c'est-à-dire en somme à l'un des constituants (lequel est sélectionné par les contraintes de précédence héritées par la présente construction).

C'est lui, et non plus la Tête, qui fournit les indications syntaxiques de mode, temps, personne et nombre (sachant que lui-même a pu être construit avec des auxiliaires, comme on l'a vu précédemment), mais c'est bien le verbe principal qui le régit car c'est lui qui dirige la transitivité de la construction verbale (il rejoint en cela le fonctionnement des auxiliaires).

### 1.5.2 Verbe infinitif avec verbe à montée (V-imm)

Cette construction (fig. 27), qui hérite de la précédente, reçoit un Verbe à montée (V-mm, construction lexicale) en plus de l'infinitif (46).

- (46) a. Je **vais / veux / peux / souhaite / viens de / semble / suis supposée** le repeindre

<b>V-imm</b>	
HERIT V-im	
Mineur	[4]V-mm

FIG. 27 – Le Verbe infinitif à verbe à montée (V-imm).

Cette construction a pour particularité le fait que c'est le verbe à montée (V-mm) qui porte le sujet (syntaxique), alors que ce même constituant fait partie uniquement des actants (sémantiques) du verbe principal. Par exemple dans (47), *il* est le sujet de *semblait* (verbe à montée du sujet) et le premier actant de *avoir été à la hauteur*, mais n'a ni relation syntaxique avec *avoir été à la hauteur* ni relation sémantique avec *semblait* <sup>156</sup>.

- (47) *il semblait ne pas avoir été à la hauteur*

Dans (48), *cette notice* est le sujet de *est à* (verbe à montée de l'objet) et le second actant de *lire*.

- (48) *cette notice est à lire attentivement* <sup>157</sup>

<sup>153</sup>fr.soc.politique, 20 décembre 2002.

<sup>154</sup>fr.misc.automoto.mecanique, 18 juillet 2006.

<sup>155</sup>fr.rec.photo.numerique, 6 juin 2006.

<sup>156</sup>Contrairement aux verbes à contrôle dont le sujet fait partie des actants.

<sup>157</sup>[Creissels, 2004].

Nous revenons sur ces spécificités dans les sections qui suivent.

Le fait que nous formions une construction constituée du verbe à montée et du verbe principal nous rapproche plus de la position de [Le Goffic, 1993] qui considère que les premiers “*tendent à former avec leur Infinitif une unité syncrétique*”, que de celle de [Riegel et al., 1999] qui ne leurs accordent qu’un “*quasi-statut d’auxiliaire modal*”.

Cela nous rapproche également (comme c’était le cas pour les verbes à contrôle) de l’analyse TAG donnée dans [Abeillé, 2002, p. 111] :

*Nous analysons les verbes auxiliaires et les verbes dits ‘à montée’ comme des verbes à projection réduite à V, ancrant un arbre auxiliaire adjoint au participe ou à l’infinitif. [...] Selon les mêmes critères, d’autres verbes déficients sont définis pour le français (verbes copules, verbes causatifs). En TAG, ils sont substitués dans des arbres spécifiques ancrés par l’infinitif ou le participe, car ces constructions mettent en jeu une redistribution des arguments [...].*

En effet l’ensemble constitué que nous construisons sous la forme du V-imm a pour tête le verbe principal. Cela implique que nous nous inscrivons également dans ce qui suit (*Ibid.*, p. 114) :

*Cette analyse rejoint la caractérisation par [...] [Blanche-Benveniste, 1991] des verbes ‘à montée’ comme verbes “sans valence, surajoutés à la construction” de l’infinitif. Elle est proche également de celle de [Jacobson, 1990], qui, dans une grammaire catégorielle, en fait des foncteurs S/S combinés par composition fonctionnelle avec l’infinitif.*

En effet, les verbes à montée auront bien une transitivity au sein du SV dont ils seront constituant, mais ne portent pas de valence en tant qu’objet lexical.

### 1.5.3 Verbe infinitif avec verbe causatif (V-ime)

Cette construction (fig. 28) reçoit un Verbe causatif (V-me) en plus de l’infinitif (49).

- (49) a. Jean **fait rire** Marie / Jean **la fait rire**  
 b. Jean **fait soigner** Marie / Jean **la fait soigner**  
 c. **Fais soigner** Marie / **Ne fais pas soigner** Marie / **Fais-la soigner** / **Ne la fais pas soigner**  
 d. Je **le fais** / **laisse repeindre**

<b>V-ime</b>	
HERIT V-im	
Mineur	4 V-me

FIG. 28 – Le Verbe infinitif avec verbe causatif (V-ime).

Tout comme le “Verbe infinitif avec verbe à montée”, elle hérite du “Verbe infinitif modalisé” (V-im) que l’on a vu en début de section. Cela signifie que nous considérons les verbes causatifs (49d) comme étant des cas particuliers de verbes modaux, mais pas comme faisant partie des verbes à montée. La tendance est partagée à ce propos, puisque p.ex. [Le Goffic, 1993] les

considère comme des verbes à montée alors que [Abeillé, 2002] les en distingue. Nous adoptons la seconde position parce que le sujet du verbe causatif (p.ex. *Jean* dans (49)) ne fait pas partie de la valence du verbe tête (*rire, soigner*). En outre nous considérons que les verbes causatifs, contrairement aux verbes à montée, ne sont pas avalents mais (au moins) bivalents : leur premier actant correspond à leur sujet, et leur second actant au verbe tête (qui, lui, porte également sa propre valence).

## 1.6 Cliticisations

Comme on l’a vu en début de chapitre, on fait compter les éventuels pronoms clitiques parmi les constituants du Verbe construit. Ceux-ci peuvent entrer dans la valence de la construction (quand ils font référence à des actants), et quand ils correspondent à des compléments de verbe, ont une incidence sur la transitivité des constructions obtenues.

Les constructions qui suivent peuvent hériter de toutes les constructions que l’on a précédemment décrites, à l’exception du “verbe à clitique nominatif” qui ne peut pas hériter des constructions modalisées (V-im et constructions qui en héritent). Elles peuvent donc avoir une constitution fort variable, allant d’un verbe seul (V) à un ensemble constitué d’un verbe, d’une négation, d’un ou plusieurs auxiliaires et d’un ou plusieurs modaux.

### 1.6.1 Verbe avec clitique nominatif (V-cn)

<b>V-cn</b> HERIT $V \vee V-n \vee V-m \vee V-a \vee V-e \vee V-ae$	
$\text{INTR.} \left[ \text{SYN} \left[ \text{SUJET} \left[ \begin{array}{l} \text{RECTEUR} \quad [5] \\ \text{DÉPENDANT} \quad [1] \end{array} \right] \right] \right]$	
Mineur	$[5]\text{Pro-pcn}$
Unicité	$[5]$
Exigence	$[1] \Rightarrow [5]$ $[5] \Rightarrow [\text{MODE} (\neg \textit{infinitif} \wedge \neg \textit{participe})]$
Précédence	$[5] < \text{Adv-n}$ $[5] < [1]$ $[5] < [2]$ $[5] < [3]$
Adjacence	$[5] \leftrightarrow \text{Adv-ng} \vee [1] \vee [2] \vee [3]$
Accord	$[5].\text{PERSONNE} \rightsquigarrow [1].\text{PERSONNE} \vee [2].\text{PERSONNE} \vee [3].\text{PERSONNE}$ $[5].\text{NOMBRE} \rightsquigarrow [1].\text{NOMBRE} \vee [2].\text{NOMBRE} \vee [3].\text{NOMBRE}$

FIG. 29 – Le Verbe avec clitique nominatif (V-cn).

Le Pronom personnel clitique nominatif (Pro-pcn, désormais indexé  $[5]$ ) doit être présent pour que cette construction soit introduite, et il se place au début de la construction. Cette assertion n’étant pas vraie pour les cas de sujet inversé (interrogatives canoniques notamment) (50), le verbe à sujet clitique inversé fait l’objet d’une construction qui hérite de celle-ci (section suivante).

- (50) a. **tu** ne la lui donnes (pas)  
b. ne la lui donnes-**tu** (pas)

Il est de plus adjacent au reste de la construction, c'est-à-dire à l'un de ses constituants à l'exception de l'Adv-np (*pas* et autres). Enfin, le verbe fini s'accorde en personne et en nombre avec lui.

**Interdiction de sujet pour un infinitif.** — Tous les verbes peuvent recevoir un pronom clitique nominatif qui fera office de sujet du verbe fini de la construction (qu'il s'agisse d'un modal, ou du verbe principal sous une forme simple ou composée suivant les cas), à l'exception des constructions verbales infinitives puisqu'on a dit précédemment que l'on considèrerait que ceux-ci ne recevaient jamais de sujet. Or, comme on l'a vu, le fait qu'une construction soit infinitive ne dépend pas uniquement du mode du verbe tête, mais peut dépendre aussi du mode des éventuels auxiliaires et modaux qui en dépendent : les exemples (51) ne peuvent recevoir de clitique nominatif.

- (51) a. manger  
b. ne pas être mangé  
c. avoir été mangé  
d. pouvoir (le manger)  
e. les en faire (avoir mangé)

Afin de gérer l'interdiction de recevoir un clitique nominatif pour ces cas, il serait donc erroné de préciser un mode infinitif pour la tête, puisque ça n'est pas toujours le cas (51b, 51c, 51e). On peut par contre préciser que le Pro-pcn (et par là même la construction) exige un objet (quelconque) qui ne soit ni au mode infinitif ni au participe, que celui-ci corresponde indifféremment à la tête (51a), à un auxiliaire (51b, 51c) ou à un modal (51d, 51e). Il est en effet inutile de préciser à quel verbe doit correspondre cette caractéristique puisque, en tenant compte des propriétés héritées on se retrouve nécessairement dans l'une des configurations suivantes :

1. La tête est le seul verbe de la construction (dans le cas d'un héritage direct de V (52a), V-n (52b) ou V-m (52c)).

- (52) a. tu manges  
b. tu ne manges pas  
c. tu peux

Dans ce cas elle est le seul objet sur lequel peut porter la contrainte, ce qui signifie que l'on interdit les verbes au participe (\* *je mangé*) et à l'infinitif (\* *je manger*).

2. La construction contient l'auxiliaire *avoir* (dans le cas d'un héritage de V-a (53a) ou V-ae (53b)).

- (53) a. tu as mangé  
b. tu as été mangé



Dans ce cas, la tête est nécessairement au participe, de même que l’auxiliaire *être* dans le cas d’un V-ae, et par conséquent le seul verbe à pouvoir satisfaire la contrainte est l’auxiliaire *avoir*, ce qui nous permet d’éviter \**je avoir mangé* et \**je avoir été mangé*.

3. La construction contient l’auxiliaire *être* et pas l’auxiliaire *avoir* (dans le cas d’un héritage de V-e (54)).

(54) tu es mangé

Dans ce cas, la tête est nécessairement au participe, et par conséquent le seul verbe à pouvoir satisfaire la contrainte est l’auxiliaire *être*, ce qui nous permet d’éviter \**je être allé*.

**Dépendance.** — Le pronom nominatif occupe la fonction syntaxique de **sujet** de la construction, ce que l’on fait figurer dans ses informations intrinsèques.

Notons bien au passage que dans la relation de dépendance entre le sujet et le verbe, on considère que c’est le sujet qui gouverne : en effet, c’est bien lui qui, du point de vue syntaxique, impose les accords en personne et nombre, et éventuellement en genre. La relation de dépendance sémantique serait exprimée dans l’autre sens, le prédicat régissant son sujet (comme dans les représentations logiques où l’on aurait par exemple *manger(Pierre, pomme)*).

### 1.6.2 Verbe avec clitique nominatif inversé (V-cni)

Cette construction (fig. 30) hérite de la précédente et consiste simplement à modifier l’ordre des objets de la construction précédente, le clitique sujet venant dans ce cas après le verbe fini et l’adverbe *ne*.

<b>V-cni</b> HERIT V-cn	
Précédence	Adv-ng $\prec$ [5] [MODE ( $\neg$ infinitif $\wedge$ $\neg$ participe)] $\prec$ [5]

FIG. 30 – Le Verbe avec clitique nominatif inversé (V-cni).

### 1.6.3 Verbe avec clitique accusatif (V-ca)

Cette construction peut hériter de toutes les constructions vues dans les sections précédentes au choix, et ajoute à leur constituance la présence (exigée) d’un Pronom personnel clitique accusatif (Pro-pca).

**Ordre relatif.** — Le clitique accusatif est soumis à des contraintes de position par rapport aux autres constituants possibles de la construction en fonction de son héritage (55).

(55) a. Pierre (ne) le reste / peut / semble (pas)

<b>V-ca</b> HERIT V ∨ V-n ∨ V-m ∨ V-a ∨ V-e ∨ V-ae ∨ V-im ∨ V-imm ∨ V-imc	
Mineur	[6]Pro-pca
Unicité	[6]
Exigence	[1] ⇒ [6]
Précédence	Adv-ng < [6] [6] < [SCAT ¬modal] (Adn-ng ∧ [6]) < [MODE impératif] Adv-np < ([6] ∧ [MODE infinitif]) [4] < [6]
Adjacence	[6] ↔ Adv-n
Accord	[6].GENRE ∼ [1].GENRE [6].NOMBRE ∼ [1].NOMBRE

FIG. 31 – Le Verbe avec clitique accusatif (V-ca).

- b. Pierre (ne) l'est (pas) resté
- c. Pierre (ne) l'a (pas) (eu) donné
- d. Pierre peut / semble / laisse (ne pas) le manger
- e. Pierre (ne) le / l' (a) (pas) fait manger
- f. (ne) le donne pas
- g. donne-le (pas)

Dans tous les cas, si *ne* (Adv-ng) est présent, il le suit ; il y est adjacent si le verbe n'est pas infinitif ; sinon il est adjacent à l'Adv-np. Si le verbe est infinitif, alors le Pro-pca précède l'Adv-np (qui lui-même précède le verbe). Il précède de plus tous les verbes principaux et auxiliaires qui ne sont pas à l'impératif positif (55g). Enfin, il suit le verbe modal (construit ou non), si ce dernier n'es pas la tête de la construction en question (i.e., il suit *peut / semble / laisse* dans (55d) parce qu'il est constituant de *peut / semble / laisse le manger*, mais il précède *fait* dans (55e) parce qu'il est alors constituant du V-m *le fait*).

**Accord dans le cas d'un auxiliaire *avoir*.** — Si l'auxiliaire *avoir* est présent, alors il suit le clitique accusatif. Or la règle en français veut que le participe passé du verbe s'accorde en genre et en nombre avec celui-ci dans cette circonstance :

- (56) a. Il a donné des pommes à Pierre  
b. Il les a données à Pierre

On ajoute donc la propriété d'accord correspondante.

**Héritage.** — Suivant la nature du verbe Tête, le clitique accusatif peut correspondre à deux dépendances syntaxiques distinctes : attribut du sujet ou complément d'objet direct. Nous précisons cela dans les constructions suivantes, qui toutes deux héritent de celle-ci.

### 1.6.4 Verbe avec clitique accusatif attribut du sujet (V-cas)

Parmi les deux attributs possibles seul l'attribut du sujet peut être cliticisé (57).

- (57) a. Pierre devient **gentil** → Pierre **le** devient  
 b. Pierre trouve le prix **élevé** → \* Pierre **le** trouve le prix  
 c. Pierre fait devenir Paul **gentil** → \* Pierre **le** fait devenir Paul

V-cas	
HERIT V-ca	
INTR.	$\left[ \text{SYN} \left[ \text{ATTR-SUJ} \left[ \text{DÉPENDANT} \boxed{6} \right] \right] \right]$

FIG. 32 – Le Verbe à clitique accusatif attribut du sujet (V-cas).

Dans ce cas le clitique est noté comme dépendant de la relation “attribut” s’établissant entre lui et le sujet (qui n’est localement pas (encore) affecté).

### 1.6.5 Verbe avec clitique accusatif objet direct (V-cod)

V-cod	
HERIT V-ca	
INTR.	$\left[ \text{SYN} \left[ \text{OBJ-DIR} \left[ \begin{array}{l} \text{RECTEUR} \boxed{1} \\ \text{DÉPENDANT} \boxed{6} \end{array} \right] \right] \right]$
Exclusion	$\boxed{6}, [\text{RÉFLÉCHI } oui] \not\sim \boxed{2}$ $\boxed{1}, \neg \text{ACTANT2} \not\sim \boxed{3}$
Accord	$\boxed{1}. \text{GENRE} \rightsquigarrow (\boxed{5} \wedge \boxed{8}. \text{GENRE})$ $\boxed{1}. \text{NOMBRE} \rightsquigarrow (\boxed{5} \wedge \boxed{8}. \text{NOMBRE})$

FIG. 33 – Le Verbe avec clitique accusatif objet direct (V-cod).

Le clitique accusatif joue le rôle de l’objet direct de la Tête de la construction, y compris dans le cas d’une construction causative ce qui signifie que dans (58), *le* sera objet direct de *fait*, et non de *manger*. La relation que le clitique entretient avec le verbe *manger* est une relation purement sémantique, du fait de la “montée” du pronom sur le verbe causatif.

- (58) Pierre le fait manger

Dans les dépendances syntaxiques, on ajoute donc l’attribut correspondant à la relation “objet direct” (OBJ-DIR) s’établissant entre le pronom clitique et le verbe Tête.

Ce cas de figure nous permet d’ajouter également la contrainte supplémentaire suivante à la construction :

- Si le clitique objet est un pronom réfléchi alors il interdit l’auxiliaire *avoir* (59), précédemment indexé [2]. On ajoute donc la propriété d’exclusion correspondante.

- (59) a. Pierre s’est donné une claque vs. \* Pierre s’a donné une claque  
b. Tu t’es lavée vs. \* Tu t’as lavée

### 1.6.6 Verbe avec clitique datif (V-cd)

Cette construction (fig. 34) est quasiment la même que le V-ca à ceci près que le clitique est ici datif. Evidemment, cette construction n’est pas l’objet des restrictions de sélection de constituants que l’on a vues précédemment et qui étaient spécifiques au pronom accusatif; toutefois les contraintes d’ordre s’opèrent de la même façon que ce que l’on a vu. Les verbes modaux, en tant que Tête (de V-m) interdisent les clitiques datifs, à l’exception des verbes causatifs <sup>158</sup>.

- (60) a. Pierre (n’y) va (pas)  
b. Pierre (ne) lui donne (pas) une pomme  
c. Pierre (n’y) est (pas) allé  
d. Pierre (ne) lui a (pas) (eu) donné  
e. Pierre (ne) semble / peut (pas) (ne pas) y aller  
f. Pierre (ne) lui (a) (pas) fait donner  
g. ne lui donne pas  
h. donne-lui (pas)

V-cd	
HERIT V ∨ V-n ∨ V-m ∨ V-a ∨ V-e ∨ V-ae ∨ V-im ∨ V-imm ∨ V-imc	
INTR.	$\left[ \text{SYN} \left[ \text{OBJ-IND} \left[ \begin{array}{l} \text{RECTEUR} \quad [1] \\ \text{DÉPENDANT} \quad [7] \end{array} \right] \right] \right]$
Mineur	[7]Pro-pcd
Unicité	[7]
Exigence	[1] ⇒ [7]
Exclusion	[7] $\not\Leftarrow$ [1, [TYPE-MODAL $\neg$ causatif ]]
Précédence	Adv-ng < [7] [7] < [SCAT $\neg$ modal] (Adn-ng ∧ [6]) < [MODE impératif] Adv-np < ([6] ∧ [MODE infinitif]) [4] < [6]
Adjacence	[6] ↔ Adv-n

FIG. 34 – Le Verbe avec clitique datif (V-cd).

**Dépendance.** — Dans cette construction le pronom clitique occupe la fonction de complément d’objet indirect de la Tête, comme on le note dans les informations intrinsèques de la construction.

<sup>158</sup>L’exception est due à la “montée” des clitiques du verbe Tête sur le verbe causatif.

### 1.6.7 Verbe à clitiques nominatif et accusatif objet direct (V-cnod)

Cette construction (fig. 35) hérite à la fois du verbe à clitique nominatif (V-cn ou V-cni) et du verbe à clitique accusatif complément d'objet direct (V-cod) (61).

- (61) a. je le mange  
b. la fais-tu  
c. je me soigne

<b>V-cnod</b> HERIT V-cn $\vee$ V-cni HERIT V-cod	
Adjacence	$[6] \leftrightarrow \text{Adv-ng} \vee [1] \vee [2] \vee [3] \vee [5]$
Accord	$[5].\text{GENRE} \rightsquigarrow [6].[\text{RÉFLÉCHI } oui].\text{GENRE}$
	$[5].\text{NOMBRE} \rightsquigarrow [6].[\text{RÉFLÉCHI } oui].\text{NOMBRE}$
	$[5].\text{PERSONNE} \rightsquigarrow [6].[\text{RÉFLÉCHI } oui].\text{PERSONNE}$

FIG. 35 – Le Verbe à clitiques nominatif et accusatif objet direct (V-cnod).

Le clitique accusatif, s'il est réfléchi (61c), s'accorde entre genre, nombre et personne avec le nominatif. On précise de plus la contrainte d'adjacence que l'on avait donnée pour le nominatif, en ajoutant l'accusatif dans la liste.

### 1.6.8 Verbe à clitiques nominatif et accusatif attribut du sujet (V-cn-as)

Cette construction (fig. 36) est quasiment la même que la précédente à ceci près qu'elle hérite du verbe à clitique nominatif (V-cn ou V-cni) et du verbe à clitique accusatif attribut du sujet (V-cas) (62).

- (62) je le suis

<b>V-cn-as</b> HERIT V-cn $\vee$ V-cni HERIT V-cas	
INTR. $\left[ \text{SYN} \left[ \text{ATTR-SUJ} \left[ \text{RECTEUR } [5] \right] \right] \right]$	
Adjacence	$[6] \leftrightarrow \text{Adv-ng} \vee [1] \vee [2] \vee [3] \vee [4] \vee [5]$
Exclusion	$[5] \not\rightarrow [4]$
Accord	$[5].\text{GENRE} \rightsquigarrow [6].\text{GENRE}$
	$[5].\text{NOMBRE} \rightsquigarrow [6].\text{NOMBRE}$
	$[5].\text{PERSONNE} \rightsquigarrow [6].\text{PERSONNE}$

FIG. 36 – Le Verbe à clitiques nominatif et accusatif attributif (V-cn-as).

Comme on l'a dit précédemment, les verbes modalisés ne peuvent pas recevoir de clitique sujet (puisque celui-ci est présent alors il est constituant du V-m qui lui-même est constituant du V-im) ; il nous faut donc préciser cela ici, sous la forme d'une propriété d'Exclusion.

Dans ce cas, on peut préciser le recteur dans la relation “attribut” des informations intrinsèques : il s’agit du pronom nominatif, précédemment indexé [5]. La conséquence de cela sur la forme de l’objet est que les deux clitiques doivent être accordés en genre, nombre et personne.

### 1.6.9 Verbe à clitiques nominatif et datif (V-cnd)

Cette construction (fig. 37) présente les mêmes caractéristiques que les verbes à clitiques nominatif et accusatif objet direct (V-cnod), au cas du pronom complément (et donc à l’héritage) près (63).

- (63) a. nous nous donnons (*les mains*)  
 b. y vas-tu  
 c. il lui fait (*repeindre la maison*)

<b>V-cnd</b> HERIT V-cn $\vee$ V-cni HERIT V-cd	
Exclusion	[5] $\not\leftrightarrow$ [4]
Adjacence	[6] $\leftrightarrow$ Adv-ng $\vee$ [1] $\vee$ [2] $\vee$ [3] $\vee$ [6]
Accord	[10].GENRE $\leadsto$ [9].[RÉFLÉCHI <i>oui</i> ].GENRE
	[10].NOMBRE $\leadsto$ [9].[RÉFLÉCHI <i>oui</i> ].NOMBRE
	[10].PERSONNE $\leadsto$ [9].[RÉFLÉCHI <i>oui</i> ].PERSONNE

FIG. 37 – Le Verbe à sujet et complément indirect clitiques (V-cnd).

### 1.6.10 Verbe à clitiques accusatif et datif (V-cad)

Cette construction (fig. 38) hérite des deux constructions à clitique accusatif objet direct (V-cod) et à clitique datif (V-cd), auxquelles on ajoute les propriétés d’ordre entre les deux pronoms (64).

- (64) a. la lui avais donnée  
 b. n’y en avait pas eu  
 c. prenons-le lui  
 d. ne le lui prenons pas

**Redondance des formes.** — La règle de cacophonie ([Arnauld & Lancelot, 1830], [Gross, 1967]) impose que les deux pronoms clitiques, s’ils sont présents, ne soient pas de la même forme (65d).

- (65) a. il remplit **un verre** de ce liquide  
 b. il **en** remplit **un verre**  
 c. il **en** remplit **un** de ce liquide  
 d. \* il **en en** remplit **un**

<b>V-cad</b> HERIT V-cod HERIT V-cd	
Exclusion	$(me \vee te \vee se \vee nous \vee vous \vee lui \vee leur)$ $\nexists (me \vee te \vee se \vee nous \vee vous \vee lui \vee leur)$ $\boxed{6}.[FORME\ x] \nexists \boxed{7}.[FORME\ x]$
Précédence	$(me \vee te \vee se) \prec (le \vee la \vee les \vee lui \vee leur)$ $(le \vee la \vee les \vee lui \vee leur) \prec (moi \vee toi \vee soi)$ $(me \vee te \vee se \vee le \vee la \vee les \vee lui \vee leur) \prec (y \vee en)$ $(le \vee la \vee les) \prec (lui \vee leur)$ $y \prec en$

FIG. 38 – Le Verbe à clitiques accusatif et datif (V-cad).

**Ordre relatif des pronoms.** — Généralement, l'ordre des pronoms clitiques non nominatifs est exprimé en fonction leurs formes et non de leurs catégories (p. ex. [Blanche-Benveniste, 1975]) (66); et il est vrai que, même s'il reste possible d'exprimer ces contraintes sur des groupes que l'on pointe en fonction de leurs caractéristiques syntaxiques ou lexicales, pointer leurs formes reste beaucoup plus simple.

- (66) a. il **me**<sub>acc</sub> présente à toi  
b. il **me**<sub>dat</sub> le présente  
c. \* il **me te** présente

Selon le filtre de [Perlmutter, 1970] (67), chacun des membres occupant une position donnée exclut celle des autres membres de cette position (ce qui était déjà partiellement signifié par les propriétés d'unicité dont la construction hérite, p.ex. pour *le / la / les*).

- (67) 1 = je / tu / il / elle / on / nous (nominatif) / vous (nominatif) / ils / elles  
2 = ne  
3 = me / te / se / nous (non nominatif) / vous (non nominatif)  
4 = le / la / les  
5 = lui / leur  
6 = y  
7 = en

De plus, les objets de la position 3 et de la position 5 s'excluent mutuellement (68a), sauf en cas de datif éthique (68b).

- (68) a. \* il **me lui** présente / \* tu **te leur** présentes / \* il **se lui** présente  
b. je **te lui** ai donné une de ces gifles

Toutefois, comme l'usage du datif éthique n'est pas toujours considéré comme grammatical, et qu'il ne peut être généralisé, nous n'en tenons pas compte dans l'expression de nos contraintes ici. Cela a pour conséquence qu'en l'état, la caractérisation d'un datif éthique ne sera pas satisfaite à 100%. Enfin, *y* et *en* (positions 6 et 7) s'excluent, sauf dans *il y en a* (on ne précise pas cette exclusion).

### 1.6.11 Verbe à clitiques nominatif, accusatif et datif (V-cnad)

Cette construction (fig. 39) hérite à la fois des caractéristiques du verbe à clitique nominatif (V-cn ou V-cni) et à clitiques accusatif et datif (V-cad).

- (69) a. je le lui donne  
b. je me le fais

<b>V-cnad</b> HERIT V-cad HERIT V-cn ∨ V-cni	
Exclusion	$\boxed{5} \not\rightarrow \boxed{4}$
Accord	$\boxed{5}.\text{GENRE} \rightsquigarrow [\text{RÉFLÉCHI } oui].\text{GENRE}$ $\boxed{5}.\text{NOMBRE} \rightsquigarrow [\text{RÉFLÉCHI } oui].\text{NOMBRE}$ $\boxed{5}.\text{PERSONNE} \rightsquigarrow [\text{RÉFLÉCHI } oui].\text{PERSONNE}$

FIG. 39 – Le Verbe à clitiques nominatif, accusatif et datif (V-cnad).

Elle n'est l'objet d'aucune caractéristique de plus que celles dont elle hérite, si l'accord entre les clitiques nominatif d'une part, et accusatif ou datif d'autre part si l'un de ces derniers est réfléchi, ainsi que la réprécision de l'interdiction pour un verbe modalisé de recevoir un clitique nominatif.

## 2 Syntagmes Verbaux (SV)

Le Syntagme Verbal est une construction constituée d'un Verbe et de ses éventuels compléments pleins (i.e., non clitiques), et/ou circonstants <sup>159</sup>. Suivant le Verbe qui le constitue, et plus précisément selon si ce verbe comporte un sujet clitique ou non, le SV pourra par la suite prendre un statut propositionnel ou pas. On donne en figure 40 la hiérarchie des SV dans la grammaire.

### 2.1 Constructions de base

#### 2.1.1 Le Syntagme Verbal (SV) général

On donne en fig. 41 la description d'un Syntagme Verbal qui a essentiellement vocation à introduire la liste des traits des divers SV qui en hériteront. Compte tenu de sa portée très générale, il n'a que peu de propriétés ; on y signale essentiellement que son Majeur indexé  $\boxed{1}$  est un soit un Verbe, soit un SV.

Du SV peuvent hériter deux constructions : d'une part les SV exocentriques qui rassemblent le verbe et ses compléments de verbe (objets direct et indirect le cas échéant), d'autre part les SV endocentriques qui rassemblent un SV et ses circonstants éventuels.

<sup>159</sup>Les circonstants sont parfois appelés "compléments périphériques".



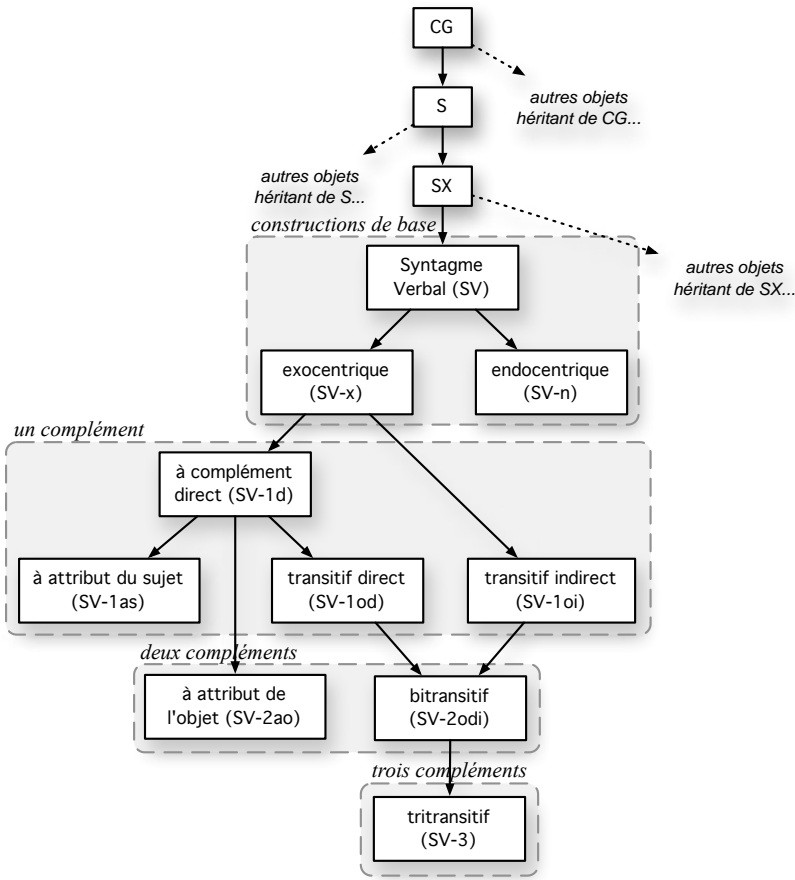


FIG. 40 – Hiérarchie des Syntagmes Verbaux (SV) dans la grammaire.

SV	
HERIT SX	
INTR.	$\left[ \begin{array}{l} \text{ID} \mid \text{NATURE} \end{array} \right]$
	$\left[ \begin{array}{l} \text{CATÉGORIE } SV \\ \text{TYPE } V \end{array} \right]$
Majeur	$\overline{1}(V \vee SV)$

FIG. 41 – Le Syntagme Verbal général (SV).

2.1.2 Le Syntagme Verbal exocentrique (SV-x)

Cette construction (fig. 42) hérite du SV général; elle en est un raffinement qui consiste à préciser que la Tête, dans ce cas, est un Verbe (et non un SV) (70).

- (70) a. (Pierre) donne un cadeau à Marie  
b. je vais aller au cinéma (demain)

Il est possible de lui ajouter un certain nombre de restrictions, comme par exemple le fait que l'élément Mineur ne peut être introduit par *parce que* (72) ([Debaisieux & Deulofeu, 2003] : les subordonnées introduites par *parce que* ne peuvent être que des modificateurs et non des compléments <sup>160</sup>).

(72) tu auras du mal de semer tes graines ou de planter tes replants **parce que la terre est trop dure** <sup>161</sup>

<b>SV-x</b>	
HERIT SV	
Majeur	<u>I</u> V
Mineur	[SPEC $\neg$ <i>parce que</i> ]

FIG. 42 – Le Syntagme Verbal exocentrique (SV-x).

Les constructions qui hériteront de celle-ci (sections suivantes) seront constituées d'un verbe et de ses compléments de verbe.

### 2.1.3 Le Syntagme Verbal endocentrique (SV-n)

Cette construction (fig. 44) hérite du SV général (SV). Elle a pour Tête un SV, c'est-à-dire un verbe qui a reçu tous les compléments réclamés le cas échéant, et qui est accompagnée alors d'un ou plusieurs circonstants (73). Elle est constituée strictement d'un SV et d'un circonstant (tel que SN, Adj, SPrep, Adv, etc.) ; si un SV a plusieurs circonstants, ils sont rattachés un à un au SV, comme l'illustre la figure 43.

- (73) a. (Pierre) donne un cadeau à Marie avec gratitude  
b. je vais aller au cinéma demain s'il pleut et que je ne sais pas quoi faire d'autre

## 2.2 Syntagmes Verbaux à un complément

Comme on l'a vu dans la hiérarchie des SV (fig. 40), les SV à complément(s) de verbe héritent, directement ou indirectement, du SV exocentrique. On décrit d'abord dans cette section les SV à un complément, puis on présentera dans les sections suivantes les SV à deux, puis à trois compléments.

<sup>160</sup>Par contre, et contrairement à ce que les auteurs proposent, nous considérons que les constructions spécifiées par *parce que* peuvent être des compléments de phrase (i.e. expansions de P et non de SV), comme p.ex. dans

(71) non c'est pas embarrassant parce que moi d'abord j'ai rien à cacher

alors qu'eux postulent dans ce cas une relation paratactique entre *c'est pas embarrassant* et *parce que moi d'abord j'ai rien à cacher* en raison de la portée de la négation plus restreinte que dans le cas d'une expansion de SV. Pour justifier de cette parataxe ils utilisent des indices non syntaxiques (et font appel notamment à l'analyse de discours — deux actes de langage — et à la sémantique — pas de relation de cause —), mais puisque nous avons décidé ici de nous limiter à ceux-ci autant que possible, alors nous posons que d'un point de vue strictement syntaxique il ne s'agit pas ici de parataxe.

<sup>161</sup>[Debaisieux & Deulofeu, 2003].

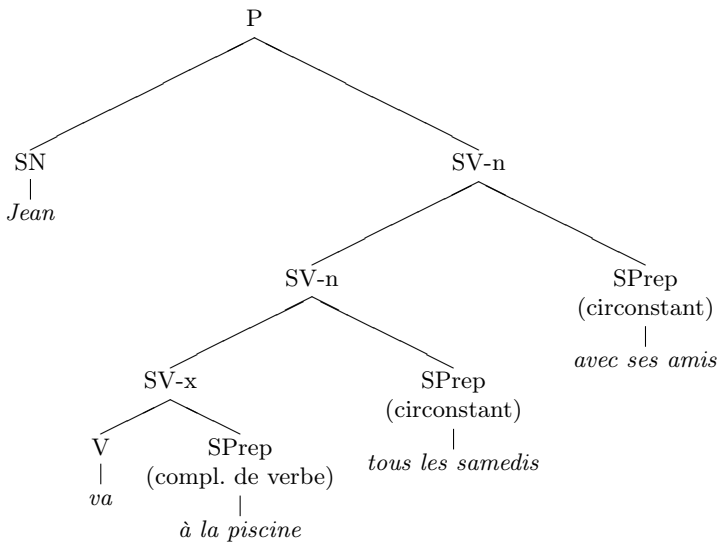


FIG. 43 – Exemple de hiérarchisation des circonstants dans un SV endocentrique.

<b>SV-n</b>	
HERIT SV	
Majeur	<u>1</u> SV

FIG. 44 – Le Syntagme Verbal endocentrique (SV-n).

2.2.1 Syntagme Verbal à complément direct (SV-1d)

Le SV à complément direct (fig. 45) hérite du SV exocentrique (SV-x). Il rassemble les caractéristiques communes aux verbes portant un complément direct (objet ou attribut du sujet) et également aux verbes portant deux compléments directs (les verbes transitifs directs à attribut de l'objet).

Il peut être constitué de deux façons différentes :

- Soit son complément direct est réalisé sous une forme non clitique (74),

- (74) a. (Pierre) mange une pomme  
b. (Sophie) devient grande  
c. (Sophie) fait dormir Pierre  
d. (Pierre) en mange trois  
e. je sais où je dois aller  
f. (Marie) dit que tout va bien  
g. (Pierre) apprend à nager

SV-1d HERIT SV-x	
Majeur	$\left[ \begin{array}{l} \text{OBJET} \left[ \begin{array}{l} \text{CATÉGORIE } V \\ \text{ÉTIQUETTE } \neg V\text{-cd} \end{array} \right] \\ \text{[1]} \text{ SYN } \left[ \text{DEP} \mid \text{MOD.} \mid \text{DÉP.} \mid \text{SYN} \mid \text{ÉTIQ.} \mid \neg(V\text{-ca} \vee V\text{-cd}) \right] \end{array} \right]$
Mineur	$\begin{array}{l} \text{[2]}(\text{SN} \vee \text{Conj-s} \vee \text{SPrep.}[\text{SPÉC.} \text{DÉPENDANT} \text{SYN} \text{MODE } \textit{infinitif}]) \\ \text{[3]Adj} \end{array}$
Unicité	$\begin{array}{l} \text{[2]} \\ \text{[3]} \end{array}$
Exigence	$\text{[1.} \neg V\text{-ca} \Rightarrow \text{[2]; [3]}$
Accord	$\text{[3.NOMBRE} \rightsquigarrow \text{[1.NOMBRE]}$

FIG. 45 – Le Syntagme Verbal à complément direct (SV-1d).

et dans ce cas la Tête est un verbe qui n’a pas reçu de complément clitique, ni direct ni indirect (c’est-à-dire  $\neg V\text{-ca}$  et  $\neg V\text{-cd}$ ). Il faut que cette contrainte soit également vraie dans le cas d’une construction causative (74c), ce qui signifie qu’on doit également interdire les verbes qui ont une dépendance de MODALISATION dont le dépendant est un V-ca ou un V-cd.

La Tête est accompagnée alors d’une (et une seule) construction faisant office de complément direct, et pouvant être un Syntagme Nominal (SN), un Adjectif (Adj, simple ou construit), une subordonnée (Conj-s) ou un SPrep dont la Tête est un SV infinitif.

Pour le cas de (74d), l’ensemble *en trois* forme un Pronom construit (qui se place sur le paradigme des SN) et est traité dans le chapitre qui leur est dédié. Même chose pour les cas suivants :

- (75) a. Nous **en** connaissons **les raisons** → Nous connaissons les raisons de ce choix  
b. J’y suis **prêt** → Je suis prêt à cela

– Soit le complément direct est réalisé sous une forme clitique (76),

- (76) a. (Pierre) la mange  
b. elle le devient  
c. (Pierre) la fait dormir

et dans ce cas la Tête est un soit verbe avec clitique accusatif (i.e. un V-ca, mais sans datif c’est-à-dire  $\neg V\text{-codi}$  <sup>162</sup>), soit un verbe modalisé à verbe causatif (V-imc) dont le modal est un V-ca (76c).

Il est alors le seul constituant de la construction.

Si le complément est un Adjectif alors il doit porter le même nombre que le verbe.

<sup>162</sup>Le V-codi hérite du V-cd, or on a déjà éliminé le V-cd des têtes possibles dans la configuration précédente : il est donc inutile de préciser cette étiquette ici.

### 2.2.2 Syntagme Verbal à attribut du sujet (SV-1as)

Le SV à attribut du sujet (fig. 46) hérite du SV à complément direct (SV-1d). Il regroupe les objets constitués d'un verbe et d'un attribut (74b, 76b).

SV-1as	
HERIT SV-1d	
INTR. $\left[ \text{SYN} \left[ \text{ATTR-SUJ} \left[ \text{DÉPENDANT} \begin{smallmatrix} \boxed{3} \end{smallmatrix} \right] \right] \right]$	
Exclusion	$\boxed{1} \not\sim \boxed{2}$
Accord	$\boxed{1}.\text{SUJET}   \text{RECTEUR}   \text{ID}   \text{NATURE}   \text{ACCORD}   \text{NOMBRE} \rightsquigarrow \boxed{3}.\text{NOMBRE}$ $\boxed{1}.\text{SUJET}   \text{RECTEUR}   \text{ID}   \text{NATURE}   \text{ACCORD}   \text{GENRE} \rightsquigarrow \boxed{3}.\text{GENRE}$

FIG. 46 – Le Syntagme Verbal à attribut du sujet (SV-1as).

Dans ce cas, le constituant mineur est réduit à un Adjectif, et l'on exclut donc les SN (précédemment indexés  $\boxed{2}$ ). On ajoute les informations de dépendance syntaxique correspondant à l'attribut du sujet, et l'accord entre l'adjectif et le sujet si celui-ci est présent (sous la forme d'un clitique sujet (76b)).

### 2.2.3 Syntagme Verbal transitif direct (SV-1od)

Le SV transitif direct (fig. 47) hérite du SV à complément direct (SV-1d) ; il a la particularité de recevoir un complément d'objet direct (77) (contrairement au SV-1as qui reçoit un attribut).

- (77) a. il a fini ses devoirs  
 b. La table mesure un mètre cinquante <sup>163</sup>  
 c. (Pierre) mangera une pomme  
 d. il fait dormir Paul  
 e. Il est arrivé un accident

SV-1od	
HERIT SV-1d	
INTR. $\left[ \text{SYN} \left[ \text{OBJ-DIR} \left[ \begin{smallmatrix} \text{RECTEUR} & \boxed{1}, \neg V\text{-}imc ? : \boxed{1}.\text{MODAL.}   \text{DÉP.}   \text{SYN}   \text{TÊTE} \\ \text{DÉPENDANT} & \boxed{2} \end{smallmatrix} \right] \right] \right]$	
Exclusion	$\boxed{1} \not\sim \boxed{3}$

FIG. 47 – Le Syntagme Verbal transitif direct (SV-1od).

Il n'a aucune caractéristique supplémentaire que ce dont il hérite, à part la réduction des objets mineurs à des SN uniquement (c'est-à-dire que l'on exclut les Adj, précédemment indexés  $\boxed{3}$ ),

<sup>163</sup>Il s'agit là d'un complément qui ne se pronominalise pas.

et la mention de la dépendance “objet direct” (OBJ-DIR) dans les informations intrinsèques. Le recteur de cette relation est la Tête si elle n’est pas causative (i.e., si elle n’est pas un V-imc), et le verbe causatif (donc la tête du V-m dépendant de la relation de modalisation) dans le cas contraire : dans (77d), *Paul* est l’objet direct de *fait* et non de *dormir* (ni de *fait dormir*), tout comme on avait posé plus haut que dans *Pierre le fait dormir*, *le* était objet direct du verbe causatif.

## 2.2.4 Syntagme Verbal transitif indirect (SV-1oi)

Cette construction (fig. 48) hérite du SV exocentrique (SV-x). Elle peut être constituée de deux manières différentes :

SV-1oi	
HERIT SV-x	
INTR.	$\left[ \text{SYN} \left[ \text{OBJ-IND} \left[ \begin{array}{ll} \text{RECTEUR} & \boxed{1}. \text{TÊTE} \\ \text{DÉPENDANT} & \boxed{4} \end{array} \right] \right] \right]$
Majeur	$\boxed{1} \left[ \begin{array}{ll} \text{CATÉGORIE} & V \\ \text{ÉTIQUETTE} & \neg V\text{-}ca \end{array} \right]$
Mineur	$\boxed{4}(\text{SPrep} \vee \text{Conj-s} \vee \text{Adv-d} \vee \text{SV}.\text{[MODE infinitif]})$
Unicité	$\boxed{4}$
Exigence	$\boxed{1}.\neg V\text{-}cd \Rightarrow \boxed{4}$
Exclusion	$(\boxed{1}.V\text{-}cd) \vee (\boxed{1}.V\text{-imc}.\text{[MODALISATION DÉPENDANT OBJET ÉTIQUETTE } V\text{-}cd]) \not\Rightarrow \boxed{4}$

FIG. 48 – Le Syntagme Verbal transitif indirect (SV-1oi).

– Soit le complément indirect est réalisé sous une forme non clitique (78),

- (78) a. (Marie) part à la plage  
b. Paul est au chômage  
c. Je compte sur votre aide  
d. Pierre essaie de vendre sa voiture <sup>164</sup>  
e. je vais bien  
f. on part manger

et dans ce cas la Tête est un verbe qui n’a pas reçu de complément clitique, ni direct ni indirect (c’est-à-dire  $\neg V\text{-}ca$  et  $\neg V\text{-}cd$ ). Elle est accompagnée alors d’une (et une seule) construction faisant office de complément d’objet indirect, et pouvant être un Syntagme Prépositionnel (SPrep) (78a, 78d), une Subordonnée (Conj-s), un Adverbe non clitique (Adv-d) (78e) ou un SV à l’infinitif (78f).

Pour ce cas on ajoute l’information intrinsèque de dépendance (OBJ-IND) de l’objet Mineur (donc, le complément non clitique) envers la Tête du verbe qui est constituant immédiat de

<sup>164</sup>Les verbes à contrôle portent donc l’infinitif en tant que complément indirect de verbe.

la présente construction (c'est-à-dire que le recteur de cette relation OBJ-DIR n'est pas la totalité du V tête de la construction, mais le verbe tête du verbe construit). P.ex., dans (79),

(79) (Marie) a voulu aller à la plage

le SV-1oi *a voulu aller à la plage* a pour constituants immédiats le V-im *a voulu aller* et le SPrep *à la plage* ; selon ce qu'on a vu précédemment, la tête du V-im est *aller* ; donc le recteur de la relation OBJ-IND est *aller* (c'est-à-dire la "tête de la tête").

- Soit le complément indirect est réalisé sous une forme clitique (80),

(80) a. (Marie) y part  
b. elle lui fait réparer

et dans ce cas la Tête est un soit verbe avec clitique datif (80a) (V-cd, sans clitique accusatif, i.e.  $\neg$ V-codi<sup>165</sup>), soit un verbe modalisé à verbe causatif (V-imc) dont le modal est un V-cd (80b).

Il est le seul constituant de la construction.

## 2.3 Syntagmes Verbaux à deux compléments

### 2.3.1 Syntagme Verbal transitif direct avec attribut de l'objet (SV-2ao)

Cette construction (fig. 49) hérite du SV à complément direct (SV-1d). Elle peut être constituée, soit d'un "verbe avec clitique accusatif objet direct" (V-cod, §1.6.5) et d'un Adjectif (81a), soit d'un verbe sans clitique accusatif (i.e.,  $\neg$ V-ca), d'un Adjectif et d'un complément d'objet direct (81b).

(81) a. (Jean) la trouve salée  
b. (Jean) trouve l'addition salée

Dans ce type de construction, l'attribut s'accorde en genre et nombre avec l'objet du SV (81).

### 2.3.2 Syntagme Verbal bitransitif (SV-2odi)

Le SV bitransitif (fig. 50) hérite à la fois des caractéristiques du SV transitif direct (SV-1od) et transitif indirect (SV-1oi). Il peut être constitué de plusieurs façons différentes :

- Soit sa Tête est un verbe dont la transitivité est nulle (i.e., n'ayant pas reçu de complément clitique, donc  $\neg$ V-ca et  $\neg$ V-cd), et elle est alors accompagnée d'un complément d'objet direct

<sup>165</sup>Le V-codi hérite du V-cod, qui lui-même hérite du V-ca, or on a déjà éliminé le V-ca des têtes possibles dans la configuration précédente : il est donc inutile de préciser cette étiquette ici dans les propriétés.

SV-2ao HERIT SV-1d	
INTR.	<div style="display: flex; align-items: center;"> <div style="margin-right: 10px;"> <div style="border-left: 1px solid black; padding-left: 5px;">ID</div> <div style="border-left: 1px solid black; padding-left: 5px;">SYN</div> </div> <div> <div style="border-left: 1px solid black; padding-left: 5px;">OBJ-DIR</div> <div style="border-left: 1px solid black; padding-left: 5px;">ATTR-OBJ</div> </div> <div style="margin-left: 10px;"> <div style="border-left: 1px solid black; padding-left: 5px;"> RECTEUR [1].¬V-ime ? : [1].MODAL.   DÉP.   SYN   TÊTE  DÉPENDANT [2] ? : [1].OBJ-DIR   DÉPENDANT </div> <div style="border-left: 1px solid black; padding-left: 5px;"> RECTEUR [2] ? : [1].OBJ-DIR   DÉPENDANT  DÉPENDANT [3] </div> </div> </div>
Exigence	[1] ⇒ [3] [1].¬V-ca ⇒ [2]
Accord	[2].GENRE ∼ [3].GENRE [2].NOMBRE ∼ [3].NOMBRE

FIG. 49 – Le Syntagme Verbal transitif direct avec attribut de l'objet (SV-2ao).

SV-2odi HERIT SV-1od HERIT SV-1oi	
Majeur	[1]V.[ÉTIQUETTE ¬V-cas]
Exigence	[1].¬V-ca ⇒ [2] [1].¬V-cd ⇒ [4]
Exclusion	[1].V-cd ≠ [4] [1].V-ca ≠ [2]
Adjacence	[1] ↔ [2] ; [4]

FIG. 50 – Le Syntagme Verbal bitransitif (SV-2odi).

(précédemment indexé [2]) et d'un complément d'objet indirect (précédemment indexé [4]). Les trois constituants doivent être adjacents les uns aux autres, sans ordre de préférence (82).

- (82) a. je donne un cadeau à Marie  
b. (Marie) aide Paul à travailler  
c. j'entends les oiseaux chanter  
d. (Marie) empêche Paul de travailler  
e. (Pierre) rend à Paul sa trousse  
f. (Marie) sent arriver la pluie  
g. (Pierre) en donne un à Marie  
h. (Pierre) peut rendre sa trousse à Paul  
i. Paul apprend à nager à sa petite soeur  
j. elle fait manger sa soupe à Pierre

– Soit sa Tête est un verbe ayant reçu (uniquement) un complément direct clitique (donc un V-cod mais pas un V-codi), et elle est alors accompagnée uniquement d'un complément indirect (83).

- (83) a. (Pierre) le donne à Marie  
b. Paul l'apprend à sa petite soeur



- c. elle la fait manger à Pierre
- Soit sa Tête est un verbe ayant reçu (uniquement) un complément indirect clitique (donc un V-cd mais pas un V-codi), et elle est alors accompagnée uniquement d'un complément direct (84).
- (84) a. (Pierre) lui donne un cadeau  
b. Paul lui apprend à nager  
c. elle lui fait manger sa soupe
- Soit enfin, sa Tête est un verbe ayant reçu à la fois un complément direct et indirect clitiques (donc un V-codi), et elle est figure alors comme unique constituant (85).
- (85) a. (Pierre) le lui donne  
b. Paul le lui apprend  
c. elle la lui fait manger

## 2.4 Syntagmes Verbaux à trois compléments

### 2.4.1 Syntagme Verbal tritransitif (SV-3)

Cette construction (fig. 51) hérite du SV bitransitif (SV-2odi). Elle représente les constructions causatives dont le verbe tête est un verbe trivalent dont les trois actants sont réalisés (86).

- (86) il fait donner un cadeau à Marie par Pierre

SV-3 HERIT SV-2odi	
$\text{INTR.} \left[ \text{SYN} \left[ \text{OBJ-IND2} \left[ \begin{array}{l} \text{RECTEUR} \quad [1].\text{MODAL.} \mid \text{DÉP.} \mid \text{SYN} \mid \text{TÊTE} \\ \text{DÉPENDANT} \quad [5] \end{array} \right] \right] \right]$	
Majeur	$[1]V.[\text{ÉTIQUETTE } (V\text{-imc} \wedge \neg V\text{-cas})]$
Mineur	$[5]\text{SPrep.}[\text{SPEC.} \mid \text{RECTEUR} \mid \text{OBJ.} \mid \text{FORME} \mid \text{GRAPH } par]$
Unicité	$[5]$
Exigence	$[1] \Rightarrow ([4]; [5])$
Exclusion	$[1].V\text{-cd} \not\Leftarrow ([4] \vee [5])$
Adjacence	$[1] \leftrightarrow [2]; [4]; [5]$

FIG. 51 – Le Syntagme Verbal tritransitif (SV-3).

Dans tous les cas, donc, sa Tête est un verbe causatif (V-imc), et ne peut avoir reçu d'attribut du sujet ( $\neg V\text{-cas}$ ). La construction peut en outre être constituée de plusieurs façons différentes :

- Si sa Tête est un verbe n'ayant reçu aucun clitique complément (donc  $\neg V\text{-ca}$  et  $\neg V\text{-cd}$ ), alors elle est accompagnée d'un complément direct (un SN, précédemment indexé [2]) et de

deux compléments indirects (dont le premier indexé précédemment [4] et le second est un SPrep introduit par *par*, indexé [5]) (86).

- Si sa Tête est un verbe ayant reçu (uniquement) un clitique accusatif (donc V-ca et  $\neg$ V-cd), alors elle exclut la présence d'un autre complément direct ([2]) et réclame celle de deux compléments indirects ([4] et [5]) (87).

(87) il le fait donner à Marie par Pierre

- Si sa Tête est un verbe ayant reçu (uniquement) un clitique datif (donc V-cd et  $\neg$ V-ca), alors elle réclame la présence d'un complément direct ([2]) et d'un complément indirect ([5]) (88).

(88) a. il lui fait donner un cadeau à Marie  
b. il lui fait donner un cadeau par Pierre

- Si sa Tête est un verbe ayant reçu un clitique accusatif et un clitique datif (V-codi), alors elle réclame la présence d'un complément indirect ([5]) uniquement (89).

(89) a. il le lui fait donner à Marie  
b. il le lui fait donner par Pierre

### Résumé

Dans ce chapitre nous avons traité des constructions de type verbal, c'est-à-dire des Verbes syntaxiquement construits (V), et des Syntagmes Verbaux (SV). Des éléments de grammaire que l'on présente dans ce travail, il s'agit de la hiérarchie la plus complexe actuellement.

Parmi les **Verbes** construits, nous faisons figurer les constructions dont la tête est un verbe principal ou modal (i.e., non auxiliaire) et qui comptent parmi leurs constituants (sans exclusion mutuelle) une négation, un ou plusieurs auxiliaires, un ou plusieurs modaux, un ou plusieurs pronoms clitiques. Nous avons montré (assez brièvement, en guise de première approche) comment l'on représente dans la grammaire les verbes à contrôle, à montée et causatifs.

Parmi les **Syntagmes Verbaux**, on a vu deux grands types de constructions : les SV exocentriques, dont le constituant Majeur est un verbe et les Mineurs des compléments, et les SV endocentriques dont le constituant Majeur est un SV et les Mineurs des circonstants.

Pour compléter ce chapitre, le lecteur peut se référer au chapitre 13 dans lequel on illustre le traitement des constructions verbales à pronoms clitiques en donnant quelques exemples d'analyses pas à pas.



## Chapitre 8

# Constructions propositionnelles

Dans ce chapitre nous présentons les constructions non pas de TYPE, mais de CATÉGORIE *P*<sup>166</sup>. On y rassemble les objets qui portent un statut phrastique **autonome** du fait de leur **constitution syntaxique** (90).

- (90) a. Chaque jour grandissait son angoisse<sup>167</sup>  
b. oui  
c. A la maison, c'est plus intime qu'au restaurant.<sup>168</sup>  
d. Je suis parti. Avec Marie. Parce qu'il le fallait bien.<sup>169</sup>  
e. mon frère sa moto le guidon il est cassé  
f. c'est la Nature qui me donne mes valeurs<sup>170</sup>  
g. Voilà ma maison  
h. Noël au scanner, Pâques au cimetierre<sup>171</sup>  
i. Heureusement qu'il est venu

Cela signifie tout d'abord que nous n'y incluons pas les subordonnées (qui sont traitées dans le chapitre 9) : en effet celles-ci ne sont pas sur le même paradigme que les propositions<sup>172</sup>.

Cela signifie également que nous n'y incluons pas les objets qui acquièrent leur statut "phrastique" (ou équivalent, cf. à ce propos les discussions sur le statut de "phrase" à l'oral, notamment [Berrendonner, 2002], [Blanche-Benveniste, 2002], [Kleiber, 2003]) uniquement du fait de leur contexte linguistique (typiquement, à l'écrit, commencer par une majuscule et finir par un

---

<sup>166</sup>Le type *P* n'existe que dans un cas particulier : celui des pro-propositions *oui* et *non*.

<sup>167</sup>[Le Goffic, 1993, p. 132].

<sup>168</sup>[Le Goffic, 1993, p. 136].

<sup>169</sup>[Deulofeu, 2003].

<sup>170</sup>francom.esoterisme - Jul 19 2001.

<sup>171</sup>Pierre Desproges.

<sup>172</sup>A la rigueur, on aurait pu considérer que la subordonnée est à la proposition ce que le syntagme nominal est au nom, i.e. une construction syntaxique dont le premier constituant est un spécifieur, et le second un objet du paradigme de la proposition, la différence étant que N et SN portent le même type (N) alors que ce n'est pas le cas de proposition et subordonnée, et que la Proposition hérite de SX et non de X.

point <sup>173</sup>) (91).

- (91) a. La revendication qu’aucun théoricien ne crée puisqu’elle est seule à fonder la création poétique. <sup>174</sup>  
b. – Tiens, j’ai croisé Alice hier  
– **Qui ?**  
c. – Je vais faire les courses, est-ce que tu veux quelque chose ?  
– **Des oeufs**

On différencie les deux sur la base du fait que les premières, celles que l’on considère comme étant des Propositions, peuvent être constituants immédiats de Subordonnées (92),

- (92) a. il me dit que c’est la Nature qui lui donne ses valeurs  
b. je pense qu’heureusement il est venu

alors que ce n’est pas le cas des secondes (93).

- (93) a. \* il me dit que qui  
b. \* je pense que des oeufs

Enfin, cela signifie que parmi ce que nous classons dans les Propositions, nous ne donnons pas aux structures à tête verbale un statut spécifique, qui aurait pour conséquence que les phrases averbales se voient attribuer un statut particulier (“atypiques” selon [Riegel *et al.*, 1999], aux structures propres selon [Lefeuve, 1999]). En ce sens, nous partageons l’avis de [Deulofeu, 2004], qui fait l’hypothèse que les énoncés averbaux n’ont pas de structure syntaxique spécifique.

Ainsi, nous distinguons le fait d’être une **proposition**, qui requiert de satisfaire un ensemble de contraintes syntaxiques que nous développons dans ce qui suit, et le fait d’être une “**phrase**” <sup>175</sup>, ce qui peut être le cas d’une proposition (mais pas toujours, e.g. pas dans le cas où celle-ci est insérée dans une phrase plus grande via un processus de subordination) mais pas uniquement, et dépend d’un ensemble de contraintes bien plus hétérogène que ce que l’on a pour but de présenter ici (cf. par exemple à ce propos [Morel, 2003]). Autrement dit, nous considérons que les Propositions sont des constructions syntaxiques, ce qui n’est pas le cas des “phrases”, et que les deux entités ne sauraient être confondues.

L’une des conséquences de cela est que la proposition ne prend pas le statut des “structures initiales” utilisées par les théories générativistes (cf. à ce propos chapitre 1, section 1). En d’autres termes, le fait d’introduire une proposition en tant que construction ultime d’une analyse, couvrant la totalité de l’entrée à traiter, n’est non seulement pas le but recherché, mais n’est pas non plus systématique : en effet les énoncés de (91) ne porteront pas ce type puisqu’ils ne répondent pas aux propriétés syntaxiques des propositions telles que définies ci-après, ce qui n’en empêchera pas l’analyse.

On donne la hiérarchie des constructions propositionnelles en fig. 52.

---

<sup>173</sup>Notons d’ailleurs que les formes appelées *epexégèses* (90d) peuvent être considérées comme constituant une seule Proposition en dépit des ponctuations “fortes” séparant leurs constituants.

<sup>174</sup>Raoul Vaneigem (1967), *Traité de savoir-vivre à l’usage des jeunes générations*, Folio (Paris), p. 144.

<sup>175</sup>Nous n’utilisons ce terme que par simple commodité, tout en gardant à l’esprit les limites qui sont opposées à la notion de “phrase”.

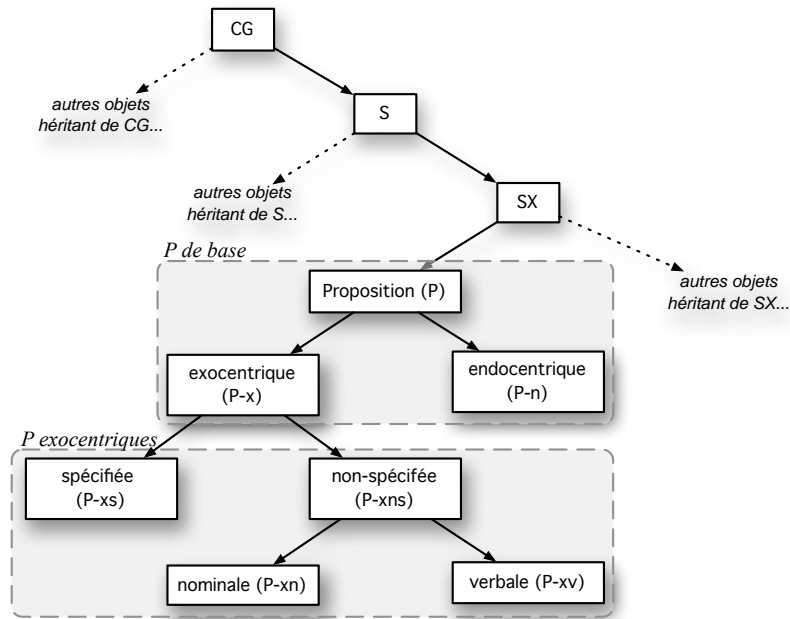


FIG. 52 – Hiérarchie des constructions Propositionnelles dans la grammaire.

## 1 Propositions

### 1.1 Propositions de base

#### 1.1.1 Proposition (P)

Cette construction (fig. 53) hérite du Syntagme X (SX). On y indique qu'il s'agit d'une construction de CATÉGORIE *P*. Elle hérite de SX le fait que son TYPE est celui de sa TÊTE (donc il ne sera pas *P* dans la majorité des cas).

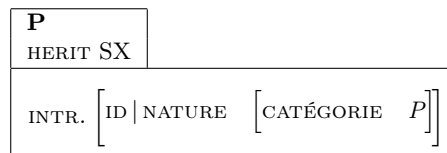


FIG. 53 – La Proposition (P).

#### 1.1.2 Proposition endocentrique (P-n)

La Proposition endocentrique (fig. 54) hérite de la Proposition (P).

<b>P-n</b>	
HERIT P	
INTR.	[ SYN   EXPANSION [ RECTEUR [1] DÉPENDANT [2] ] ]
Majeur	[1]P
Mineur	[2]~P

FIG. 54 – La Proposition endocentrique (P-n).

Elle a pour constituants une Proposition (elle-même exo- ou endocentrique) et un complément de phrase (et un seul à la fois, comme on l'a vu plus haut) (90c, 90e, 94).

- (94) a. cette idée bof <sup>176</sup>  
 b. ça il y a de la marge  
 c. Hélas, j'aurais aimé vous trouver caricatural (...) <sup>177</sup>  
 d. il y a des gens on peut pas mettre de nom mais il y en a deux là on peut mettre des noms tout à fait <sup>178</sup>

Ce complément de phrase peut être de toute nature, sauf être une Proposition (auquel cas on entrerait dans la description de l'entassement paradigmatique, cf. chapitre 11). Dans le cas où le complément de phrase est placé entre les deux constituants de la proposition (95), alors la seule conséquence est que la P exocentrique précédemment introduite (*le temps était frisquet*) n'aura pas été satisfaite à 100%.

- (95) a. Le temps, ce matin, était frisquet <sup>179</sup>

### 1.1.3 Proposition exocentrique (P-x)

Cette construction (fig. 55) hérite de la Proposition (P).

<b>P-x</b>
HERIT P

FIG. 55 – La Proposition exocentrique (P-x).

Elle peut être de deux sortes très différentes, qui donnent lieu à l'introduction de deux constructions héritant de celle-ci, et qui expliquent qu'elle ne présente aucune caractéristique spécifique : la première correspond aux exemples (90f, 90g) et fait l'objet de la Proposition exocentrique spécifiée (§1.2.1) et la seconde correspond aux exemples (90a 90d 90e 91c 90i 90a) et fait l'objet de la Proposition non-spécifiée (§1.2.2).

<sup>176</sup>Cet exemple et le suivant : [Deulofeu, 2004].

<sup>177</sup>fr.education.divers, 27 décembre 2003.

<sup>178</sup>[Cappeau & Deulofeu, 2001]. On remarque ici que l'on fait une distinction d'analyse syntaxique entre les formes *il y (en) a ... qui* (proposition exocentrique dont *il y a* est le spécifieur) et *il y (en) a ... ils* (construction endocentrique dont les deux constituants immédiats sont *il y a ...* d'une part et *ils ...* d'autre part).

<sup>179</sup>[Riegel *et al.*, 1999].



En fait, c'est à ce niveau précis que l'on pourrait rajouter (au moins une partie) des informations de domaines autres que syntaxiques et qui sont communes aux propositions exocentriques, comme par exemple les structurations en *topique* et *commentaire*, etc.

## 1.2 Propositions exocentriques

### 1.2.1 Proposition exocentrique spécifiée (P-xs)

Cette construction (fig. 56) hérite de la Proposition exocentrique (P-x). Sa caractéristique principale est d'être constituée d'un deuxième objet Majeur, qui fait office de SPÉCIFIEUR (96). Cet objet est une Préposition propositionnelle (Prep-p), qui peut être lexicale (*voilà*, *voici*) ou syntaxiquement construite (*c'est*, *il y a*). Le cas échéant, la préposition s'accorde en nombre avec le SN (elle ne s'accorde pas en genre puisqu'elle ne porte jamais ce trait).

- (96) a. c'est gentil de donner des fleurs! <sup>180</sup>  
 b. C'est bien de connaître la géographie, la géologie et aussi l'agriculture, mais il est encore plus important de connaître nos terres intérieures et d'apprendre comment éviter les unes et pénétrer dans les autres pour les entretenir et les cultiver. <sup>181</sup>  
 c. c'est qu'il commence à me connaître, celui-là <sup>182</sup>  
 d. Voilà que les autrichiens s'y mettent aussi <sup>183</sup>  
 e. il y a une mouche dans ma soupe  
 f. il y a des gens qui travaillent le dimanche  
 g. il y en a qui travaillent le dimanche  
 h. à malin, malin et demi  
 i. à grand navire longue navigation <sup>184</sup>

L'autre constituant Majeur (celui qui occupe la fonction de TÊTE) peut être un SN (90g, 96e, 96f, 96g, 96i), un Adjectif (96a, 96h), un Adverbe (96b) ou une Subordonnée conjonctive (96c, 96d).

La construction peut en outre recevoir un constituant supplémentaire, Mineur cette fois, qui peut être un SPrep à tête verbale (96a, 96b, 96e) ou bien une construction de même catégorie que la tête à l'exclusion des subordonnées (96h, 96i) si le spécifieur est *à*. Ce Mineur est exigé par la préposition *à*, mais reste facultatif pour les autres cas ; il est par contre toujours unique.

Notons que pour (96f) nous considérons que *qui travaillent le dimanche* n'est pas un troisième constituant immédiat de P-xs, mais un constituant immédiat du N endocentrique *gens qui travaillent le dimanche* ; nous posons, en effet, que la structure syntaxique conserve cette forme, bien que la structure informationnelle, elle, exprime que *des gens* sera le thème et *qui travaillent le dimanche* le rhème, ce qui est dû à la présence du spécifieur (et donc à la catégorie Proposition) et non à la constitution syntaxique du SN lui-même.

<sup>180</sup>fr.rec.jeux.jdr.par-forum, 8 février 2001.

<sup>181</sup>fr.com.esoterisme, 30 juin 2001.

<sup>182</sup>fr.rec.moto, 7 août 2003

<sup>183</sup>fr.misc.securite routiére, 6 octobre 2005.

<sup>184</sup>[Deulofeu, 2004].

<b>P-xs</b> HERIT P-x	
INTR.	$\left[ \text{SYN} \left[ \begin{array}{l} \text{SPÉCIFIEUR} \quad \boxed{2} \\ \text{COMPLÉMENT} \quad \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{1} \\ \text{DÉPENDANT} \quad \boxed{3} \end{array} \right] \end{array} \right] \right]$
Majeur	$\boxed{1}(\text{SN} \vee \text{Adj} \vee \text{Adv} \vee \text{Conj-sc})$ $\boxed{2}\text{Prep-p}$
Mineur	$\boxed{3}(\text{SPrep}[\text{TÊTE } V] \vee \text{SN} \vee \text{Adj} \vee \text{Adv})$
Unicité	$\boxed{3}$
Exigence	$(\boxed{2} \hat{a} \wedge \boxed{1}\text{SN}) \Rightarrow \boxed{3}\text{SN}$ $(\boxed{2} \hat{a} \wedge \boxed{1}\text{Adj}) \Rightarrow \boxed{3}\text{Adj}$ $(\boxed{2} \hat{a} \wedge \boxed{1}\text{Adv}) \Rightarrow \boxed{3}\text{Adv}$
Exclusion	$\boxed{2} \hat{a} \not\sim \boxed{1}\text{Conj-sc}$
Précédence	$\boxed{2} \prec \boxed{1}$ $\boxed{2} \prec \boxed{3}$ $\boxed{1} \prec \boxed{3}$
Adjacence	$\boxed{1} \leftrightarrow \boxed{2}$ $\boxed{2} \leftrightarrow \boxed{3}$
Accord	$\boxed{1}.\text{NOMBRE} \sim \boxed{2}.\text{ACCORD}   \text{NOMBRE}$

FIG. 56 – La Proposition exocentrique spécifiée (P-xs).

### 1.2.2 Proposition exocentrique non-spécifiée (P-xns)

Cette construction (fig. 57) rassemble les propositions exocentriques qui ne portent pas de spécifieur (i.e., propositions exocentriques à tête verbale, §1.2.4 et à tête nominale, §1.2.3). Tout comme la précédente elle hérite de la proposition exocentrique, mais ne porte aucune caractéristique spécifique.

<b>P-xns</b> HERIT P-x
---------------------------

FIG. 57 – La Proposition exocentrique non-spécifiée (P-xns).

### 1.2.3 Proposition exocentrique à tête nominale (P-xn)

Cette construction (fig. 58) hérite de la Proposition exocentrique non-spécifiée (P-xns). Elle a pour tête une construction nominale, N (91c, 97a) ou SN (97b, 97c).

- (97) a. Téléviseur couleur très peu servi  
b. pas folle la guêpe  
c. l'addition pour la six

Elle porte également un second constituant immédiat, Mineur, Unique et Exigé, qui peut prendre la forme des compléments de N ou de SN (mis à part celle du clitique *en*). Les deux objets doivent être adjacents, sans contrainte sur leur ordre relatif.

<b>P-xn</b> HERIT P-xns	
INTR. [SYN   SPÉCIFIEUR [2]]	
Majeur	[1][TYPE N]
Mineur	[2](N ∨ Adj ∨ Adv ∨ SPrep ∨ Conj-s)
Unicité	[2]
Exigence	[1] ⇒ [2]
Adjacence	[1] ↔ [2]
Accord	[1].GENRE ∼ [2].ACCORD   GENRE
	[1].NOMBRE ∼ [2].ACCORD   NOMBRE

FIG. 58 – La Proposition exocentrique nominale (P-xn).

L'objet Mineur, s'il est présent et porte les valeurs d'ACCORD correspondantes, s'accorde en GENRE et en NOMBRE avec la tête de la construction.

#### 1.2.4 Proposition exocentrique verbale (P-xv)

Cette construction (fig. 59) est la seconde qui hérite de la Proposition exocentrique non-spécifiée (P-xns).

<b>P-xv</b> HERIT P-xns	
INTR. [SYN   SUJET [RECTEUR [2] DÉPENDANT [1]]]	
Majeur	[1]SV.[MODE ¬( <i>infinitif</i> ∨ <i>modal</i> )]
Mineur	[2](SN ∨ SV.[MODE <i>infinitif</i> ] ∨ Conj-sc ∨ Conj-sr)
Unicité	[2]
Exigence	$[1]. \left[ \begin{array}{l} \text{MODE} \quad \neg \textit{impératif} \\ \text{SYN} \quad \left[ \begin{array}{l} \text{SUJET} \mid \text{RECTEUR} \quad \neg \textit{Pro-pcn} \\ \text{MODAL.} \mid \text{DÉP.} \quad \neg \textit{V-cn} \end{array} \right] \end{array} \right] \Rightarrow [2]$ $[2] \Rightarrow [1].[\text{PERSONNE } 3]$
Exclusion	$[1].[\text{TÊTE } \textit{V-cn}] \not\sim [2]$ $[1].[\text{MODAL.} \mid \text{DEP} \mid \text{ÉTIQUETTE } \textit{V-cn}] \not\sim [2]$ $[1].[\text{MODE } \textit{impératif}] \not\sim [2]$
Accord	$[2].\text{GENRE} \sim [1].\text{GENRE}$ $[2].\text{NOMBRE} \sim [1].\text{NOMBRE}$

FIG. 59 – La Proposition exocentrique verbale (P-xv).

La tête de la construction est un SV tensé, qui répond aux conditions suivantes :

- Il figure seul s'il porte un clitique nominatif, c'est-à-dire soit si sa tête est un V-cn (98a), soit s'il porte un modal qui est lui-même un V-cn (98b).

- (98) a. J'ai pas besoin de grande théorie et que le post fasse de 50 à 100 lignes pour démonter la votre <sup>185</sup>  
b. il fait manger sa soupe à Marie
- Il figure seul également s'il est de mode impératif (99),
- (99) viens ici
- Dans les autres cas il doit être accompagné d'un sujet (non clitique), qui peut prendre la forme d'un SN (100a) ou d'une Subordonnée infinitive (100b), complétive (100c), relative (100d) ou circonstancielle (100e).
- (100) a. Toutes les mesures seront reconduites <sup>186</sup>  
b. Vivre à la campagne est aujourd'hui un luxe qui se paie très cher. <sup>187</sup>  
c. Qu'il démissionne sur le champ serait la meilleure des solutions <sup>188</sup>  
d. Qui dort dîne  
e. Comment Paul a pu s'en sortir reste un mystère <sup>189</sup>

Il s'accorde alors en genre et nombre avec son sujet, et doit être à la 3<sup>e</sup> personne.

---

<sup>185</sup>fr.reseaux.internet.fournisseurs, 8 juillet 2002.

<sup>186</sup>[Riegel *et al.*, 1999].

<sup>187</sup>fr.reseaux.telecoms.adsl, 17 septembre 2004.

<sup>188</sup>[Riegel *et al.*, 1999].

<sup>189</sup>[Le Goffic, 1993, p. 136].

### Résumé

Dans ce chapitre on a traité des constructions propositionnelles. On a vu que l'on ne considère comme telles que les constructions dont ce sont les caractéristiques syntaxiques qui leur confèrent ce statut (et non leur contexte, notamment), ce qui fait que l'on ne confond pas la notion de "proposition" avec celle de "phrase".

On a vu également que l'on ne fait pas non plus figurer les subordonnées parmi les propositions. Celles-ci font l'objet de constructions à part, de type spécifique (et dont l'un des constituants immédiats est une proposition).

On a vu enfin que l'on ne confère pas aux propositions à tête verbale un statut de norme, qui aurait pour conséquence que les constructions propositionnelles averbales se verraient conférer un statut spécial ; dans cette description, les propositions à tête verbale ont un statut équivalent à celui des propositions à tête nominales, qui à eux deux constituent une partie des propositions possibles, partie complétée par les propositions spécifiées d'une part, et par les propositions endocentriques d'autre part.



## Chapitre 9

# Constructions de subordination

Les constructions de [TYPE *Sub*] <sup>190</sup> rassemblent les **Subordonnants** (Sub, §1) et les **Conjoints de subordination** (Conj-s, §2) : dans ce que les grammaires classiques nomment “subordination”, nous appelons les différents éléments de la façon suivante :

- On appelle **Subordonnant** la construction qui fait office de spécifieur de la construction traditionnellement appelée “subordonnée”. P.ex., dans (101), le subordonnant est *quand*.

(101) je vais à la piscine **quand j’en ai le temps**

- On appelle **Subordonné** la construction qui est introduite par le subordonnant. Dans (101), le subordonné est *j’en ai le temps*.
- Enfin, on appelle **Conjoint de subordination** l’ensemble constitué des deux constructions sus-citées. Dans (101), le Conjoint de subordination est *quand j’en ai le temps*.

Cette classification est parallèle à celle de la coordination que l’on propose au chapitre 11 (où le Coordonnant correspond à ce que l’on appelle souvent “conjonction de coordination”, le Coordonné est l’objet introduit par le coordonnant, et le Conjoint de coordination est la construction constituée des deux précédents) <sup>191</sup>.

On donne la hiérarchie des constructions de subordination en fig. 60.

---

<sup>190</sup> Précisons que nous n’utilisons le terme de “subordonné” que pour faire référence aux classiques “propositions subordonnées” et non à toutes les relations littéralement de subordination qui peuvent exister en syntaxe.

<sup>191</sup> C’est d’ailleurs la raison pour laquelle on a appelé *quand j’en ai le temps* “Conjoint de subordination” et non simplement “Subordination” : sinon, dans le parallèle avec la coordination, ce que l’on aurait nommé “Coordination” aurait correspondu à une partie seulement de la coordination, e.g. seulement à *et Paul* dans *Pierre et Paul*.

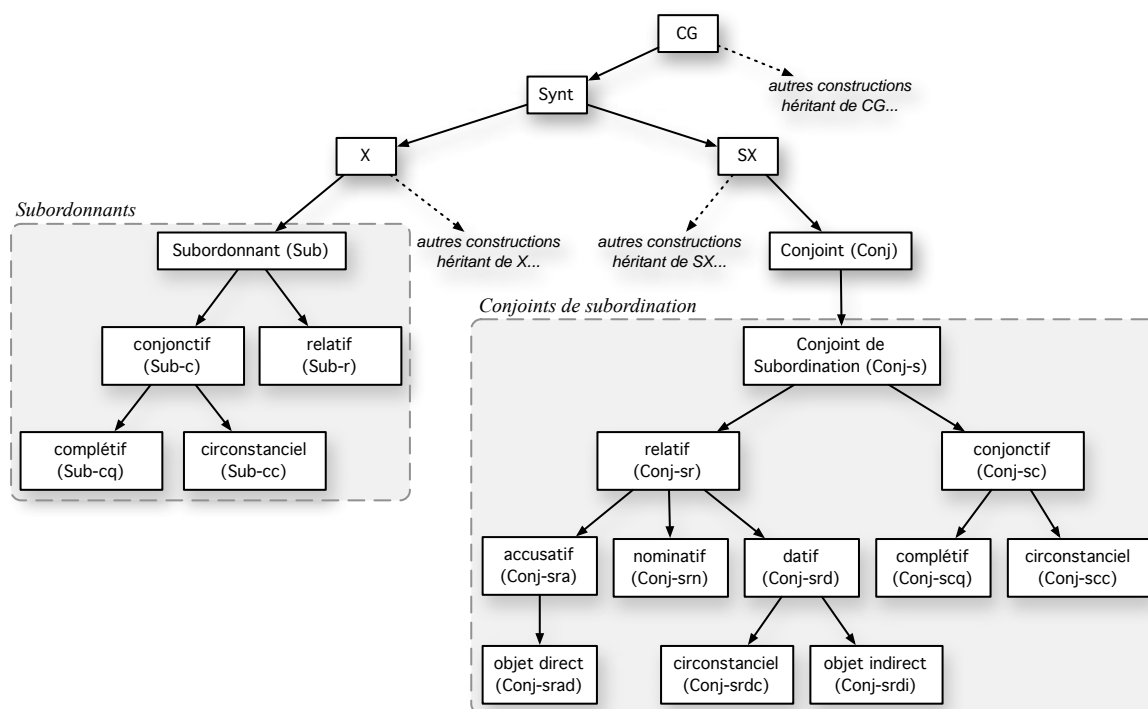


FIG. 60 – Hiérarchie des constructions de subordination dans la grammaire.

## 1 Subordonnants (Sub)

Les Subordonnants (fig. 61) regroupent les objets servant à introduire un SV ou une P en tant qu'élément constitutif d'une unité syntaxique plus grande, i.e. les conjonctions de subordination et les pronoms relatifs.

Sub	
HERIT X	
INTR. $\left[ \text{ID} \mid \text{NATURE} \left[ \text{TYPE} \text{ Sub} \right] \right]$	
Majeur	$\mathbb{I}(\text{Cj-s} \vee \text{Pro-r})$

FIG. 61 – Le Subordonnant (Sub).

Cette construction hérite de X. Elle est de  $[\text{TYPE Sub}]$ , et porte la catégorie de sa tête, qui peut être, comme on l'a dit, soit une conjonction de subordination (Cj-s), soit un pronom relatif (Pro-r). De cette construction générale en héritent deux autres : le Subordonnant conjonctif (Sub-c, §1.1) et le Subordonnant relatif (Sub-r, §1.2).



## 1.1 Subordonnant conjonctif (Sub-c)

Les conjonctions de subordination (Cj-s) (102) constituent des subordonnants conjonctifs (fig. 62).

- (102) a. Se nourrir est un art et un plaisir, je trouve bien triste **que** tu veuilles t'en priver. <sup>192</sup>  
 b. c'est donc ce qui doit se passer **quand** je suis en butée à gauche <sup>193</sup>

<b>Sub-c</b>	
HERIT Sub	
Majeur	□Cj-c

FIG. 62 – Le Subordonnant conjonctif (Sub-c).

De cette construction en héritent deux autres : le Subordonnant complétif (§1.1.1) et le circonstanciel (§1.1.2).

### 1.1.1 Subordonnant complétif (Sub-cq)

La conjonctions de subordination de la forme *que* (étiquetés Cj-sq) constituent, dans leur emploi canonique (103), les subordonnants complétifs (fig. 63).

- (103) a. Partant du principe **que** celui qui s'est penché sur un pb dans un sens est mieux placé que quiconque pour le regarder différemment. <sup>194</sup>  
 b. J'espère **que** ça n'empêche personne de respirer <sup>195</sup>

<b>Sub-cq</b>	
HERIT Sub-c	
Majeur	□Cj-sq

FIG. 63 – Le Subordonnant complétif (Sub-cq).

Ils héritent de Sub-c et ont pour caractéristique d'être des connecteurs syntaxiques dénués de signification propre (contrairement aux subordonnants circonstanciels) et ne portant pas de fonction dans l'objet qu'ils introduisent (contrairement aux relatifs) : ils sont, typiquement, des morphèmes grammaticaux.

### 1.1.2 Subordonnant circonstanciel (Sub-cc)

La construction Sub-cc (fig. 64) regroupe les autres conjonctions de subordination. Celles-ci ont une valeur sémantique propre, mais ne correspondent pas à un actant de la tête de la construction qu'ils introduisent.

<sup>192</sup>fr.sci.philo, 8 octobre 2004.

<sup>193</sup>fr.rec.moto, 11 janvier 2004.

<sup>194</sup>fr.sci.psychanalyse, 2 mars 2005.

<sup>195</sup>*ibid.*.

La construction peut prendre une forme lexicale simple qui peut être celle d’une conjonction de subordination circonstancielle (Cj-sc) (104a) ou alors d’une conjonction de subordination complétive employée circonstancielle (104b), ou alors être syntaxiquement construite et dans ce cas être constituée d’une Cj-sq et d’une préposition (104c), plus éventuellement du pronom personnel *ce* (104d).

- (104) a. sur ma voiture **lorsque** le moteur est éteint et que j’appuie sur la pédale d’embrayage alors tout va bien je peux changer les vitesses sans problèmes. <sup>196</sup>  
 b. la mort nous prend **que** nous sommes encore pleins de nos misères <sup>197</sup>  
 c. Ideus observe nonchalamment le tout, prêt à intervenir **avant que** le stylo n’explose <sup>198</sup>  
 d. **Jusqu’à ce que** Joe Nickell (...) réussisse à redessiner le condor de Nazca avec les instruments de l’époque, sans aucune supervision aérienne. <sup>199</sup>

Notons que dans ce dernier cas il ne s’agit pas du même pronom *ce* que dans la construction adverbiale *ce faisant* ou la nominale *ce que tu as dit* puisque dans ce deux derniers cas, le pronom correspond à un actant de la tête de la subordonnée introduite, alors que dans (104d) ce n’est pas le cas. La différence entre les deux est marquée par la présence, dans la construction introduite par le subordonnant, du mode subjonctif. Ce mode, d’ailleurs, est en réalité un affixe grammatical qui devrait faire partie du Subordonnant et non d’une caractéristique du Subordonné (qui, en toute rigueur, devrait être *Joe Nickel réussit à redessiner (...)*, sans le subjonctif). Mais comme l’on n’a pas encore traité la morphologie séparément dans les constructions telles qu’on les décrit actuellement, on s’est abstenu de le mentionner ici pour des raisons de cohérence avec le reste de la grammaire (et l’on pourra ajouter, pour l’instant, la contrainte de MODE, en tant que TRAIT et non que morphème, dans le subordonné correspondant).

Sub-cc	
HERIT Sub-c	
Mineur	$\boxed{2}$ Prep $\boxed{3}$ <i>ce</i>
Unicité	$\boxed{2}$ $\boxed{3}$
Exigence	$\boxed{2} \Rightarrow \boxed{1}$ Cj-sq $\boxed{3} \Rightarrow \boxed{1}$ Cj-sq $\boxed{3} \Rightarrow \boxed{2}$
Exclusion	$\boxed{2} \not\Rightarrow \boxed{1}$ Cj-sc $\boxed{3} \not\Rightarrow \boxed{1}$ Cj-sc
Précédence	$\boxed{2} \prec \boxed{1}$ $\boxed{2} \prec \boxed{3}$ $\boxed{3} \prec \boxed{1}$
Adjacence	$\boxed{2} \leftrightarrow (\boxed{3} \vee \boxed{1})$ $\boxed{3} \leftrightarrow \boxed{1}$

FIG. 64 – Le Subordonnant circonstanciel (Sub-cc).

Puisque les deux sortes de conjonctions de subordination (complétive et circonstancielle)

<sup>196</sup>fr.misc.automoto.mecanique, 10 décembre 2005.

<sup>197</sup>[Deulofeu, 1999].

<sup>198</sup>fr.rec.jeux.jdr.par-forum, 27 septembre 2006.

<sup>199</sup>FrancoSeniors\_amities, 30 juin 2006.

peuvent être tête de cette construction, on ne re-précise pas ici la propriété Majeur héritée de Sub-c. On ajoute toutefois les propriétés d'Exigence et d'Exclusion qui indiquent que les expansions de la conjonction ne peuvent être ajoutées qu'à une tête complétive. On ajoute également les propriétés de Précédence et d'Adjacence contraignant l'ordre et la contiguïté des constituants.

## 1.2 Subordonnant relatif (Sub-r)

Le subordonnant relatif (fig. 65), enfin, est représenté par les pronoms relatifs (Pro-r). Ils ont la particularité de porter un CAS et un certain nombre de traits d'ACCORD qui leur permet de marquer la fonction qu'ils portent dans la construction qu'ils introduisent (105).

- (105) a. J'ai travaillé avec une boîte de Marseille **qui** faisait des électroencéphalographes (...) <sup>200</sup>  
 b. Ce minable écrit et dirige un journal minable **dont** le but est de faire croire à des minables qu'ils savent utiliser leurs cerveaux. <sup>201</sup>  
 c. l'insatisfaction (...) te mène là **où** elle est impossible (...) <sup>202</sup>

Ils peuvent figurer seuls, ou alors précédés d'une Préposition (106).

- (106) a. Avez-vous déjà réfléchi **à qui** profitent les CPE ? <sup>203</sup>  
 b. (...) mais il ne peut pas choisir l'adresse **avec laquelle** il voudrait envoyer ces mails. <sup>204</sup>

Dans ce cas, la préposition précède immédiatement la tête de la construction.

Sub-r	
HERIT Sub	
Majeur	Recteur( <u>1</u> Pro-r)
Mineur	<u>2</u> Prep
Unicité	<u>2</u>
Précédence	<u>2</u> < <u>1</u>
Adjacence	<u>2</u> ↔ <u>1</u>

FIG. 65 – Le Subordonnant relatif (Sub-r).

La construction hérite de Sub; on précise localement que la Tête, ici, ne peut être qu'un pronom relatif et non plus une conjonction de subordination.

Le pronom peut dépendre de la tête du Subordonné et peut alors soit en représenter un actant (105a) soit un circonstant (105c), ou alors dépendre de l'un de ses constituants (105b). En fait, il est parfois considéré que le pronom relatif correspond à l'amalgame d'un pronom personnel non clitique (p.ex. *lui* dans *qui*) et d'une conjonction de subordination complétive (p.ex. *que*, dans *qui*) : cette position mènerait à ne pas analyser le pronom relatif comme une

<sup>200</sup>fr.comp.os.linux.debats, 21 mai 2004.

<sup>201</sup>fr.sci.philo, 1er novembre 2005.

<sup>202</sup>fr.sci.philo, 2 juin 2005.

<sup>203</sup>qc.politique, 6 décembre 2005.

<sup>204</sup>microsoft.public.fr.exchange.clients, 29 juillet 2005.

construction syntaxique, mais à construire la partie pronominale comme constituant immédiat d'une P, puis de poser que cette P est le second constituant immédiat d'un Conjoint de subordination introduit par la conjonction de subordination. Dans l'immédiat, compte tenu du fait que l'on ne fait pas encore d'analyse morphématique dans notre grammaire, il nous est impossible d'envisager cette possibilité dans notre description sans lui ajouter un ensemble *ad hoc* d'informations, mais nous n'ignorons pas son existence et étudierons cette possibilité dès qu'on aura ajouté une véritable dimension morphologique à nos descriptions.

## 2 Conjoints de subordination (Conj-s)

Ce que l'on appelle habituellement "(propositions) subordonnées" correspond dans notre classification aux constructions que l'on nomme Conjoints de subordination, qui ne sont selon nous ni des "propositions" (puisqu'elles ne sont pas sur le paradigme des Propositions, cf. chapitre 8) ni des "subordonnées" (puisque selon nos définitions les subordonnées sont les constructions introduites par les Subordonnants, i.e. les seconds constituants immédiats du Conjoint de subordination).

Cette construction hérite de la construction générale Conjoint (qui rassemble les caractéristiques communes aux Conjoints de subordination auxquels on s'intéresse ici, et de coordination que l'on décrit dans le chapitre 11 §2.3.2. Nous commençons donc ici par donner la description de la construction Conjoint (§2.1), avant de nous intéresser plus précisément au Conjoint de subordination (§2.2).

### 2.1 Conjoint (Conj)

<b>Conj</b> HERIT SX	
INTR. $\left[ \text{SYN} \left[ \text{SPEC} \quad \boxed{2} \right] \right]$	
Majeur	$\boxed{2}(\text{Cj} \vee \text{Pro-r})$
Précédence	$\boxed{2} \prec \boxed{1}$
Adjacence	$\boxed{1} \leftrightarrow \boxed{2}$

FIG. 66 – Le Conjoint (Conj).

Le Conjoint (fig. 66) est une construction qui hérite de SX ; elle est constituée de deux objets Majeurs, la premier faisant office de tête (et restant sous-spécifié ici), et le second de spécifieur de la construction. Ce second objet peut être une conjonction (Cj, de coordination ou de subordination) ou bien un Pronom relatif (Pro-r). Le spécifieur précède immédiatement l'objet qu'il introduit.

### 2.2 Conjoint de subordination (Conj-s)

<b>Conj-s</b>	
HERIT Conj	
INTR.	$\left[ \text{ID} \mid \text{NATURE} \left[ \begin{array}{cc} \text{TYPE} & \text{Sub} \\ \text{CAT} & \text{Conj-s} \end{array} \right] \right]$
Majeur	$\begin{array}{l} \textcircled{1}(\text{SV} \vee \text{P}) \\ \textcircled{2}(\text{Cj-s} \vee \text{Pro-r}) \end{array}$
Mineur	$\textcircled{3}(\text{SN} \vee \text{Adj} \vee \text{Adv-d} \vee \text{SPrep} \vee \text{Conj-s})$
Unicité	$\textcircled{3}$

FIG. 67 – Le Conjoint de subordination (Conj-s).

Le Conjoint de subordination (fig. 67) hérite du Conjoint. Il porte deux constituants immédiats, le premier étant le Subordonnant (qui constitue son spécifieur), et le second le Subordonné (qui porte sa tête). Le Subordonné peut être un SV ou une P. Il peut également recevoir des expansions, décrits comme des objets Mineurs, et pouvant prendre un certain nombre de formes différentes, comme p.ex. celle d'un SPrep *sans aucune supervision aérienne* dans (104d), ou bien celle d'un Adv dans (107).

- (107) a. Applis qui restent dans le Dock (**même** quand on en veut pas) <sup>205</sup>  
 b. J'ai toujours lu des CD audios sur mon Mac, **y compris** depuis que je suis en 7.5.5. <sup>206</sup>

Les Conjoints de subordination peuvent être relatifs (§2.4), ou conjonctifs circonstanciels (§2.3.2) ou complétifs (§2.3.1). La construction générale présentée ici (Conj-s, fig. 67) rassemble les informations générales à toutes ces constructions, qui en héritent.

### 2.3 Conjoint de subordination conjonctif (Conj-sc)

Le Conjoint de subordination conjonctif (fig. 68) est constitué d'un subordonnant conjonctif (Sub-c), immédiatement suivi d'une proposition P (108).

- (108) a. Se nourrir est un art et un plaisir, je trouve bien triste **que tu veuilles t'en priver**. <sup>207</sup>  
 b. c'est donc ce qui doit se passer **quand je suis en butée à gauche** <sup>208</sup>

<b>Conj-sc</b>	
HERIT Conj-s	
Majeur	$\begin{array}{l} \textcircled{1}\text{P} \\ \textcircled{2}\text{Sub-c} \end{array}$

FIG. 68 – Le Conjoint de subordination conjonctif (Conj-sc).

Deux constructions en héritent : le complétif et le circonstanciel.

<sup>205</sup>fr.comp.os.mac-os.x, 8 janvier 2004.

<sup>206</sup>fr.comp.sys.mac, 17 mai 1997.

<sup>207</sup>fr.sci.philo, 8 octobre 2004.

<sup>208</sup>fr.rec.moto, 11 janvier 2004.

### 2.3.1 Conjoint de subordination conjonctif complétif (Conj-scq)

Le Conjoint de subordination complétif (fig. 69) reçoit un subordonnant complétif (Sub-cq) (108a).

<b>Conj-scq</b>	
HERIT Conj-sc	
Majeur	<u>2</u> Sub-cq

FIG. 69 – Le Conjoint de subordination complétif (Conj-scq).

### 2.3.2 Conjoint de subordination conjonctif circonstanciel (Conj-scc)

Le Conjoint de subordination circonstanciel (fig. 70) est constitué d'un subordonnant circonstanciel (Sub-cc) (108b).

<b>Conj-scc</b>	
HERIT Conj-sc	
Majeur	<u>2</u> Sub-cc

FIG. 70 – Le Conjoint de subordination circonstanciel (Conj-scc).

## 2.4 Conjoint de subordination relatif (Conj-sr)

Le Conjoint de subordination relatif (fig. 71) est constitué d'un subordonnant relatif, immédiatement suivi d'un syntagme verbal (non propositionnel) (105).

<b>Conj-sr</b>	
HERIT Conj-s	
Majeur	<u>1</u> SV
	<u>2</u> Sub-r

FIG. 71 – Le Conjoint de subordination relatif (Conj-sr).

On fait hériter de ce Conjoint de subordination autant de constructions que le pronom relatif peut recevoir de fonctions différentes.

### 2.4.1 Conjoint de subordination relatif nominatif (Conj-srn)

Ce Conjoint de subordination (fig. 72) est constitué d'un pronom relatif nominatif qui occupe la fonction sujet (109). Il réclame un SV ne portant pas déjà de sujet, i.e. n'ayant pas pour tête un V-cn.

(109) Une pizza est offerte pour ce ou celle **qui m'aide à trouver une solution** <sup>209</sup>

Conj-srn	
HERIT Conj-sr	
INTR.	$\left[ \text{SYN} \mid \text{SUJET} \left[ \begin{array}{l} \text{RECTEUR} \quad [2] \\ \text{DÉPENDANT} \quad [1] \end{array} \right] \right]$
Majeur	$[1]\text{SV}.[\text{TÊTE} \neg V\text{-cn}]$ $[2]\text{Sub-r.}[\text{CAS } \textit{nominatif}]$
Accord	$[2].\text{PERSONNE} \rightsquigarrow [1].\text{PERSONNE}$ $[2].\text{NOMBRE} \rightsquigarrow [1].\text{NOMBRE}$

FIG. 72 – Le Conjoint de subordination relatif nominatif (Conj-srn).

### 2.4.2 Conjoint de subordination relatif accusatif (Conj-sra)

Ce Conjoint de subordination (fig. 73) est constitué d'un pronom relatif accusatif (Sub-ra), qui occupe la fonction d'objet direct (110), mais pas d'attribut.

(110) a. La conviction **que je défends** et pour laquelle je n'ai pas la moindre tolérance <sup>210</sup>

Conj-sra	
HERIT Conj-sr	
INTR.	$\left[ \text{SYN} \mid \text{OBJ-DIR} \left[ \begin{array}{l} \text{RECTEUR} \quad [1] \\ \text{DÉPENDANT} \quad [2] \end{array} \right] \right]$
Majeur	$[2]\text{Sub-r.}[\text{CAS } \textit{accusatif}]$

FIG. 73 – Le Conjoint de subordination relatif accusatif (Conj-sra).

### 2.4.3 Conjoint de subordination relatif datif (Conj-srd)

Ce Conjoint de subordination (fig. 74) est constitué d'un pronom relatif datif, qui peut occuper la fonction d'objet indirect ou de circonstant (111).

- (111) a. nous pouvons réaliser une réplique TOTALE de l'individu **à qui nous avons plélevé cette cellule**, AVEC ses souvenirs, sa personnalité, son caractère, etc. <sup>211</sup>  
b. une petite erreur dans mon explication et je m'en excuse, ce n'est pas de l'énergie **dont je voulais parler** mais de la puissance en Watts (...) <sup>212</sup>  
c. l'insatisfaction (...) te mène là **où elle est impossible** (...)

<sup>209</sup>microsoft.public.fr.access, 26 juillet 2004.

<sup>210</sup>fr.sci.philo, 9 juil 2003.

<sup>211</sup>soc.culture.quebec, 2 juillet 2000.

<sup>212</sup>fr.soc.economie, 23 octobre 2000.

<b>Conj-srd</b>	
HERIT Conj-sr	
Majeur	[2]Sub-r.[CAT <i>datif</i> ]

FIG. 74 – Le Conjoint de subordination relatif datif (Conj-srd).

Pour le premier cas, la construction réclame un SV ne portant pas déjà cette fonction, i.e. qu'elle exclut les SV-1-oi. Ceci donne lieu à deux constructions qui héritent de celle-ci : le Conj-srd objet indirect (Conj-srdi, fig. 75) et le Conj-srd circonstantiel (Conj-srdc, fig. 76).

<b>Conj-srdi</b>	
HERIT Conj-srd	
INTR. $\left[ \text{SYN} \mid \text{OBJ-IND} \left[ \begin{array}{l} \text{RECTEUR} \quad [1] \\ \text{DÉPENDANT} \quad [2] \end{array} \right] \right]$	
Exclusion	[2] $\not\rightarrow$ [1]SV-1oi

FIG. 75 – Le Conjoint de subordination relatif datif objet indirect (Conj-srdi).

<b>Conj-srdc</b>	
HERIT Conj-srd	
INTR. $\left[ \text{SYN} \mid \text{CIRCONSTANT} \left[ \begin{array}{l} \text{RECTEUR} \quad [1] \\ \text{DÉPENDANT} \quad [2] \end{array} \right] \right]$	

FIG. 76 – Le Conjoint de subordination relatif datif circonstanciel (Conj-srdc).



### Résumé

Dans ce chapitre nous avons présenté les constructions de subordination. Elles se découpent en deux catégories : les Subordonnants et les Conjoint de subordination.

Les Subordonnants rassemblent les Conjonctions de subordination et les Pronoms relatifs, pouvant se présenter soit seuls, sous une forme lexicale, soit dans certains cas construits avec une préposition (et éventuellement un pronom personnel *ce*).

Les Conjoint de subordination représentent les constructions traditionnellement appelées “(propositions) subordonnées”. Ils héritent de la même construction que les Conjoint de coordination, et sont constitués d’un Subordonnant et d’un Subordonné qui peut être une P ou un SV, plus ou moins contraint, suivant les cas.

A travers ces descriptions on a vu le rapprochement terminologique (et pas là, descriptif) que l’on a effectué entre la coordination et la subordination, qui sont décrits parallèlement : le Conjoint de subordination, le Subordonnant et le Subordonné sont à la subordination ce que le Conjoint de coordination, le Coordonnant et le Coordonné sont respectivement à la coordination.



# Chapitre 10

## Constructions adjectivales

Contrairement aux deux types de constructions que l'on a vus dans les chapitres précédents (nominales et verbales), les constructions de [TYPE *Adj*] ne contiennent que ce que l'on appelle, sur la base des autres dénominations que l'on a adoptées dans la grammaire, des **Adjectifs** syntaxiquement construits, i.e. des constructions endocentriques dont la tête est un Adjectif, et qui restent sur le paradigme de l'adjectif lexical <sup>213</sup>. Nous en présentons la hiérarchie en figure 77.

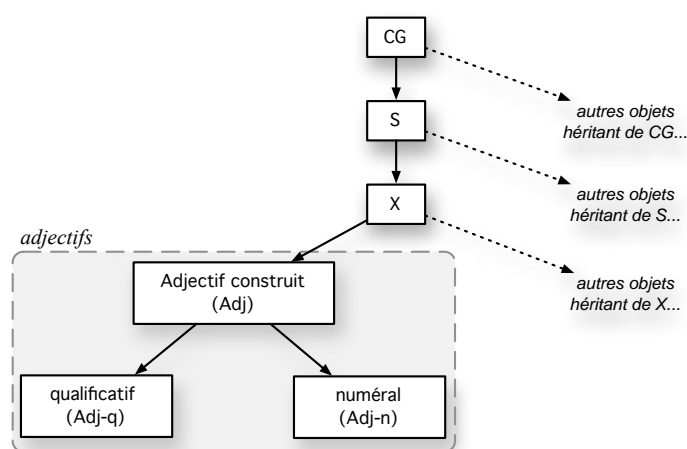


FIG. 77 – Hiérarchie des Adjectifs construits dans la grammaire.

Concrètement, bien qu'il existe plusieurs types d'adjectifs lexicaux <sup>214</sup>, tous ne peuvent prendre une forme syntaxiquement construite : seuls le peuvent les qualificatifs et les numéraux. C'est la raison pour laquelle dans notre hiérarchie eux seuls héritent de l'Adjectif construit, et sont décrits dans ce chapitre.

<sup>213</sup>Dans de très nombreuses grammaires on appelle cela des Syntagmes Adjectivaux.

<sup>214</sup>Les types d'adjectifs lexicaux que l'on utilise ici sont les suivants : qualificatif (Adj-q), relationnel (Adj-r), numéral (Adj-n) et comparatif (Adj-c).

## 1 Adjectifs construits (Adj)

Les Adjectifs syntaxiquement construits (fig. 78) héritent du Syntagme endocentrique (S-n). Ils sont constitués d'un Adjectif et de leur complément (112).

- (112) a. ce n'est pas non plus pour pour se faire reprocher de légitimer de la censure lorsqu'on émet l'hypothèse que Thierry était **légèrement provocateur** <sup>215</sup>  
 b. l'on ne va pas se priver de "parapluies" permettant d'ignorer des infinités de questions simultanément, sauf à se faire **délibérément plus bêtes que nécessaire**. <sup>216</sup>  
 c. Maturité du potiron "**Rouge vif** d'Etampes" ? <sup>217</sup>  
 d. la dernière raie de lumière du soleil couchant apparait d'un beau **vert émeraude** <sup>218</sup>  
 e. Le problème c que j'ai toujours entenu dire que les lapins nains c'était **gentil comme tout**. <sup>219</sup>

Adj HERIT X	
INTR.	$\left[ \begin{array}{l} \text{ID   NATURE} \left[ \begin{array}{l} \text{TYPE } Adj \\ \text{CATÉGORIE } Adj \end{array} \right] \\ \text{SYN} \left[ \begin{array}{l} \text{CPLT} \left[ \begin{array}{l} \text{RECTEUR } [1] \\ \text{DÉPENDANT } [2] \end{array} \right] \end{array} \right] \end{array} \right]$
Majeur	[1](Adj-q ∨ Adj-n ∨ V-p ∨ N ∨ SPrep ∨ Adv-d)
Mineur	[2]SPrep
Unicité	[2]
Exigence	[1].Adj ⇒ [2]
Adjacence	[1] ↔ [2]
Accord	[1].GENRE ∼ [2].GENRE
	[1].NOMBRE ∼ [2].NOMBRE

FIG. 78 – L'Adjectif construit (Adj).

Dans cette construction générale, on ne peut préciser que les compléments possibles à la fois pour les Adj-q et les Adj-n, c'est-à-dire uniquement les Syntagmes Prépositionnels. Les autres compléments possibles pour le qualificatif sont donnés dans la construction Adj-q.

### 1.1 Adjectifs qualificatifs (Adj-q)

Les adjectifs qualificatifs construits (fig. 79) rassemblent d'une part les objets passant sur le paradigme de l'Adjectif (Noms, Adverbes, etc.) et d'autre part les Adjectifs qualificatifs recevant une complément.

La nature de la tête est variable ; elle peut être :

<sup>215</sup>fr.usenet.abus.d, 20 avril 2001.

<sup>216</sup>fr.sci.philo, 16 mars 2004.

<sup>217</sup>fr.rec.jardinage, 13 septembre 2005.

<sup>218</sup>fr.sci.astronomie, 27 août 1998.

<sup>219</sup>fr.rec.animaux, 6 janvier 2003.

Adj-q HERIT Adj	
Majeur	①(Adj-q $\vee$ V-p $\vee$ N $\vee$ SPrep $\vee$ Adv-d)
Mineur	②(Conj-scq; Conj-sr; Adv-d; Adj-q; N-c; SPrep)
Précédence	Adv $\prec$ ①
	① $\prec$ $\neg$ Adv

FIG. 79 – L'Adjectif qualificatif (Adj-q).

- un Adjectif qualificatif (Adj-q) (112),
  - un Verbe au participe (V-p) présent (113a) ou passé (113b),
- (113) a. coupant, bouleversant, etc.  
b. blasé, usé, vieilli, etc.
- un Nom, construit (114a) ou non (114b, 114c),
- (114) a. vieille France, bon enfant, bon marché, etc.  
b. une tarte **maison**, un air **canaille**, etc.  
c. il est très **vache**, **joueur**, **moqueur**, **menteur**, etc.
- un Syntagme Prépositionnel (115),
- (115) a. de bonne humeur  
b. sur ses gardes, en colère, etc.
- un Adverbe non clitique, simple (116a) ou construit (116b, 116c). Dans le cas de (116c), le complément de l'adverbe prend la forme du clitique *en*, situé à distance de l'adverbe puisqu'il “monte” sur le verbe.
- (116) a. il est **bien**, **super**, etc.  
b. pas mal, etc.  
c. j'**en** suis **loin**

Il n'y a que dans le cas où la tête est un Adjectif que la construction exige un complément (bien qu'elle puisse en recevoir même s'il n'est pas réclamé). Ce complément peut lui-même prendre plusieurs formes différentes :

- un Adverbe non clitique (Adv-d) (112a, 112b, 117),
- (117) a. La soupe est **un peu** / **assez** / **très** / **trop froide** (**pour en manger** / **pour la saison** / **pour de la bonne soupe** / **pour qu'on en mange**)<sup>220</sup>  
b. La soupe est **plus** / **moins** / **aussi chaude** (**que le café** / **que salée** / **que hier soir** / **que chez Yvonne**)<sup>221</sup>
- un Adjectif qualificatif (Adj-q) (112c),

<sup>220</sup>D'après [Riegel *et al.*, 1999].<sup>221</sup>[Riegel *et al.*, 1999].

– un Nom commun (N-c) (112d),

– un SPrep (112e, 118) <sup>222</sup>,

(118) a. savoir si ceux qui se sont livré à ces actes sont vraiment **aptes à éduquer nos enfants**. <sup>223</sup>

b. La France est-elle tellement **réfractaire à la nouveauté** ? <sup>224</sup>

c. Étiez-vous en mesure de rencontrer des femmes russes **contentes de leur choix** ? <sup>225</sup>

d. Si vous souhaitez être **attentive à tout ce que vous faites**, pourquoi pas, mais en même temps, des tas d'informations vous échapperont. <sup>226</sup>

– ou une Subordonnée complétive (Conj-scq) (119).

(119) moi je suis méga **content que tu soies contente que Marko soit content de te revoir** <sup>227</sup>

## 1.2 Adjectifs numéraux (Adj-n)

Les adjectifs numéraux (fig. 80) sont constitués d'un Adjectif numéral et de son complément, qui ne peut être qu'un SPrep (120) (ce qui est hérité de la construction Adj).

(120) une élève **première de la classe**

<b>Adj-n</b>	
HERIT Adj	
Majeur	$\boxed{\text{I}}\text{Adj-n}$

FIG. 80 – L'Adjectif numéral (Adj-n).

On se contente donc de préciser la nature du constituant Majeur de la construction.

---

<sup>222</sup>Le fait d'accepter ce type de modificateur dépend de la sémantique de l'adjectif; p.ex., *apte* le réclame, *fier* le permet, et *intelligent* l'interdit. Ceci dépend de la valence de l'adjectif, et est une information qui doit donc, au même titre que la valence verbale, figurer dans le lexique. Ce n'est actuellement pas le cas, donc nous n'en tenons pas compte ici.

<sup>223</sup>fr.soc.economie, 30 juin 2003.

<sup>224</sup>fr.rec.tv.series.sf, 29 mars 2000.

<sup>225</sup>soc.culture.belgium, 21 mars 2002.

<sup>226</sup>fr.sci.philo, 30 décembre 2004.

<sup>227</sup>fr.rec.arts.musique.rock, 4 janvier 2002.

### Résumé

On a vu dans ce chapitre les constructions de type adjectival. Celles-ci, n'héritant que de X, ne peuvent être que des **Adjectifs** construits (autrement dit, on n'a pas de Syntagmes Adjectivaux dans la grammaire). En outre, d'un point de vue syntaxique, il n'y a que les Adjectifs qualificatifs et les numéraux qui peuvent être construits ; ce sont donc les seuls à figurer dans cette description, les autres (comparatif et relationnel) ne font pas l'objet d'une construction syntaxique.





## Chapitre 11

# Entassements paradigmatiques

L’objet de ce chapitre est de présenter une description de ce que l’*Approche Pronominale* a appelé **entassements paradigmatiques** ([Blanche-Benveniste, 1987]), et d’en proposer une représentation homogène qui en permette l’intégration théorique et formelle à notre grammaire.

Nous avons choisi de suivre ce cadre de description en cherchant à affiner le traitement de la coordination<sup>228</sup> dans la grammaire. Les propositions classiques de la linguistique formelle que nous avons trouvées ne répondaient pas assez clairement à notre besoin de cohérence descriptive : les descriptions de la coordination sont régulièrement confrontées à des problèmes de considérable variabilité des formes, de contraintes de natures très diverses que la syntaxe à elle seule ne saurait résoudre (sauf à faire de la suranalyse), et de cohérence dans la représentation formelle par rapport aux autres constructions syntaxiques plus canoniques. Si bien qu’une hypothèse a progressivement émergé de l’observation des corpus, selon laquelle il serait possible de traiter la coordination non pas comme un type de syntagme classique, mais comme un phénomène proche de celui des disfluences (tel que décrit dans [Guénot, 2005b]), en ce qu’ils ont tous deux la particularité de constituer un ensemble d’unités partageant une unique fonction syntaxique, comme l’affirme notamment [Bilger, 1997].

Nous nous éloignons en cela des assumptions les plus courantes à ce sujet : la plupart des études que l’on peut trouver n’opèrent pas de rapprochement initial entre disfluences et coordination, traitant les secondes comme des syntagmes classiques et les premières comme des particularités de l’oral dont le statut dans l’analyse est fort variable (quand elles sont traitées). Toutefois ce rapprochement a déjà été envisagé, selon deux approches différentes. Tout d’abord, [Levelt, 1983] a proposé de traiter les disfluences sur le modèle des coordinations. Or, ceci revient à considérer que les premières seraient une forme des secondes, ce qui est bien sûr trop réducteur pour fonctionner comme le montrent [De Fornel & Marandin, 1996], ou comme l’avait déjà fait remarquer [Blanche-Benveniste, 1987, p. 140-141] :

*W. [Levelt, 1983] avait déjà observé que les corrections d’erreurs étaient organisées comme des coordinations ; les sujets qu’il observait, dans une expérience de psycho-linguistique, devaient*

---

<sup>228</sup>Pour des raisons de simplicité, nous utilisons le terme *coordination* pour faire référence à la fois à la coordination et à la *juxtaposition*, parfois nommée *coordination asyndétique* [Huddleston & Pullum, 2002].

désigner des points de couleurs et des croisements ; lorsqu'ils corrigent leurs mauvaises réponses, c'est sous forme de listes, par exemple :

- (121) to the right is a green, a blue node  
at the right side of an orange dot, orange dot.

*W. Levelt en tirait la conclusion que les corrections d'erreurs prenaient la la [sic] forme des coordinations. Je crois qu'on peut aller plus loin et estimer que les coordinations ne sont que des effets particuliers du grand procédé de listing [i.e., notre "entassement paradigmatique"], qui couvre aussi bien ces corrections d'erreurs que la recherche du mot approprié, ou des effets rhétoriques assez variés.*

C'est d'ailleurs elle-même qui a proposé la seconde approche, que l'on a suivie. Celle-ci consiste à traiter disfluences et coordinations comme des phénomènes distincts, mais partageant des caractéristiques communes.

L'éloignement des descriptions habituelles ne s'arrête pas là, puisque comme on a pu le dire en introduction de cette partie, toutes les constructions de notre grammaire ne sont pas des syntagmes. C'est ici que l'on trouve l'application de ceci : les entassements paradigmatiques ne sont pas des constructions syntagmatiques. Il s'agit là de la différence la plus notable d'avec les traitements habituels, et qui demande quelques explications que nous présentons au fil de ce chapitre.

Dans ce qui suit, nous commençons par définir et décrire ce que l'on appelle précisément Entassement Paradigmatique (EP, §1), phénomène qui rassemble disfluences et coordination comme l'indique la fig. 81. Nous verrons que la prise en compte d'un tel type de construction a fait émerger un certain nombre de questions générales sur le statut des objets représentés dans la grammaire, que nous développons au §1.1. Après avoir introduit cette construction générale, nous présentons les spécificités des Coordinations dans ce cadre (§2), puis celles des Disfluences (§3). Enfin, nous proposons des informations à ajouter à ces descriptions, provenant d'autres domaines que la syntaxe, et qui permettront de remédier au manque d'indications syntaxiques pour le traitement de ces phénomènes, qui se situent en réalité à l'interface de plusieurs domaines (§4).

**Représentation en grille.** — Avant de commencer ces descriptions, précisons que pour représenter les entassements paradigmatiques nous utilisons la **représentation en grille** des énoncés afin de rendre plus lisible l'opposition entre les deux axes d'organisation (cf. à ce propos [Blanche-Benveniste *et al.*, 1979], [Bilger, 1982], ou [Blanche-Benveniste, 1995]). Celle-ci consiste à représenter non seulement l'enchaînement syntagmatique, linéaire, des énoncés, mais également l'accumulation d'objets sur une même position paradigmatique en exploitant l'axe vertical, comme on peut le lire p.ex. dans [Blanche-Benveniste, 1987, p. 133] :

*(...) sur une ligne graphique continue, nous avons l'habitude de lire la prose en enchaînant des syntagmes linéarisés ; toute succession graphique est supposée avoir une interprétation syntaxique pertinente ; cette procédure ne permet pas de traiter des successions comme un un un ; c'est du reste une des raisons pour lesquelles les transcriptions de langue orale sont insupportables à lire. Comme la succession de un un un n'est pas interprétable sur l'axe horizontal du déroulement des syntagmes, nous proposons de la disposer (à des fins d'étude) sur un axe vertical, ce qui permet de rattraper la linéarité :*

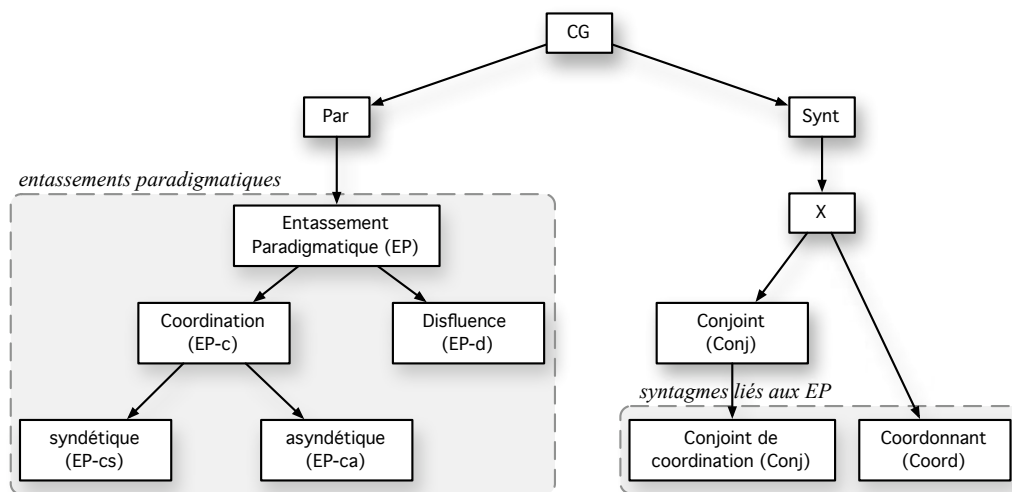


FIG. 81 – Hiérarchie des Entassements Paradigmatiques dans la grammaire.

(122)	<div>le Dôme de Milan est    un</div> <div>un</div> <div>un monument regrettable.</div>
-------	---

On représente donc à la verticale les uns des autres les objets qui occupent la même place syntagmatique. Ceci est valable, même pour les cas où les objets occupant la même place ne sont pas adjacents :

*Il existe un autre type de bribe, liée au phénomène d’“anticipation” : dans le déroulement de l’énoncé, un élément semble “venir trop tôt” ; il est, très souvent, repris à “sa bonne place” dans une séquence suivante. (...)*

*Pour représenter graphiquement ces anticipations, nous prendrons pour base la séquence la plus longue — qui vient souvent en second — et nous situerons le morceau anticipé, verticalement, par rapport à cette séquence :*

(123) a.	<div>on est amené            plusi</div> <div>à se voir            plusieurs fois</div>
b.	<div>un jeune homme me l'avait            dans</div> <div>lancé            dans la figure</div> <div>lancé</div>

*Le morceau anticipé apparaît ainsi séparé de ce qui précède par un blanc ; ce blanc ne représente aucun “trou” dans le déroulement réel de ce qui a été dit ; il représente un “trou” dans l’arrangement syntaxique de la séquence où il figure (...). Ibid., p. 134-135.*

**Constructions présentées.** — Comme on a pu le voir en fig. 81, dans ce chapitre on va étudier deux types de constructions : tout d’abord les **Entassements Paradigmatiques** eux-mêmes, ainsi que les deux constructions qui en héritent, i.e. **Coordination** et **Disfluence** ;

mais on étudiera aussi deux constructions syntagmatiques qui sont des constituants remarquables des EP : les **Coordonnants** d’une part (souvent appelés “conjonctions de coordination” dans la terminologie classique) et les **Conjoints de coordination** d’autre part (qui sont des syntagmes constitués d’un Coordonnant + un objet coordonné). Ces deux derniers, comme on vient de le dire, ne sont pas des EP mais sont des syntagmes constituants des EP ; toutefois on les décrit ici même compte tenu de leur lien évident avec les constructions paradigmatiques.

## 1 Entassement Paradigmatique (EP)

On définit l’Entassement Paradigmatique comme étant une relation entre plusieurs objets qui **partagent une unique fonction syntaxique** au sein de l’énoncé. Cette relation est **paratactique** (selon la définition p.ex. de [Le Goffic, 1993] ou de [Riegel *et al.*, 1999]<sup>229</sup>) au sens où elle n’établit aucun lien dépendancier entre les objets concernés, et à ce titre elle ne constitue pas non plus de syntagme(s) à proprement parler. Les entassements paradigmatiques “*interrompent le déroulement syntagmatique (...), piétinent sur le même emplacement syntaxique*” [Blanche-Benveniste, 1987, p. 137], permettant ainsi de réaliser plusieurs occurrences d’une même place syntagmatique, et par là même de la fonction syntaxique correspondante. Par exemple, en (124)<sup>230</sup> les deux déterminants en italiques (qui font l’objet d’une disfluence) se partagent l’unique fonction spécifieur : il n’y a pas deux fonctions spécifieur dans le même SN, chacune remplie par une occurrence de déterminant, mais c’est l’ensemble des deux déterminants qui occupe l’unique occurrence de la fonction ; de même les deux objets en gras, coordonnés cette fois, se partagent l’unique fonction expansion.

(124)	tu as fait				
		<i>des</i>			
		<i>des</i>	boulots		
				<b>quelque part</b>	
			ou	<b>autres qu’ici</b>	je sais pas

**Phénomènes concernés.** — Deux phénomènes en français entrent dans la définition des EP : les disfluences et les coordinations. En effet, qu’il s’agisse d’une accumulation dont les divers constituants ajoutent à l’énoncé des informations sémantiques (coordinations) ou non (disfluences), dans les deux cas on doit traiter un ensemble d’objets qui n’ont qu’une seule (et unique) fonction syntaxique.

Il est à noter que l’interprétation de la notion d’EP est assez variable, et que concrètement certains linguistes ne considèrent pas que les coordinations en fassent partie. Pourtant la notion d’“*allées et venues sur l’axe syntagmatique*” sur laquelle se base Claire Blanche dans sa

<sup>229</sup>Par contre, il ne s’agit pas exactement de l’acception de “parataxe” que l’on trouve chez [Koch, 1995] puisque l’on considère pour notre part que les entassements paradigmatiques ne recouvrent pas la totalité des phénomènes de parataxe. Lui appelle la coordination “parataxe syndétique” et la juxtaposition “parataxe asyndétique” ; nous considérons que la parataxe rassemble également les disfluences, ainsi que des phénomènes linguistiques qui n’ont pas de conséquences sur l’analyse syntaxique.

<sup>230</sup>Dans la suite de ce chapitre, sauf mention contraire les exemples sont tirés du *Corpus d’Interactions Dialogiques*, [Bertrand & Priego-Valverde, 2005].

définition des EP est une notion purement syntaxique. Sur ce point nous rejoignons [Kracht, 2005b] qui considère que l'une des caractéristiques qui permettent d'opposer informations syntaxiques et sémantiques, est le fait que la structure syntaxique se déroule sur un plan *linéaire* contrairement à la structure sémantique. Ce qui signifie que cette dernière n'est pas sujette à l'opposition syntagmatique / paradigmatique dans l'énoncé <sup>231</sup>.

Notre choix trouve une autre justification dans l'observation des corpus : il est bien souvent malaisé de trancher sur la nature exacte d'un EP, entre disfluece ou coordination, ou les deux mêlées, ce qui montre bien à quel point les deux phénomènes sont apparentés (125).

(125)	il y a	des conflits
	soit	des conflits d'intérêt
	soit	des gens qui savent pas que tu es là

Dans (125) on voit que *soit des conflits d'intérêt* et *soit des gens qui savent pas que tu es là* sont coordonnés, mais on ne dispose pas de suffisamment d'information pour décider clairement si ces deux éléments sont liés à *des conflits* en une disfluece, ni si la portée de cette disfluece regroupe uniquement *des conflits* et *soit des conflits d'intérêt*, ou alors *des conflits* avec la coordination en entier.

Précisons ici que la description que l'on propose n'a pas vocation à permettre une telle désambiguïsation, mais au contraire à permettre que ces constructions soient reconnues comme faisant partie d'un ensemble, même s'il reste imprécis, de façon à exploiter le mieux possible les informations que l'on peut tirer de l'énoncé. En l'occurrence, notre proposition permet *a minima* d'introduire des constructions génériques EP pour (125) faute de pouvoir préciser de quel type il s'agit, au lieu de proposer plusieurs solutions en analyse non-déterministe (disfluece, ou coordination, ou les deux), ou pire de devoir choisir l'une ou l'autre en se basant sur des critères externes en analyse déterministe<sup>232</sup>.

## 1.1 Intégration dans la grammaire

Le fait de vouloir intégrer un phénomène tel que celui-ci dans une grammaire met en avant une différence fondamentale entre les descriptions linguistiques et leur formalisation : là où les premières s'attachent à décrire finement le fonctionnement interne et les propriétés d'un phénomène donné, le modèle formel qui a pour tâche de le représenter doit également intégrer, outre sa description propre et indépendante, les propriétés plus générales du phénomène, *i.e.* conserver un point de vue plus global sur l'interaction entre celui-ci et les autres, sur la façon dont le tout s'articule et conserve une cohérence générale.

Ce que l'on a pu voir jusqu'ici met en avant le fait que pour traiter des EP il est nécessaire de commencer par répondre à une question d'ordre plus général : *Quelle est leur place dans la grammaire ?* Et pour répondre à cette question, la première chose que l'on doit se demander est

<sup>231</sup> Ceci n'exclut pas qu'il puisse y avoir des relations hypotactiques et paratactiques en sémantique, mais qui ne s'opposent pas sur les deux axes syntaxiques.

<sup>232</sup> Notons que les deux cas de figure nous intéressent, puisque la grammaire que nous développons a vocation à être implantée dans des analyseurs des deux types.

ce que l'on représente au juste dans une grammaire : *Quels sont les objets que l'on y manipule ?* Y représente-t-on des relations entre les **occurrences possibles** d'un énoncé, ou alors entre les **places syntagmatiques** occupées par ces occurrences ? Dans le cas d'énoncés sans EP ces questions ne sont pas si évidentes ; ainsi, dans un énoncé tel que (126), les "occurrences possibles" et les "places syntagmatiques" sont confondues, puisque chaque place n'est occupée que par une unique occurrence.

(126) quelque part je vais être impardonnable

Ainsi, comme l'illustre la fig. 82 <sup>233</sup>, les relations (figurées *a* et *b* pour les besoins de l'exemple) peuvent porter indifféremment sur les objets en tant qu'occurrences possibles ou en tant que places syntaxiques.

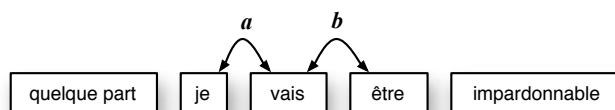


FIG. 82 – Relations dans un énoncé sans EP.

Par contre, dès lors que l'on traite des énoncés avec des entassements paradigmatiques (127), ces questions prennent toute leur importance.

(127) 

quelque part	je vais
	je vais être impardonnable

Considérons d'abord que l'on représente dans une grammaire des **relations entre des occurrences possibles**. L'analyse d'une disfluece peut alors être représentée selon l'illustration de la figure 83.

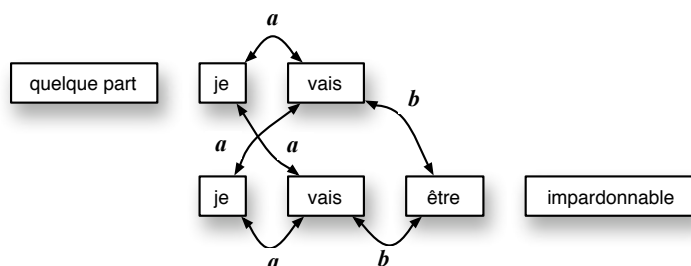


FIG. 83 – Relations entre occurrences.

<sup>233</sup>Les "relations" que l'on représente ici sont des représentations intuitives, et n'illustrent pas une théorie donnée.

On voit que dans ce cas on multiplie le nombre de chaque relation par le nombre d'occurrences de la même place syntagmatique : la relation *a* figure quatre fois et *b* deux fois, au lieu d'une seule dans la figure 82. La conséquence de cela est que l'ensemble de caractéristiques, spécifique au syntagme qui contient la disfluen­ce (ici, *je vais je vais être impardonnable*), varie non seulement en fonction de la présence, mais aussi de la forme d'une disfluen­ce : l'ensemble  $\{a, b\}$  de la figure 82 devient  $\{a, a, a, a, b, b\}$  pour la disfluen­ce précise de la figure 83, comme elle serait devenue  $\{a, a, b\}$  pour *je je vais être impardonnable*, ou bien  $\{a, a, a, a, a, a, b, b\}$  pour *je je vais je vais être impardonnable*, etc. De plus, un certain nombre de propriétés définitoires du syntagme (e.g. l'unicité du pronom clitique nominatif, ou l'ordre linéaire entre ce même pronom et le verbe) sont faussées par la présence de la répétition, et la définition dans la grammaire doit tenir compte de ces variations de caractéristiques. Pourtant les caractéristiques spécifiques à la présence d'un entassement paradigmatique n'ont pas d'incidence sur l'analyse syntaxique d'un énoncé, et donc ne devraient pas avoir d'incidence sur la définition (syntaxique) d'un syntagme.

Représenter des relations entre occurrences ne semble donc pas être le fait d'une grammaire ; considérons alors que l'on y représente des **relations entre des places syntagmatiques**. Dans ce cas on peut illustrer le traitement des disfluen­ces comme dans la figure 84. Ici, chaque occurrence de la même place syntagmatique est d'abord analysée (on constitue deux objets de la forme *je vais* en vertu de la relation *a*), puis ces deux objets sont rassemblés en un objet "disfluent" qui à son tour est relié à *être* suivant la relation *b*.

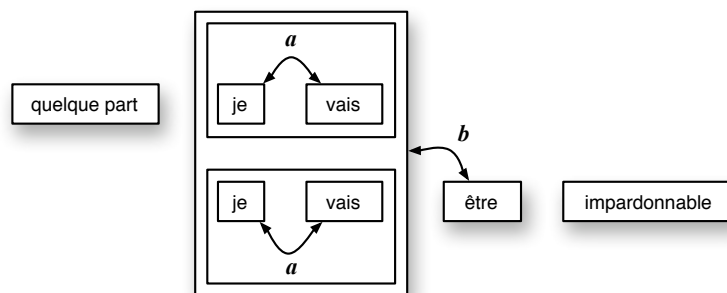


FIG. 84 – Relations entre places syntagmatiques.

Ici l'on résout le problème de la multiplication injustifiée des caractéristiques, mais on se trouve face à un autre problème : pour pouvoir faire une analyse syntaxique il faut qu'à chaque place corresponde une catégorie (linguistiquement viable), et dans l'exemple quelle est la catégorie de *je vais je vais*, ou de *c'était je crois qu'il était* dans l'ex. (128) ?

(128) **c'était**  
**je crois qu'il était** autrichien ou un truc comme ça

En effet, s'il est simple de traiter de cette façon les reprises simples telle que *des des* dans l'exemple (124) en lui affectant la place de "déterminant", il devient plus difficile de s'accorder sur le statut d'un groupe constitué d'un fragment de début de syntagme, qui ne correspond

à aucune étiquette syntaxique. Comment intégrer dans une analyse syntaxique, des objets qui ne sont pas des unités syntaxiques ? Il faudrait pour cela les ajouter artificiellement à la grammaire, intégrer ces groupes qui ne sont pas vraiment des syntagmes, mais dont la seule raison d’y figurer est qu’ils peuvent apparaître en tant qu’occurrence. Au-delà du problème linguistique de fond que ce type d’artefact suppose (quelle analyse fait-on ? quelle est la nature des objets que l’on introduit ?), on en revient également au problème posé par la technique précédente : les disfluences ne sont pas décrites en tant que phénomène, mais chacune des possibilités de disfluence devra être l’objet d’une construction particulière, et l’on sera limité à un moment ou à un autre par l’itération limitative des possibilités *a priori* infinies.

Une façon de remédier à ce problème sans fabriquer de catégories *ad hoc* est de ne pas rassembler les différentes occurrences d’une disfluence en un groupe unique, mais de considérer que chaque reprise est une occurrence, achevée ou non, du syntagme complet (figure 85).

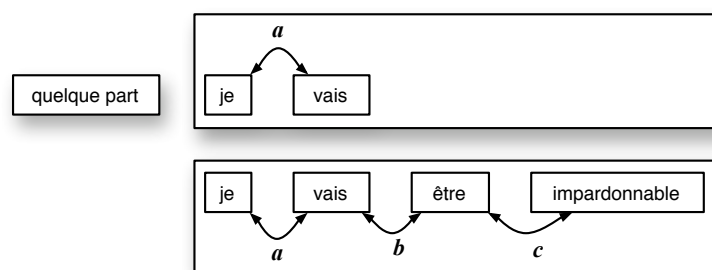


FIG. 85 – Les entassements paradigmatiques comme occurrences complètes de syntagmes.

Pour ne pas recourir à des catégories vides <sup>234</sup>, on peut tout simplement considérer (en référence au fonctionnement de GP) que la caractérisation des occurrences inachevées de syntagme est le reflet de leur constitution : un certain nombre de propriétés sont, à juste titre, non évaluées, et d’autres sont évaluées et non-satisfaites, en comparaison avec l’occurrence complète. Il s’agit ensuite, pour ne pas se contenter de déplacer au niveau supérieur le problème posé par la première possibilité envisagée ici, de mettre en relation ces occurrences du même syntagme au sein de la grammaire, en tant que “phénomène d’entassement paradigmatique”. Il est possible de le repérer par un ensemble de caractéristiques dont, en syntaxe, l’occupation de la même fonction.

Ce qui pose un problème plus délicat dans cette dernière représentation, c’est l’expression de relations entre parties différentes d’occurrences différentes. Observons par exemple l’énoncé (129) :

(129)	lesquels registres	sont très	euh
		sont	
	doivent	être	
	doivent pouvoir	être	contrôlés

<sup>234</sup>Puisque nous développons une grammaire qui n’a pas de recours au postulat de catégories vides, que ce soit pour ce cas ou pour tous les autres.



tiré de [Blanche-Benveniste *et al.*, 1990], p. 24, où l’auteur dit que l’on

*[...] vise à dégager la séquence maximale qui a été donnée par le locuteur, en tenant compte de toutes les bribes qu’il a fournies; dans l’exemple précédent, on retiendra comme séquence maximale : lesquels registres doivent pouvoir être très contrôlés, bien qu’il n’ait jamais prononcé cet énoncé (il a dit très la première fois, mais sans le reprendre avec doivent pouvoir)*<sup>235</sup>.

ce qui implique de pouvoir établir des relations entre *très* apparu dans la première partie, et *contrôlés* apparu dans la dernière. On a le même type de configuration dans (130),

(130)	les Anglais	qui ont quand-même beaucoup d’humour	euh
	les journaux anglais		enfin
	les médias britanniques		ont fini par me répondre

où la séquence maximale est *les médias britanniques qui ont quand-même beaucoup d’humour ont fini par me répondre*. Or, bien qu’une analyse en GP repérera cette relation, il paraît par contre bien difficile de la représenter au sein même de la grammaire et donc d’en tenir compte dans la description de la “séquence maximale”. Difficile, parce que cela demanderait d’utiliser une relation entre constituants de syntagmes différents, lorsque ces syntagmes sont liés entre eux par une relation d’entassement paradigmatique, et que l’on n’utilise pas de telles relations dans d’autres cas. On devrait donc créer artificiellement, non pas de nouveaux syntagmes comme précédemment, mais de nouvelles propriétés uniquement pour traiter le phénomène.

La représentation que l’on propose finalement conserve la vision des entassements paradigmatiques en tant que construction à part entière, telle que montrée ci-dessus, mais ne les considère pas systématiquement comme mettant en relation des syntagmes : on reliera en tant qu’EP chaque occurrence répétée d’une même place syntagmatique, qu’elle soit construite ou non. En l’occurrence, dans notre exemple (127) *je vais* (ou *je vais je vais*) ne constitue pas une place syntagmatique, mais par contre *je* (ou *je je*) d’une part et *vais* (ou *vais vais*) d’autre part, occupent bel et bien des places syntagmatiques : celles de sujet et de verbe. On aura alors une représentation telle qu’illustrée en fig. 86.

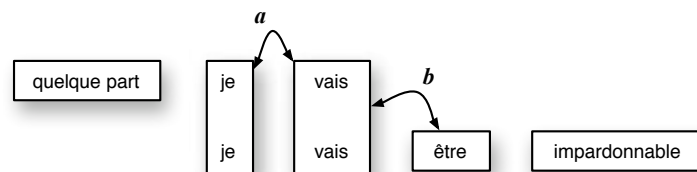


FIG. 86 – Les EP comme accumulation de places syntagmatiques, 1.

Cela ne signifie pas que l’on “entasse” systématiquement les objets les plus petits possible ; cela signifie que l’on entasse les constructions linguistiquement viables, les plus grandes possibles.

<sup>235</sup>Dans [Blanche-Benveniste & Jeanjean, 1986, p. 102-015, 159-162], il a été signalé que la prise en compte de cette séquence maximale et non uniquement de la séquence finale n’est pas un simple *challenge* de linguiste, mais se justifie par des études psycholinguistiques.

En l'occurrence, *je vais* (ou *je vais je vais*) ne constitue pas un syntagme dans (127), et c'est la raison pour laquelle on entasse indépendamment les *je* d'une part, et les *vais* d'autre part ; par contre dans le cas d'une énumération de plusieurs objets constituant bel et bien des syntagmes (i.e., occupant des places syntagmatiques construites) (131) <sup>236</sup>, on aura une représentation telle que celle de la figure 87.

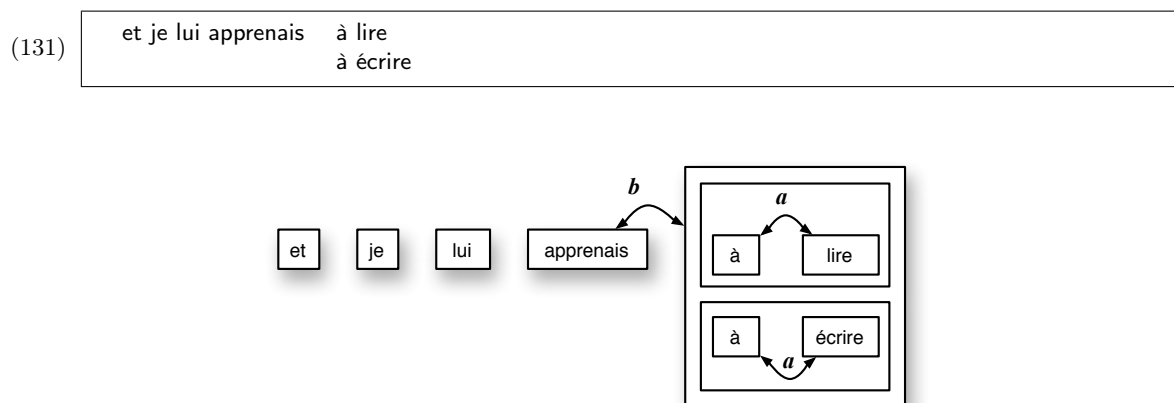


FIG. 87 – Les EP comme accumulation de places syntagmatiques, 2.

A la suite de ce traitement, chaque EP (e.g. *je je* dans (127), ou *à lire à écrire* dans (131)) peut être considéré exactement comme une simple construction lexicale (e.g. *je* dans (126)) ou syntagmatique (e.g., *à lire*, ou bien *à écrire*), bénéficiant du même traitement grammatical que celle-ci ; la seule différence tient à sa forme (entassée), mais les relations que chacune entretient avec le reste de l'énoncé (et notamment avec la “reprise”, i.e. le *repair* de [Shriberg, 1994]) ne demande aucun traitement particulier et est traité dans la grammaire exactement au même titre qu'une construction non entassée.

De plus, toutes les places syntaxiques qui doivent être mises en relation lors de l'introduction du syntagme qui les contient (le “syntagme maximal” de Blanche) sont ainsi séparées les unes des autres et peuvent être prises en compte par le reste de la grammaire de la même façon, qu'elles soient entassées ou non, sans introduire ni formes de constructions ni relations *ad hoc* (fig. 88).

Comme on le voit en comparant avec la figure 89, on a le même ensemble de relations que si l'on avait eu la séquence maximale ; c'est uniquement la forme des objets qui change.

## 1.2 Caractéristiques des EP

Dans la fig. 90 qui représente l'Entassement Paradigmatique dans la grammaire, on remarque qu'il n'est nullement fait état explicitement de l'unicité fonctionnelle. Ceci est normal, puisque selon ce que l'on a expliqué au cours des chapitres précédents, les fonctions des objets leurs sont

<sup>236</sup>[Blanche-Benveniste *et al.*, 1990].

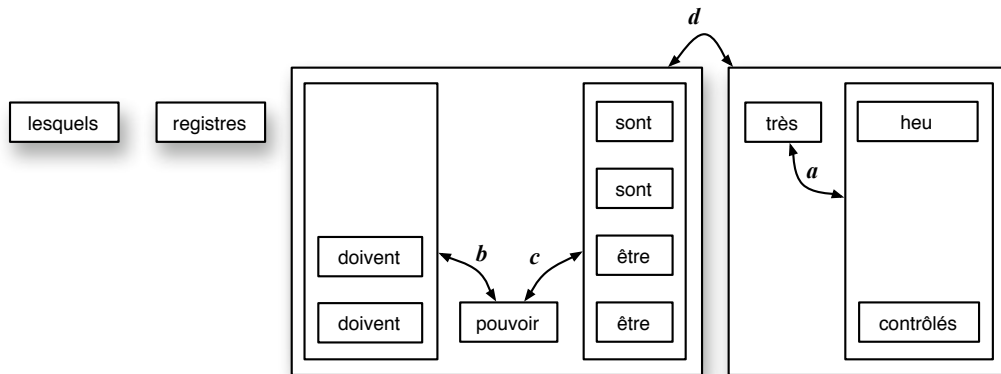


FIG. 88 – Les EP comme accumulation de places syntagmatiques, 3.

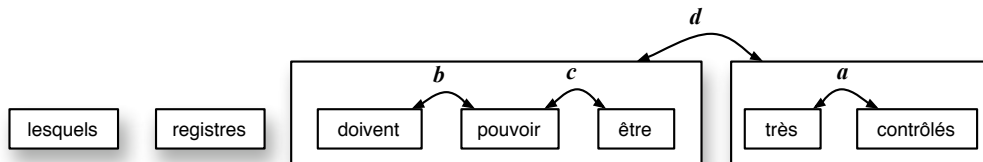


FIG. 89 – Relations dans la séquence maximale dans (129).

affectés en tant que constituants d’une construction de niveau supérieur. Ici, les constituants de l’EP n’ont pas de fonction les uns relativement aux autres ; par contre, c’est l’entassement paradigmatique ainsi constitué qui occupera une (et une seule) fonction donnée au sein de la construction dont il sera constituant, et c’est ainsi que l’unicité fonctionnelle est garantie. Dans le cas contraire, i.e. si l’EP complet ne peut pas recevoir de fonction (unique) au sein d’une construction dont il est constituant, alors cela signifie que cet EP introduit n’est pas pertinent, et il ne sera donc pas réutilisé dans la suite de l’analyse.

**Catégorie de l’EP.** — La construction EP hérite de la construction Paradigme (Par). Elle est de type EP, et porte comme catégorie l’ensemble de celles de ses constituants. Nous nous sommes sur ce point inspirée de [Huddleston & Pullum, 2002] à propos des coordinations, qui expliquent que ce qui permet de rassembler plusieurs éléments comme étant coordonnés (ou plus généralement, pour ce qui nous intéresse, entassés), ce n’est pas leur similitude catégorielle mais leur identité fonctionnelle. En d’autres termes, on peut très bien entasser des constructions qui ne sont pas de la même nature (132),

(132)

il a pas eu l’occasion de voir	ce qu’ils euh
	qui ils sont
	comment ils fonctionnent
	tout ça
	(...)

EP	
HERIT Par	
INTR. $\left[ \text{ID} \mid \text{NATURE} \left[ \text{TYPE} \quad EP \right] \right]$	
Exclusion	<p> <math>\text{Sub} \neq \neg\text{Sub}</math>  <math>\text{Det} \neq \neg\text{Det}</math>  <math>\text{Prep} \neq \neg\text{Prep}</math>  <math>\text{V} \neq \neg\text{V}</math>  <math>\text{Sup} \neq \neg\text{Sup}</math>  <math>[ \text{TYPE } \textit{relationnel} ] \neq [ \text{TYPE } \textit{qualificatif} ]</math>  <math>\text{SV-i} \not\in \{ \text{Adj} ; \text{Adv} \}</math>  <math>\{ \text{SPrep} ; \text{Adj} ; \text{Adv} \} \not\in \{ \text{Conj-sr} ; \text{SV} ; \text{Pro} ; \text{N} \}</math>  <math>\{ \text{Conj-sc} \} \not\in \{ \text{SV} ; \text{Pro} \}</math>  <math>\text{SV-tens} \neq \neg\text{SV-tens}</math>  <math>\text{P-SV} \not\in \neg\text{P-SV}</math>  <math>\text{N} \not\in \{ \text{SN-nd} ; \text{Conj-s} ; \text{SV} \}</math> </p>

FIG. 90 – L’Entassement Paradigmatique (EP).

si bien que dans notre figure 90, les objets Majeurs (hérités de Par) peuvent être des objets de catégories différentes. Pour cette raison les auteurs n’affectent pas de catégorie syntaxique “classique” (simple) aux coordinations, mais les appellent *NP-coordination* (et non *NP* tout court) pour plusieurs NP coordonnés, ou bien *NP-PP-coordination* pour une coordination entre un NP et un PP. Nous nous éloignons de leur proposition uniquement sur ce dernier point, puisque nous ne souhaitons pas leur affecter des catégories *ad hoc*, nos EP ne portent pas une unique catégorie complexe (mentionnant leur nature d’entassement), mais une liste de catégories syntaxiques simples.

Cette liste joint les catégories par des “et” logiques, si bien que dans la suite de l’analyse, l’EP ainsi formé devra satisfaire les propriétés de l’ensemble de ses catégories (et non pas celles d’une seule uniquement).

**Constituance et cooccurrence.** — Puisque comme on l’a dit plus haut les EP ne constituent pas des syntagmes, la construction EP n’hérite pas de la construction Syntagme. D’ailleurs on notera que cette construction ne contient pas d’attribut TÊTE, puisqu’entre les différents constituants d’un EP il n’y a pas de relation de dépendance.

Comme on peut le voir, il ne figure dans cette construction ni contrainte d’ordre ni d’adjacence ; deux choses uniquement sont imposées pour l’introduction d’un EP, d’une part la présence de plusieurs constituants (l’EP ne pouvant pas par définition être une construction unaire — d’où la présence de l’Exigence dans la construction Par), et d’autre part la possibilité d’une unicité fonctionnelle.

En effet, bien que les objets accumulés puissent être de natures différentes (raison pour laquelle on n’en contraint pas la nature dans la propriété Majeur), tous les objets ne peuvent porter toutes les fonctions (p.ex., un Coordonnant ne peut pas être sujet). Donc, sans spécifier que

les constituants des EP doivent être de même nature, ce qui serait une vision sévèrement réductrice, il est possible de lister des groupes de constituants qui peuvent être “entassés” les uns avec les autres (et inversement, de constituants qui ne le peuvent pas). Cette information ne figurant pas explicitement dans les grammaires sur lesquelles nous nous sommes basée <sup>237</sup>, nous avons établi des listes de catégories pouvant occuper des fonctions identiques en nous basant sur le contenu actuel de notre grammaire <sup>238</sup>, listes qui sont utilisées dans la formalisation de l’EP présentée en fig. 90 en un ensemble de propriétés d’Exclusion.

**Imprécision et limitation de la surgénération.** — Il est bien évident que la construction proposée en fig.90 est remarquablement imprécise, ce qui signifie que les informations qu’elle contient sont *nécessaires* à la reconnaissance des EP, mais pas *suffisantes* pour ne pas surgénérer inutilement un nombre conséquent d’EP là où ils n’ont pas lieu d’être <sup>239</sup>.

De ceci nous pouvons dire deux choses. Premièrement, comme on l’a dit plus haut cette description ne contient que des informations syntaxiques. Pour fournir une description des EP qui soit réellement précise et pertinente, il est indispensable d’ajouter des informations provenant d’autres domaines, tels que la sémantique et la prosodie (ou des informations sur la mise en page et la ponctuation pour l’analyse de l’écrit). On revient sur ce point au §4.

Deuxièmement, il est important de rappeler ici que notre grammaire n’est pas un mécanisme génératif mais une **ressource descriptive**, et c’est là que l’on peut en observer une conséquence fondamentale : notre grammaire n’a pas pour but de limiter les introductions de constructions superflues, mais de permettre d’introduire le plus de constructions pertinentes possible en fonction des informations disponibles ; la gestion de la surgénération n’est donc pas un problème adressé à la grammaire. Cependant deux limitations à la surgénération sont à considérer :

- La première est une limitation *de facto* : toute entrée à analyser, quel que soit son degré d’écart à la norme, pour peu qu’elle soit une production attestée, offrira suffisamment de concordances avec les descriptions de la grammaire pour que celle-ci limite ses introductions superflues par le simple fait de la structuration de l’énoncé. Ce point a été évoqué dans le rapport entre complexité théorique et réelle des analyseurs GP dans [Vanrullen, 2005].
- La seconde est liée à la distinction entre la **grammaire** et les **analyseurs** qui l’utilisent. Bien que notre grammaire soit conçue comme une ressource homogène et ne nécessitant pas d’adaptations en fonction des phénomènes décrits, il en est autrement pour les parseurs qui eux, suivant leurs objectifs, peuvent recourir à un certain nombre d’heuristiques visant à modérer plus ou moins fortement l’introduction de constructions superflues. Par exemple, un analyseur déterministe fera intervenir un nombre de mécanismes de modération plus

<sup>237</sup>Notamment [Riegel *et al.*, 1999], [Le Goffic, 1993], [Chevalier *et al.*, 1997], [Wilmet, 1998]. On trouve ce type d’information dans [Huddleston & Pullum, 2002], mais pour l’anglais. Il serait sans doute intéressant d’observer les classifications du lexique *Dicovalence* (ex-*Proton*, cf. p.ex. [van den Eynde & Mertens, 2003]), basées sur l’Approche Pronominale, pour établir nos listes d’exclusions en nous basant sur les pronominalisations possibles des constructions.

<sup>238</sup>Cette grammaire étant actuellement incomplète, les listes le sont probablement également. Toutefois elles seront progressivement affinées et complétées à mesure du développement de la grammaire.

<sup>239</sup>Concrètement, telle qu’elle est présentée ici elle permet d’introduire des EP constitués de toutes les combinaisons possibles d’objets linguistiques, à tous les niveaux.

important qu'un parseur non-déterministe, puisque ce dernier n'aura pas besoin de fournir un résultat unique. De plus, ceci permet aux analyseurs de choisir de faire intervenir de tels mécanismes à différentes étapes et selon des choix qui leur sont propres (cf. par exemple [Balfourier *et al.*, 2005] concernant les divers analyseurs GP).

Dans la perspective d'une limitation automatisée des introductions d'EP, on peut noter une chose dans les énoncés contenant des EP : l'énoncé les contenant ne peut pas être analysé syntaxiquement sans une description qui leur confère une place (syntagmatique) unique : on ne pourrait pas construire d'analyse syntaxique de l'énoncé de l'exemple (132) sans avoir, sous une forme ou une autre, de description de la coordination (au moins) ; on ne pourrait pas les considérer comme autant de constructions différentes, chacune ayant une place distincte dans l'énoncé. En d'autres termes, la construction *ce qu'ils euh qui ils sont comment ils fonctionnent tout ça* ne pourrait en aucun cas être analysée comme une succession classique de constituants d'un énoncé (*qui ils sont comment ils fonctionnent* par exemple ne peut pas constituer une catégorie syntagmatique simple, leur relation est purement paratactique et ne peut être traitée comme hypotactique, et la relation qui unit *voir* et *qui ils sont* d'une part, et *voir* et *comment ils fonctionnent* d'autre part, est censée être unique selon la valence du verbe *voir*). Cette simple constatation peut éventuellement permettre une introduction "raisonnable" des EP dans les analyses d'énoncés : il ne pourrait y avoir un EP que quand une analyse sans EP est impossible.

## 2 Coordination (EP-c)

Comme on l'a vu dans la figure 81, deux constructions héritent de l'EP que l'on vient de décrire : une construction décrivant les spécificités des disfluences (EP-d, §3), et l'autre celles des coordinations (EP-c, *i.e.* éléments de description différents des disfluences, et communs aux coordinations et juxtapositions). Nous nous intéressons ici à cette dernière (fig. 91) et aux constructions qui en héritent.

**Terminologie.** — La terminologie employée au sujet des coordinations étant très variable en linguistique, avant de présenter la description du phénomène nous jugeons utile de préciser les termes que nous utiliserons ici avec leur signification :

- Nous appelons **Coordination** l'ensemble des phénomènes d'accumulation paradigmatique dont chacune des occurrences ajoute des informations sémantiques à l'énoncé <sup>240</sup>. Ce phénomène rassemble aussi bien les coordinations qui ne sont pas lexicalement marquées, souvent appelées juxtapositions et que nous appelons ici **Coordinations asyndétiques**, que les coordinations marquées, souvent appelées coordinations (tout court) et nous appelons ici **Coordinations syndétiques**.
- Nous appelons **Coordonnant** l'ensemble des objets qui sont des marqueurs (le plus souvent

---

<sup>240</sup>Il ne s'agit donc pas là d'une définition strictement syntaxique. Toutefois nous pensons que la coordination est un phénomène linguistique qui se situe typiquement à l'interface entre plusieurs domaines, et qu'il serait impossible de la définir uniquement à l'aide de critères syntaxiques ; et les difficultés que l'on peut voir dans les grammaires qui essaient de le faire semblent corroborer notre idée.

lexicaux) de la coordination. Ceci rassemble à la fois les traditionnelles “conjonctions de coordinations” et les adverbes conjonctifs tels que *sans*, *alias*, *autrement dit*, *c’est-à-dire*, *à savoir*, *plus*, *moins*, *sauf*, *excepté*, *outre*, *au lieu de*, *partant*, *même*, *jusqu’à*, *par-dessus le marché*, etc. (133) <sup>241</sup>.

(133)	a.	<table><tr><td></td><td>Sidonie</td><td></td></tr><tr><td>avec</td><td>son père</td><td>formeront une équipe imbattable (...)</td></tr></table>				Sidonie		avec	son père	formeront une équipe imbattable (...)						
		Sidonie														
avec	son père	formeront une équipe imbattable (...)														
	b.	<table><tr><td></td><td>Michel</td><td></td></tr><tr><td>alias</td><td>Gouretzky</td><td></td></tr><tr><td>alias</td><td>porte parole de Brian Green</td><td></td></tr><tr><td>alias</td><td>copié-collé</td><td>est en fait ton père.</td></tr></table>				Michel		alias	Gouretzky		alias	porte parole de Brian Green		alias	copié-collé	est en fait ton père.
	Michel															
alias	Gouretzky															
alias	porte parole de Brian Green															
alias	copié-collé	est en fait ton père.														

- Nous appelons **Coordonné** chacun des objets coordonnés, sans leurs éventuels marqueurs, e.g. dans (133b), *Michel*, puis *Gouretzky*, puis *porte-parole de Brian Green*, etc.
- Nous appelons **Conjoint de coordination** le syntagme constitué d’un Coordonnant et d’un Coordonné, e.g. dans (133b), *alias Gouretzky*, puis *alias porte-parole de Brian Green*, etc., mais pas *Michel* qui ne porte pas de coordonnant.

Le Coordonnant (*alias*) peut être une construction lexicale ou un syntagme ; le Coordonné (*copié-collé*) et le Conjoint de coordination (*alias copié-collé*) sont des Syntagmes ; la Coordination (*Michel alias Gouretzky (...)* *alias copié-collé*) est un Paradigme.

## 2.1 Coordination (EP-c)

La construction générale décrivant la Coordination (fig. 91) hérite de l’Entassement Paradigmatique et y ajoute les contraintes propres aux coordinations à la fois syndétiques et asyndétiques.

EP-c	
HERIT EP	
INTR.	$\left[ \begin{array}{c} \text{ID} \mid \text{NATURE} \\ \left[ \begin{array}{cc} \text{TRAITS} & \text{[1].TRAITS} \vee \text{[2].TRAITS} \vee \dots \vee \text{[n].TRAITS} \\ \text{ACCORD} & \text{[1].ACCORD} \vee \text{[2].ACCORD} \vee \dots \vee \text{[n].ACCORD} \end{array} \right] \end{array} \right]$
Unicité	Pro-pc
Exclusion	$\frac{[\text{CAS } x] \not\equiv [\text{CAS } y]}{P \not\equiv \neg(\text{SN} \vee P)}$

FIG. 91 – La Coordination (EP-c).

D’une manière générale, les objets portant un CAS ne peuvent être coordonnés <sup>242</sup> si cet attribut porte des valeurs différentes.

<sup>241</sup>Liste d’adverbes et premier exemple empruntés à [Wilmet, 1998]. Le second vient de fr.sci.philo, 4 juin 2003.

<sup>242</sup>Mais ils peuvent faire l’objet d’une disfluece ; c’est la raison pour laquelle on mentionne cette contrainte ici et non dans l’EP.

Les Pronoms personnels clitiques (Pro-pc) ne peuvent pas être coordonnés entre eux (134) : on les mentionne donc dans l'Unicité.

- (134) a. **tu** la vois et *moi* aussi  
b. \* **il** et **elle** viendront

Il existe une restriction sur ce qui peut être coordonné avec une proposition. Cela se borne à ce que [Wilmet, 1998] appelle **prédication incomplète** (*sentential coordination* chez [Huddleston & Pullum, 2002]). Dans de tels cas, seuls des SN peuvent être entassés avec des P (135).

- (135) a. Encore une remarque comme cela et je m'en vais  
b. Un pas de plus et tu es mort

Enfin, le fait que (136) soit interdit, alors que (137) est possible est géré par les contraintes s'appliquant sur l'utilisation des pronoms clitiques en tant que constituants de constructions verbales : ils doivent être adjacents à un verbe. Dans ce cas, il n'est donc pas nécessaire d'ajouter ici des contraintes d'ordre entre les constituants pronominaux clitiques et non clitiques dans la coordination.

- (136) a. \* **tu** et *moi* la voyons  
b. \* **tu** la vois et **je** aussi

- (137) a. *toi* et *moi* la voyons  
b. **tu** la vois et **je** la vois aussi

Soulignons que dans l'analyse de (137b) ce ne sont pas les deux clitiques *tu* et *et je* qui sont coordonnés (donc, cela ne va pas à l'encontre de l'Unicité des Pro-pc), mais ce sont les SV *tu la vois* et *et je la vois* qui constituent un EP-c.

## 2.2 Coordination asyndétique (EP-ca)

La Coordination asyndétique (EP-ca, fig. 92) hérite de la Coordination que l'on vient de voir, et se distingue de la syndétique par une absence de coordonnant. Puisqu'on a posé en début de chapitre que les constructions contenant des Coordonnants étaient les Conjointes de coordination, cela signifie qu'elle est constituée des mêmes objets que l'EP-c, à l'exclusion des Conjointes.

## 2.3 Coordination syndétique (EP-cs)

Avant de décrire l'autre construction qui hérite de l'EP-c, c'est-à-dire la Coordination syndétique (§2.3.3), il est nécessaire de commencer par présenter deux objets qui en sont caractéristiques : les Coordonnants construits (§2.3.1) tout d'abord, puis les Conjointes de coordination (§2.3.2).



EP-ca	
HERIT EP-c	
Majeur	①→Conj-c
	②→Conj-c
	...
	②→Conj-c

FIG. 92 – La Coordination asyndétique (EP-ca).

### 2.3.1 Coordonnant construit (Coord)

La plupart des coordonnants sont des objets non syntaxiquement construits (138a), et dans la majorité des cas ils ne peuvent pas se combiner avec d'autres objets. Ils peuvent aussi prendre une forme syntagmatiquement construite (138b), et c'est à cela que l'on s'intéresse ici.

- (138) a. *et, car, donc, pourtant, soit, etc.*  
 b. *ainsi que, c'est pourquoi, de plus, en revanche, c'est-à-dire, et (...) aussi, etc.*

Notons qu'ils peuvent aussi être constitués d'un objet qui change de paradigme, comme *que* dans (139) :

- (139) a. *il me le demanderait à genoux **que** je ne céderais pas* <sup>243</sup>  
 b. *il a été habiter à côté de chez Rosalie **que** Rosalie elle savait pas*

Nous n'en tenons pas encore compte dans la partie de grammaire décrite ici, parce que ceci nécessiterait une étude approfondie de tous les objets pouvant occuper ce paradigme, et de leurs caractéristiques d'alors.

Certains coordonnants peuvent se combiner entre eux (140) pour former des coordonnants complexes.

- (140) a. **Ainsi donc** vous ne perdez jamais pied ? <sup>244</sup>  
 b. Je veux installer plusieurs programmes, install ou setup, jusqu'à 99% **et puis** plus rien <sup>245</sup>  
 c. Ça pourrait être les pro-coalition contre les pro-parti. **Ou bien donc** les pro-épilation contre les pro-pois. <sup>246</sup>  
 d. Vous allez me dire : c'est le but du jeu. Certes. **Mais ensuite**, on a un style, on fait des choix tactiques et on cherche à les imposer. <sup>247</sup>

D'après [Wilmet, 1998, p. 583] c'est uniquement le cas de ceux qu'il appelle **coordonnants consécutifs** dans sa classification, c'est-à-dire *donc, puis, ainsi, ensuite, alors,...*, qui peuvent se combiner soit entre eux (140a), soit avec les membres d'autres classes (140b, 140c, 140d). L'auteur oppose cette classe (sémantique) de coordonnants à celles des "copulatifs" (*et, ni,*

<sup>243</sup>Cet exemple et le suivant : [Deulofeu, 1999].

<sup>244</sup>fr.sci.psychologie, 30 avril 2004.

<sup>245</sup>microsoft.public.win2000.beta.general, 16 janvier 2000.

<sup>246</sup>qc.politique, 20 septembre 2001.

<sup>247</sup>fr.rec.sport.rugby, 3 novembre 2003.

*aussi, avec, outre, plus ... plus, moins ... moins*), aux “disjonctifs” (*ou, voire, soit ... soit*), aux “adversatifs” (*mais, cependant, néanmoins, toutefois, au contraire, en revanche, par contre*), et aux “causaux” (*car, en effet*), qui eux ne peuvent pas se combiner entre eux (ni avec d’autres classes sauf les copulatifs).

Bien que cette classification soit d’ordre sémantique, puisqu’elle a une conséquence syntaxique importante nous l’intégrons dès à présent à la grammaire (comme nous l’avions fait pour la valence précédemment). Nous attribuons à nos objets Coordonnants un trait de SCAT qui pourra recevoir les valeurs *consécutif, copulatif, disjonctif, adversatif* ou *causal*. Ceci nous permet d’exprimer des contraintes sur l’introduction de Coordonnants construits en fonction de ce trait.

Nous élargissons la contrainte énoncée par Wilmet en posant que *et* et *ou*, dont le contenu sémantique est bien plus variable que celui des autres coordonnants <sup>248</sup>, peuvent eux aussi être construits avec des coordonnants d’autres classes que les consécutifs (141).

- (141) a. Pour que ce soit insultant (**et cependant** vrai), il faudrait dire, par exemple (...) <sup>249</sup>  
b. La construction dramatique du livret se retrouve complètement dans la musique, avec une utilisation systématique, **et toutefois** subtile, des leitmotifs. <sup>250</sup>  
c. j’en connais pas mal qui prennent la bagnole pour un oui ou pour un non et qui vont faire du jogging **ou voire même** qui se payent des abonnements coûteux dans des salles de remise en forme. <sup>251</sup>  
d. je cherche un emploi dans un service Qualité, environnement **et/ou** sécurité. <sup>252</sup>  
e. C’est, derrière, des brevets sur le vivant, l’interdiction **ou/et** l’impossibilité de produire et d’utiliser ses propres semences, (...) <sup>253</sup>

Dans tous ces cas, les Coordonnants ainsi construits présentent toutes les caractéristiques de coordinations asyndétiques : la construction n’est pas syntagmatique dans la mesure où il n’existe pas de relation de dépendance entre les constituants (dans (140a) on ne saurait affirmer ni que *ainsi* dépend de *donc* ni le contraire), qu’ils occupent la même et unique fonction au sein de la construction dont ils seront à leur tour constituant immédiat en tant que Coordonnant construit, et que chacun d’eux apporte une information sémantique à la construction (à la différence des disfluences). Il n’est donc pas nécessaire de tenir compte de ces caractéristiques dans notre Coord construit, puisque ces cas de figure sont (à juste titre) pris en considération dans la définition des Coordinations asyndétiques.

Par contre il existe une seconde possibilité de construction des Coordonnants, qui n’est pas un EP cette fois, mais un syntagme endocentrique dont le Coordonnant constituant est la tête (142). C’est à ce cas précisément que l’on s’intéresse ici. Notre Coordonnant construit (fig. 93) est composé d’un Coordonnant (Tête) et d’un Adverbe <sup>254</sup>. L’adverbe se place après le Coordonnant, mais peut figurer à une certaine distance de celui-ci (plus exactement, cette distance

---

<sup>248</sup>Cette variabilité de contenu sémantique est comparable à celle des prépositions *à* et *de*.

<sup>249</sup>fr.soc.sectes, 8 octobre 2005.

<sup>250</sup>fr.rec.arts.musique.classique, 30 avril 2000.

<sup>251</sup>fr.misc.transport.velo, 17 mars 2001.

<sup>252</sup>fr.emplois.demandes, 14 février 1997.

<sup>253</sup>fr.sci.philo, 2 juin 2004.

<sup>254</sup>Tous les adverbes ne peuvent pas entrer dans cette construction, toutefois ici aussi il faudrait tenir compte d’informations de sémantique lexicale, donc nous préférons ne pas affiner plus avant dans un premier temps et laisser tous les adverbes pouvoir prendre cette place tout en sachant qu’en réalité une restriction est nécessaire.

est bornée par le coordonné auquel le coordonnant se rapporte, au-delà duquel l’adverbe ne peut être déplacé) (142a).

- (142) a. **Mais** ou **donc** le fils apprendra-t-il jamais à nager s’il ne peut se baigner ??? <sup>255</sup>  
 b. Son action biocide s’effectue sur toutes les parties, y compris ( **et même** surtout) sur le feuillage. <sup>256</sup>

Coord	
HERIT X	
Majeur	[1]Coord
Mineur	[2]Adv
Exigence	[1] $\Rightarrow$ [2]
Précédence	[1] $\prec$ [2]

FIG. 93 – Le Coordonnant construit (Coord).

Le coordonnant (syntagmatiquement) construit occupe dans la suite de l’analyse exactement le même statut que le non construit (et que le paradigmatiquement construit), et c’est la raison pour laquelle il porte la même catégorie et la même étiquette.

### 2.3.2 Conjoint de coordination (Conj-c)

Suivant les théories linguistiques, la structuration interne des coordinations syndétiques peut être conçue de deux manières différentes :

- soit comme une structure plate, de la forme  $[X_1 \text{ Coord } X_2]$ , (fig. 94 à gauche)
- soit comme une structure hiérarchisée, de la forme  $[X_1 [\text{Coord } X_2]]$ . (fig. 94 à droite)



FIG. 94 – Deux structurations possibles de la Coordination.

Sur ce point nous rejoignons [Mouret, 2003] qui montre que la forme hiérarchisée est la plus probable. Nous décrivons donc ici la construction  $[\text{Coord } X]$ , que l’on appelle Conjoint de coordination. Comme ce constituant est lui-même un constituant immédiat des coordinations, il est nécessaire de commencer par décrire cette construction avant d’arriver aux Coordinations syndétiques.

Le Conjoint de coordination (fig. 95) est une construction syntagmatique endocentrique constituée d’une construction et d’un Coordonnant (SPÉCIFIEUR de la construction). Comme le

<sup>255</sup>Alt.support.loneliness, 27 juillet 1999.

<sup>256</sup>fr.rec.jardinage, 23 juillet 2006.

Coordonnant construit, le Conjoint n'est pas un Entassement Paradigmatique. En effet, bien qu'elle soit présentée ici pour des raisons contextuelles, cette construction ne répond pas à la définition des EP donnée plus haut.

<b>Conj-c</b>	
HERIT Conj	
Majeur	2Coord

FIG. 95 – Le Conjoint de coordination (Conj-c).

Le Conj-c hérite de la construction Conjoint, décrite au chapitre 9 §2.1 (p. 208), qui rassemble les caractéristiques communes aux Conjoints de coordination (qui nous intéressent ici) et de subordination. On n'ajoute ici que la précision concernant la catégorie du spécifieur, réduite ici aux Coordonnants.

Remarquons que la construction peut être constituant d'autres constructions que des Coordinations (143) <sup>257</sup>.

- (143) a. Heu, et les améliorations de l'environnement, il les comptabilise où, bertrand ? <sup>258</sup>  
 b. et vous vous en êtes tirés comment vous a- vous avez continué ou

Ceci nous permet de répondre à une question fréquente dans le traitement de la coordination (cf. par exemple [Abeillé, 2003]) : celle de la fonction des objets coordonnés. Comme nous l'avons dit plus haut, puisque l'on considère que la coordination est une forme d'entassement paradigmatique, cela signifie que les éléments coordonnés partagent une unique fonction. Or, dans le cas d'un conjoint dans des contextes tels que (143) il n'est nullement question de partage de fonction, puisqu'il n'y a pas de rupture syntagmatique. Il ne s'agit donc pas, selon notre approche, d'une coordination <sup>259</sup>.

**Ordre.** — Un Conjoint est toujours introduit par son spécifieur, qui précède la tête. Ils sont toujours adjacents. Signalons toutefois que d'après [Wilmet, 1998] il existe une sorte de "sous-classe" de *coordonnants occasionnels* qui sont déplaçables : *donc, en effet, alors, au contraire, ainsi, par conséquent, toutefois*, etc. Ce dernier point n'est (pour l'heure) pas pris en considération dans la représentation formelle : de tels Conjoints pourront être introduits, mais sans être satisfaits à 100%.

### 2.3.3 Coordination syndétique (EP-cs)

Maintenant que nous avons défini les Coordonnants (simples et construits), constituants des Conjoints, puis les Conjoints eux-mêmes, nous pouvons présenter les Coordinations syndétiques (fig. 96), qui sont les Coordinations dont l'un au moins des constituants est un Conjoint (144).

<sup>257</sup>Ce qui nous fait concéder que la dénomination "Conjoint de coordination" n'est pas la meilleure possible, et qu'elle serait à modifier.

<sup>258</sup>fr.soc.environnement, 27 juillet 2006.

<sup>259</sup>Nous suivons ici également la position de [Wilmet, 1998].

- (144) a. Les auteurs du rapport félicitent les États-Unis pour le rôle de pionnier qu'ils jouent en ce qui concerne **les solutions fondées sur le libre jeu des forces du marché, les mesures novatrices et les partenariats destinés à améliorer l'environnement.** <sup>260</sup>
- b. (...) il est hasardeux de proclamer qu'**une position ou son contraire** représente une valeur québécoise. <sup>261</sup>
- c. Non tout ne va pas bien mais il ne faut pas se voiler la face **le progrès avance et ni vous ni moi ne pourrons l'infléchir.** <sup>262</sup>

EP-cs	
HERIT	EP-c
Majeur	$\boxed{1}$ Conj-c
Exclusion	$\boxed{\text{SPEC} \text{ID} \text{NATURE} \text{SCAT } x} \not\Rightarrow \boxed{\text{SPEC} \text{ID} \text{NATURE} \text{SCAT } \neg x}$
Précédence	$\neg\text{Conj-c} \prec \text{Conj-c}$

FIG. 96 – La Coordination syndétique (EP-cs).

La construction obéit aux contraintes suivantes :

- Par définition, puisqu'il s'agit d'une coordination syndétique (donc marquée), elle doit contenir au moins un objet qui soit d'étiquette Conjoint de coordination. En revanche, la présence d'un non-conjoint n'est pas obligatoire, d'où la précision que l'un des constituants ( $\boxed{1}$ , en l'occurrence) doit être (au moins) un Conjoint de coordination.
- Concrètement, elle peut recevoir un ou plusieurs Conjoints et un ou plusieurs non-conjoints, mais les Conjoints suivent toujours les non-conjoints. [Huddleston & Pullum, 2002] considèrent que les formes  $X, X, \dots, \text{et } X, \text{et } X$  (i.e. plusieurs conjoints et plusieurs non-conjoints) est agrammaticale (ou doit être analysée en plusieurs étapes successives et non en une seule coordination). Nous ne partageons pas cet avis, puisque l'on peut trouver des occurrences de cette forme.
- Quand plusieurs Conjoints sont présents, ils doivent tous être introduits par des Coordonnants, soit identiques (*et* avec *et*, *soit* avec *soit*, mais pas *et* avec *soit*), soit corrélés (145); nous exprimons cela en exprimant que les Coordonnants ne peuvent pas être de SCAT différente (telle que définie précédemment).

- (145) a. à la fois + et  
b. et + ainsi que / de même que  
c. ou + ou bien + ou alors

### 3 Disfluence (EP-c)

Les disfluences constituent un phénomène non négligeable dans l'étude de l'oral spontané, puisqu'on peut compter une amorce toutes les 50 secondes dans un corpus de français spontané

<sup>260</sup>fr.soc.environnement, 13 janvier 2006.

<sup>261</sup>qc.politique, 8 décembre 2005.

<sup>262</sup>fr.soc.environnement, 22 janvier 2006. Il y a deux EP-cs dans cet énoncé :  $[ni\ vous] / [ni\ moi]$  d'une part, et  $[le\ progrès\ avance] / [et\ ni\ vous\ ni\ moi\ ne\ pourrons\ l'infléchir]$  d'autre part.

d'après [Henry & Pallaud, 2003], alors que les amorces elles-mêmes ne constituent, comme on va le voir plus bas, qu'une partie des disfluences possibles.

Elles sont considérées par [Blanche-Benveniste *et al.*, 1990] comme des traces du travail de formulation lors de la production, et plus précisément comme le disent [Pallaud & Henry, 2004], comme des traces d'hésitation et non des corrections d'erreurs.

Phénomène non négligeable puisque très fréquent en oral spontané, la linguistique de corpus en a fourni un certain nombre d'études fines, présentant son organisation interne et ses caractéristiques. Cependant ces descriptions, quoique très précises dans leurs propositions, ne sont souvent pas exploitées en TALN ni même formalisées, sans doute en partie parce que le statut des disfluences dans une grammaire n'y est pas défini de manière claire et formalisable. En effet les applications symboliques de traitement automatique qui s'efforcent d'analyser des données orales font appel à des techniques différentes pour traiter les disfluences, techniques qui sont pourtant basées pour la plupart sur les mêmes descriptions initiales. Pour notre part, nous intégrons les Disfluences en tant qu'Entassements Paradigmatiques, en les opposant aux Coordinations que l'on vient de voir dans la section précédente. Nous commençons ici par donner une typologie de ce phénomène avant de présenter un aperçu des divers traitements qui en sont faits en TALN (§3.2), puis nous en présentons la formalisation (§3.3).

### 3.1 Typologie(s) des disfluences

Comme on l'a vu précédemment, dans la littérature linguistique une disfluence est un endroit dans un énoncé où "*le déroulement syntagmatique est brisé*" ([Blanche-Benveniste *et al.*, 1990]) : on occupe une même place syntaxique avec plusieurs objets (146).

(146)	il			
	il a quand-même	un :		
		une fibre pédagogique	assez :	
			assez euh enfin réelle	quoi

Ce mécanisme n'est pas propre aux disfluences, mais aux entassements paradigmatiques ; en revanche, alors que dans les coordinations, chaque occurrence de la même place ajoute un élément à la sémantique de l'énoncé, l'accumulation de *il*, de *un une* ou de *assez* dans (146) n'en modifie pas les caractéristiques sémantiques.

Parmi ces disfluences, on distingue deux grandes classes générales : les **bribes** qui sont des reprises à partir de syntagmes inachevés (146), et les **amorces** qui sont des reprises à partir de morphèmes inachevés (*paran-* dans (147)).

(147)	s'il n'y a pas d'éléments à mon avis	euh		
		il		
		il	tombe	dans la paran-
				dans la parano quoi

Au sein des amorces [Pallaud & Henry, 2004] identifient trois formes différentes (formes que

l'on peut, d'après elles, appliquer également aux bribes) :

- celles qui sont **complétées** (147), c'est-à-dire que les diverses étapes de reformulation reprennent un contenu identique en l'enrichissant progressivement,
- celles qui sont **modifiées** (148), c'est-à-dire que la reprise n'a pas la même forme que la (les) formulation(s) précédente(s), mais reste de la même catégorie (les deux verbes en gras dans l'exemple),

(148)	ils ont des ouvriers	euh	
		<b>payés</b>	
		<b>spécialisés</b>	sup-
			sur les chantiers de fouille

- et celles qui sont **inachevées** (149), c'est-à-dire que l'intégralité du syntagme est repris et modifié (*dans mon champs visuel (...)* à la place de *j'ai v-* ; c'est également le cas entre *sup-* et *sur* dans (148)).

(149)	tu sais	<b>j'ai v-</b> enfin	
		dans mon champ visuel	il y a eu quelque chose tu vois

Pour sa part, [Shriberg, 1994], inspirée par [Levelt, 1983], a décrit l'organisation interne des disfluences en un ensemble d'espaces distincts : le **reparandum** qui est le lieu de la première production, inachevée au niveau du point d'interruption (**interruption point**), suivi de l'**interregnum** au sein duquel il peut se produire soit rien, soit une marque d'hésitation, soit une à plusieurs nouvelles tentatives de formulation (inachevées), jusqu'au **repair** qui correspond à la reprise du déroulement syntagmatique.

Toutes ces études, qui décrivent l'organisation interne des disfluences, peuvent être prises en considération dans le développement d'une formalisation. Elles les présentent comme un phénomène unique, avec des caractéristiques régulières (l'entassement paradigmatique, le partage d'une unique fonction syntaxique et d'une unique fonction sémantique, les espaces internes), et des caractéristiques plus spécifiques à certains cas (bribes *vs.* amorces, inachèvement *vs.* complétion *vs.* modification, composition de l'interregnum). Cependant elles n'indiquent pas comment l'on différencie une disfluence d'une autre construction, ni comment l'on doit les traiter lors de l'analyse d'un énoncé.

### 3.2 Les disfluences en TALN

Le traitement des disfluences prend une place de plus en plus importante en TALN, d'une part avec la constitution et l'utilisation grandissantes de *treebanks* rassemblant des données de natures variées, dont quantité de productions orales transcrites ([Cotton & Bird, 2002]) ; d'autre part avec les recherches en Communication Homme-Machine, dont les techniques sont très variées et le plus souvent privilégient la robustesse et la rapidité à la finesse ([Zue & Glass, 2000]), mais nécessitent toutefois de tenir compte de ce phénomène. Observons donc comment

les disfluences sont traitées en TALN <sup>263</sup> : en dépit de la prise en considération, dans la plupart des cas, de tout ou partie des descriptions exposées ci-dessus, les solutions concrètes proposées pour le traitement automatique sont nettement différentes suivant la tâche à accomplir et le type d’approche.

La première technique que l’on peut rencontrer (p.ex. dans le système *Gemini*, [Dowding *et al.*, 1999]) consiste à **“effacer” les disfluences** de l’entrée qui sera analysée, en effectuant un pré-traitement des données dont l’objet est de reconnaître les disfluences et de les remplacer par une forme considérée comme “équivalente” ne présentant pas de rupture du déroulement syntagmatique. On peut se demander quel est précisément le niveau d’“équivalence” recherché, et quelles sont les limites imposées par des résultats d’analyse pour des utilisations ultérieures, si ceux-ci sont basés sur des entrées qui ne contiennent plus la totalité des informations linguistiques produites.

Une deuxième technique (cf. par exemple chez [Pérennou, 1996] ou chez [Seneff, 1992]) consiste en quelque sorte à **“ignorer” les disfluences**, *i.e.* à ne pas les prendre en compte lors de l’analyse. Ceci permet d’obtenir un résultat de parsing dit “robuste”, mais ne pose pas la question du statut des unités qui n’ont pas été considérées dans l’analyse : bien que les disfluences n’aient pas de fonction syntaxique en tant que telles, chaque élément qui occupe une place syntagmatique remplit en lui-même la fonction syntaxique de cette place, et il semble difficile d’admettre dans ce cas que seule une occurrence de chaque place sera considérée dans l’analyse. En d’autres termes, si l’on n’analyse que le *repair*, quel est le statut des constituants des autres espaces de la disfluence ?

Une troisième technique consiste à **regrouper les disfluences** en un groupe. [Antoine *et al.*, 2003] proposent dans cette perspective des analyseurs qui forment des disfluences en rassemblant des chunks (chacun d’eux devant être une occurrence de la même place syntaxique) en vertu de “*relations de dépendance sémantico-pragmatiques*”.

Dans un même ordre d’idées, dans le *Switchboard corpus of conversational speech* (une extension du *Penn Treebank*, [Taylor *et al.*, 2004]), [Godfrey *et al.*, 1992] proposent une méthode d’annotation des disfluences qui consiste à effectuer un parenthésage de la totalité de l’accumulation paradigmatique. Dans ce cas on peut se demander comment est déterminée la catégorie syntagmatique de cet ensemble (indispensable à l’analyse), qui peut être constitué de plusieurs répétitions ne comptant pas toujours les mêmes constituants, et qui ne correspond pas systématiquement à un syntagme complet. Ceci ne permet en outre que d’annoter les disfluences contiguës.

On voit que les techniques qui prennent en compte les disfluences diffèrent nettement dans leurs traitements, qui sont tous limités par la nature de leur représentation du phénomène. C’est la raison pour laquelle nous avons commencé par exposer clairement dans ce chapitre le statut que nous lui conférons dans la grammaire, ce qui nous permet de ne pas avoir à recourir à des mécanismes ou des représentations *ad hoc*.

---

<sup>263</sup>Parce que l’on se place nous-même dans la perspective générale de la représentation formelle de la langue, on ne s’intéresse ici qu’aux méthodes de TALN qui sont basées sur des descriptions linguistiques, et non sur les techniques probabilistes, qui certes proposent des approches intéressantes, mais ne font pas partie de notre cadre de recherche.



### 3.3 Disfluence (EP-d)

Dans la grammaire, la construction Disfluence (EP-d) prend la forme donnée en fig. 97. Ces quelques caractéristiques nous permettent de signifier qu’une disfluence consiste en la mise en relation de plusieurs objets (Majeurs hérités de EP) dont chacun des TRAITS des constituants sont de valeurs identiques un à un (propriété d’Accord).

<b>EP-d</b>	
HERIT EP	
Accord	$\boxed{x}. \text{TRAITS} \rightsquigarrow \boxed{y}. \text{TRAITS}$

FIG. 97 – La Disfluence (EP-d).

La satisfaction complète de ces propriétés caractérise une bribe **complétée** ; si un certain nombre de contraintes d’accord sont évaluées et non satisfaites alors on caractérise une bribe **modifiée**, et si les constituants sont des syntagmes (qui peuvent eux-mêmes ne pas être satisfaits à 100%) alors on a affaire à une disfluence **inachevée**. La propriété d’Accord telle qu’exprimée ici permet de traiter avec cette même description tout aussi bien les disfluences dont la modification est syntaxique (*un une* dans (146)), que celles dont la modification est sémantique (150).

(150)	ils sont pas à l’abri de ça quoi mais c’est	<b>un peu</b>	d’hypocrisie quand-même à ce niveau-là
		<b>pas mal</b>	

## 4 Précisions des descriptions

Ici on voit très clairement que les informations syntaxiques ne permettent pas, à elles seules, de traiter les phénomènes de parataxe avec finesse ; force est de constater qu’elles ne suffisent pas à obtenir des analyses (pourtant syntaxiques) suffisamment détaillées pour comprendre (ni pour expliquer, ni même pour représenter) les différences entre certaines coordinations dont les comportements s’opposent, comme par exemple les coordinations distributives (151a) *vs.* collectives (152a, 152c), symétriques (153a) ou réciproques (154a) <sup>264</sup>.

- (151) a. Jean et Paul sont grands  
b. Jean est grand et Paul est grand
- (152) a. Jean et Paul ont déplacé mon coffre-fort de 300kg  
b. ? Jean a déplacé mon coffre-fort et Paul a déplacé mon coffre-fort  
c. Pierre entre et sort de sa chambre sans arrêt  
d. \* Pierre entre de sa chambre et sort de sa chambre sans arrêt <sup>265</sup>
- (153) a. Pierre et Paul sont différents

<sup>264</sup>Exemples et dénominations tirés de [Riegel *et al.*, 1999].

<sup>265</sup>[Wilmet, 1998, p. 587].

- b. Pierre est différent de Paul et Paul est différent de Pierre
- c. \* Pierre est différent et Paul est différent

- (154)
- a. Pierre et Paul se détestent
  - b. Pierre déteste Paul et Paul déteste Pierre
  - c. \* Pierre déteste et Paul déteste
  - d. \* Pierre déteste de Paul et Paul déteste de Pierre

Cela vient notamment du fait que les EP, et parmi eux les coordinations, ne sont pas des phénomènes purement syntaxiques : ils relèvent d'un ensemble de caractéristiques, qui prennent place dans une analyse et ne peuvent être réglés par des informations lexicales et syntaxiques.

Il est donc évidemment nécessaire de compléter ces descriptions par des informations de niveaux différents. Ceci, dans notre approche, n'intervient pas au travers de traitement séparés ni même de ressources séparées (grammaires distinctes [Kahane & Lareau, 2005], représentations distinctes [Kordoni, 2003], etc.), mais au sein de notre unique grammaire, dont le mode de représentation a été développé précisément pour répondre à ce type de besoin : permettre aux différentes informations linguistiques d'interagir entre elles en une analyse globale. Ainsi, il est tout à fait envisageable d'ajouter aux propriétés utilisées dans les descriptions proposées ci-dessus des informations supplémentaires, et parmi elles notamment :

- Pour l'oral, des informations sur le contour prosodique des objets qui entrent dans un entassement paradigmatique : on peut penser que l'intonation d'une disflue (continuation) ne sera pas la même que celle d'une coordination (répétition d'un même patron) ; pour l'écrit, des informations sur la forme des objets qui entrent dans l'EP : listes, retours à la ligne, ponctuation, etc.
- Des informations sur la signification des objets qui entrent dans l'EP : au delà de l'unité fonctionnelle, l'une des choses qui contraignent une coordination est que les objets soient plausiblement coordonnables, sous peine de formation de zeugmes (155a), de tautologies (155b) ou de contradictions (155c)<sup>266</sup>.

- (155)
- a. Les invités ont été introduits par la grande porte et par le majordome
  - b. La femme de Jean travaille dans un ministère et Jean est marié
  - c. Jean est célibataire et il est marié depuis dix ans

Ces exemples sont syntaxiquement satisfaits, ce n'est que sémantiquement qu'ils enfreignent un certain nombre de contraintes. C'est la raison pour laquelle la grammaire dans son état actuel permet de les introduire sans les distinguer de coordinations qui sont sémantiquement plausibles.

Il est important enfin de remarquer que les descriptions que l'on propose, pour généralistes qu'elles soient, permettent de ce fait et grâce à la flexibilité du modèle de traiter les coordinations les plus canoniques aussi bien que des coordinations elliptiques telles que (156)<sup>267</sup>.

---

<sup>266</sup>Exemples et dénominations tirés de [Riegel *et al.*, 1999].

<sup>267</sup>[Riegel *et al.*, 1999].

(156)	Pierre	a écrit	le début	
	et Paul		la fin	de mon discours

Dans ce cas la propriété d'adjacence de la Coordination ne sera pas satisfaite, pour les deux constructions qui seront introduites, *Pierre et Paul* d'une part, *le début et la fin* d'autre part. Nous développons cette analyse dans le chapitre 14.

Il reste bien entendu nécessaire d'affiner les descriptions en ajoutant aux constructions présentées ici d'autres constructions plus fines, présentant des spécificités plus remarquables (contraintes sur la concordance des temps, sur les inversions de sujets, etc.).

### Résumé

On a proposé ici une description de la coordination qui s'approche de celle proposée dans l'*Approche Pronominale* : on se base sur la proximité structurelle de ce phénomène avec celui des disfluences, pour introduire dans notre grammaire une représentation générale des deux, appelée Entassement Paradigmatique, dont héritent d'une part les Coordinations, d'autre part les Disfluences.

Les phénomènes dont nous traitons ici ont ceci de particulier qu'ils ne constituent pas des constructions syntagmatiques (leurs constituants n'entretiennent pas de relations hypotactiques, et ils partagent une même place syntaxique, ainsi que sémantique dans le cas des disfluences). Ces entassements paradigmatiques peuvent être constitués d'objets parfaitement identiques ou partiellement différents, mais ont toujours un certain nombre de caractéristiques communes, et l'on peut en décrire une organisation interne relativement précise.

Cependant la formalisation <sup>a</sup> de ces phénomènes, quand elle est proposée, montre qu'en plus de tout cela, il est nécessaire avant même de les formaliser de répondre aux questions suivantes : Quelle est leur place dans une grammaire ? Quand et comment les analyse-t-on ? Nous avons donc ici proposé une réflexion sur la place des entassements paradigmatiques dans une grammaire formelle, à la suite de laquelle nous avons introduit notre représentation de ce phénomène. Au delà du problème posé par cette représentation particulière, nous avons mené une réflexion plus générale sur la place de ce type de phénomènes, propres à l'oral, et par là nous avons présenté une vision globale du développement de grammaire et de ces spécificités. Nous avons proposé ensuite la description des Coordinations et des Disfluences en tant qu'Entassements Paradigmatiques, qui nous permet de mettre en avant leurs relations, et de les intégrer à la grammaire sans l'intervention de représentations spécifiques.

---

<sup>a</sup>Ainsi que l'exploitation très variable que l'on peut en voir en TALN.

# Conclusion

Dans cette partie, nous avons présenté des éléments de grammaire du français, formalisés en GP. L'ensemble de ces éléments a vocation à constituer un squelette permettant une analyse relativement complète de la plupart des phénomènes syntaxiques du français, tant écrit qu'oral. Par rapport aux autres grammaires GP existantes, celle-ci est la première ressource exploitant pleinement les caractéristiques de GP, décrites dans la première partie.

L'ensemble des constructions syntaxiques que l'on y propose est inégalement développé : certains types de constructions ont une description très grossière qui permet uniquement, en l'état, d'en effectuer une analyse sommaire, d'autres types sont décrits plus finement, comme notamment les verbes à montée ou les entassements paradigmatiques. Toutefois, l'ensemble de ces constructions présente l'avantage d'être homogénéisé et de ne pas présenter de contradictions internes. Chaque partie le nécessitant pourra par la suite être développée.

Les descriptions des constructions de cette grammaire ne contiennent (quasiment) que des informations syntaxiques : il s'agit là d'une limitation volontaire que l'on s'est imposée en raison du temps imparti, et pour pouvoir proposer à son issue une ressource permettant malgré le manque inévitable de complétude, de faire des analyses d'énoncés. Ces analyses sont donc, pour l'instant, des analyses strictement syntaxiques, bien qu'il ne tienne qu'aux développements ultérieurs d'ajouter des informations et des constructions d'autres ordres pour modifier cet état de fait qui n'est finalement qu'une étape, insuffisante mais nécessaire, de développement.

La part personnelle de la description de ces constructions est irrégulière : certaines constructions sont des reproductions de descriptions linguistiques empruntées à d'autres auteurs, d'autres sont le fruit d'une synthèse personnelle basée sur plusieurs théories comme c'est le cas pour les Syntagmes Nominaux, d'autres encore sont issues d'emprunts de descriptions auxquels on a dû ajouter des recherches pour compléter l'intégration dans la grammaire, comme les Entassements Paradigmatiques pour lesquels on disposait d'une typologie relativement détaillée mais on n'avait pas de description syntaxique claire, et d'autres enfin sont le résultat de nos propres recherches, comme p.ex. l'organisation des Verbes construits.

L'une des conséquences notables de cela est que la grammaire ainsi proposée est finalement le produit de plusieurs théories syntaxiques, différentes mais ici rassemblées. Plusieurs points généraux concernant cette grammaire émergent des descriptions que l'on a faites au cours des chapitres précédents :

- Toutes les constructions de la grammaire ne sont pas des syntagmes. Et ce, malgré le fait que l’on se soit bornée à ne mentionner que des informations syntaxiques. Les Entassements Paradigmatiques, par exemple, sont des constructions dont les constituants entretiennent des relations paratactiques, et d’après la définition qui en est donnée aussi bien dans l’approche spécifique de [Blanche-Benveniste, 1975] que dans celles généralistes de [Le Goffic, 1993] ou de [Riegel *et al.*, 1999], ces relations ne donnent pas lieu à la constitution de syntagmes (dont les constituants, comme son nom l’indique, entretiennent des relations syntagmatiques, i.e. hypotactiques).
- De nombreux syntagmes de la grammaire ne sont pas endocentriques. Ceci va à l’encontre de la position de la syntaxe X-barre et de ses descendants (y compris HSPG), pour lesquels tout syntagme possédant une tête est un syntagme endocentrique (selon la définition de l’endocentricité de [Bloomfield, 1961]). C’est pour se détacher de cette idée que l’on a redéfini ici, avant toute chose, le terme de “tête”. Nous faisons, dans cette grammaire, une différence entre les “X construits” (équivalents de  $\bar{X}$ ) dont le recteur est X et qui sont sur le même paradigme que X, et les “Syntagmes X” (équivalents de  $\bar{\bar{X}}$ ) dont le recteur est X mais qui ne sont pas sur le même paradigme que X. Les “X construits” sont des syntagmes endocentriques, mais pas les “Syntagmes X”.
- On ne fait pas usage de *Principes* dans notre grammaire. Les informations qui peuvent être généralisées sont remontées au point le plus haut possible de la hiérarchie d’héritage, quitte à ce qu’elles soient, au besoin, localement complétées ou modifiées pour le traitement de certains épiphénomènes. Cela nous permet une certaine flexibilité dans l’affectation des traits des constructions en fonction de leurs constituants, tout en étant en accord avec les idées théoriques directrices de cette grammaire, qui ne souhaite pas postuler de généralisations sur le fonctionnement syntaxique des langues alors même que seul le français a été étudié ici.
- Un certain parallélisme est notable entre les hiérarchies syntagmatiques par types de constructions : ceci est volontaire et est une manifestation d’une idée que l’on se fait de l’organisation syntaxique, selon laquelle malgré sa complexité apparente elle fait l’objet d’une recherche d’économie, tout comme c’est le cas pour les autres domaines. Ainsi, nous pensons que même si les constructions auront tendance à être de plus en plus spécifiques à mesure que l’on entre dans le détail des descriptions (et que l’on descend dans la hiérarchie), nous pensons que la structuration de base reste globalement parallèle pour la grande majorité des types de constructions. C’est ce que l’on a voulu illustrer par le parallélisme assez clair entre les hiérarchies locales de types de constructions.

Les choix théoriques que l’on a faits dans cette grammaire ont été guidés non pas essentiellement par la volonté de suivre une théorie donnée plutôt qu’une autre (puisque l’on a fait usage de nombreuses d’entre elles), mais de favoriser l’homogénéité des descriptions. Nous concédons que cet objectif aurait pu être mené à bien en faisant d’autres choix que ceux que l’on a faits ici, et que ceux-ci se justifient tout autant par une conception de l’organisation syntaxique des informations qui devient personnelle à un certain degré. Nous ne prétendons certainement pas que nos choix soient ni les meilleurs, ni les seuls possibles ; nous avons uniquement essayé à travers ces éléments de grammaire de les mettre en pratique afin de montrer ce qu’ils peuvent apporter.

---

Maintenant que nous avons proposé cette grammaire, nous verrons dans la partie qui suit comment celle-ci permet de traiter un certain nombre d'énoncés présentant des phénomènes différents, et ainsi de montrer quel est son intérêt pour l'analyse linguistique.





Troisième partie

Exemples d'analyses



# Introduction

Dans la première partie de ce manuscrit, nous avons présenté les positions théoriques sous-jacentes au modèle que nous utilisons. Dans la deuxième partie, nous avons proposé des éléments de description syntaxique ayant pour objet de constituer un squelette de grammaire sur lequel articuler des recherches ultérieures. Dans cette troisième partie, nous souhaitons illustrer notre propos en présentant l'application de cette grammaire à l'analyse d'énoncés.

## 1 Structure produite

Les analyses en GP produisent des structures hiérarchiques représentant la constitution progressive des constructions telles que définies dans la grammaire. Ces constructions étant (idéalement) des ensembles forme-sens, la hiérarchie produite est censée représenter une hiérarchie de la construction de la signification des productions linguistiques.

Néanmoins et comme on l'a souligné dans la partie précédente, du fait que la grammaire que nous avons présentée ici est une grammaire à vocation essentiellement syntaxique (dans le premier temps que représente ce travail), la hiérarchie qui sera construite à partir de cette grammaire est une hiérarchie syntaxique, en constituants <sup>268</sup>. Toutefois il ne s'agit pas d'un arbre syntagmatique (à la Chomsky) pour autant, puisque notre hiérarchie ne représente (en elle-même) que les informations de constituance, et non celles d'ordre linéaire comme c'est le cas dans les autres <sup>269</sup>; à titre de comparaison on pourrait dire que l'on représente ici un arbre strictement *syntaxique* et non *syntagmatique*.

En l'état de la grammaire, les structures construites sont donc des structures syntaxiques. Il faut malgré cela imaginer que celles-ci, une fois enrichies d'informations provenant d'autres domaines, ne seront plus nécessairement strictement syntaxiques, dans la mesure où la syntaxe ne sera plus l'unique base sur laquelle des constructions seront décrites dans la grammaire.

---

<sup>268</sup>Et non en dépendances comme les stemmas de [Tesnière, 1959] et autres structures dérivées.

<sup>269</sup>Dans notre cas elles sont fournies à l'intérieur des objets constituant la structure mais n'ont pas d'incidence sur sa forme.

## 2 Analyse *vs.* Parsing

Les exemples d'analyse que nous présentons dans les chapitres qui suivent ne sont pas des représentations de ce qu'un parseur ferait, et ce pour deux raisons.

La première raison est que ce que l'on souhaite montrer ici n'est pas ce qu'un parseur peut faire à l'aide de GP, mais plus généralement comment une grammaire GP est appliquée à l'analyse d'un énoncé, c'est-à-dire concrètement comment elle fonctionne, ce qu'elle permet de faire, la façon dont elle permet d'envisager et d'effectuer des analyses. Ce fonctionnement a bien entendu une relation que l'on souhaite la plus étroite possible avec le fonctionnement des analyses automatisées, mais ce que l'on montre ici ressemble plus à une sorte de "fonctionnement idéal" qu'à un algorithme déjà existant <sup>270</sup>.

La seconde raison est que, comme on l'a dit précédemment, chaque mise en application automatisée de GP peut répondre à des enjeux fort différents et que l'on ne pourrait pas tous détailler ici. L'on devrait dans ce cas tenir compte à la fois d'une perspective déterministe et alors envisager toutes les possibilités de détermination des analyses, et également d'une perspective non déterministe où il peut s'agir soit de chercher à obtenir les meilleurs résultats possibles (en termes de qualité - encore faudrait-il déterminer si la notion de "qualité" est la même dans tous les cas) parmi toutes les possibilités offertes par la grammaire, ou à l'opposé de chercher à obtenir tous les résultats rendus possibles par la grammaire sans distinction de qualité de manière à se faire une idée de sa couverture. Toutes ces démarches sont différentes par essence, et l'on ne saurait ni les décrire toutes dans une unique présentation, ni en choisir une à défaut des autres.

De plus, les enjeux particuliers de chaque parseur ne sont pas du ressort de la grammaire développée. Celle-ci possède son enjeu propre, qui est en l'occurrence de permettre une analyse syntaxique cohérente et homogène de toutes sortes de productions linguistiques.

En conséquence, nous ne détaillerons pas ici toutes les analyses rendues possibles par la grammaire pour chaque énoncé. Nous nous concentrerons sur une sélection des caractérisations successives nous permettant d'arriver à l'analyse que l'on a voulu permettre pour chaque phénomène. Il va de soi que, puisque la grammaire proposée n'est pas un mécanisme définitoire de la langue, elle n'interdit que très peu de constructions (et d'autant moins qu'en son état actuel elle n'est pas encore suffisamment développée et ne tient pas compte de suffisamment d'informations); alors pour éviter de noyer nos analyses sous le bruit engendré par cet état de fait, et donc pour rendre nos exemples plus lisibles, nous ne représenterons pas ici ce qui n'est pas pertinent pour l'analyse des phénomènes auxquels nous nous sommes intéressée. On pourrait considérer cela comme une façon d'ignorer les problèmes inhérents à nos positions théoriques, mais nous préférons le voir simplement comme une façon de mettre en exergue le traitement rendu possible des phénomènes qui nous intéressent, reléguant les questions de décidabilité et de surgénération au niveau de l'automatisation de l'analyse, et non à celui de l'analyse elle-même. De fait, toute théorie syntaxique offre des solutions alternatives à l'analyse d'un énoncé donné, et notre grammaire n'échappe pas à cette règle.

---

<sup>270</sup> Celui s'en rapprochant le plus actuellement est celui du *SeedParser* de [Vanrullen, 2005].

### Résumé

Dans cette introduction nous avons présenté la structure produite par une analyse GP, qui en l'état de la grammaire utilisée constitue une structure syntaxique (mais non uniquement syntagmatique).

Nous avons également expliqué que les exemples d'analyse qui sont présentés dans cette partie ne représentent pas les résultats de ce que ferait un parseur, mais ce qui est possible de faire avec une analyse GP en général. On a fait ce choix d'une part parce que l'analyse en GP peut correspondre à plusieurs algorithmes de parseurs aux caractéristiques fort différentes (déterministe ou non p.ex.), et parce que ces caractéristiques ne dépendent pas (et ne doivent pas dépendre) de la grammaire utilisée, qui elle-même répond à des besoins qui lui sont propres (les positions théoriques qu'elle met en jeu).



# Chapitre 12

## Un énoncé simple

Dans ce chapitre nous détaillons l’analyse d’une phrase très simple (157), ceci afin de présenter le déroulement d’une analyse basée sur GP.

(157) Jean admire la maison

Pour les besoins de l’exemple, nous supposons que cet énoncé soit produit par un seul locuteur, sous forme écrite.

### 1 Début de l’analyse : Informations disponibles

Une fois l’énoncé segmenté et étiqueté en vertu d’un lexique <sup>271</sup>, nous disposons de l’ensemble des quatre constructions lexicales représentées dans les figures 1 à 4, à partir desquelles on va effectuer notre analyse <sup>272</sup>.

Dans ces constructions on trouve les informations intrinsèques suivantes (il s’agit de celles qui ont été définies au chapitre 3, que nous rappelons ici) :

- **ancree** : les index de début et de fin dans l’énoncé.
- **Forme** : la forme (graphique, en l’occurrence) des objets.
- **ID** : l’identité de la construction : sa catégorie, son type, sa sous-catégorie le cas échéant, ses traits et marques d’accords.

---

<sup>271</sup>Ces étapes ne font pas l’objet du présent travail et nous ne les détaillons donc pas ici. Selon son objectif principal, la grammaire que nous utilisons permettra à terme d’effectuer l’analyse morphématique des énoncés, mais elle ne le peut pas en son état actuel ; nous partons donc, comme la plupart des grammaires formelles, d’une entrée segmentée et étiquetée.

<sup>272</sup>Notons que pour les besoins de l’exemple nous utilisons des descriptions très simples de chacun des objets, afin de ne pas rendre l’exemple trop compliqué à lire. Il va de soi cependant que les analyses que nous proposons dans les chapitres suivants sont nettement plus fins que celui exposé ici.

N	
INTR.	ANCRE
	INDEX
	DÉBUT 0
	FIN 1
FORME	ORTH Jean
	CAT N
ID   NATURE	SCAT propre
	TRAITS
	GENRE masculin
	NOMBRE singulier

FIG. 1 – Construction lexicale *Jean* dans (157).

V	
INTR.	ANCRE
	INDEX
	DÉBUT 1
	FIN 2
FORME	ORTH admire
	ID   NATURE
	CAT V
	SCAT principal
TRAITS	MODE indicatif
	TEMPS présent
ACCORD	PERSONNE 3
	NOMBRE singulier

FIG. 2 – Construction lexicale *admire* dans (157).

Det	
INTR.	ANCRE
	INDEX
	DÉBUT 2
	FIN 3
FORME	ORTH la
	CAT Det
ID   NATURE	ACCORD
	GENRE féminin
	NOMBRE singulier

FIG. 3 – Construction lexicale *la* dans (157).

## 2 Premier niveau d'analyse

### 2.1 Mise en relation des unités

260

Lors de la première étape où il s'agit de mettre en relation les objets présents en fonction des informations contenues dans la grammaire, on peut considérer que l'on réduit la grammaire



N							
INTR.	ANCRE	INDEX	<table><tr><td>DÉBUT</td><td>3</td></tr><tr><td>FIN</td><td>4</td></tr></table>	DÉBUT	3	FIN	4
	DÉBUT	3					
	FIN	4					
	FORME	ORTH	<i>maison</i>				
	ID   NATURE	CAT	<i>N</i>				
	SCAT	<i>commun</i>					
TRAITS	GENRE	<i>féminin</i>					
	NOMBRE	<i>singulier</i>					

FIG. 4 – Construction lexicale *maison* dans (157).

à sa partie pertinente pour l'analyse des objets présents, i.e. on relève, parmi toutes les propriétés figurant dans les définitions de toutes les constructions dans la grammaire, celles qui mentionnent les objets présents dans l'énoncé à analyser et qui remplissent leur **condition d'évaluation** (on construit alors l'ensemble  $\mathcal{P}_{G,1,1}^{disp}$  des propriétés disponibles <sup>273</sup>). Notons bien que, comme on l'a dit précédemment, cette étape ne tient pas compte des structures auxquelles les propriétés appartiennent ; le lien entre ces propriétés et les constructions correspondantes sera effectué au cours du deuxième mouvement. Une fois la grammaire réduite à un sous-ensemble pertinent pour les objets présents, on évalue chacune des propriétés (i.e. on vérifie leurs **conditions de satisfaction**). A l'issue de cette étape on obtient le réseau complexe de relations  $\mathcal{P}_{G,1,1}^{eval}$ , dont on a une illustration en fig. 5.

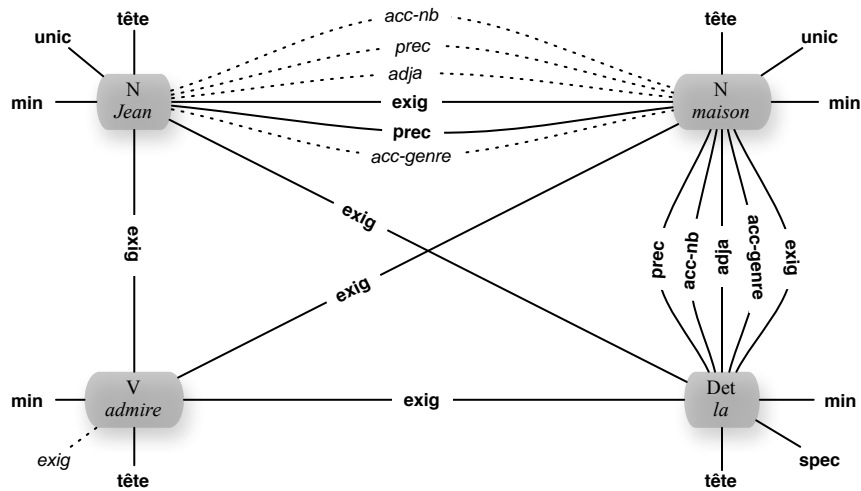


FIG. 5 – Mise en relation des unités, niveau 1.

<sup>273</sup>Pour simplifier les représentations, et parce que ce sont les seules qui nous serviront pour l'analyse, on n'a tenu compte ici que des constructions nominales, verbales et propositionnelles de la grammaire.

Ce graphe représente l'ensemble des informations relationnelles que l'on a pu évaluer pour les trois constructions présentes <sup>274</sup>. Chaque noeud du graphe correspond à une construction (le nom *Jean*, le déterminant *la*,...), et chaque arc à une propriété évaluée. Les propriétés satisfaites sont représentées avec un trait plein (p.ex. la précédence entre *la* et *maison*), et les propriétés violées avec un trait pointillé (p.ex. l'accord en genre entre *la* et *Jean*). Les propriétés dont l'évaluation met en relation deux membres (adjacence, exigence satisfaite, unicité violée,...) relient les deux noeuds correspondants, celles dont l'évaluation met en relation un seul membre pointent uniquement sur le noeud concerné (majeur, mineur, unicité satisfaite,...).

## 2.2 Interprétation du système de contraintes

Le second mouvement du processus d'analyse consiste à tirer des conclusions du réseau de relations que l'on vient de fabriquer. A première vue, en observant le graphe qui représente ce réseau, on constate que le déterminant *la* et le nom commun *maison* entretiennent de nombreuses relations satisfaites. C'est le cas aussi de *maison* et *Jean*, cependant plusieurs propriétés qui les lient sont violées. On peut assez facilement déduire de tout cela que la première construction proposée sera constituée de *la* et *maison*.

Plus formellement maintenant, l'étape d'interprétation consiste à retrouver à quelle(s) construction(s) correspond(ent) les relations construites, puis en fonction du degré de grammaticalité que l'on souhaite conserver lors de l'analyse à introduire les nouvelles constructions. En remplaçant les arcs représentant les propriétés par un arc pour chaque construction que l'on peut introduire à partir de leur évaluation, et en notant pour chaque arc le **degré de satisfaction local** (DSL) de la construction correspondante (suivi du ratio nombre de propriétés satisfaites / nombre de propriétés évaluées), on obtient la figure 6, qui représente  $C_{\hat{G},1,1}$ .

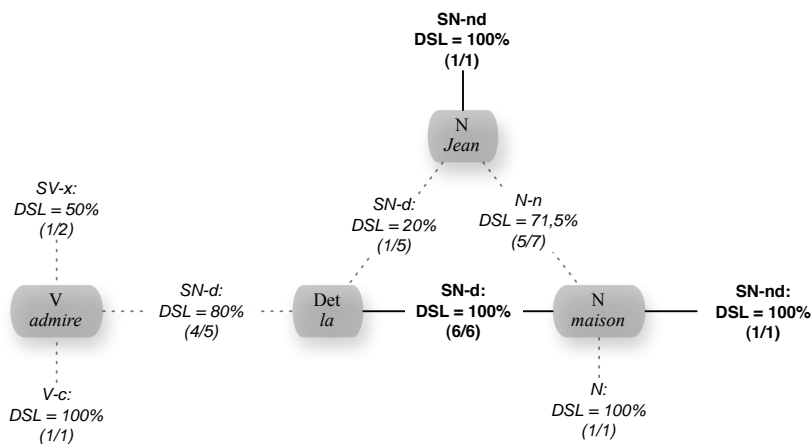


FIG. 6 – Interprétation du système de contraintes, niveau 1.

<sup>274</sup>Pour des raisons de lisibilité, plutôt que de mentionner “Majeur” pour les propriétés Majeur étant évaluées, on a dans cette figure et dans les suivantes indiqué “Tête” pour les propriétés Majeur qui indiquent la tête de la construction, et “Spec” pour celles indiquant son spécifieur.

Dans cette figure on a représenté en trait continu les constructions dont aucune propriété n'est violée (DSL = 100%), et en pointillés les constructions dont au moins une propriété est non satisfaite (DSL < 100%) ; en d'autres termes, on considère ici que l'on n'accepte d'introduire que les constructions qui sont parfaitement grammaticales (i.e., conformes à la grammaire)<sup>275</sup>. On a également représenté en pointillés le N construit uniquement constitué de *maison*, et le V construit uniquement constitué de *admire*, dans la mesure où ces deux constructions seraient inutiles, puisque renferment les mêmes informations que les objets dont elles sont constituées.

A l'issue de cette étape on introduit donc trois nouvelles constructions, de niveau 2, qui correspondent au SN non déterminé *Jean* (fig. 7), au SN déterminé *la maison* (fig. 8) et au SN non déterminé *maison*. On ne représente pas ce dernier, d'une part parce qu'il a la même caractérisation que le SN-nd *Jean*, d'autre part parce qu'il ne nous servira pas dans la suite de l'analyse.

SN-nd		
INTR.	FORME	[ ORTH [1].FORME ]
	ID   NATURE	[ CAT SN TYPE [1].TYPE ]
	TRAITS	[ GENRE [1].GENRE NOMBRE [1].NOMBRE ]
	SYN	[ TÊTE [1] ]
Majeur		[1]N +

FIG. 7 – Niveau 1 : Le SN-nd *Jean*.

Aucune autre construction ne pouvant être introduite à partir du réseau de relations dont on dispose, on reprend à partir de ce point le premier mouvement, en utilisant cette fois les nouvelles constructions introduites pour établir les mises en relation.

### 3 Deuxième niveau

#### 3.1 Mise en relation des unités

Lors de la mise en relation des objets dont on dispose à présent, on obtient le réseau de relations présenté dans la figure 9.

<sup>275</sup>Notons encore une fois qu'il s'agit d'un choix, et que l'on aurait pu accepter d'introduire les constructions dont le degré de satisfaction était supérieur ou égal à un pourcentage différent de 100%. Par exemple si l'on avait choisi d'accepter les constructions dont le DSL était supérieur à ou égal à 60% alors on aurait pu introduire également le SN-d *admire la* (d'un DSL de 80%) et le N-n *Jean maison* (DSL = 71,5%), mais pas le SN-d *Jean la* (d'un DSL de 20%) ni les autres.

SN-d		
INTR.	FORME	[FORME   ORTH <i>la maison</i> ]
	ID   NATURE	CAT <i>SN</i> TYPE [1]. <i>cat</i> TRAITS [             GENRE [1]. <i>genre</i> NOMBRE [1]. <i>nombre</i> ]
	SYN	TÊTE [1] SPEC [2]
Majeur	[1]N [2]Det	+ +
Précédence	[2] < [1]	+
Adjacence	[2] ↔ [1]	+
Accord	[1].GENRE ∼ [2].GENRE	+
	[1].NOMBRE ∼ [2].NOMBRE	+

FIG. 8 – Niveau 1 : Le SN-d *la maison*.

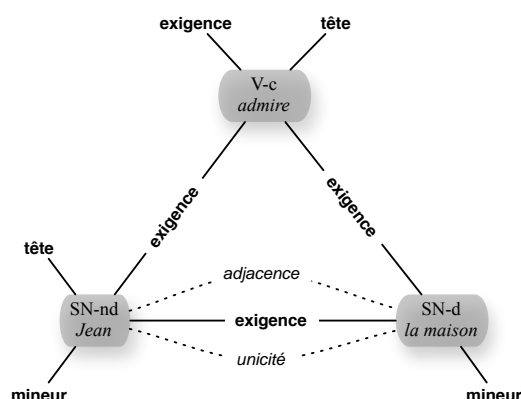


FIG. 9 – Mise en relation des unités, niveau 2.

### 3.2 Interprétation du système de contraintes

Une fois ce nouveau réseau de relations constitué, on observe que l'on peut introduire les constructions représentées dans la figure 10.

Parmi les constructions qui apparaissent, deux ne seront pas introduites, ayant un DSL trop bas : le SV à complément direct *Jean admire la maison* et le SN endocentrique *Jean la maison* ; les trois autres constructions seront introduites : le SV intransitif *admire* (qui ne devrait plus être introduit une fois que des informations de sémantique lexicale seront ajoutées aux entrées, puisque *admire* est bivalent, ce qui impliquera qu'il ne sera plus satisfait à 100%) et les deux SV transitifs directs *Jean admire* (lui pourra ne pas être introduit inutilement quand on aura affiné les descriptions syntaxiques des constructions) et *admire la maison*. Parmi ces trois

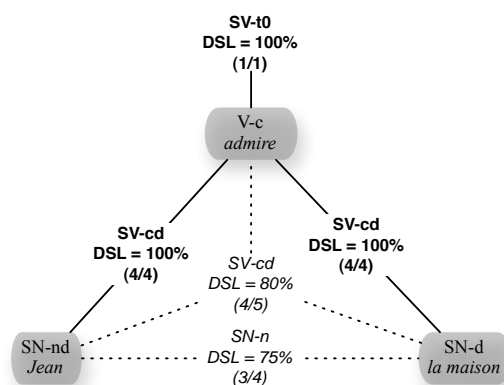


FIG. 10 – Interprétation du système de contraintes, niveau 2.

constructions dont le DSL est de 100%, une seule nous sera utile par la suite, le SV *admire la maison* (fig. 11).

SV-cd				
INTR.	FORME	[ ORTH <i>admire la maison</i> ]		
	ID   NATURE	CAT	SV	
		TYPE	[1].TYPE	
		TRAITS	[ MODE    [1].MODE TEMPS    [1].TEMPS ]	
	ACCORD	PERSONNE	[1].PERSONNE	
NOMBRE		[1].NOMBRE		
SYN	TÊTE	[1]		
	OBJ-DIR	[ RECTEUR            [1] DÉPENDANT    [2] ]		
	Majeur	[1]V-c	+	
	Mineur	[2]SN	+	
	Unicité	[2]	+	
	Exigence	[1] ⇒ [2]	+	

FIG. 11 – Le SV-cd *admire la maison*.

Puis on repasse une nouvelle fois au premier mouvement, pour mettre en relation les objets que l'on vient d'introduire.

## 4 Troisième niveau

### 4.1 Mise en relation des unités

On obtient cette fois le réseau de relations présenté dans la figure 12.

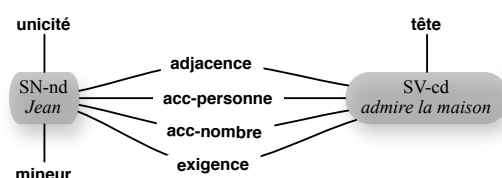


FIG. 12 – Mise en relation des unités, niveau 3.

### 4.2 Interprétation du système de contraintes

On peut introduire à ce niveau les constructions représentées dans la figure 13.

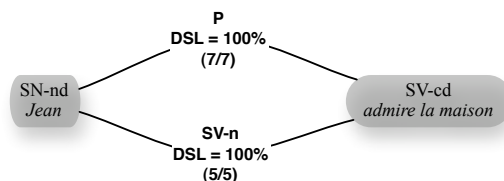


FIG. 13 – Interprétation du système de contraintes, niveau 3.

On peut donc introduire deux constructions à ce niveau, qui représentent toutes les deux *Jean admire la maison* : une première qui serait un SV endocentrique, où *Jean* serait un complément circonstanciel, et une seconde qui serait une Proposition exocentrique verbale, et qui correspond à la bonne analyse. Nous présentons cette dernière construction en figure 14.

A la suite de l'introduction de cette nouvelle construction, aucune nouvelle relation ne peut être construite pour l'énoncé. L'analyse s'arrête donc à ce point.

P		
INTR.	FORME	$\left[ \text{FORME} \mid \text{ORTH} \quad \textit{Jean admire la maison} \right]$
	ID   NATURE	$\left[ \begin{array}{l} \text{CAT} \quad P \\ \text{TYPE} \quad \boxed{1}.\text{TYPE} \end{array} \right]$
	SYN	$\left[ \begin{array}{l} \text{SUJET} \quad \left[ \begin{array}{l} \text{RECTEUR} \quad \boxed{2} \\ \text{DÉPENDANT} \quad \boxed{1} \end{array} \right] \end{array} \right]$
	Majeur	$\boxed{1}\text{SV} \quad +$
	Mineur	$\boxed{2}\text{SN} \quad +$
	Unicité	$\boxed{2} \quad +$
	Exigence	$\boxed{1} \Rightarrow \boxed{2} \quad +$
	Adjacence	$\boxed{2} \leftrightarrow \boxed{1} \quad +$
	Accord	$\boxed{2}.\text{PERSONNE} \rightsquigarrow \boxed{1}.\text{PERSONNE} \quad +$
		$\boxed{2}.\text{NOMBRE} \rightsquigarrow \boxed{1}.\text{NOMBRE} \quad +$

FIG. 14 – La P *Jean admire la maison*.

## Résumé

Dans ce chapitre on a montré le déroulement pas à pas de l'analyse d'un énoncé très simple, afin de montrer le fonctionnement de la caractérisation en GP.

On a vu que chaque niveau d'analyse successif se construit en deux temps, le premier où l'on construit un réseau de relations entre les objets via les propriétés de la grammaire, et le second où l'on évalue ce réseau de contraintes afin d'en faire émerger les constructions à introduire.





# Chapitre 13

## Constructions verbales

Dans ce chapitre nous souhaitons montrer quelles analyses nous permet la description des constructions verbales (V et SV) que l'on a faite dans la grammaire de la Partie 2 (chapitre 7). A cette fin nous présenterons l'analyse détaillée de plusieurs exemples. La première section illustre la différence de structure syntaxique entre constructions verbales simples, à auxiliaire, à contrôle, à modal et à montée. La deuxième section illustre la différence de structure syntaxique entre l'analyse d'un verbe à sujet plein et celle d'un verbe à sujet clitique. La troisième, enfin, illustre les différents traitements des clitiques non-nominatifs dans différents types de constructions verbales.

### 1 Différentes constructions verbales

La façon dont nous décrivons les différents types de verbes dans la grammaire (principaux, auxiliaires, modaux), engendre des différences de traitement (et par conséquent de structure syntaxique) entre les constructions verbales simples, composées (i.e., à auxiliaire), et modalisées (dont causatives).

Dans le cas d'une **construction verbale simple**, le Verbe, qui est alors une construction lexicale (i.e. ce n'est pas un Verbe construit) est constituant immédiat d'un Syntagme Verbal, dont le(s) autre(s) constituant(s) immédiat(s) est (sont) son (ses) complément(s) de verbe (pleins <sup>276</sup>). Par exemple, pour un énoncé tel que (158),

(158) Marie mange une pomme

on obtient la structure présentée en figure 15 (nous encadrons la partie de l'arbre qui nous intéresse).

Dans le cas d'une **construction verbale composée** (c'est-à-dire, puisque la description de notre grammaire reste sur ce point très sommaire, à la fois les constructions à temps composé et

---

<sup>276</sup>On illustre le cas des compléments clitiques en §3.

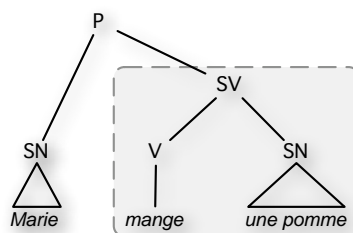


FIG. 15 – Exemple de structure de l'analyse d'une construction verbale simple.

les constructions passives), les Verbes (auxiliaire(s) et principal) sont constituants immédiats d'un Verbe construit, lequel est par la suite lui-même constituant immédiat d'un Syntagme Verbal, dont le(s) autre(s) constituant(s) immédiat(s) est (sont) son (ses) complément(s) de verbe tout comme pour les constructions verbales simples. Par exemple, pour un énoncé tel que (159),

(159) Marie a mangé une pomme

on obtient une structure telle que celle présentée en figure 16.

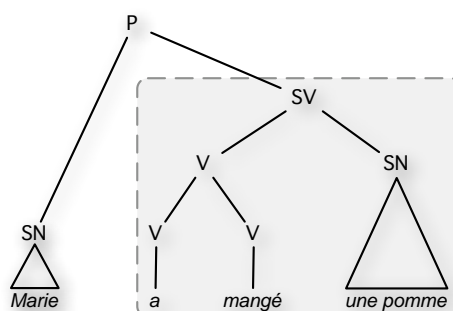


FIG. 16 – Exemple de structure de l'analyse d'une construction verbale composée.

Notons que l'analyse des **verbes à contrôle** entre dans ce que l'on vient de voir (suivant s'il est auxilié ou non), et que dans ce cas le verbe principal est le verbe à contrôle ; le verbe "contrôlé" fait partie de son complément de verbe. Par exemple pour un énoncé tel que (160),

(160) Marie essaie de manger une pomme

la structure obtenue prendra la forme présentée en fig. 160.

Cette analyse signifie que l'on considère que ce sont les relations sémantiques entre les objets qui varient dans le cas d'un verbe à contrôle (considéré comme un verbe principal), mais pas les relations syntaxiques.

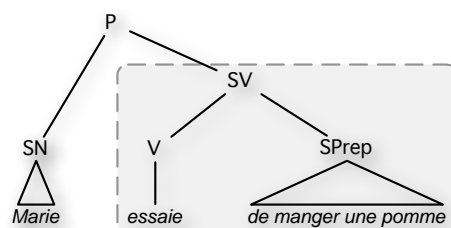


FIG. 17 – Exemple de structure de l'analyse d'une construction verbale à contrôle.

L'analyse des constructions à contrôle est différente de celle que l'on fait des verbes à montée et des constructions causatives, qui entrent dans le cas des **constructions verbales modali-sées** : les Verbes (modal et principal) sont alors constituants immédiats d'un Verbe construit, lequel est par la suite lui-même constituant immédiat d'un Syntagme Verbal, dont le(s) autre(s) constituant(s) immédiat(s) est (sont) son (ses) complément(s) de verbe. Par exemple, pour un énoncé tel que (161),

(161) Marie peut manger une pomme

on obtient une structure telle que celle présentée en figure 18. Nous développons le traitement des clitiques pouvant porter sur l'un ou l'autre verbe au §3.

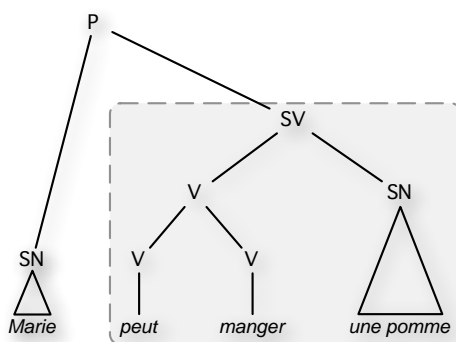


FIG. 18 – Exemple de structure de l'analyse d'une construction verbale à modal.

Nous présentons dans ce qui suit le détail de l'analyse de quelques exemples, afin de voir comment ceci est rendu possible par la grammaire. Notons que l'analyse d'une construction verbale simple a été développée dans le chapitre précédent, et qu'il est possible de s'y référer pour comparer avec les traitements qui suivent.

## 1.1 Construction verbale à auxiliaire

Dans cette section nous présentons l'analyse de l'énoncé (159). Au début de l'analyse les unités de l'énoncé sont les suivantes :

- *Marie* est un Nom propre féminin singulier, comme illustré en fig. 19.

N	
INTR.	FORME
	GRAPH <i>Marie</i>
	CAT <i>N</i>
	SCAT <i>propre</i>
ID   NATURE	TRAITS
	GENRE <i>féminin</i>
	NOMBRE <i>singulier</i>

FIG. 19 – L'entrée lexicale correspondant à *Marie* dans (159).

- *A* est un Verbe auxiliaire conjugué à la 3<sup>e</sup> personne du singulier à l'indicatif présent (fig. 20).

V	
INTR.	FORME
	GRAPH <i>a</i>
	CAT <i>V</i>
	SCAT <i>auxiliaire</i>
ID   NATURE	TRAITS
	MODE <i>indicatif</i>
	TEMPS <i>présent</i>
ACCORD	PERSONNE <i>3</i>
	NOMBRE <i>singulier</i>

FIG. 20 – L'entrée lexicale correspondant à *a* dans (159).

- *Mangé* est un Verbe principal au participe passé singulier de genre indéterminé (fig. 21). A la différence du verbe auxiliaire *a* que l'on vient de voir, celui-ci, parce qu'il est au mode participe, porte un trait d'ACCORD en GENRE, et en revanche ne porte pas d'information morphologique de PERSONNE.
- *Une* est un Déterminant indéfini féminin singulier (fig. 22). A la différence des noms, le déterminant porte le GENRE en tant qu'information d'ACCORD (et non pas comme un TRAIT) puisqu'il s'agit d'une marque d'accord avec l'objet dont il dépend.
- Enfin, *Pomme* est un Nom commun féminin singulier (fig. 23).

V	
INTR.	FORME
	GRAPH <i>mangé</i>
	CAT <i>V</i>
	SCAT <i>principal</i>
ID   NATURE	TRAITS
	MODE <i>participe</i>
ACCORD	TEMPS <i>passé</i>
	GENRE <i>masculin ∨ féminin</i>
	NOMBRE <i>singulier</i>

FIG. 21 – L'entrée lexicale correspondant à *mangé* dans (159).

Det	
INTR.	FORME
	GRAPH <i>une</i>
	CAT <i>Det</i>
	SCAT <i>indéfini</i>
ID   NATURE	ACCORD
	GENRE <i>féminin</i>
	NOMBRE <i>singulier</i>

FIG. 22 – L'entrée lexicale correspondant à *une* dans (159).

N	
INTR.	FORME
	GRAPH <i>pomme</i>
	CATÉGORIE <i>N</i>
	SCAT <i>commun</i>
ID   NATURE	TRAITS
	GENRE <i>féminin</i>
	NOMBRE <i>singulier</i>

FIG. 23 – L'entrée lexicale correspondant à *pomme* dans (159).

### 1.1.1 Introduction du SN *Marie*

Au premier niveau, la première construction issue de notre analyse est un SN non-déterminé (SN-nd, décrit au chapitre 6, §2.2.2) constitué du seul Nom propre *Marie*. On en donne la caractérisation en fig. 24.

Cette construction étant constituée d'un unique objet, il est évident que son "réseau de relations" est réduit au minimum possible : la seule propriété disponible pour la caractérisation de ce SN est celle de Majeur concernant la tête (qui est satisfaite).

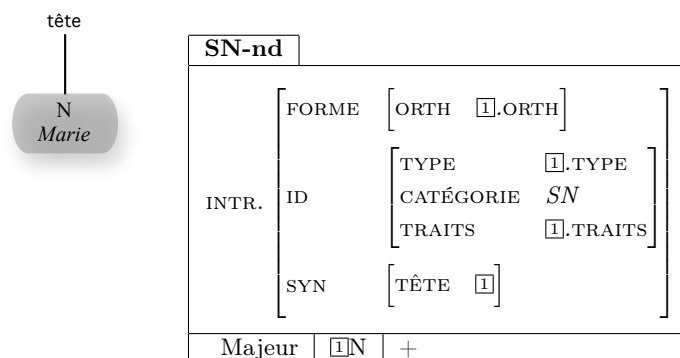


FIG. 24 – Introduction du SN *Marie* dans (158).

### 1.1.2 Introduction du SN *une pomme*

Une seconde construction est introduite à ce premier niveau d'analyse : il s'agit du SN déterminé (SN-d, décrit au chapitre 6, §2.2.1) constitué du Déterminant *une* et du Nom commun *pomme*. On en donne la caractérisation en fig. 25.

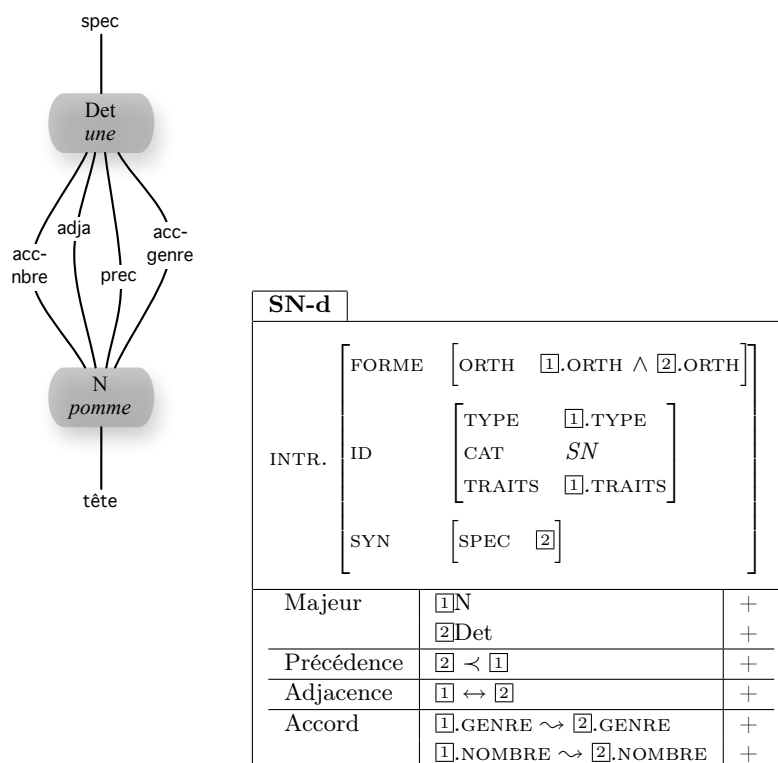


FIG. 25 – Introduction du SN *une pomme* dans (158).

Etant constitué de deux objets, le réseau de relations de ce SN est bien sûr un peu plus

conséquent que le précédent : on y voit que le N satisfait la propriété de Tête et le Det celle de Spécifieur, et que les deux entretiennent quatre relations également satisfaites : leurs Accord en genre et en nombre ainsi que leur Adjacence et la Précédence du déterminant par rapport au nom.

Le SN-d introduit à partir de ce réseau de relations est de type nominal et de catégorie SN ; il récupère les TRAITS de sa tête (donc du N), et porte une dépendance syntaxique que l'on appelle SPÉCIFIEUR.

### 1.1.3 Introduction du V *a mangé*

A ce même niveau est caractérisé un “Verbe au participe passé avec auxiliaire avoir” (V-a, décrit au chapitre 7, §1.4.1), constitué de l’auxiliaire *à* et du verbe principal *mangé* (fig. 26).

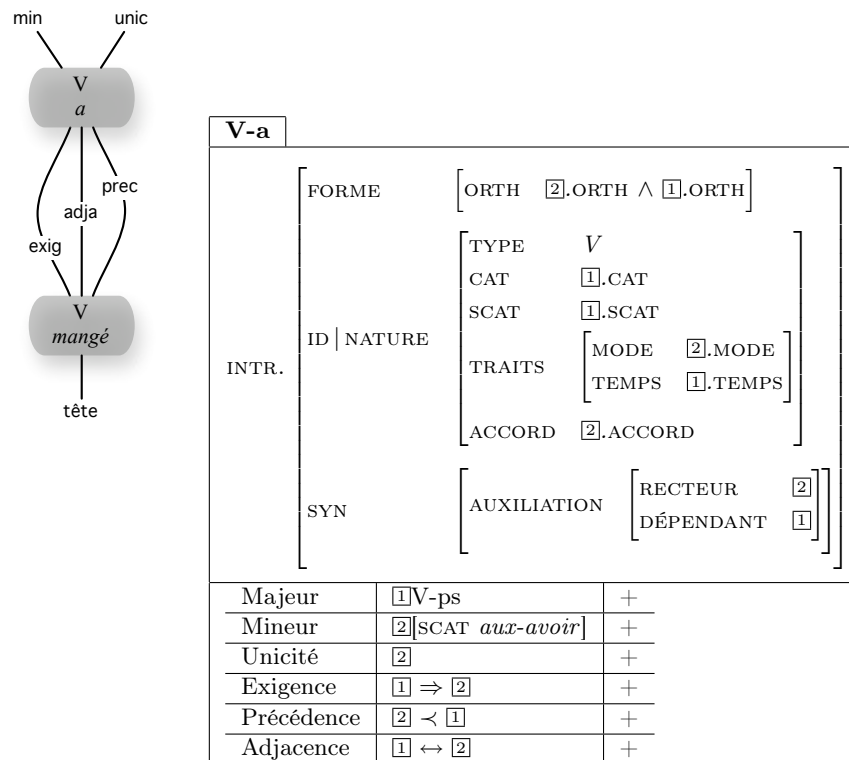


FIG. 26 – Introduction du V *a mangé* dans (159).

Dans la caractérisation de cette construction, on voit que l’auxiliaire satisfait les propriétés de Mineur et d’Unicité tandis que le verbe principal satisfait celle de Tête. Les deux sont liés par trois propriétés satisfaites : l’Exigence de l’auxiliaire par le verbe, ainsi que les contraintes de Précédence et d’Adjacence entre les deux constituants.

La construction introduite est de type et de catégorie V. De sa tête, elle prend la sous-catégorie (SCAT), ainsi que les valeurs de TEMPS et ACCORD ; de son auxiliaire elle prend les valeurs de

MODE. Elle porte en outre une dépendance syntaxique appelée AUXILIATION dont le recteur est l’auxiliaire, et le dépendant est le verbe tête.

#### 1.1.4 Introduction du SV *a mangé une pomme*

Au second niveau d’analyse est caractérisé un Syntagme Verbal transitif direct (SV-1od, décrit au chapitre 7, §2.2.3) constitué du Verbe *a mangé* et du SN *une pomme* que l’on vient d’introduire (fig. 27).

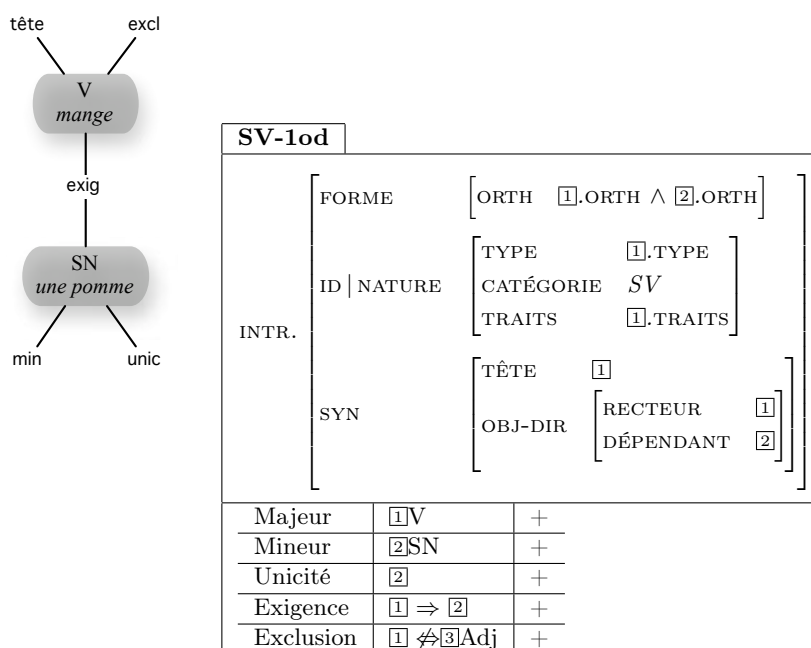


FIG. 27 – Introduction du SV *mange une pomme* dans (159).

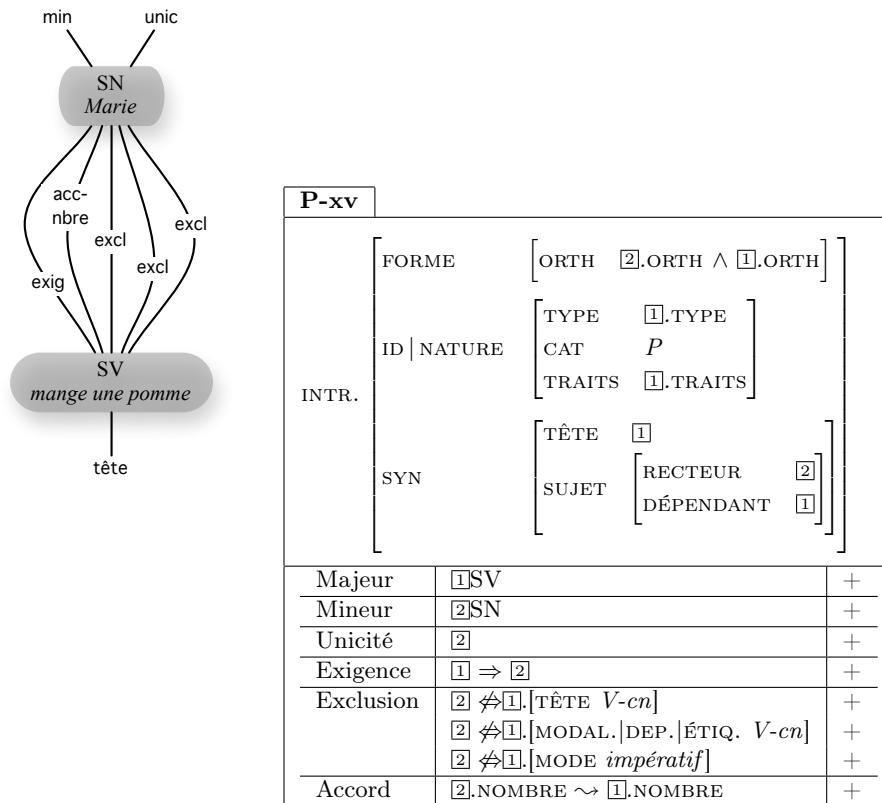
On voit dans la fig. 27 que le verbe satisfait la propriété Tête ainsi qu’une Exclusion (qui concerne un Adjectif). Le SN quant à lui satisfait les propriétés de Mineur et d’Unicité. Enfin, les deux objets sont liés par l’Exigence du SN par le Verbe.

La construction obtenue est de type V et de catégorie SV. Elle récupère les TRAITS de sa tête, et porte une dépendance syntaxique appelée OBJET DIRECT dont le recteur est le verbe et le dépendant, le SN.

#### 1.1.5 Introduction de la P *Marie a mangé une pomme*

Enfin, au troisième (et dernier) niveau d’analyse est caractérisée une “Proposition exocentrique verbale” (P-xv, décrite au chapitre 8, §1.2.4) constituée du SV *a mangé une pomme* et du SN *Marie* (fig. 28).



FIG. 28 – Introduction de la P *Marie mange une pomme* dans (158).

On voit dans la fig. 28 que le SV satisfait la propriété Tête et le SN satisfait les propriétés de Mineur et d'Unicité. Les deux objets sont liés par l'Exigence du SN par le Verbe, leur accord en nombre, ainsi que trois Exclusions entre SN et V qui sont satisfaites parce que le Verbe ne porte pas les traits correspondants <sup>277</sup>.

La construction obtenue est de type et de catégorie P. Elle récupère les TRAITS de sa tête, et porte une dépendance syntaxique appelée SUJET dont le recteur est le SN et le dépendant, le Verbe tête.

## 1.2 Construction verbale à contrôle

Dans cette section nous présentons l'analyse de l'énoncé (160). Encore une fois, au début de l'analyse les unités de l'énoncé portent des informations très proches de celles décrites précédemment :

<sup>277</sup> On aurait pu les représenter comme pointant uniquement sur le SN puisque le V, en l'occurrence n'est pas vraiment pointé par ces Exclusions. On les a représentées sous cette forme parce que les propriétés faisaient explicitement référence à ces deux objets. Mais qu'on les représente d'une façon ou de l'autre, cela ne modifie en rien la caractérisation.

- *Marie*, *essaie*, *une* et *pomme* correspondent respectivement aux constructions présentées en fig. 19, 20, 22 et 23. Notons tout de même que le verbe à contrôle est considéré comme un verbe principal (sa différence, i.e. le fait d’être “à contrôle”, s’effectue au niveau sémantique et non au niveau syntaxique).
- *De* est une Préposition (fig. 29).

Prep	
INTR.	FORME [ORTH <i>de</i> ]
	ID   NATURE [CAT <i>Prep</i> ]

FIG. 29 – L’entrée lexicale correspondant à *de* dans (160).

- *Manger* est un Verbe principal infinitif (fig. 30). Contrairement aux verbes précédents, il ne porte ni TEMPS, ni informations d’ACCORD.

V-i	
INTR.	ID   NATURE [CAT <i>V</i>
	SCAT <i>principal</i>
	TRAITS [MODE <i>infinitif</i> ]

FIG. 30 – L’entrée lexicale correspondant à *manger* dans (160).

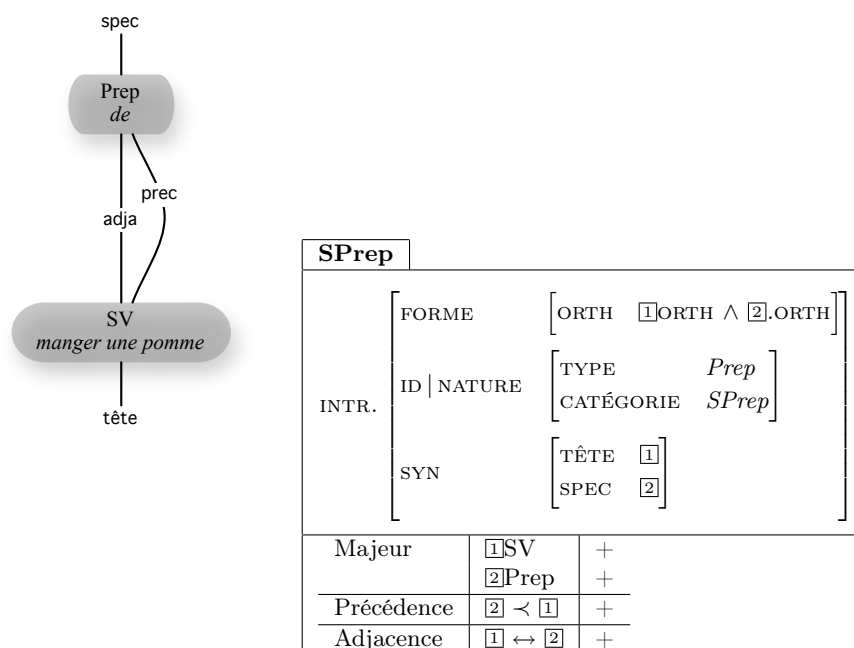
Au premier niveau d’analyse sont introduits les mêmes SN que précédemment, cf. fig. 24 pour *Marie* et fig. 25 pour *une pomme* (à l’index près pour ce dernier).

Au niveau suivant est introduit un Syntagme Verbal transitif direct (SV-1od) constitué du verbe *manger* et du SN *une pomme*, dont la caractérisation est semblable à celle du SV *a mangé une pomme* (fig. 27).

### 1.2.1 Introduction du SPrep *de manger une pomme*

Au troisième niveau est introduit un Syntagme Prépositionnel dont les constituants immédiats sont le SV que l’on a précédemment introduit *manger une pomme* et la préposition *de* (fig. 31).

La préposition est son Spécifieur et le SV sa Tête; l’Adjacence et la Précédence des deux constituants sont en outre satisfaites. Le SPrep introduit à partir de cette caractérisation est de type Prep.

FIG. 31 – Introduction du SPrep *de manger une pomme* dans (160).

### 1.2.2 Introduction du SV *essaie de manger une pomme*

Au niveau suivant est introduit un Syntagme Verbal transitif indirect (SV-1oi, décrit au chapitre 7, §2.2.4) dont les constituants immédiats sont le SPrep que l'on a précédemment introduit *de manger une pomme*, et le verbe *essaie* (fig. 32).

Sa caractérisation est assez semblable à celles des SV transitifs directs vus précédemment. Ce qui diffère dans cette construction est évidemment la dépendance syntaxique qui est ici OBJET INDIRECT, et dont le recteur est la tête et le dépendant, le SPrep.

Enfin, en dernier niveau il est introduit une Proposition dont les constituants immédiats sont le SN *Marie* et le SV *essaie de manger une pomme*. Cette construction est semblable à celle présentée en figure 28.

## 1.3 Construction verbale à modal

Dans cette section nous présentons l'analyse de l'énoncé (161).

Au début de l'analyse les unités de l'énoncé sont semblables aux unités décrites précédemment, à l'exception de *peut* qui est un verbe modal (fig. 33).

Aux premiers niveaux d'analyse, comme dans les cas précédents sont introduits les SN-nd *Marie* (fig. 24) et SN-n *une pomme* (fig. 25).

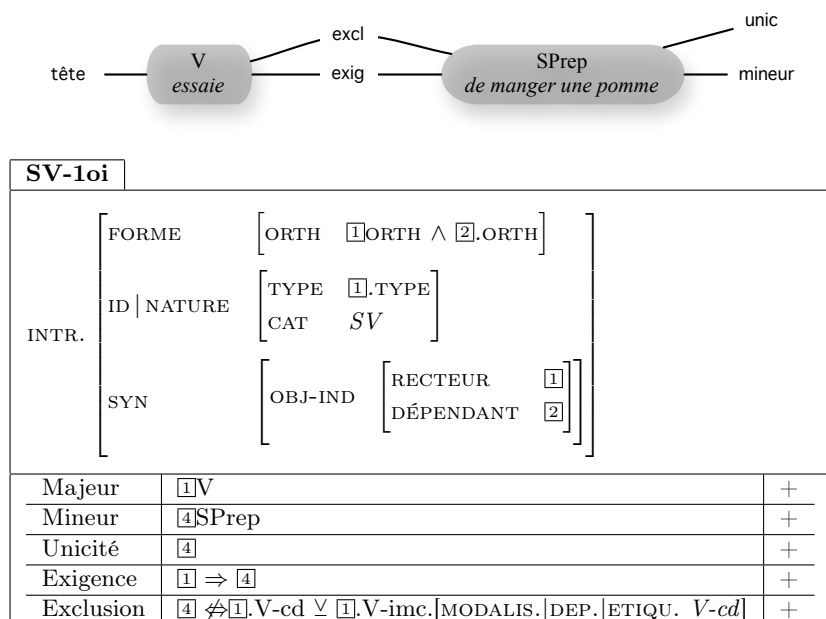


FIG. 32 – Introduction du SV *essaie de manger une pomme* dans (160).

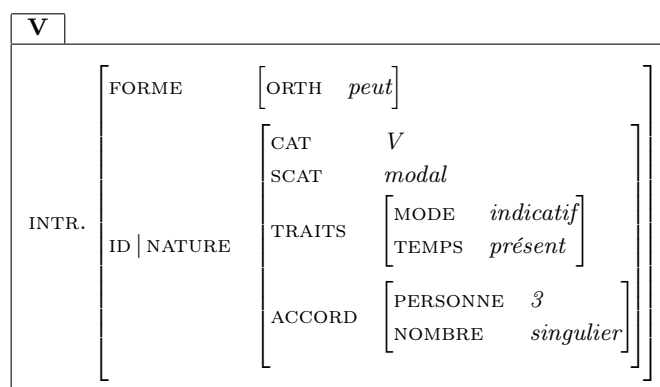


FIG. 33 – L'entrée lexicale correspondant à *peut* dans (161).

### 1.3.1 Introduction du V *peut manger*

Cette construction de catégorie verbale et d'étiquette V-im ("Verbe infinitif à modal", décrit à chapitre 7, §1.5.1) et constituée du modal *peut* et du verbe principal *manger* (fig. 34).

Il a la même caractérisation que le verbe à auxiliaire *avoir* (*a mangé*), bien que les caractéristiques des unités mises en relation dans les deux cas soient bien sûr différentes (la tête étant un participe dans un cas, un infinitif dans l'autre, et l'objet Mineur étant un auxiliaire dans un cas, un modal dans l'autre). La seule différence entre les deux réseaux de relations est que celui-ci ne comporte pas de mention de l'Adjacence (parce que le modal n'est pas nécessairement adjacent au verbe auquel il se rapporte, cf. (162)).

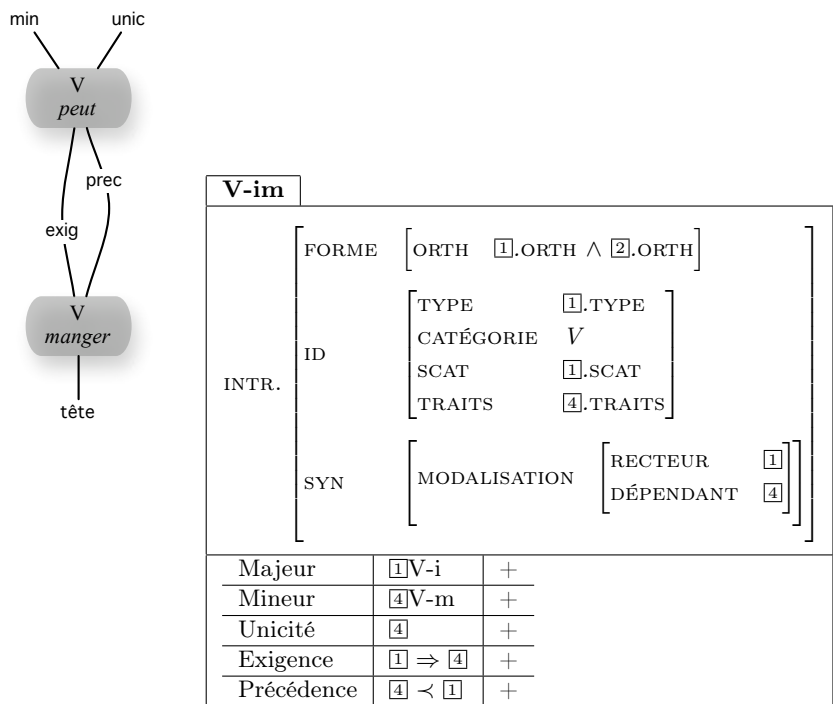


FIG. 34 – Introduction du V *peut manger* dans (161).

- (162) a. Marie **entend** Paul **arriver**  
 b. Marie **laisse** Paul **manger** une pomme

La fin de l'analyse est la même que précédemment : un SV transitif direct *peut manger une pomme* est construit (semblable à celui de la figure 27), puis une P *Marie peut manger une pomme* (id. fig. 28).

## 2 Sujet plein vs. sujet clitique

La façon dont nous décrivons les pronoms clitiques, c'est-à-dire comme des "affixes verbaux", engendre une différence de traitement (et par conséquent de structure syntaxique) entre les verbes à sujet plein (SN, Pro disjoint,...) et ceux à sujet clitique.

Dans le premier cas, un syntagme verbal est tout d'abord construit, constitué du verbe et de ses compléments le cas échéant ; à la suite de quoi, le sujet est attaché à ce SV en une proposition. Par exemple pour (163),

- (163) Marie aime le théâtre

on obtient une structure telle que celle présentée en figure 35.

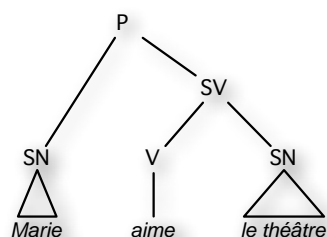


FIG. 35 – Exemple de structure de l'analyse d'une proposition à sujet plein.

Dans le second cas, le sujet étant un pronom clitique, il est alors constituant d'un verbe construit, et y est alors directement rattaché avant même ses éventuels compléments pleins, qui eux sont les constituants d'un SV de niveau supérieur. Par exemple pour (164),

(164) elle aime le théâtre

cela donne une structure telle que présentée en figure 36.

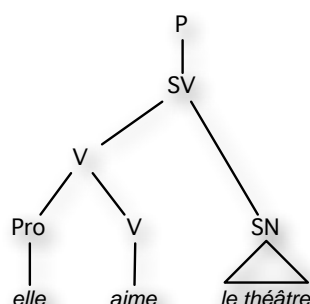


FIG. 36 – Exemple de structure de l'analyse d'une proposition à sujet clitique.

Nous présentons dans ce qui suit le l'analyse de ces deux structures différentes, afin de voir comment ceci est rendu possible par la grammaire.

## 2.1 Constructions à sujet plein

Tous les énoncés de la section précédente ((158) à (162)) ainsi que l'exemple (163) ci-dessus sont des exemples de constructions verbales à sujet plein. On ne revient donc pas dessus, mais on note simplement ici que dans ce cas, le sujet (un SN en l'occurrence, mais il aurait aussi bien pu s'agir d'un Pronom disjoint ou de n'importe quel objet occupant le même paradigme, comme

par exemple une Subordonnée relative (165a) ou un SV infinitif (165b)) est un constituant immédiat d'une Proposition exocentrique verbale, aux côtés d'un Syntagme Verbal.

- (165) a. **Qui dort** dîne  
 b. **Faire de la plongée** m'angoisse complètement

## 2.2 Constructions à sujet clitique

Contrairement aux sujets pleins, les sujets clitiques sont les constituants immédiats du Verbe construit.

### 2.2.1 Introduction du V *elle aime*

Ainsi dans l'analyse de (164), au premier niveau l'on introduit un "Verbe à clitique nominatif" (V-cn, décrit au chapitre 7, §1.6.1) dont les deux constituants sont le verbe principal *aime* et le clitique *elle* (fig. 37).

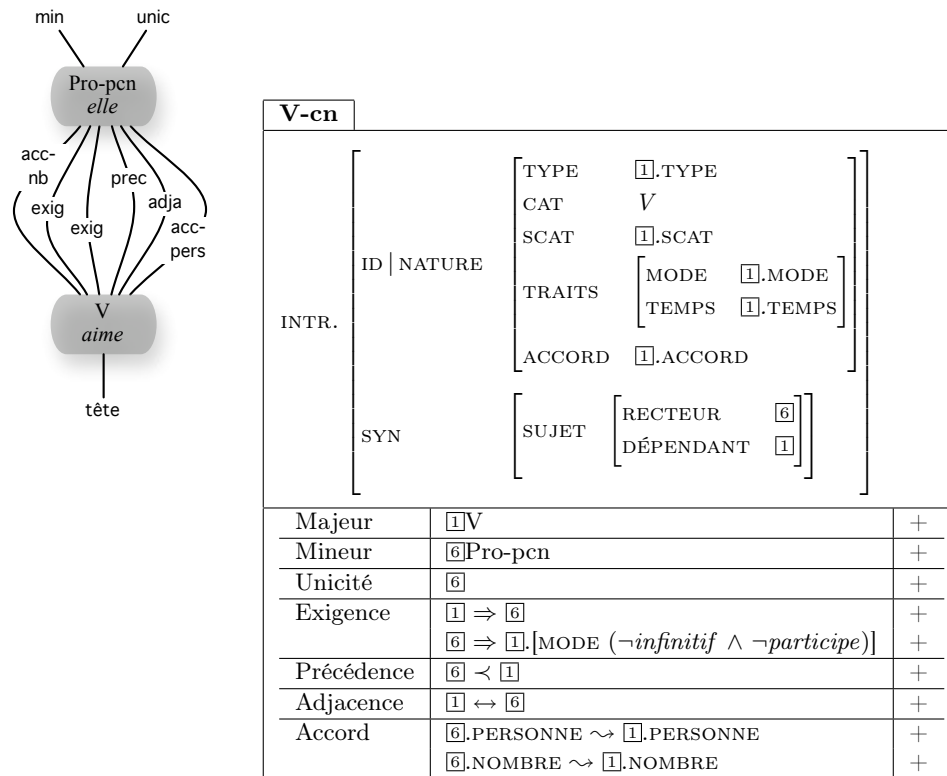


FIG. 37 – Introduction du V *elle aime* dans (164).

La relation existant entre le pronom clitique et le verbe étant plus contrainte que celle entre un sujet plein et le verbe (en termes linguistiques : contraintes de position, etc.), le réseau

de relations existant entre les deux premiers est plus dense que celui des seconds. Il permet d'introduire la dépendance SUJET dès ce niveau puisque les deux objets sont en présence.

A la suite de cela est introduit le SV transitif direct (SV-1od) *elle aime le théâtre*, dont la caractérisation est la même que celle du SV *a mangé une pomme* que l'on avait donnée dans la section précédente (fig. 27).

### 2.2.2 Introduction de la P *elle aime le théâtre*

Enfin, le SV que l'on vient d'introduire est le constituant unique d'une Proposition, dont la caractérisation est donnée en fig. 38.

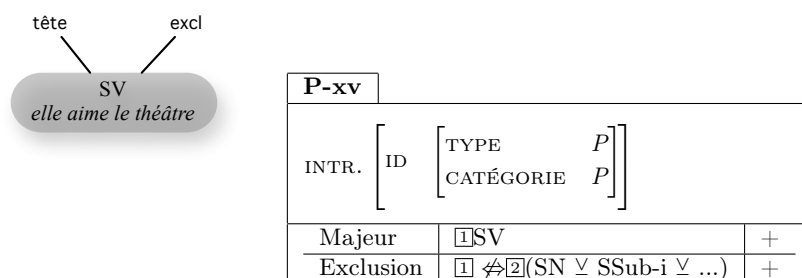


FIG. 38 – Introduction de la P *elle aime le théâtre* dans (164).

Etant constituée d'un unique objet, cette proposition présente un ensemble de relations moins important que les propositions que l'on avait pu voir précédemment. On y note la présence de la propriété d'Exclusion, satisfaite, qui indique qu'un SV dont la Tête est un V-cn (i.e., un verbe à clitique nominatif) excluent la présence d'un objet pouvant être sujet (SN ou Subordonnée, indexé  $\boxed{2}$  dans la grammaire).

## 3 Clitiques non-nominatifs

Cette dernière section est consacrée à illustrer le traitement que la grammaire nous permet des clitiques non-nominatifs dans les constructions verbales.

### 3.1 Dans une construction verbale simple

Les clitiques non-nominatifs, tout comme les nominatifs que l'on vient de voir, sont rattachés au verbe en un Verbe construit. Par exemple, dans un énoncé tel que (166),

(166) Pierre le lui donne

la structure obtenue sera celle présentée en fig. 39 (on a encadré la partie qui nous intéresse).



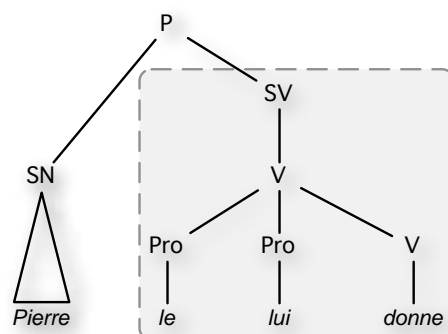


FIG. 39 – Exemple de structure de l'analyse d'un énoncé présentant des compléments de verbe clitiques.

La caractérisation du V *le lui donne* est présentée en fig. 40.

Ceci est différent de l'analyse obtenue pour des compléments de verbe pleins, dont la structure serait celle de la fig. 41, et où l'on voit que le verbe et les compléments de verbe sont constituants immédiats d'un SV et non d'un V.

La différence est plus marquée encore dans l'exemple (167),

(167) Pierre le donne à Marie

dont la structure est présentée en figure 42 : on y voit que le complément clitique est constituant immédiat d'un V construit avec le verbe, et le complément plein est à un niveau supérieur constituant immédiat d'un SV dont le second constituant immédiat est le V construit *le donne*.

Il va sans dire que ceci se combine avec l'analyse des clitiques nominatifs que l'on vient de voir, et que donc pour un énoncé tel que (168),

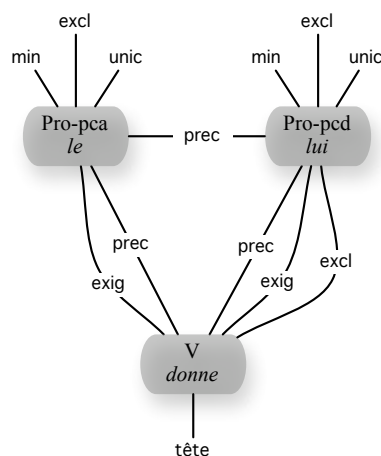
(168) il le donne à Marie

la structure obtenue prendra la forme de la figure 43, où l'on voit que les deux pronoms clitiques sont constituant immédiat du Verbe construit avant que le SPrep complément plein soit rattaché au verbe en un SV (en d'autres termes, ici le sujet est rattaché au verbe en même temps que le complément clitique et avant l'un de ses compléments de verbe).

On vient de voir à quel niveau les clitiques non-nominatifs se rattachent au verbe, par rapport à leur équivalents non-clitiques. Voyons maintenant à quel verbe ils se rattachent dans les divers cas de constructions verbales que l'on a pu voir au début de ce chapitre.

### 3.2 Dans une construction verbale à auxiliaire

Dans le cas d'une construction verbale à auxiliaire (169),



V-cad		
INTR.	ID   NATURE	$\left[ \begin{array}{ll} \text{TYPE} & \boxed{1}.\text{TYPE} \\ \text{CAT} & V \\ \text{SCAT} & \boxed{1}.\text{SCAT} \\ \text{TRAITS} & \boxed{1}.\text{TRAITS} \\ \text{ACCORD} & \boxed{1}.\text{ACCORD} \end{array} \right]$
	SYN	$\left[ \begin{array}{ll} \text{OBJ-DIR} & \left[ \begin{array}{ll} \text{RECTEUR} & \boxed{1} \\ \text{DÉPENDANT} & \boxed{6} \end{array} \right] \\ \text{OBJ-IND} & \left[ \begin{array}{ll} \text{RECTEUR} & \boxed{1} \\ \text{DÉPENDANT} & \boxed{7} \end{array} \right] \end{array} \right]$
Majeur	$\boxed{1}V$	+
Mineur	$\boxed{6}Pro-pca$	+
	$\boxed{7}Pro-pcd$	+
Unicité	$\boxed{6}$	+
	$\boxed{7}$	+
Exigence	$\boxed{1} \Rightarrow \boxed{6}$	+
	$\boxed{1} \Rightarrow \boxed{7}$	+
Exclusion	$lui \not\leq (me \vee te \vee se \vee nous \vee vous \vee lui \vee leur)$	+
	$le \not\leq le$	+
	$\boxed{7} \not\leq \boxed{1}.\text{[TYPE-MODAL } \neg \textit{causatif}]$	+
Précédence	$le < lui$	+
	$\boxed{6} < \boxed{1}$	+
	$\boxed{7} < \boxed{1}$	+

FIG. 40 – Introduction du V *le lui donne* dans (166).

(169) Pierre l'a donné à Marie

les compléments clitiques sont les constituants immédiats du même verbe construit que l'auxiliaire et le verbe principal (fig. 44).

La caractérisation du V *l'a donné* est donnée en figure 45.

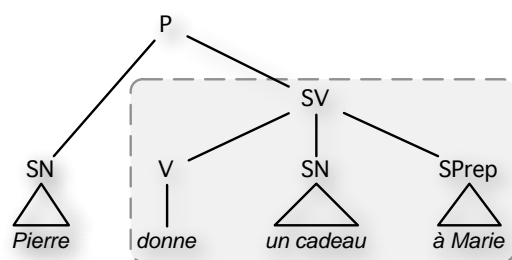


FIG. 41 – Exemple de structure de l'analyse d'un énoncé présentant des compléments de verbe plein.

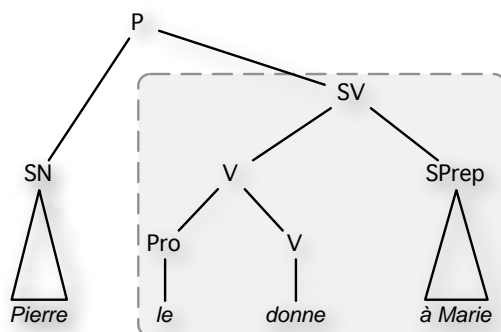


FIG. 42 – Exemple de structure de l'analyse d'un énoncé présentant un complément de verbe clitique et un plein.

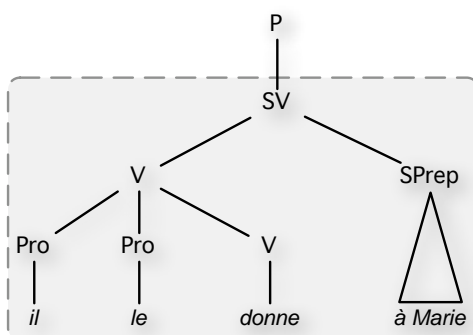


FIG. 43 – Exemple de structure de l'analyse d'un énoncé présentant un sujet et un complément de verbe clitiques.

### 3.3 Dans une construction verbale à modal

#### 3.3.1 Sans “montée” de clitiques

Dans le cas d'une construction verbale à modal (170),

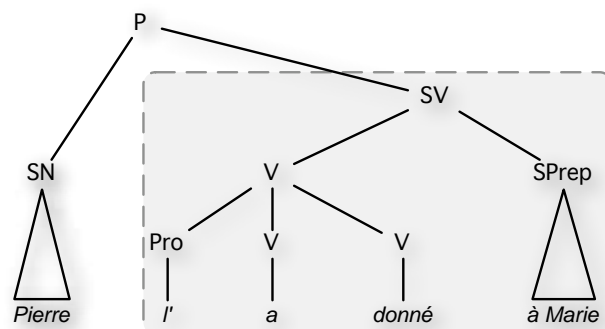


FIG. 44 – Structure de l'analyse d'un énoncé présentant un auxiliaire et un complément de verbe clitique.

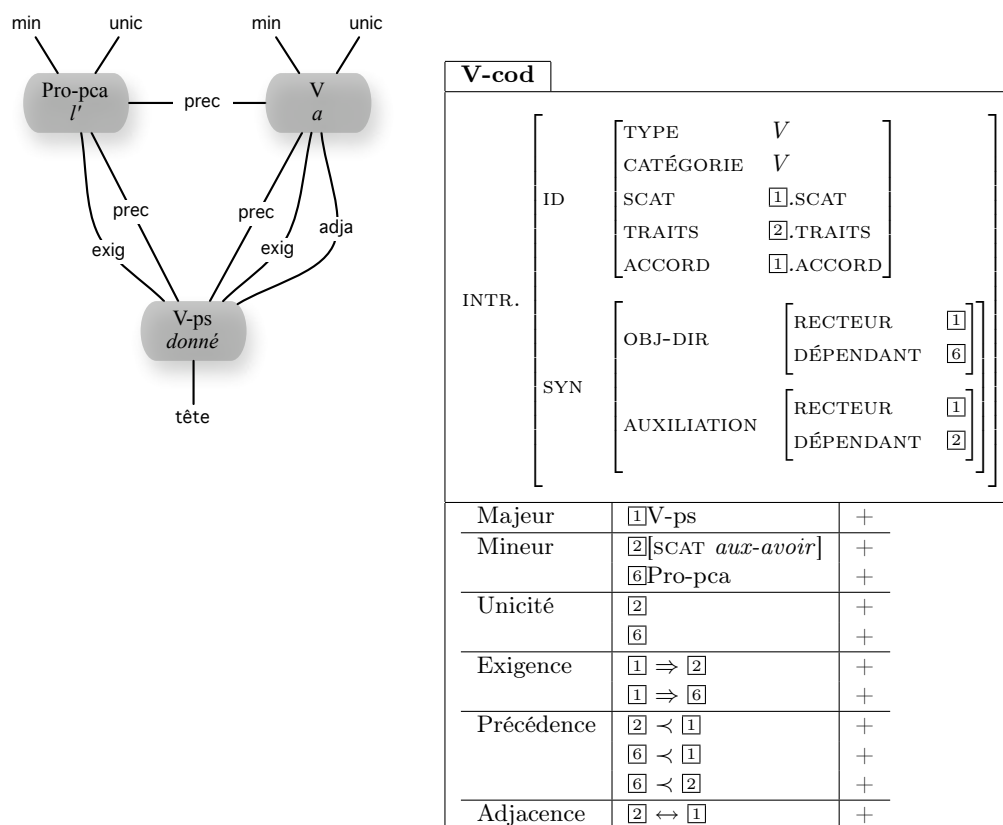


FIG. 45 – Introduction du V l'a donné dans (169).

(170) Pierre peut le donner à Marie

les compléments clitiques sont les constituants immédiats du même verbe construit que l'auxiliaire et le verbe principal (fig. 46).

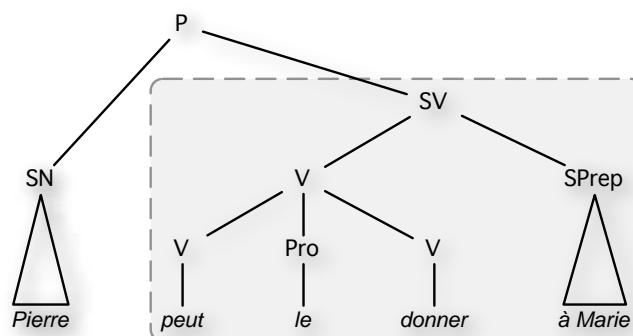


FIG. 46 – Structure de l’analyse d’un énoncé présentant un modal et un complément de verbe clitique.

La caractérisation du V *peut le donner* est assez proche de celle du V *l’a donné*, mis à part que les constituants n’étant pas les mêmes, *l’a donné* constituait un V-cod (“verbe à clitique accusatif complément d’objet direct”) héritant de V-a (“verbe à auxiliaire *avoir*”), tandis que le V-cod *peut le donner* hérite pour sa part du V-im (“verbe infinitif à modal”).

### 3.3.2 Avec “montée” de clitiques

Si le modal porte lui-même un clitique alors un V construit les contenant est d’abord introduit, qui lui-même devient constituant immédiat du V construit dont le verbe principal est la tête. P.ex., dans un énoncé tel que (171),

(171) Pierre le laisse donner à Marie

la structure correspondra à la fig. 47.

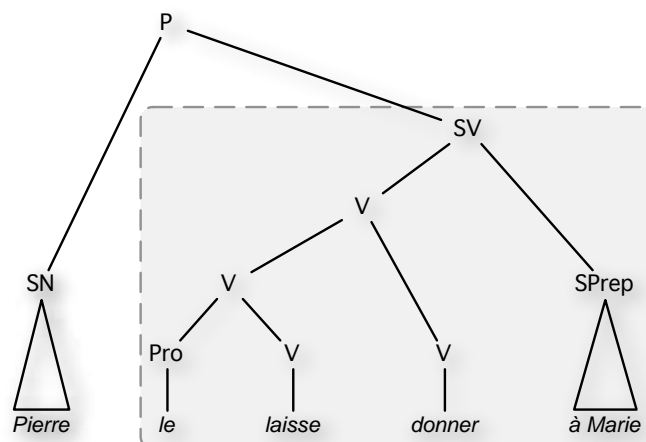


FIG. 47 – Structure de l’analyse d’un énoncé présentant un modal et un complément de verbe clitique.

Chaque verbe (modal et principal) peut donc porter ses propres clitiques. P.ex., pour (172),

(172) Pierre le laisse lui donner

la structure correspondra à la fig. 48.

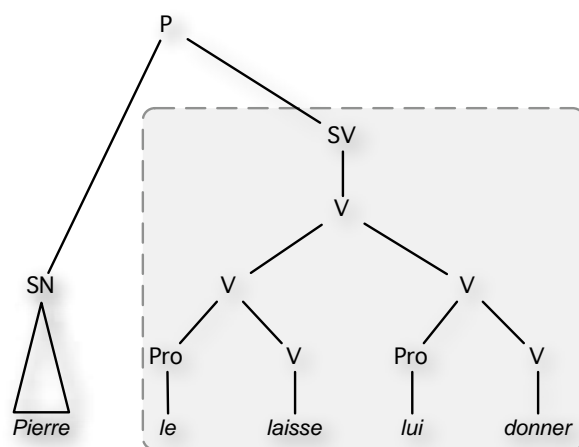


FIG. 48 – Structure de l’analyse d’un énoncé présentant un modal et un complément de verbe clitique.

Cela signifie que suivant à quel verbe sont rattachés les clitiques, la structure sera différente. Ainsi dans une construction causative telle que (173) où tous les clitiques sont “montés” sur le modal,

(173) Pierre le lui fait donner

les deux clitiques portent sur le verbe causatif et non sur le verbe principal (fig. 49).

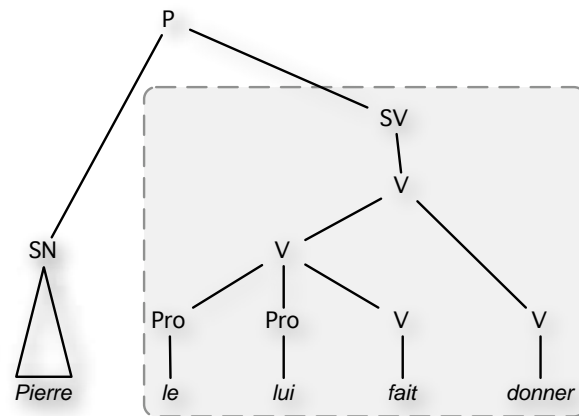


FIG. 49 – Structure de l'analyse d'un énoncé présentant un modal et un complément de verbe clitique.

### **Résumé**

Dans ce chapitre on a vu comment les descriptions proposées pour les constructions verbales nous permettent de traiter les pronoms clitiques.

On a d'abord montré quelles sont les différences d'analyse entre différentes constructions verbales (simple, à auxiliaire, à contrôle, à modal). On a ensuite montré comment est traité le pronom clitique nominatif dans les constructions verbales, par rapport à un sujet non clitique. Enfin, on a donné quelques exemples de traitement des clitiques non nominatifs dans diverses constructions verbales.



# Chapitre 14

## Coordinations

Dans ce chapitre nous présentons des exemples de traitements des entassements paradigmatiques. Nous y traitons deux exemples : tout d'abord une accumulation d'objets de catégories différentes (§1), puis une coordination elliptique (§2).

### 1 Coordination de catégories différentes

Comme on l'a noté dans le chapitre 11 qui présentait les entassements paradigmatiques, ce qui rend possible la coordination de plusieurs objets n'est pas leur identité catégorielle, mais leur identité fonctionnelle. Autrement dit, des objets de catégories différentes peuvent être coordonnés dès lors qu'ils sont capables de porter la même fonction. Pour illustrer le fonctionnement de notre grammaire sur ce type de phénomène, nous décrivons ici l'analyse de l'énoncé (174).

(174)

Pierre est	président
et	fier de l'être.

Au niveau d'analyse qui nous intéresse pour illustrer le traitement de la coordination, nous considérons que les objets à considérer sont les suivants :

- *Pierre* est un SN non déterminé (SN-nd) (fig. 50).
- *est* est un verbe principal conjugué à la 3<sup>e</sup> personne du singulier de l'indicatif présent (fig. 51).
- *président* est un SN non déterminé (SN-nd) masculin singulier (fig. 52).
- *et* est un Coordonnant (fig. 53).
- *fier de l'être* est un Adjectif construit (Adj) masculin singulier (fig. 54).

SN-nd													
INTR.	<table> <tr> <td>ID   NATURE</td><td> <table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table> </td></tr> </table>	ID   NATURE	<table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>N</i>	CAT	<i>SN</i>	TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>
ID   NATURE	<table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>N</i>	CAT	<i>SN</i>	TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>		
TYPE	<i>N</i>												
CAT	<i>SN</i>												
TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>								
GENRE	<i>masculin</i>												
NOMBRE	<i>singulier</i>												

FIG. 50 – Le SN *Pierre* dans (174).

V-c																			
INTR.	<table> <tr> <td>ID   NATURE</td><td> <table> <tr> <td>TYPE</td><td><i>V</i></td></tr> <tr> <td>CAT</td><td><i>V</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table> </td></tr> <tr> <td>ACCORD</td><td> <table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table> </td></tr> </table>	ID   NATURE	<table> <tr> <td>TYPE</td><td><i>V</i></td></tr> <tr> <td>CAT</td><td><i>V</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table> </td></tr> <tr> <td>ACCORD</td><td> <table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>V</i>	CAT	<i>V</i>	TRAITS	<table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table>	MODE	<i>indicatif</i>	TEMPS	<i>présent</i>	ACCORD	<table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	PERSONNE	<i>3</i>	NOMBRE	<i>singulier</i>
ID   NATURE	<table> <tr> <td>TYPE</td><td><i>V</i></td></tr> <tr> <td>CAT</td><td><i>V</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table> </td></tr> <tr> <td>ACCORD</td><td> <table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>V</i>	CAT	<i>V</i>	TRAITS	<table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table>	MODE	<i>indicatif</i>	TEMPS	<i>présent</i>	ACCORD	<table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	PERSONNE	<i>3</i>	NOMBRE	<i>singulier</i>		
TYPE	<i>V</i>																		
CAT	<i>V</i>																		
TRAITS	<table> <tr> <td>MODE</td><td><i>indicatif</i></td></tr> <tr> <td>TEMPS</td><td><i>présent</i></td></tr> </table>	MODE	<i>indicatif</i>	TEMPS	<i>présent</i>														
MODE	<i>indicatif</i>																		
TEMPS	<i>présent</i>																		
ACCORD	<table> <tr> <td>PERSONNE</td><td><i>3</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	PERSONNE	<i>3</i>	NOMBRE	<i>singulier</i>														
PERSONNE	<i>3</i>																		
NOMBRE	<i>singulier</i>																		

FIG. 51 – Le Verbe *est* dans (174).

SN-nd													
INTR.	<table> <tr> <td>ID   NATURE</td><td> <table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table> </td></tr> </table>	ID   NATURE	<table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>N</i>	CAT	<i>SN</i>	TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>
ID   NATURE	<table> <tr> <td>TYPE</td><td><i>N</i></td></tr> <tr> <td>CAT</td><td><i>SN</i></td></tr> <tr> <td>TRAITS</td><td> <table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table> </td></tr> </table>	TYPE	<i>N</i>	CAT	<i>SN</i>	TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>		
TYPE	<i>N</i>												
CAT	<i>SN</i>												
TRAITS	<table> <tr> <td>GENRE</td><td><i>masculin</i></td></tr> <tr> <td>NOMBRE</td><td><i>singulier</i></td></tr> </table>	GENRE	<i>masculin</i>	NOMBRE	<i>singulier</i>								
GENRE	<i>masculin</i>												
NOMBRE	<i>singulier</i>												

FIG. 52 – Le SN *président* dans (174).

Coord							
INTR.	<table> <tr> <td>ID   NATURE</td><td> <table> <tr> <td>TYPE</td><td><i>Coord</i></td></tr> <tr> <td>CAT</td><td><i>Coord</i></td></tr> </table> </td></tr> </table>	ID   NATURE	<table> <tr> <td>TYPE</td><td><i>Coord</i></td></tr> <tr> <td>CAT</td><td><i>Coord</i></td></tr> </table>	TYPE	<i>Coord</i>	CAT	<i>Coord</i>
ID   NATURE	<table> <tr> <td>TYPE</td><td><i>Coord</i></td></tr> <tr> <td>CAT</td><td><i>Coord</i></td></tr> </table>	TYPE	<i>Coord</i>	CAT	<i>Coord</i>		
TYPE	<i>Coord</i>						
CAT	<i>Coord</i>						

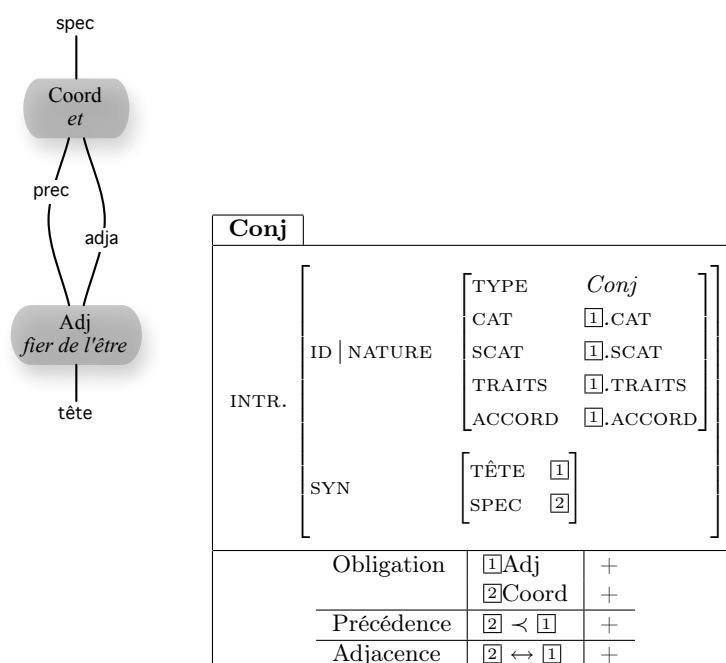
FIG. 53 – Le Coordonnant *et* dans (174).

## 1.1 Introduction du Conjoint *et fier de l'être*

Au niveau d'analyse auquel on se trouve est caractérisé un Conjoint (Conj, décrit au chapitre 11, §2.1) dont les deux constituants immédiats sont le Coord *et* et l'Adj *fier de l'être* (fig. 55).

Notons que comme on l'avait défini, la construction ainsi introduite est de type Conjoint,

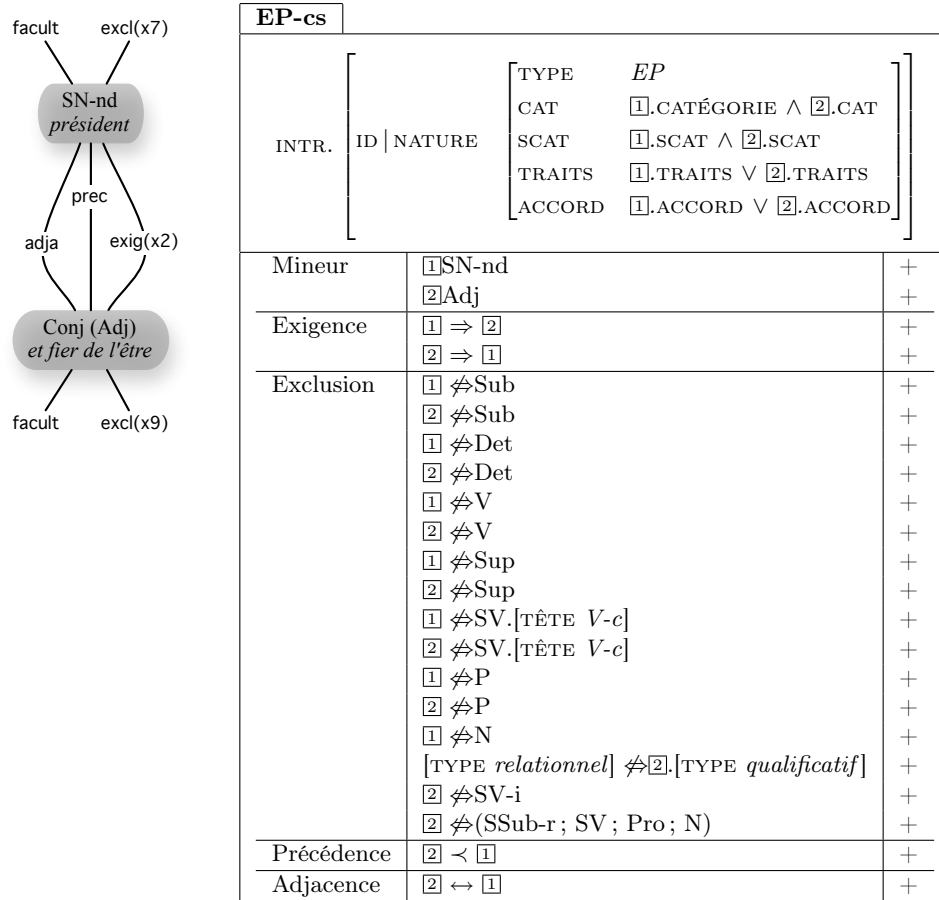
Adj	
INTR.	ID   NATURE
	TYPE
	CAT
	SCAT
	ACCORD
	GENRE
	NOMBRE
	TÊTE
	OBJ-IND
	RECTEUR
	DÉPENDANT

FIG. 54 – L'Adjectif construit *fier de l'être* dans (174).FIG. 55 – Le Conjoint *et fier de l'être* dans (174).

mais porte la catégorie de sa Tête, en l'occurrence Adjectif. Elle porte également les traits et accords de celui-ci.

## 1.2 Introduction de la Coordination *président et fier de l'être*

Au niveau d'analyse suivant, le Conjoint *et fier de l'être* que l'on vient d'introduire est mis en relation avec le SN-nd *président* (fig. 56), formant à eux deux une Coordination syndétique (EP-cs, décrite au chapitre 11, §2.3.3).


 FIG. 56 – La Coordination syndétique *président et fier de l'être* dans (174).

Pour simplifier le réseau de relations dans le graphe, on n'a représenté qu'une seule relation de chaque type pour chaque objet, en notant à côté du type le nombre de propriétés qui sont en fait évaluées (et satisfaites) ; p.ex., la mention *excl(x7)* pointant sur le SN-nd signifie que cet objet satisfait 7 propriétés d'Exclusion. Le détail des propriétés évaluées est donné dans la construction à côté.

Conformément à sa définition dans la grammaire, la construction que l'on vient d'introduire est de type Entassement Paradigmatique (EP), et porte les deux catégories de ses constituants :  $SN \wedge Adj$ , ainsi que les scat, traits, accords, et informations morphologiques des objets qui la constituent.

### 1.3 Introduction du SV *est président et fier de l'être*

A l'étape suivante est introduit le SV à attribut du sujet (SV-1as, décrit au chapitre 7, §2.2.2) *est président et fier de l'être* (fig. 57).

Son attribut étant constitué d'un EP portant deux catégories ( $SN \wedge Adj$ ), les propriétés

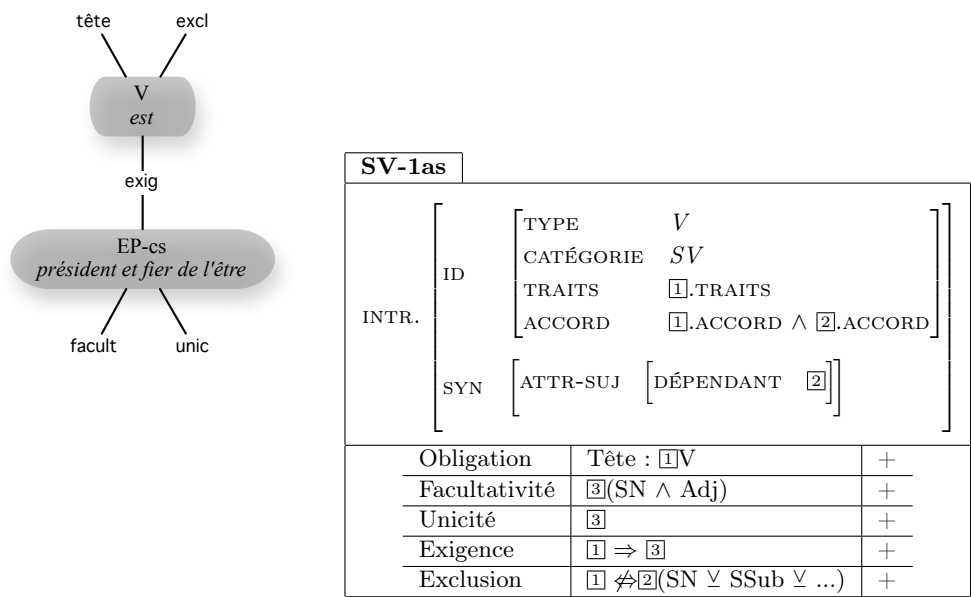


FIG. 57 – Le SV attributif *est président et fier de l'être* dans (174).

évaluées pour cette construction doivent correspondre à ces deux catégories la fois pour être satisfaites. Comme on le voit dans la description de la construction, une dépendance syntaxique ATTRibut du SUJet est introduite, dont on donne le dépendant (mais pas le recteur puisqu'on n'a pas encore de sujet) : il s'agit de l'entassement paradigmatique *président et fier de l'être* considéré dans sa globalité (ce qui nous permet, à juste titre, de faire partager une unique fonction à la totalité de la construction, et non pas deux fonctions identiques à ses deux constituants).

Enfin, la dernière étape de l'analyse permet de caractériser et d'introduire la Proposition exocentrique (P-x) *Pierre est président et fier de l'être*, dont la caractérisation est tout à fait standard.

Finalement, la structure (syntagmatique) que nous construisons peut être représentée sous la forme arborescente de la figure (58).

2 Coordination elliptique

Un autre phénomène qui peut engendrer des problèmes d'analyse syntaxique est celui de la coordination elliptique. Nous présentons ici l'étude détaillée de l'analyse de l'énoncé (175) afin d'illustrer comment ce problème est résolu par les propositions faites dans notre grammaire.

(175)	Pierre	aime	le cinéma
	et Marie		le théâtre

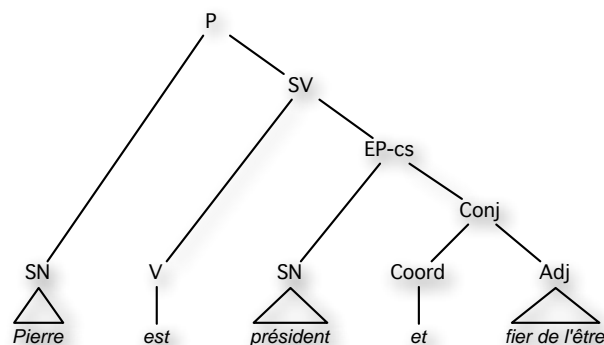


FIG. 58 – Une structure arborescente représentant notre analyse de (174).

Comme pour l'exemple précédent, nous ne commençons pas ici l'analyse à partir des unités lexicales mais considérons que les objets suivants ont déjà été introduits :

- Les SN non déterminés (SN-nd) *Pierre* et *Marie*, qui ont la même caractérisation que celle donnée en figure 50 (p. 294), à ceci près que *Marie* varie en forme et en genre. Nous ne reproduisons donc pas ces constructions ici.
- Les SN déterminés (SN-d) *le cinéma* et *le théâtre*. Les deux SN-d étant identiques à leur ancrage et leur forme près, nous n'en représentons qu'un seul (fig. 59).

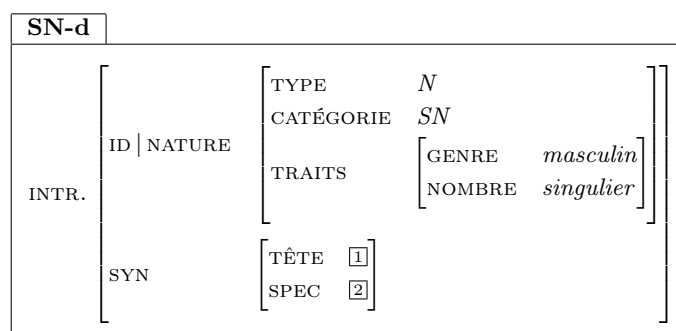


FIG. 59 – Le SN-d *le cinéma* dans (175).

- Le Verbe *aime*. Lui même a la même caractérisation que le V *est* en figure 51 (p. 294), à la forme près ; nous ne le reproduisons donc pas ici.
- Le Coordonnant *et*, équivalent à celui présenté en fig. 53 p. 294.

Selon notre proposition (chap. 11), cet énoncé sera analysé comme présentant non pas une coordination, mais deux, en l'occurrence d'une part *Pierre et Marie* et d'autre part *le cinéma le théâtre* (et non pas *le cinéma et le théâtre* : on considère que le Coordonnant *et* porte uniquement sur l'objet qu'il précède, en l'occurrence le SN *Marie*). Cela implique que la seule construction Conjoint qui est introduite correspond à *et Marie* ; elle a un réseau de relations

semblable au Conjoint *et fier de l'être* que l'on a montré en fig. 55.

## 2.1 Introduction de la Coordination syndétique *Pierre et Marie*

Une fois le Conjoint introduit on peut caractériser la Coordination syndétique *Pierre et Marie* (fig. 60) <sup>278</sup>.

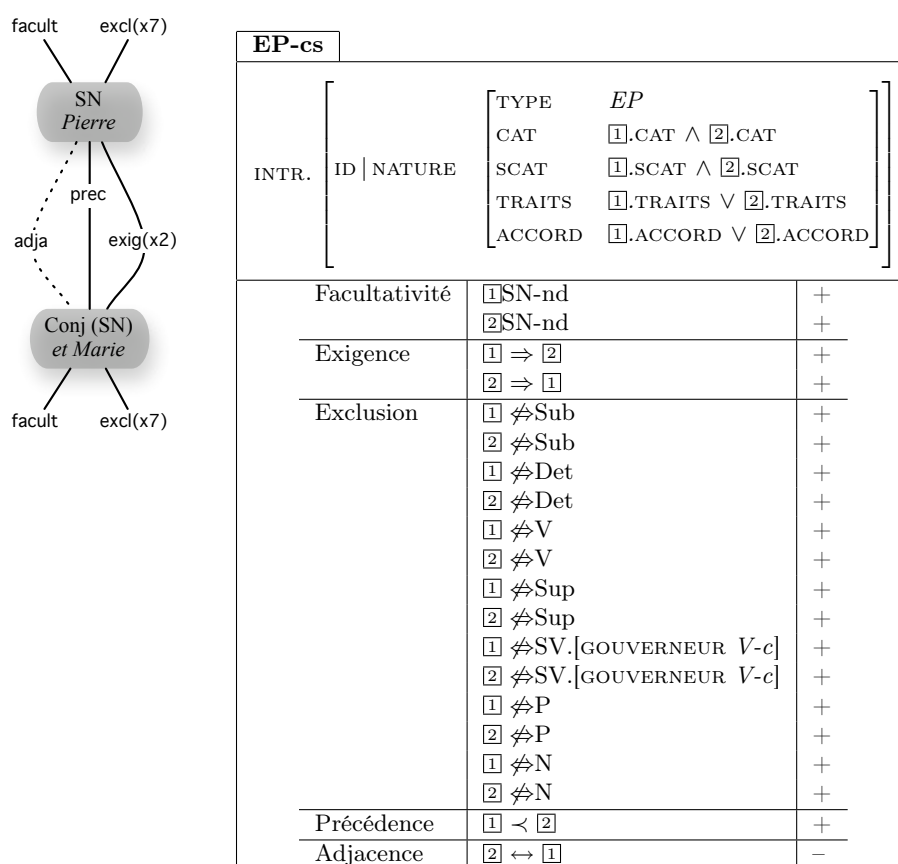


FIG. 60 – La Coordination syndétique *Pierre et Marie* dans (175).

On voit dans la caractérisation de cette coordination que l'une des propriétés évaluées n'est pas satisfaite : les deux constituants ne sont pas adjacents. Cela étant, la non-satisfaction de cette contrainte ne fait baisser le degré de satisfaction local que de 5%, étant donné nombre d'autres propriétés satisfaites par ailleurs (1 propriété non satisfaite sur 20 évaluées = degré de satisfaction local de 95%).

<sup>278</sup>Ici comme précédemment on a simplifié le graphe en n'y représentant qu'un seul arc pour chaque type de propriété portant sur un objet donné, alors qu'il devrait y en avoir autant que de propriétés évaluées.

## 2.2 Introduction de la Coordination asyndétique *le cinéma le théâtre*

Parallèlement à cette première coordination est introduite une seconde coordination, asyndétique (EP-ca) cette fois : *le cinéma le théâtre* (fig. 61).

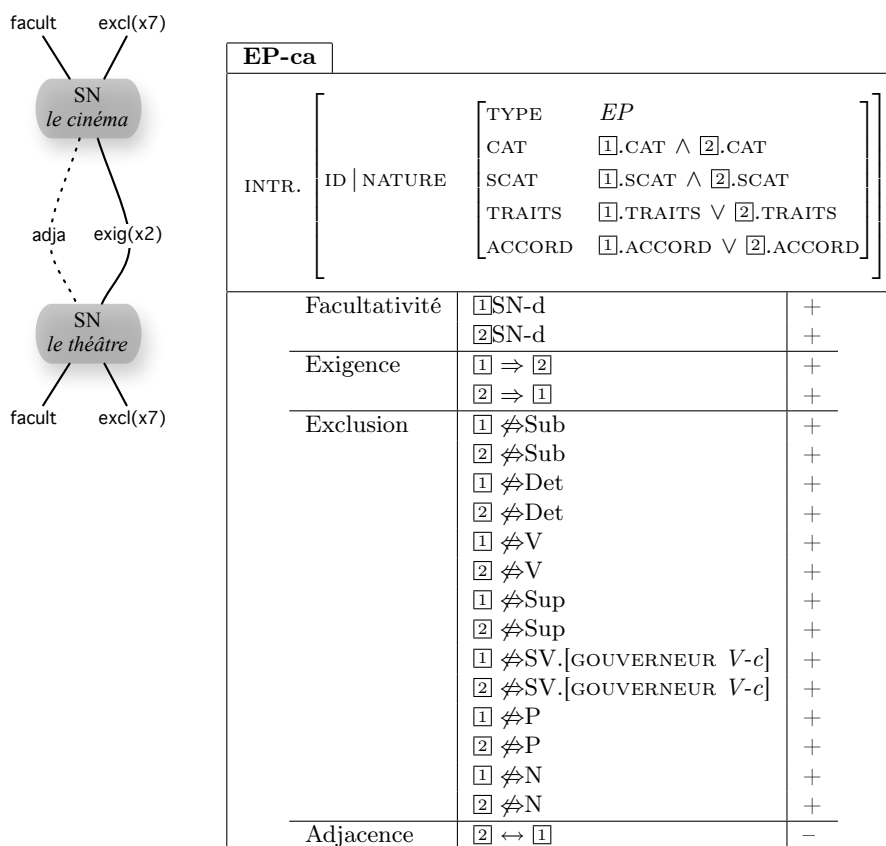


FIG. 61 – La Coordination asyndétique *le cinéma le théâtre* dans (175).

Comme la précédente, cette coordination n'est pas satisfaite à 100% puisque sa contrainte d'Adjacence est violée. Ce qui la différencie de la précédente est qu'elle ne comporte pas de contrainte de Précédence.

## 2.3 Introduction du Syntagme Verbal *aime le cinéma le théâtre*

Une fois la Coordination précédente introduite, le SV transitif direct (SV-1od) *aime le cinéma le théâtre* peut être construit (fig. 62).

On peut remarquer qu'il s'agit exactement de la même caractérisation que celle du SV-1od *mange une pomme* que l'on avait décrit au chapitre précédent (fig. 27 p. 276), ce qui est normal puisque c'est le but de cette description de faire en sorte que l'analyse d'un entassement paradigmatique reste la même que celle d'un objet équivalent non entassé.



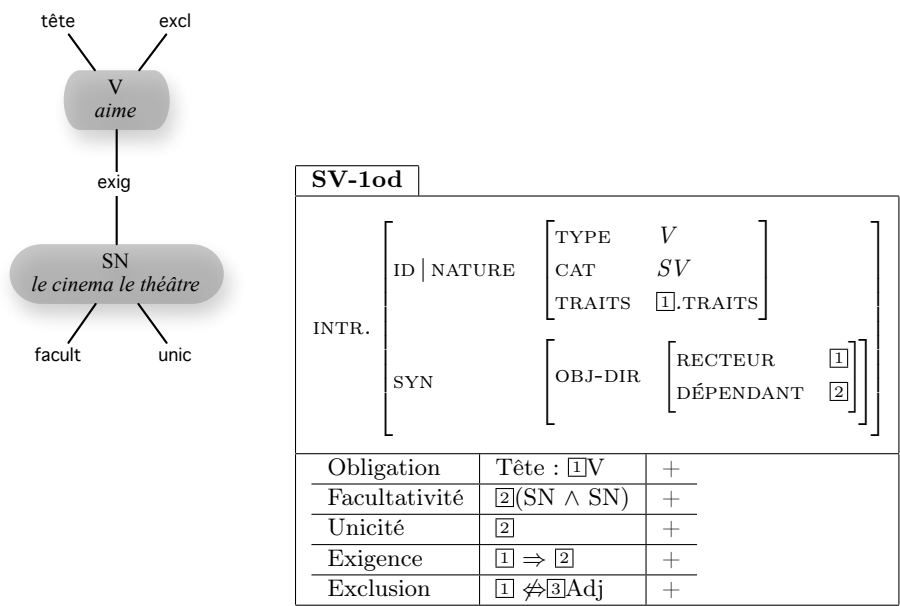


FIG. 62 – Le SV transitif direct *aime le cinéma le théâtre* dans (175).

2.4 Introduction de la Proposition

Enfin, le SV précédent ayant été introduit, le réseau de relations créé permet de caractériser la Proposition exocentrique (P-x) *Pierre aime le cinéma et Marie le théâtre* (ou, dans son ordre de constituants *Pierre et Marie aime le cinéma le théâtre*) (fig. 63).

**Contraintes d’Accord.** — Dans le cas d’une coordination sans ellipse (176a), les éventuels accords se font localement au sein des constituants coordonnés (*vois* s’accorde avec *je* et *voit* avec *Pierre*). En revanche, dans le cas de coordinations elliptiques (176b), les constituants de la coordination (*je* et *Pierre* uniquement) ne sont pas adjacents, et l’accord s’opère de manière linéaire : *vois* s’accorde avec son sujet, i.e. l’élément qui le précède, mais pas avec l’EP complet (*je Pierre*).

(176)	a.	<div>je te vois et Pierre te voit aussi</div>
	b.	<div>je            te vois et Pierre            aussi</div>

Ceci est permis par la façon dont on affecte les traits d’ACCORD dans l’EP, la valeur est constituée de la liste des valeurs des constituants, mais qui ne sont pas joints par des “et” comme celles de CATÉGORIE mais par des “ou inclusifs”, ce qui signifie que la propriété d’accord entre les objets sera satisfaite dès lors que l’un des constituants de l’EP porte la valeur adéquate. Ainsi dans (176b) l’EP *je Pierre* est de [PERSONNE 1; 3], ce qui lui permet de satisfaire l’Accord entre lui et *vois* qui est de [PERSONNE 1].

P-x		
INTR.	FORME	$\left[ \begin{array}{l} \text{ORTH} \text{ Pierre aime le cinéma} \\ \text{et Marie le théâtre} \end{array} \right]$
	ID   NATURE	$\left[ \begin{array}{l} \text{TYPE} \text{ [1].TYPE} \\ \text{CAT} \text{ P} \end{array} \right]$
	SYN	$\left[ \begin{array}{l} \text{SUJET} \left[ \begin{array}{l} \text{RECTEUR} \text{ [2]} \\ \text{DÉPENDANT} \text{ [1]} \end{array} \right] \end{array} \right]$
Majeur	Recteur([1]SV-td)	+
Mineur	[2]SN	+
Unicité	[2]	+
Adjacence	[1] $\leftrightarrow$ [2]	+
Exigence	[1] $\Rightarrow$ [2]	+
Exclusion	SV.[GOUVERNEUR V-cn] $\not\leftrightarrow$ [2]	+
	SV.[MODE impératif] $\not\leftrightarrow$ [2]	+
	P $\not\leftrightarrow$ [2]	+
Accord	[2].NOMBRE $\prec$ [1].NOMBRE	+
	[2].PERSONNE $\prec$ [1].PERSONNE	+
DE = 87,5%, DSL = 100%, DSP = 100%		

FIG. 63 – La Proposition *Pierre aime le cinéma et Marie le théâtre*.

Finalement, la structure que nous construisons peut être représentée sous la forme arborescente de la figure (64).

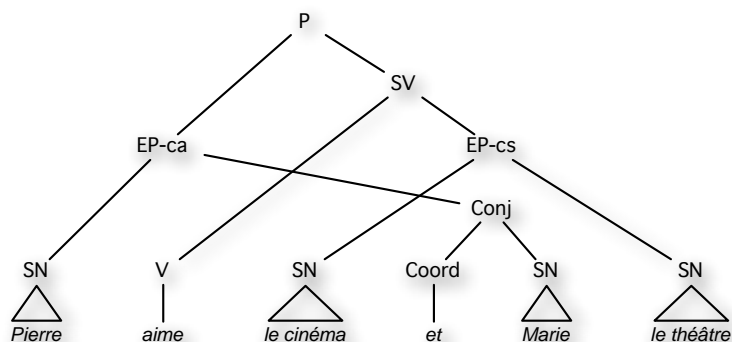


FIG. 64 – Une structure arborescente syntagmatique représentant notre analyse de (175).

C'est là que l'on mesure ce que l'ont avait eu l'occasion de dire précédemment, quand on disait que les structure que l'on construisait n'étaient pas à proprement parler des structures syntagmatiques telles que les arbres "à la Chomsky" mais plutôt des structures strictement syntaxiques, puisque le représentation que l'on devrait en donner pour qu'elle soit plus lisible et plus cohérente devrait être la suivante (fig. 65).

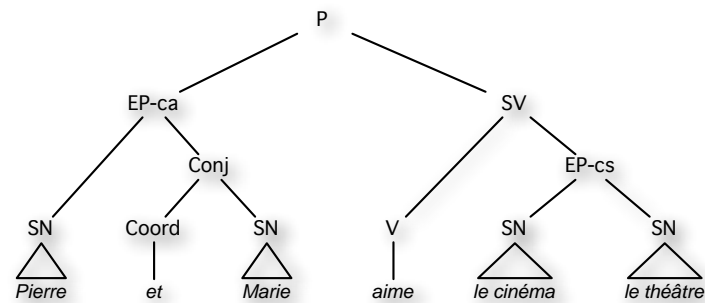


FIG. 65 – Une structure arborescente syntaxique représentant notre analyse de (175).

Dans cette dernière, comme on l’avait dit, les branches de l’arbre ne sont pas ordonnées en fonction de l’ordre des mots dans l’énoncé : l’information sur leur position existe, et elle est située à l’intérieur des constructions auxquelles les noeuds font référence, mais elle n’est pas explicitée ici (pas plus que ne le sont les informations d’accord, d’exigence, ou toute autre relation entre objets qui font l’objet de propriétés dans notre modèle).

Une version “réduite” de cette même structure pourrait être représentée comme en fig. 66, où l’on a rassemblé les deux coordinations en un seul noeud réduit, et où l’on a représenté les constituants immédiats verticalement à la manière d’une mise en grille. Cela fait bien ressortir la simplicité de la structure, et le partage de fonctions des objets entassés.

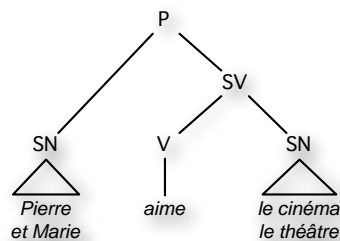


FIG. 66 – Une structure arborescente syntaxique à deux dimensions représentant notre analyse de (175).

### Résumé

Dans ce chapitre nous avons présenté deux exemples de traitement d'entassements paradigmatiques : d'abord la coordination d'objets de natures différentes, puis une coordination elliptique.

A travers le premier exemple, on a montré comment il est possible de vérifier l'unicité fonctionnelle des constituants d'un entassement paradigmatique, alors même que ce partage de fonction n'est pas spécifié dans la construction elle-même, puisqu'il s'agit d'une information de syntaxe externe.

Puis à travers le second exemple, on a montré comment l'on traite les coordinations elliptiques dans sans recours à des objets vides, et comment notre analyse s'intègre à l'analyse syntagmatique.

# Conclusion

Dans cette dernière partie, l'on a souhaité présenter l'application de notre grammaire à l'analyse de productions linguistiques. A cette fin on a présenté le déroulement étape par étape des analyses de plusieurs énoncés.

On a tout d'abord montré comment se déroule l'analyse en GP, d'une manière générale, afin d'illustrer le fonctionnement de la grammaire (chapitre 12). Ceci nous a permis de présenter le fonctionnement pratique de GP, la façon dont les structures sont construites en fonction des relations établies entre les objets présents.

On s'est ensuite intéressés plus précisément à l'analyse de deux phénomènes qui ont retenu notre attention lors du développement de notre grammaire. Le premier est typiquement syntaxique et concerne l'analyse des pronoms clitiques dans les constructions verbales (chapitre 13), le second est à l'interface de plusieurs domaines d'analyse linguistique, bien que nécessaire à l'analyse syntaxique puisque sans sa prise en compte l'on ne peut proposer de structure syntaxique cohérente : il s'agit du traitement des entassements paradigmatiques (chapitre 14).

Ces exemples montrent que les analyses rendues possibles par la combinaison du formalisme GP présenté en première partie et de la grammaire proposée en deuxième partie, permettent un traitement élégant, c'est-à-dire homogène et cohérent de phénomènes divers, et dont la complexité rend habituellement le traitement difficile, et nécessitant le plus souvent un recours à des artefacts formels spécifiques, ou tout au moins nécessitant de modifier le modèle pour qu'il puisse les prendre en compte. Ici, les exemples de traitement que l'on développe présentent un avantage notable : ils ne représentent qu'une application normale de la grammaire, sans heuristique, sans représentation *ad hoc*, et n'alourdissent donc ni la grammaire ni son application. Ils sont traités avec la même efficacité et la même simplicité que des phénomènes plus simples et plus évidents.

Il s'agit précisément de l'objectif de la grammaire que l'on propose : permettre de traiter d'une manière simple et homogène un ensemble de phénomènes plus ou moins spécifiques, plus ou moins complexes, plus ou moins précis, sans demander de modification du reste de la grammaire ou du formalisme utilisé.



# Conclusion générale

Dans cette thèse nous avons proposé des éléments de description grammaticale du français, après avoir présenté leurs cadres théorique et formel généraux. Nous avons par la suite illustré nos propositions par des exemples de leur application à l'analyse de quelques phénomènes linguistiques.

## 1 Retour sur le développement

### 1.1 Pour une théorie descriptive et formelle de la langue

#### 1.1.1 Théorie

Nous avons commencé par dresser un bilan des approches théoriques existantes afin de mettre nos positions en perspective avec les autres travaux contemporains. Plutôt que de nous concentrer sur les quelques théories dominantes nous avons préféré proposer un panorama plus large, évoquant aussi bien celles-ci que d'autres propositions moins répandues mais qui sont souvent motivées, justement, par leur divorce plus ou moins marqué d'avec les approches les plus fréquentes ; ceci afin de souligner la considérable diversité des positions actuelles dans le domaine, et les relations pouvant néanmoins s'établir entre celles-ci en dépit de leurs apparentes divergences.

A travers ceci nous avons fait émerger quatre grands axes sur lesquels on s'est basée pour définir notre démarche : le *non-générativisme* de l'approche (on conçoit la grammaire comme une ressource descriptive et non comme un mécanisme de dérivation), la *non-présomption* de la grammaire (on ne présuppose pas de la structure dans laquelle un objet est impliqué), la *non-modularité* du modèle (on ne constitue qu'une seule grammaire pour représenter toute la structuration linguistique) et la *multi-dimensionnalité* des domaines (la structure que l'on construit tient compte simultanément de plusieurs axes d'organisation).

### 1.1.2 Formalisme

Une fois présentées nos positions théoriques mises en contraste avec les courants linguistiques coexistants, nous avons donné une présentation de l'architecture du formalisme utilisé pour représenter notre grammaire : les *Grammaires de Propriétés* ([Blache, 2005]). Nous y avons présenté les grands principes de sa forme et de son fonctionnement, qui contient dans sa présente expression à la fois ce qui préexistait à nos travaux ainsi que les compléments de description que nous lui avons apportés à la lumière de nos recherches personnelles.

Toutes les parties qui ont été ajoutées à la description du formalisme ne constituent pas des modifications de sa forme d'origine mais au contraire des précisions portant tant sur son expression que son contenu qui, nous l'espérons, assoient GP en tant que formalisme pertinent et original pour la représentation de grammaires, et suffisamment flexible pour limiter autant que faire se peut les contraintes qu'il pourrait imposer aux théories linguistiques l'utilisant.

### 1.1.3 Modèle(s)

Afin d'illustrer ce dernier point, nous avons présenté à la suite de cela, tout d'abord la représentation formelle constituant le modèle de notre grammaire, puis un certain nombre de modélisations de théories différentes de la nôtre.

Dans la présentation de notre modèle nous avons donné la structuration des informations des objets que nous manipulons (*informations intrinsèques* en GP), puis nous avons présenté le jeu de contraintes relationnelles que nous avons utilisé dans notre grammaire (*informations extrinsèques*) ainsi que leur fonctionnement (*spécification des propriétés*). Parce que nous avons fait le choix de ne pas utiliser un certain nombre des propriétés qui avaient été définies dans les grammaires GP précédentes, nous avons comparé nos propriétés avec celles pré-existantes en expliquant les raisons de nos modifications.

Ensuite nous avons développé la modélisation en GP de trois extraits de théories qui présentaient l'avantage d'être éloignées de la nôtre sur un certain nombre de points : nous avons tout d'abord illustré la formalisation d'une grammaire modulaire à travers l'exemple de la *Grammaire d'Unification Sens-Texte* ([Kahane & Lareau, 2005]) ; puis celui d'une grammaire catégorielle et procédurale à travers l'exemple de *Dynamic Syntax* ([Kempson *et al.*, 2001]) ; et enfin, celui d'une *Métagrammaire* à travers l'exemple de la grammaire TAG de [Crabbé, 2005b].

## 1.2 Éléments de grammaire du français

Dans la seconde partie nous avons développé un certain nombre de descriptions de constructions syntaxiques du français conformément au modèle présenté précédemment.



### 1.2.1 Syntaxe

Nous nous sommes volontairement pour ce travail limitée à la description de phénomènes syntaxiques, en usant uniquement d'informations syntaxiques et lexicales. Nous avons fait ce choix de manière à pouvoir proposer un ensemble constituant un système au sens Saussurien, c'est-à-dire dont les objets constitutifs s'opposent et se complètent. Nous avons également espéré illustrer par là même que, même pour constituer des descriptions syntaxiques qui soient pertinentes, il est très souvent nécessaire de recourir à des informations provenant de domaines différents. Cela conforte notre position selon laquelle l'information linguistique est *éparse et inconstante*, et qu'il est donc primordial de pouvoir accéder au plus d'informations possible pour pouvoir proposer une analyse des productions, quelle que soit la structure recherchée. En l'occurrence, les informations proprement syntaxiques ne suffisent pas à proposer une analyse syntaxique suffisamment approfondie : il serait nécessaire de leur adjoindre des informations sémantiques notamment, et prosodiques (ou équivalent pour l'écrit).

### 1.2.2 Extraits

Ces éléments décrits n'ont pas prétention à constituer une grammaire : ils ne forment qu'une partie de celle-ci qui n'est que le début d'un développement plus large et plus complet.

Premièrement, parce que les descriptions que l'on propose ne constituent que des constructions syntaxiques alors que notre grammaire a vocation à constituer une ressource non-modulaire, c'est-à-dire dont les objets constitutifs ne sont pas nécessairement des constructions syntaxiques, mais des *constructions* au sens de CxG c'est-à-dire des ensembles forme-sens ([Goldberg, 1995]) dont l'existence ne repose ni uniquement ni primitivement sur la syntaxe.

Deuxièmement, parce que même la partie de grammaire que l'on a décrite ici peut elle-même être largement affinée en y ajoutant des descriptions de plus en plus précises de phénomènes de plus en plus contraints. On n'a ici présenté, dans la plupart des cas, que les parties de la hiérarchie grammaticale les plus générales, sans entrer dans la description de phénomènes plus fins. Là aussi il s'agit d'une limitation volontaire *a priori* dont la raison d'être principale est que l'un des enjeux à court terme de ce travail était de proposer une première grammaire qui soit couvrante (quitte à rester grossière), et qui exploite réellement toutes les capacités de GP.

Nous avons donc dans ces extraits présenté les descriptions des grands types des constructions syntaxiques de notre grammaire, en mettant l'accent sur ce qui en fait l'originalité principale, i.e. l'introduction, parallèlement aux syntagmes, de constructions paradigmatiques qui nous permettent de traiter non seulement des disfluences mais également des coordinations.

## 1.3 Exemples d'analyses

Enfin, nous avons proposé quelques exemples des analyses permises par notre grammaire et qui, finalement, en constituent l'intérêt principal. En effet, en plus d'autoriser une certaine flexibilité grâce au mécanisme de satisfaction de contraintes sur lequel repose intégralement

l'analyse en GP, la combinaison de ceci avec nos positions théoriques (et plus particulièrement celle de multi-dimensionnalité) permet de proposer une analyse simple et homogène avec le reste de la grammaire, pour des phénomènes dont la considérable variabilité aussi bien descriptions que des traitements reflète l'absence de consensus quant à leur statut et à la place qui leur revient.

### **1.3.1 Flexibilité**

L'une des caractéristiques du formalisme que nous avons choisi d'utiliser est le fait que celui-ci est entièrement basé sur les contraintes. Cette spécificité de GP rend possible d'envisager l'analyse comme quelque chose qui va au-delà du strict respect de la *compétence* linguistique, pour permettre d'extraire des informations des productions réelles des locuteurs, y compris quand ces informations sont partielles, hétérogènes, ambiguës voire parfois en partie erronées. Et il ne s'agit pas ici d'en faire un mécanisme de compensation que l'on surajouterait à l'analyse, mais bien d'un processus qui soit partie intégrante de l'analyse elle-même, et qui sache tenir compte dès le départ des aléas de la production linguistique humaine.

### **1.3.2 Proposition(s) alternative(s)**

Sans aucunement prétendre que les propositions que l'on fait ici constitueraient une réponse qui soit unique ou ultime à des problèmes syntaxiques récurrents, ce que nous souhaitons mettre en avant par les différences de notre grammaire avec les autres modèles est que le nombre de pistes envisageables pour répondre aux questions syntaxiques problématiques est plus large que ce que l'on envisage habituellement dans le domaine formel. De fait, un nombre conséquent de recherches descriptives, sur corpus pour la plupart, et non automatisées ne sont bien souvent pas formalisées, et ce pour de multiples raisons.

Nous avons ici tenté un rapprochement entre certaines de ces descriptions syntaxiques jusque là non formalisées, et un modèle de grammaire formelle qui soit comparable aux modèles coexistants, et qui vient en complément d'autres démarches du même ordre qui peuvent être faites par ailleurs. Nous espérons par cela élargir le champ des recherches en syntaxe formelle en apportant des propositions alternatives de solutions, tout en faisant passer le test de la validation par une expression régulière aux propositions théoriques que nous avons retenues.

## **2 Retour sur les objectifs**

En introduction nous avons attribué trois objectifs à ce travail. Nous revenons sur chacun d'eux à la lumière de ce que nous avons développé au fil de cette thèse.

## 2.1 Elaboration d'une théorie alternative

Nous avons proposé un ensemble de positions théoriques qui fait de notre approche une proposition originale au sein des recherches en syntaxe formelle. Par cela et à travers notre soin à mettre en perspective de nos travaux avec d'autres propositions contemporaines, nous adhérons à l'idée selon laquelle la diversité des pistes envisagées, aussi bien concernant les objectifs que les points de vue et les méthodes, conduit par leur complémentarité à un enrichissement global des travaux du domaine.

## 2.2 Elaboration d'un modèle mettant cette théorie en oeuvre

Nous avons proposé un nouveau modèle de grammaire, que nous avons pour l'instant appliqué à la description de phénomènes syntaxiques du français. Ce modèle a pour objectif à court terme de constituer la première grammaire GP exploitant pleinement les possibilités du formalisme, et à long terme de constituer une ressource à large couverture des phénomènes linguistiques du français. En l'état, il est d'ores et déjà novateur en ce qu'il permet la prise en compte d'informations syntaxiques qui sont le plus souvent ignorées ou déformées malgré leur intérêt pour l'analyse.

## 2.3 Formalisation du modèle

Nous avons représenté notre modèle de grammaire selon le formalisme de GP. Ceci nous a permis de préciser la description de l'architecture du formalisme. Il constitue un outil permettant de représenter une grammaire, et par là même d'en valider les propositions, de l'utiliser à échelle automatisée <sup>279</sup>, d'aider à son développement en autorisant de faire varier un ensemble de caractéristiques, et de la comparer de manière approfondie avec d'autres sur la base d'une représentation commune.

---

<sup>279</sup>Moyennant l'implémentation du modèle.



# Perspectives

Comme on n’a cessé de l’affirmer tout au long du manuscrit, cette thèse a plus vocation à rapporter les premiers jalons d’un travail de recherche de longue haleine plutôt que d’être une fin en soi. Selon ce point de vue beaucoup d’études restent à faire, et nous proposons ici, de manière non exhaustive, quelques pistes de recherches visant à figurer les développements ultérieurs possibles, et par là le devenir de la grammaire que l’on a présentée. Parmi ces pistes, certaines peuvent (voire, doivent) être faites à une échelle individuelle, d’autres nécessitent des collaborations pluridisciplinaires.

On commence ici par présenter quelques pistes de développements possibles de la face théorique de ce travail, puis de la grammaire elle-même, avant de présenter un certain nombre d’approfondissements du formalisme, et enfin de son exploitation informatique.

## 3 Perspectives théoriques

La théorie ici proposée fait typiquement partie des recherches dont l’initiative reste personnelle. Il s’agit là de reconnaître que, puisque l’on souhaite à cette grammaire de se développer à une échelle la plus large possible, ceci nécessitera évidemment, pour être viable scientifiquement et réalisé dans une échelle de temps raisonnable, un travail d’équipe. Et il est inévitable que ceci mène à l’émergence progressive d’un consensus théorique général plutôt qu’à l’imposition de l’idée individuelle présente.

Ceci ne constitue pas une limitation au travail ; bien au contraire, plus important sera le nombre de personnes impliquées dans le développement d’un projet, plus complètes seront les connaissances et donc plus stables et riches seront les perspectives théoriques qui s’en dégageront. Mais cela signifie également que ce qui est proposé ici sera probablement appelé à évoluer au fil des participations, ainsi qu’au fil des évolutions personnelles des participants impliqués.

En outre, pour être complète la théorie manque d’un positionnement concernant un certain nombre de points ; dans la plupart des cas auxquels nous pensons (la question des universaux linguistiques pour ne citer que celui-ci), cela nécessite un élargissement de la comparaison avec un plus grand nombre de théories, un approfondissement des théories déjà abordées, ainsi que l’apport de connaissances linguistiques concrètes les plus diversifiées, notamment en termes de familles de langues. Plus on perfectionnera la connaissance des travaux et des phénomènes

linguistiques existants, plus on sera à même de prendre position sur des questions précises.

En guise d'exemple de connaissances théoriques à approfondir, prenons celui de la *Systemic Functional Linguistics* de [Halliday, 1985] : bien que n'étant pas majoritaire dans les recherches formelles, cette approche est porteuse d'un certain nombre d'informations que l'on a soulignées ici, à commencer par la *systémique* justement, selon laquelle "le tout est différent de la somme de ses parties". C'est une idée que l'on retrouve dans les approches dites "non-compositionnelles" telles que CxG, sans que l'on ait trouvé de parallèle fait entre les deux, ni même que l'on ait vu mentionné le terme de systémique dans ces derniers contextes. Ce qui distingue les deux approches tient à première vue à la différence de représentation grammaticale puisque les grammaires fonctionnelles constituent presque un troisième type de grammaires à opposer aux syntagmatiques et aux dépendanciennes.

Et concernant les connaissances des langues à approfondir, quelques notions des études menées en langue des signes par exemple permettent d'apporter une relativisation d'un certain nombre d'éléments que l'on pourrait considérer comme généraux (la définition de la linéarité dans l'organisation syntaxique par exemple), et la confirmation d'autres choses par ailleurs (comme par exemple le découpage en morphèmes – qui certes demande une redéfinition pour tenir compte de ce que l'on nomme ainsi en langue des signes, mais qui reste utilisable, cf. p.ex. [Stokoe, 1972]).

## 4 Perspectives grammaticales

Il s'agit là du point sur lequel on aurait le plus à dire ; on se contentera néanmoins de quelques pistes en guise de possibilités. Toute personne ayant participé au développement d'une grammaire sait à quel point l'investissement dans la description d'un phénomène linguistique, quel qu'il soit, ouvre un nombre considérable de pistes complémentaires à étudier par ailleurs ; on ne saurait par conséquent raisonnablement toutes les lister ici et l'on se contentera de donner quelques exemples de directions possibles (et ne s'excluant pas les unes des autres, au contraire).

Les améliorations de la grammaire peuvent se faire dans trois directions :

- En augmentant la **profondeur de la hiérarchie** existante (§4.1), c'est-à-dire en ajoutant aux descriptions existantes la suite de leur hiérarchie en descendant progressivement vers la description de constructions de plus en plus contraintes, pour en arriver finalement à l'intégration des figements ;
- En augmentant le **contenu des constructions** existantes (§4.2), c'est-à-dire en leur ajoutant les informations qui leur sont caractéristiques et qui proviennent de domaines différents de la syntaxe ;
- Et en augmentant la **largeur de la hiérarchie** existante (§4.3), c'est-à-dire en ajoutant à la présente hiérarchie des constructions dont l'existence n'est pas basée sur les caractéristiques syntaxiques, mais sur d'autres critères tels que la prosodie notamment (qui constitue des objets distincts des constructions syntaxiques et souvent appelés "syntagmes" par com-

paraison).

## 4.1 Profondeur de la hiérarchie

Parmi les directions à suivre, la première consiste à approfondir la hiérarchie de constructions que l’on a déjà proposée, en y ajoutant progressivement les descriptions de constructions de plus en plus fines et contraintes. Considérons l’exemple de *il y a* ; il s’agit d’un Verbe construit constitué du verbe principal *avoir*, du clitique *y* et du pronom personnel clitique nominatif *il* dans son usage neutre (souvent appelé “impersonnel”).

### 4.1.1 Introduction d’une nouvelle construction dans la hiérarchie

Ce verbe construit est semi-figé puisqu’il accepte un certain nombre de variations (en temps, en mode mais pas en nombre ni en personne), et accepte quelques constituants supplémentaires précis (négation, modalisation, auxiliaire *avoir* et pronom clitique *en*). Telle qu’elle se présente actuellement, la grammaire ne décrit pas cette construction d’une manière précise et ne rend pas compte du degré de figement nécessaire à l’introduction précise d’une telle construction. Pour ce faire, on peut envisager de créer une construction que l’on pourrait étiqueter *il y a* (sachant que celle-ci, malgré son nom restrictif, rassemblerait toutes les variations sus-citées) et qui hériterait de la construction déjà décrite “Verbe à clitiques nominatif et datif” (V-cnd, cf. chap. 7 §1.6.9). En plus des caractéristiques héritées, on se contenterait d’ajouter une propriété d’Exclusion de l’auxiliaire *être* (qui avait été indexé [3] plus haut la hiérarchie), et puis l’on préciserait trois propriétés préexistantes : le Majeur [1] demanderait un verbe de lexème *avoir*, le Mineur [5] un pronom clitique nominatif de la forme *il*, et le Mineur [7] un pronom clitique datif de la forme *en* ; les autres contraintes (de précedence, d’adjacence, d’exigence et d’accord) figurent dans les constructions héritées, qui traitent également (suivant comment se fait l’héritage) de la négation, de la modalisation et de l’auxiliation. On aurait donc une construction telle que dans la figure 1.

il y a HERIT V-cnd	
Majeur	[1] $\left[ \begin{array}{l} \text{ID} \mid \text{NATURE} \mid \text{SCAT} \quad \textit{principal} \\ \text{MORPH} \mid \text{LEXÈME} \quad \textit{avoir} \end{array} \right]$
Mineur	[5] <i>il</i> [7] <i>y</i>
Exclusion	[1] $\not\rightarrow$ [3]

FIG. 1 – Exemple de construction pour *il y a*.

Reste alors l’introduction du clitique *en*, qui ferait l’objet d’une autre construction, appelons-la *il y en a*, qui hériterait à la fois de celle-ci et du “Verbe à clitiques nominatif, accusatif et datif” (V-cnad, cf. chap. 7 §1.6.11). A ces deux héritages on n’aurait pas besoin d’ajouter d’information locale autre que la précision de la forme du clitique datif, *en*. On représente cela en figure 2.

il y en a	
HERIT V-cnad	
HERIT il y a	
Mineur	6en

FIG. 2 – Exemple de construction pour *il y en a*.

#### 4.1.2 Référence à cette nouvelle construction dans le reste de la grammaire

Ces deux Verbes construits, en plus d’être contraints dans leur constitution, imposent également des restrictions quant aux objets avec lesquels il se combinent et les constructions dont ils peuvent être constituants. *Il y a* ne peut être constituant que d’un “Syntagme Verbal exocentrique bitransitif” (SV-2odi, cf. chap. 7 §2.3.2), au même titre que les V-cnd dont il hérite : on ne précise donc rien à ce propos. Il faut par contre ajouter dans ces SV-2odi qu’ils excluent les *il y en a*, qui héritent aussi, indirectement, de V-cnd.

Les constructions *il y en a*, quant à elles, constituent à elles seules des “SV bitransitifs”, tout comme leur héritage (V-cnad) le permet : on n’ajoute donc rien dans ces constructions <sup>280</sup>.

En dehors de l’emploi comme constituant verbal, la construction *il y a* pourra en outre passer sur le paradigme prépositionnel ([Blanche-Benveniste, 2000], [Le Goffic, 1993]) :

(177) il est arrivé il y a trois ans

Afin d’en rendre compte on pourra ajouter la construction dans la liste des objets Majeurs de la Préposition construite, tout en précisant que celui-ci, s’il est présent, exclut la présence de tout autre constituant dans la construction.

Ceci est un exemple simple de la façon dont on intègre un nouvel objet dans la grammaire, mais il est bien entendu possible de développer cet exemple en le précisant encore en s’appuyant par exemple sur la classification donnée dans [Cappeau & Deulofeu, 2001]. Il nous permet de montrer de quelle façon l’on peut intégrer des constructions nouvelles à la hiérarchie existante, grâce à l’utilisation de l’héritage qui nous permet de nous contenter d’ajouter les contraintes supplémentaires à celles déjà présentes pour les constructions plus hautes dans la hiérarchie, c’est-à-dire moins figées.

## 4.2 Contenu des constructions

La deuxième direction de développement de la grammaire concerne la complétion du contenu des constructions, c’est-à-dire l’intégration au sein même de celles déjà décrites d’informations provenant d’autres domaines que la syntaxe, et ce aussi bien au niveau de leurs informations intrinsèques qu’extrinsèques.

<sup>280</sup>Les formes du type *il y en a du givre sur le pare-brise* constitueront, pour leur part, des SV endocentriques dont les deux constituants immédiats seront *il y en a* d’une part, et *du givre* d’abord, puis *il y en a du givre* et *sur le pare-brise* ensuite.



### 4.2.1 Informations intrinsèques

Dans la structure de traits il est possible d'ajouter toute une série de traits qui permettent de recenser les informations provenant de domaines dont on n'a encore tenu compte dans la version actuelle. L'utilisation de cela dans la grammaire impliquera d'ajouter ces informations d'une manière homogène, c'est-à-dire de tenir compte de la façon dont ces informations sont diffusées d'une construction à une autre, de la même façon que l'on a intégré à différents niveaux de la hiérarchie la remontée des traits syntaxiques (p.ex. tous les Syntagmes portent une tête, mais parmi eux les X construits portent la catégorie de leur tête).

Pour donner un exemple en morphologie, on pourrait ajouter à la structure de traits un attribut MORPH situé au même niveau que SYN et qui contiendrait l'information sur les relations de dépendance morphologique entre les constituants (morphématiques) de la construction. Par exemple, dans l'objectif de proprement séparer les traits des morphèmes à signifiant discontinu dans les constructions (on se base ici sur les propositions de [Touratier, 2002]), on pourrait représenter une construction lexicale <sup>281</sup> contenant une partie d'un morphème à signifiant discontinu, comme p.ex. *maisons*, comme selon la fig. 3 <sup>282</sup>.

<b>N fem plur</b>																			
HERIT N																			
INTR.	<table> <tr> <td>ID</td><td> <table> <tr> <td>NATURE   CAT</td><td><i>N</i></td></tr> <tr> <td>C.I.</td><td><math>\boxed{1} \wedge \boxed{2}</math></td></tr> <tr> <td>TRAITS</td><td><math>\boxed{1}.</math>TRAITS</td></tr> <tr> <td><i>etc.</i></td><td></td></tr> </table> </td></tr> <tr> <td>MORPH</td><td> <table> <tr> <td>BASE</td><td><math>\boxed{1}</math></td></tr> <tr> <td>LEXÈME</td><td><math>\boxed{1}.</math>LEXÈME</td></tr> <tr> <td>AFFIXES</td><td><math>\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]</math></td></tr> </table> </td></tr> </table>	ID	<table> <tr> <td>NATURE   CAT</td><td><i>N</i></td></tr> <tr> <td>C.I.</td><td><math>\boxed{1} \wedge \boxed{2}</math></td></tr> <tr> <td>TRAITS</td><td><math>\boxed{1}.</math>TRAITS</td></tr> <tr> <td><i>etc.</i></td><td></td></tr> </table>	NATURE   CAT	<i>N</i>	C.I.	$\boxed{1} \wedge \boxed{2}$	TRAITS	$\boxed{1}.$ TRAITS	<i>etc.</i>		MORPH	<table> <tr> <td>BASE</td><td><math>\boxed{1}</math></td></tr> <tr> <td>LEXÈME</td><td><math>\boxed{1}.</math>LEXÈME</td></tr> <tr> <td>AFFIXES</td><td><math>\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]</math></td></tr> </table>	BASE	$\boxed{1}$	LEXÈME	$\boxed{1}.$ LEXÈME	AFFIXES	$\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]$
ID	<table> <tr> <td>NATURE   CAT</td><td><i>N</i></td></tr> <tr> <td>C.I.</td><td><math>\boxed{1} \wedge \boxed{2}</math></td></tr> <tr> <td>TRAITS</td><td><math>\boxed{1}.</math>TRAITS</td></tr> <tr> <td><i>etc.</i></td><td></td></tr> </table>	NATURE   CAT	<i>N</i>	C.I.	$\boxed{1} \wedge \boxed{2}$	TRAITS	$\boxed{1}.$ TRAITS	<i>etc.</i>											
NATURE   CAT	<i>N</i>																		
C.I.	$\boxed{1} \wedge \boxed{2}$																		
TRAITS	$\boxed{1}.$ TRAITS																		
<i>etc.</i>																			
MORPH	<table> <tr> <td>BASE</td><td><math>\boxed{1}</math></td></tr> <tr> <td>LEXÈME</td><td><math>\boxed{1}.</math>LEXÈME</td></tr> <tr> <td>AFFIXES</td><td><math>\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]</math></td></tr> </table>	BASE	$\boxed{1}$	LEXÈME	$\boxed{1}.$ LEXÈME	AFFIXES	$\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]$												
BASE	$\boxed{1}$																		
LEXÈME	$\boxed{1}.$ LEXÈME																		
AFFIXES	$\left[ \begin{array}{l} \text{NOMBRE} \quad ? \boxed{2} : \textit{singular} \end{array} \right]$																		
<table> <tr> <td>Majeur</td><td><math>\boxed{1}N\text{-base}</math></td></tr> <tr> <td>Mineur</td><td><math>\boxed{2}N\text{ombre}</math></td></tr> <tr> <td>Précédence</td><td><math>\boxed{1} \prec \boxed{2}</math></td></tr> <tr> <td>Adjacence</td><td><math>\boxed{1} \leftrightarrow \boxed{2}</math></td></tr> </table>		Majeur	$\boxed{1}N\text{-base}$	Mineur	$\boxed{2}N\text{ombre}$	Précédence	$\boxed{1} \prec \boxed{2}$	Adjacence	$\boxed{1} \leftrightarrow \boxed{2}$										
Majeur	$\boxed{1}N\text{-base}$																		
Mineur	$\boxed{2}N\text{ombre}$																		
Précédence	$\boxed{1} \prec \boxed{2}$																		
Adjacence	$\boxed{1} \leftrightarrow \boxed{2}$																		

FIG. 3 – Exemple de construction pour *maisons*.

Sans entrer dans ses propriétés extrinsèques (on y reviendra dans ce qui suit), on voit que l'on placerait sous le trait MORPH un trait indiquant la base du lexème (*maison* p.ex., ou toute autre base possible acceptant le nombre) et le lexème correspondant (*maison*), puis une série de traits présentant les affixes portés le cas échéant (*s* en l'occurrence, qui est une partie du signifié du morphème de pluriel porté en tant qu'affixe). Le genre par contre, qui reste un trait du nom, figurera toujours sous ID|TRAITS.

Il est bien entendu possible de développer ce type d'informations pour d'autres domaines

<sup>281</sup>Qui serait alors, plus exactement, une construction correspondant à un "mot" plutôt que proprement lexicale.

<sup>282</sup>Pour les exemples qui suivent, nous affectons des étiquettes qui n'auraient pas grand intérêt en réalité, mais qui ont l'avantage d'être plus compréhensibles que si l'on leur donnait des étiquettes sous forme d'abréviations

également, et l'on voit que dans tous les cas, la notation et les mécanismes restent les mêmes, on ne fait qu'ajouter du contenu à ce qui existe déjà sans avoir à le modifier en profondeur.

#### **4.2.2 Informations extrinsèques**

De la même façon on peut ajouter des informations extrinsèques aux constructions déjà existantes, et nous permettant de relever les indices relationnels relevant d'autres domaines que celui de la syntaxe. C'est là par exemple que l'on pourra faire intervenir les informations de restriction de sélection entre constituants imposées par le contenu sémantique des objets.

Suivant le type de relations qui seront nécessaires pour les autres domaines, on pourra définir de nouvelles propriétés pour en rendre compte, dont on précisera alors la spécification. On pourra également utiliser les mêmes propriétés que celles déjà définies, puisqu'elles ne sont pas limitées à un domaine particulier (p.ex., les restrictions de sélection lexicale se feront à l'aide de propriétés d'Exigence et d'Exclusion, tandis que celles portant sur la place des adjectifs qualificatifs s'exprimeront en termes de Précédence).

### **4.3 Largeur de la hiérarchie**

La troisième direction de développement, enfin, consiste à élargir la hiérarchie grammaticale, c'est-à-dire ajouter des constructions appartenant à d'autres domaines que la syntaxe et qui sont nécessaires à une analyse non-modulaire. En effet telle qu'elle est actuellement, la grammaire ne permet que l'analyse d'une partie d'un énoncé, puisqu'elle débute au niveau d'unités syntaxiques minimales (qui sont pourtant déjà des unités construites), et qu'elle se termine aux unités syntaxiques maximales alors que l'analyse peut continuer de s'effectuer sur une structuration à plus gros grain. Cela implique donc de développer non seulement des constructions plus grosses que celles que l'on a déjà, mais aussi des constructions plus petites.

#### **4.3.1 Constructions plus grandes**

Un exemple simple et rapidement utile de constructions plus grandes que celles que l'on peut baser sur la syntaxe sont les constructions relevant du domaine énonciatif. Celles-ci, dont les constituants ne sont justement pas marqués par des relations syntaxiques, mais souvent par un assemblage de marques prosodiques (contours) et sémantiques (références partagées, notamment), pourront être ajoutées à la grammaire non pas en les intégrant à la hiérarchie déjà existante, mais en introduisant une nouvelle hiérarchie dont la racine sera assez probablement distincte de celle des constructions syntaxiques (on avait déjà évoqué le fait que notre grammaire n'ait qu'une unique racine n'étant qu'un effet secondaire de sa couverture restreinte), ou en tout cas assurément des constructions syntaxiques de base (syntagme et paradigme, quoi qu'il semble possible que ce parallèle se retrouve dans certains cas).

Notons que ces constructions, bien qu'on les présente ici comme exemples de constructions plus grosses que les constructions syntaxiques, peuvent également être de portée plus petite et,

pourquoi pas, entrer elles-mêmes comme constituants immédiats de constructions aux caractéristiques syntaxiques : c'est l'une des possibilités qu'offre notre modèle non-modulaire. Nous ne sommes pas ici en train de distinguer plusieurs étapes d'une analyse qui s'effectueraient systématiquement dans un ordre donné (morphologie, puis syntaxe, puis énonciation), mais nous ne faisons que recenser quelles sont les constructions pouvant entrer dans la grammaire et dont on n'a pas encore tenu compte. Que ces constructions soient introduites avant, pendant ou après celles que l'on a déjà décrites lors d'une analyse n'a aucune importance puisqu'elles sont décrites dans la même ressource et n'y sont pas ordonnées.

### 4.3.2 Constructions plus petites

Les constructions que l'on a manipulées dans nos exemples de la troisième partie comme s'il s'agissait de constructions minimales sont en réalité la résultat d'une autre partie de l'analyse qui s'est effectuée avant d'en arriver à obtenir la base de nos constructions syntaxiques, et celles-ci doivent être intégrées à la grammaire au même titre que celles que l'on a déjà décrites ainsi que celles dont les constituants sont des constructions syntaxiques.

Reprenons notre exemple de construction morphologique du nom *maisons* de la figure 3. On a vu à travers cet exemple qu'une unité syntaxique minimale telle que *maisons* est en réalité constituée d'une base lexicale et d'une partie d'un morphème à signifiant discontinu. Pour permettre à la grammaire de décrire de telles constructions morphématiques, il est nécessaire d'introduire des constructions d'un tel niveau, et celles nécessaires à la construction *maison* pourraient ressembler à celles présentées en figures 4 et 5.

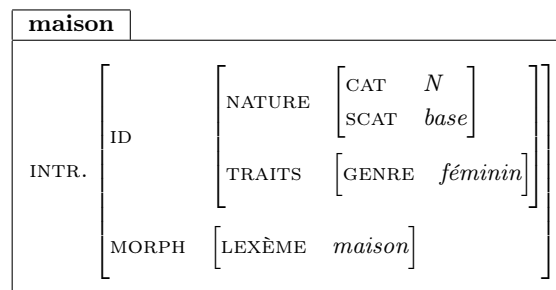


FIG. 4 – Exemple de construction morphématique pour la base *maison*.

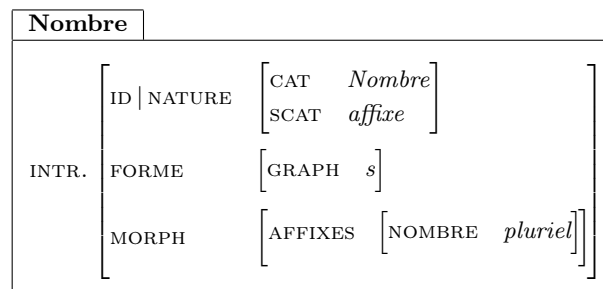


FIG. 5 – Exemple de construction morphématique pour un affixe pluriel.

On pourrait raffiner encore plus en proposant une construction N-base qui porterait une propriété Majeur qui réclamerait une autre construction dont les seules informations seraient la forme graphique et le lexème (p.ex.), lesquelles viendraient instancier les traits correspondants dans N-base. Il est possible de descendre dans la construction des signes linguistiques de cette façon jusqu'à aboutir aux signes minimaux, et donc au début d'une analyse. Ceci permettra, en outre, de traiter les figements lexicaux selon le même modèle de construction que les figements syntaxiques, et une fois les informations sémantiques ajoutées aux descriptions de leur affecter leur sens de la même façon.

Toujours selon ce même exemple, pour décrire un syntagme nominal tel que *les maisons*, on n'utiliserait plus la propriété d'Accord pour mettre en relation les deux parties du morphème (unique) de pluriel, mais une propriété d'Exigence selon laquelle un affixe pluriel réclame un affixe pluriel (en d'autres termes, une partie du signifiant réclame l'autre partie), et Exclut les affixes singuliers (ou exclut les absences d'affixe pluriel, si l'on considère que le singulier est une absence de morphème de pluriel).

## 5 Perspectives formelles

D'autres développements sont possibles concernant non plus la grammaire développée ou la théorie associée, mais le formalisme lui-même cette fois. Ceux-ci portent principalement sur des compléments d'informations qu'il sera intéressant d'ajouter à la spécification des propriétés d'une part, et aux constructions introduites lors des analyses d'autre part.

### 5.1 Index de grammaticalité

La question d'affecter une importance aux propriétés les unes relativement aux autres au sein de la description d'une construction s'est posée très tôt lors du développement de GP. En effet, il semblait assez évident que les contraintes s'appliquant sur un objet ne pesaient pas toutes le même poids dans la justification ou non de l'existence de cet objet : on peut facilement imaginer par exemple que le fait qu'un déterminant ne précède pas un nom soit un fait rédhibitoire à l'introduction d'un SN constitué de ces deux objets, alors que comparativement, le fait qu'un déterminant immédiatement suivi d'un nom qui s'accordent en genre mais pas en nombre montre certes une certaine inadéquation entre les deux mais n'empêche pas de postuler un SN (même s'il n'est pas complètement conforme à ce que la norme grammaticale voudrait).

On a donc très vite imaginé affecter un poids relatif aux propriétés au sein des informations extrinsèques des objets afin de faire, *a minima*, une distinction entre les propriétés dont la satisfaction (si toutefois elles sont évaluées) qui sont indispensables à l'introduction d'une construction (comme la précédence du Det sur le N dans le SN), et celles qui lui sont accessoires (comme les Accords). Cela a été par exemple le cas d'un premier parseur de GP qui imposait (sans l'avouer clairement toutefois) que les propriétés d'Obligation soient satisfaites pour que les constructions correspondantes soient introduites, alors que les autres propriétés n'étaient pas primordiales. Mais ceci était insuffisant et pour aller plus loin et gagner en précision et ainsi en justesse ; on ne pouvait pas affecter ces poids sans avoir une idée précise de

leur fonctionnement : Combien proposer de degrés d'importance différents?, sur quelle base en décider?, ces poids étaient-ils réguliers d'une propriété à l'autre? (i.e., est-ce que la Précédence entre un spécifieur et la tête est toujours plus importante que celle entre la tête et un complément?), ou alors d'un type de propriété à l'autre? (i.e., est-ce que la Précédence est toujours plus importante que l'Accord?), ou encore est-ce que ces importances variaient d'une construction à l'autre? (la Précédence étant plus importante pour une construction X, mais moins pour une Y?) Il était fondamental, pour répondre à ces questions, de mener un certain nombre de tests pour mettre cela en évidence d'une manière scientifique et non en nous basant uniquement sur nos *a priori* de linguistes <sup>283</sup>.

D'après les expériences dont les premiers résultats sont présentés dans [Blache *et al.*, 2006], il semble possible d'obtenir une idée précise du degré de grammaticalité d'une construction (et, par extension, d'un énoncé) en se référant à un ensemble d'informations fournies par l'analyse GP : le degré de bonne formation grammaticale (constitué par le rapport entre propriétés satisfaites et violées d'une caractérisation), la densité d'information (la quantité de contraintes constituant la caractérisation), et la structure elle-même (i.e., ses constituants). Sur la base de ces trois paramètres il est possible de calculer un *index de grammaticalité* qui est fortement corrélé avec les jugements d'acceptabilité des mêmes énoncés produits par des locuteurs, c'est-à-dire, par extension, de prédire l'analysabilité d'un énoncé (ainsi que des constructions qui le constituent).

Ces résultats sont maintenant à vérifier sur un ensemble de données plus larges, mais autorisent dès à présent de penser que l'importance relative portera non pas sur les propriétés elles-mêmes les unes relativement aux autres (et d'une manière pouvant être variable d'une construction à l'autre), mais sur les types de propriétés, et de manière stable dans la grammaire. En conséquence, si cela se confirme, on pourra ajouter une information de poids dans la spécification des propriétés lors de leur définition.

## 5.2 Spécification des types de propriétés

L'ajout de l'information concernant le poids relatif des types de propriétés (si elle se révèle confirmée par les tests à venir), permettant de calculer l'index de grammaticalité des constructions, n'est pas la seule information que la spécification des types de propriétés peut recevoir afin d'être complétée.

### 5.2.1 Détection des invariants

On a dit, dans la présentation de la spécification, que l'on avait progressivement fait émerger un certain nombre d'informations invariantes dans la description et le fonctionnement des propriétés; il est toutefois certain que les méthodes employées à cette fin ont manqué de rigueur, pour deux raisons : d'une part elles ont été menées sans réelle dimension formelle : il existe des méthodes clairement définies d'extraction d'invariants, dont on n'a pas encore fait

<sup>283</sup>Lesquels nous aurait certainement aidés à progresser en la matière, mais sans conteste moins vite et moins rigoureusement.

usage ; d'autre part elles n'ont été évaluées que sur la base du jeu de propriétés initiales (celles de [Blache, 2001]), ce qui ne constitue pas un panel de possibilité suffisamment large. C'est d'ailleurs ce qui fait que la formulation de cette spécification n'est pas encore stabilisée : il arrive encore à présent que lorsqu'on souhaite définir un nouveau type de propriété, celui-ci fasse émerger une nouvelle information que l'on avait considérée comme fixe jusque là parce qu'elle ne variait pas pour les autres.

Dans une perspective beaucoup plus générale que l'utilisation de GP en tant que formalisme linguistique, on peut même imaginer l'appliquer à la résolution de toutes sortes de problèmes complexes pouvant reposer sur une méthode flexible de satisfaction de contraintes (c'est par exemple ce que font [Estratat & Hencoque, 2004] en utilisant GP pour résoudre des problèmes de configuration qui n'ont de rapport que méthodologique avec l'analyse linguistique), et il sera alors tout à fait indispensable de disposer d'une spécification dont les éléments constitutifs soient réellement exhaustifs et qui permette de définir toutes sortes de contraintes pouvant être fort différentes de celles nécessaires à l'analyse linguistique.

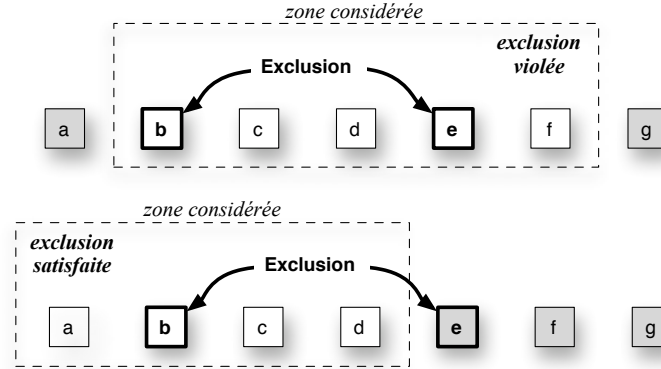
### 5.2.2 Portée de satisfaction

En plus de l'élaboration d'un ensemble de caractéristiques de définition qui soit basé sur des tests plus rigoureux, il existe déjà un certain nombre de caractéristiques des propriétés dont on connaît l'existence et dont on n'a néanmoins pas tenu compte dans la présentation faite au chapitre 2. C'est le cas par exemple de ce que l'on peut appeler la *portée de satisfaction*.

En effet, on a vu que lors de la première étape d'analyse on évalue les propriétés comme un ensemble général considéré en dehors des constructions auxquelles elles peuvent appartenir, et par là même considéré au-delà des limites de ces mêmes constructions. Cela pose un problème que l'on peut illustrer de la façon suivante : imaginons une propriété d'Exclusion  $b \not\equiv e$  contenue dans une construction C, et imaginons un énoncé constitué de la suite d'objets *abcdefg*. Lors de la première étape d'analyse de l'énoncé, la propriété d'Exclusion sera évaluée (puisque *b* est présent), et sera violée (puisque *e* est également présent). Pourtant, pour juger de la caractérisation de C, cette information sur la présence conjointe de *b* et de *e* dans l'énoncé ne suffit pas : ce qu'il faut savoir, c'est si ces deux objets figureraient précisément dans C. Si c'est le cas alors la propriété sera violée dans la caractérisation de C, mais si seul *b*, ou seul *e*, fait partie des constituants de C alors la propriété sera violée dans l'énoncé mais satisfaite dans C. La satisfaction des propriétés dépend donc, en fait, non pas de leur évaluation sur l'ensemble des objets observés, mais sur la partie de cet ensemble concerné par la construction dont on est en train de calculer la caractérisation ; et de ce point de vue, la valeur de satisfaction d'une propriété peut varier. C'est ce que l'on illustre en fig. 6.

Il est possible de tenir compte de cette relativité de la valeur de satisfaction en fonction de l'empan observé dans la spécification des types de propriétés. On pourrait en effet définir deux types de portée de satisfaction possibles :

- Une portée **générale**, qui signifierait qu'une propriété de ce type conserve sa valeur de satisfaction quel que soit l'empan observé. C'est typiquement le cas des propriétés de consti-

FIG. 6 – Relativité de la satisfaction de la propriété d'Exclusion  $b \neq e$  à la zone considérée.

tuance : Obligation, Facultativité, Majeur, Mineur <sup>284</sup> ;

- Une portée **relative**, qui signifierait que la valeur de satisfaction qui lui est attribuée pour la totalité de l'énoncé serait opposée à sa valeur locale à un empan plus petit (celui d'une construction, en l'occurrence), si tous les objets qu'elle met en relation ne figurent pas dans l'empan considéré. Ainsi, la propriété d'Exclusion que l'on a considérée dans l'exemple précédent, puisqu'elle est violée sur l'énoncé, prendra cette même valeur (violée) dans toutes les caractérisations de constructions qui contiennent à la fois  $b$  et  $e$ , et prendront la valeur inverse (satisfaite) dans toutes les caractérisations de constructions qui ne contiennent que  $b$ , ou que  $e$ . Pour une propriété d'Exigence (mettons  $b \Rightarrow e$ ), elle sera satisfaite sur l'énoncé ainsi que dans toutes les constructions qui contiennent les deux termes de la propriété, et violée dans les constructions ne contenant qu'un seul des deux termes (plus exactement, elle sera violée si elle ne contient que le premier terme, et non disponible — donc non évaluée — si elle ne contient que le second terme).

Ceci n'est qu'une esquisse de fonctionnement possible, et elle demande à être développée et améliorée ; c'est la raison pour laquelle on ne l'a pas fait figurer dans la spécification ici. Toutefois il faut rester conscient de cette caractéristique, surtout lors d'envisager l'implémentation de GP <sup>285</sup>.

<sup>284</sup>Dans ce cas ce n'est pas la satisfaction qui varie en fonction de l'empan, mais la disponibilité de la propriété : une propriété  $M(a)$  est disponible (et satisfaite) pour une construction qui contient  $a$ , et ne l'est pas pour une construction qui ne le contient pas (et dans ce cas elle n'est pas évaluée).

<sup>285</sup>D'un point de vue strictement implémentationnel, une autre possibilité existe pour prendre en compte ce phénomène. Elle consiste à n'effectuer durant la première étape que la mise en relation à partir des propriétés qui ne posent pas ce problème, et à n'évaluer les autres que lors de la deuxième étape (i.e., une fois que les constructions ont été introduites). Cette possibilité offre l'avantage d'ignorer tout simplement le problème, mais elle a pour conséquence que toute une partie des propriétés, i.e. une partie des relations que nous avons définies telles que ne prenant un sens structurel que considérées dans leur ensemble, dans leur interaction, ne pourra pas être prise en compte à l'issue de la première étape d'analyse. Cela faussera, notamment, le calcul de l'index de grammaticalité tel que défini précédemment puisqu'il dépend de la densité d'information, densité que l'on ne pourra pas évaluer sans l'évaluation de l'ensemble des propriétés.

### 5.3 Comparaison de modèles

Il s'agit là non pas d'une amélioration supplémentaire du formalisme, mais d'une utilisation qui peut être envisagée.

On a vu au cours de cette thèse qu'il était possible de modéliser à partir de GP des grammaires aux caractéristiques théoriques très différentes. GP n'est bien sûr pas le seul formalisme à permettre cela, mais c'est néanmoins suffisamment rare pour être souligné. De plus, les quelques autres formalismes dans lesquels on a essayé de modéliser notre grammaire <sup>286</sup> ne nous ont pas permis de le faire : nous nous trouvions toujours limitée, à un moment ou à un autre, par une caractéristique du modèle qui nous demandait de faire fi de l'une de nos positions théoriques.

C'est en ceci, constatant la vague de campagnes d'évaluations diverses et le nombre de Journées d'études et autres *Workshops* (tant à l'échelle nationale qu'internationale) actuellement consacrés à la réflexion sur les méthodes d'évaluation, que l'on a imaginé qu'il serait possible et intéressant d'utiliser GP comme un outil de formalisation de plusieurs théories linguistiques différentes, afin de pouvoir évaluer leurs points communs et leurs divergences en s'efforçant de tendre vers une équivalence formelle par ailleurs, afin de pouvoir se concentrer sur un point unique (ou tout au moins restreint) de variabilité, toutes choses égales par ailleurs. On pourrait alors évaluer p.ex. le degré de lexicalisation de plusieurs théories, comparer leurs choix d'organisation des informations, et toutes sortes de choses sur une base commune.

## 6 Perspectives informatiques

Ce point, le dernier des perspectives, fait partie des projets à mener évidemment dans une perspective de coopération pluridisciplinaire. Les quelques idées que nous donnons ici ne sont que quelques pistes auxquelles on a pu penser dans ce cadre, sans avoir évalué de manière précise dans quelle mesure chacune d'elles serait réalisable, quel en serait le coût opératoire, ni s'il existe des alternatives auxquelles on n'a pas pensé.

Comme on a assez souvent eu l'occasion de le dire tout au long de ce manuscrit, nous avons tenu dans ce travail à clairement séparer ce qui relève des informations linguistiques et par là de la grammaire et/ou du formalisme chargé de la représenter, et ce qui relève des heuristiques qu'il serait envisageable d'ajouter au fonctionnement du formalisme de manière à produire automatiquement des résultats dont la teneur (en qualité et en quantité et même en forme) dépend de manière directe des objectifs respectifs de chacun des outils, et est par conséquent aussi variable que peuvent l'être les besoins.

---

<sup>286</sup>Nous n'avons bien entendu pas eu le temps de tester ceci sur tous les formalismes censés être théoriquement neutres mais nous avons essayé cela avec les *Grammaires d'Unification Polarisées* et *Dynamic Syntax*. Nous aurions aimé tester également la représentation Métagrammaticale mais nous n'en avons pas eu l'occasion, parce qu'une métagrammaire suppose un formalisme dans lequel transposer les méta-informations et que nous n'avons pas trouvé de formalisme utilisant les MG et qui nous ne soit pas orienté théoriquement.



## 6.1 Implémentation de la présente grammaire

Dans son état actuel, la grammaire n'a pas été implémentée : c'est la raison pour laquelle on n'a pas évoqué cette partie du développement, qui pourtant constitue une phase cruciale en ce qu'elle permet de vérifier à grande échelle les résultats fournis par la grammaire, et ainsi de pouvoir l'améliorer. Pour avoir expérimenté cette technique d'allers-et-retours entre propositions de grammaire et tests de parsing à l'occasion de la campagne EASY (cf. annexe A), on sait bien à quel point le développement de la ressource est facilité par cette technique.

Notre grammaire a eu l'occasion d'être implantée à un certain nombre de reprises, essentiellement à but de tests de fonctionnement de l'analyseur *SeedParser* de [Vanrullen, 2005], mais la version actuelle ne l'a pas encore été. D'une part pour des raisons de temps, d'autre part pour des raisons d'affinement progressif du modèle qui n'ont pas encore été prises en compte dans le parseur en question. Le résultat de cela est que, même si l'on implantait notre grammaire dans le *SeedParser*, les résultats obtenus ne seraient pas représentatifs des possibilités de la grammaire. Le problème essentiel, actuellement, est celui de la possibilité d'introduire parallèlement des structures syntagmatiques et paradigmatiques, alors que le parseur a été conçu pour ne produire que les secondes <sup>287</sup> ; il conviendra donc de modifier cela en tout premier lieu.

## 6.2 Heuristiques variables

Afin d'utiliser cette grammaire de manière automatique, et dans la mesure où il s'agit d'une ressource de description et non pas de dérivation, il semble évident que le recours à des mécanismes de limitation de l'introduction de constructions soit essentiel. Suivant le type de résultats que l'on souhaite obtenir, on peut envisager plusieurs heuristiques qui seraient utilisables soit alternativement, soit conjointement. Pour cela, il est intéressant de s'inspirer des idées provenant des travaux en TAL qui s'efforcent de proposer des solutions robustes à des problèmes précis et concrets.

C'est par exemple en s'inspirant de [Johnson *et al.*, 2004] que l'on peut imaginer introduire des mécanismes pour choisir parmi les possibilités d'entassements paradigmatiques, entre disfluence ou coordination : eux proposent d'utiliser une liste de marqueurs qui, suivant s'ils appartiennent à l'une ou l'autre liste, permettront de trancher et de n'introduire que la solution la plus probable. La raison pour laquelle on fait figurer cela dans les heuristiques et non pas dans la grammaire elle-même est que ceci est basé sur des listes dont la fiabilité linguistique n'est pas prouvée : globalement, on introduit dans la grammaire ce qui est véritablement linguistiquement stable, et dans les heuristiques ce qui permet d'aider au choix tout en étant limité par le fait qu'il s'agit d'un moyen de trancher dans une certaine mesure et qu'il existe toujours quelques exceptions dans lesquelles le choix effectué ne sera pas le bon (mais qui permettent malgré tout d'obtenir des résultats satisfaisants dans la plupart des cas).

Dans le même ordre d'idées, un autre mécanisme auquel on pourrait penser pour aider à

<sup>287</sup>En fait, ceci fait typiquement partie des éléments qui ne devraient pas dépendre du parseur mais du formalisme et du modèle, mais pour s'en rendre compte il fallu que le besoin émerge.

décider de quelles constructions caractérisées introduire, serait quelque chose qui consisterait à choisir, entre deux constructions s'opposant et ayant les mêmes caractéristiques par ailleurs, celle dont les constituants sont les plus proches les uns des autres.

On pourrait citer encore tout un ensemble de moyens de limiter la surgénération évidente de la grammaire. Tous ceux-ci, dans l'idéal, devraient pouvoir prendre la forme d'options que l'on pourrait appliquer pour effectuer un type de résultat particulier (plus ou moins déterministe, par exemple), à l'aide d'un même parseur et suivant les besoins que l'on a pour remplir une tâche particulière. On disposerait alors d'un outil d'analyse automatique qui ait des caractéristiques de flexibilité comparables ce que l'on a cherché à permettre dans le formalisme.

## Annexe A

# La campagne d'évaluation des analyseurs syntaxiques EASY

Nous présentons ici la grammaire réalisée pour les besoins de la campagne EASY. Elle a été développée dans le cadre très précis de la mise en jeu de trois analyseurs syntaxiques, qui sont décrits dans [Balfourier *et al.*, 2005] :

*Les approches utilisées par les parseurs sont symboliques, mais utilisent des techniques différentes. Le premier est un analyseur superficiel, utilisant une grammaire simplifiée adaptée à la campagne d'évaluation dans laquelle un sous ensemble des propriétés de GP est exploité. Les second et troisième analyseurs utilisent des techniques d'analyse profonde. Ils emploient tous deux une représentation de l'information syntaxique sous la forme d'une grammaire complète et contrôlent le processus d'analyse et de détermination des résultats grâce à des algorithmes différents qu'il nous a intéressé de comparer.*

*Ces trois approches présentent bien entendu des résultats différents en termes d'efficacité et de couverture. Cette expérience permet de comparer, au sein d'une même approche symbolique et à partir de ressources identiques, les approches superficielles et profondes, ainsi que différents algorithmes d'analyse.*

*Nous présentons dans un premier temps les ressources utilisées comme base pour l'analyse, puis nous exposons les trois analyseurs syntaxiques. Enfin, nous abordons la question de l'évaluation des résultats en tentant de différencier les origines des erreurs dans les processus d'analyse.*

## 1 Les ressources utilisées

**Le lexique et l'étiquetage** <sup>288</sup>. — Les trois analyseurs prennent une entrée identique se présentant sous la forme d'un texte tokenisé et étiqueté. Chaque token est imposé dans le corpus à analyser, de sorte que tous les participants soient évalués sur les mêmes données. Chacun de ces tokens est accompagné d'un ensemble d'étiquettes possibles fournies de façon automatisée grâce à l'étiqueteur WinBrill. Nous avons réétiqueté ces tokens à l'aide de notre

---

<sup>288</sup>Cette section et une citation intégrale de [Balfourier *et al.*, 2005], que pour des raisons de lisibilité nous n'avons pas mis en exergue. Nous tenons toutefois à signaler que ce texte a été écrit à plusieurs mains.

propre étiqueteur (celui du LPL), afin de faire correspondre le jeu de traits morphosyntaxiques de ces étiquettes avec celui qu'emploie notre grammaire, ainsi que dans le but d'affiner la qualité de l'étiquetage. De fait, l'étiqueteur du LPL présente une bonne efficacité et se base sur un ensemble de traits plus fins que ceux proposés par WinBrill. Le lexique de 430000 formes utilisé par notre étiqueteur est une variante de celui du LPL (présenté dans [Vanrullen, 2005]), adaptée aux entrées du corpus d'évaluation EASY. Il reste cependant un certain nombre d'erreurs qui se répercuteront bien entendu sur le résultat de l'analyse. Ce point est discuté plus loin.

**La grammaire.** — D'un point de vue linguistique, la participation à la campagne Easy a consisté à concevoir une grammaire formelle dont les résultats visent à coïncider avec les indications fournies dans le Protocole d'Evaluation des Analyseurs Syntaxiques (ci-après PEAS, [Gendner & Vilnat, 2004]), qui servait de référence à la fois pour les annotateurs (référence pour les résultats) et pour les développeurs de grammaire (production des résultats). Le formalisme que nous utilisons pour développer des grammaires au LPL est celui des Grammaires de Propriétés (ci-après GP, cf. par exemple [Blache, 2005]).

Le guide PEAS a été conçu dans un souci de consensus entre les différentes théories linguistiques prises en considération lors de son élaboration. De fait, les choix effectués à la base du guide étant plus ou moins éloignés (suivant les cas) des thèses couramment soutenues dans le paradigme des GP, la conséquence de cela est que nous avons développé une grammaire spécifique à la tâche demandée pour Easy. Cette campagne nous a donc permis de tester, dans des circonstances concrètes, la flexibilité du modèle et de vérifier sa capacité d'expression pour une théorie significativement différente de celles habituellement adoptées.

Concrètement, le développement de la GP a consisté en les étapes suivantes :

1. suite à une première lecture du guide, un choix des propriétés utilisées, et la définition de leur sémantique (leur mode de fonctionnement, variable à volonté, cf. par exemple [Vanrullen *et al.*, 2003]),
2. interprétation approfondie des indications données dans PEAS et transcription de celles-ci en propriétés,
3. vérification de l'adéquation des résultats fournis par les trois parseurs sur un corpus-test,
4. estimation des causes des erreurs produites : méthodes de parsing, descriptions grammaticales, interférences entre propriétés, etc.,
5. ajustements de la grammaire pour les cas qui la concernent (modifications de la sémantique des propriétés, et/ou des propriétés elles-mêmes et/ou des ensembles de propriétés),
6. reprise des étapes 3, 4 et 5 jusqu'à ce que les résultats soient satisfaisants.

Ce va-et-vient entre ajustements de grammaire et tests sur corpus représentatif nous a permis d'adapter la grammaire, à la fois aux indications de PEAS en amont, et aussi aux différents traitements qui peuvent en être faits par chacun parseurs, en aval.

## 2 Les parseurs

**Analyseur superficiel** <sup>289</sup>. — L'analyseur superficiel prend en entrée un texte étiqueté et désambiguïsé. Il construit dans une première passe l'ensemble des groupes, puis les relations. La construction des groupes repose sur des informations syntaxiques partielles. Plus précisément, seules les propriétés de constituance et de linéarité ont été utilisées.

La stratégie repose sur une technique d'analyse coin-gauche. Il s'agit de repérer pour chaque token, grâce aux deux propriétés citées, sa faculté d'être coin gauche d'un groupe. Dans de nombreux cas, les informations citées sont suffisantes et l'initialisation du groupe correspondant est systématique. En revanche, certaines situations nécessitent la vérification du contexte immédiat. C'est par exemple le cas des groupes PV, qui débutent par une préposition. Cette dernière initialise habituellement un GP, mais dans un contexte verbal à droite, la préposition devient coin gauche du PV. Chaque initialisation de groupe entraîne la fermeture du groupe précédent.

Le mécanisme est donc en une passe unique et consiste à analyser successivement toutes les suites de trois tokens (le token candidat coin gauche, son contexte gauche et son contexte droit). La connaissance du type du groupe en cours permet de compléter la décision d'initialisation.

Les relations sont calculées dans une seconde passe sur la base des groupes construits. Chaque relation correspond à un traitement spécifique. Un premier traitement consiste à construire différentes tables regroupant les groupes et formes susceptibles d'être source ou cible d'une relation. Chaque relation consiste ensuite à parcourir ces tables et vérifier, en fonction des positions des candidats, leur appartenance à une relation. Dans certains cas (par exemple la relation complément-verbe) les candidats sont des ressources uniques. En d'autres termes, un candidat ne peut être complément d'un seul verbe. Cette information, correspondant à une consommation de ressource, est ajoutée dans les tables pour les items concernés. La détection des relations repose donc globalement sur des critères topologiques. Il s'agit d'une approximation qui ne permet pas d'assurer un bon contrôle ce qui entraîne une surgénération.

Les techniques utilisées par l'analyseur superficiel sont donc très simples, ce qui a bien entendu des conséquences sur le résultat obtenu (en particulier pour ce qui concerne les relations). L'avantage majeur de cette technique est sa robustesse et son efficacité : le corpus total est analysé en 4 minutes environ.

**Premier analyseur profond.** — Le premier analyseur profond utilise une version XML de la grammaire comme ressource permettant l'analyse. Deux algorithmes spécifiques sont mis en oeuvre : celui lié à l'analyse et celui destiné à déterminer la sortie. La technique d'analyse repose sur une transformation de la grammaire en un graphe de contraintes. Ce graphe, ainsi que la sémantique des contraintes sont confrontés aux corpus à analyser. L'analyse consiste à produire l'ensemble des contraintes satisfaites et enfreintes par chaque succession de tokens constituant un énoncé phrastique. En cours d'analyse, les catégories de la grammaire EASY

<sup>289</sup>Cette section est également une citation intégrale de [Balfourier *et al.*, 2005], que pour des raisons de lisibilité nous n'avons pas mis en exergue.

sont construites en fonction du nombre de contraintes qui les satisfont. Cette technique utilise la mesure de *densité de satisfaction* présentée dans [Vanrullen, 2004]. Si une construction est suffisamment satisfaite (ceci en fonction d'un seuil préétabli pour chaque corpus), elle sera conservée dans le résultat. Une fois l'analyse achevée, une détermination des résultats est réalisée en utilisant à nouveau la mesure de densité de satisfaction : les constructions définitivement choisies pour le résultat final sont celles qui maximisent la somme des densités de satisfaction pour chaque énoncé. Cette maximisation est réalisée grâce à un calcul sur des cliques (au sens de la théorie des graphes), où chaque clique présente un conflit entre plusieurs hypothèses d'analyse. Cet analyseur présente l'avantage de séparer les algorithmes et les données (le programme est indépendant de la grammaire et de sa sémantique, contrairement au troisième analyseur qui doit les inclure au sein même du programme). Son inconvénient réside par contre dans sa lenteur. La complexité moyenne reste polynomiale, mais la durée du traitement est de plusieurs jours sur plusieurs machines pour le million de mots que constitue le corpus. Pour cette raison, bien qu'il était possible de calculer aussi bien les constituants que les relations à l'aide de ce parseur, nous n'avons pu effectuer que la première tâche dans les délais impartis. Ceci confronte les contingences de la campagne d'évaluation aux possibilités réelles des analyseurs, lorsque ceux-ci sont programmés dans le cadre de la recherche expérimentale en laboratoire.

**Second analyseur profond** <sup>290</sup>. — Ce second analyseur profond, dans son principe, permet la production de tous les arcs possibles conformes à une grammaire de propriétés donnée. Cette analyse se fait en différentes passes, chaque passe engendrant un niveau hiérarchique supplémentaire dans les syntagmes produits. Les constituants EASY n'ayant qu'un niveau de hiérarchie, une seule passe est nécessaire pour les produire. Mais l'analyse a été complétée par des passes supplémentaires afin de produire les relations.

Pour cela, la grammaire EASY a été étendue par l'ajout de catégories syntagmatiques décrivant l'intégralité d'une phrase. Le choix a été fait, au sein de cette grammaire étendue, de présenter chaque relation comme une contrainte de dépendance particulière entre 2 (ou 3 pour certaines relations) constituants d'un syntagme supérieur. Ainsi, la relation <sujet-verbe> est une contrainte de dépendance entre le groupe nominal et le groupe verbal au sein de la catégorie phrase.

Une fois produit par l'analyseur l'ensemble des constituants possibles d'une phrase, on recherche sa meilleure couverture possible (le syntagme le plus englobant, si possible du genre phrase) ainsi que l'ensemble de ses constituants jusqu'aux catégories lexicales. Les constituants EASY correspondent alors au premier niveau de cet assemblage et les relations, à toutes les dépendances nécessaires à sa construction.

---

<sup>290</sup>C'est cet analyseur qui a obtenu les meilleurs résultats. C'est également celui-ci qui a fait l'usage le plus strict de la grammaire présentée en section suivante.

## 3 La grammaire

Grammaire basée sur le guide d'annotations syntaxiques de référence Peas, version 1.6. Le jeu de traits des étiquettes de catégories est celui du dictionnaire du LPL.

**Pondérations possibles.** — S'il y a des pondérations à introduire dans la grammaire, il faut essayer d'abord de mettre un maximum de poids sur les propriétés de OBLIGATION, FACULTATIVITÉ, UNICITÉ et LINÉARITÉ (i.e. si l'une d'elles est violée, **ne pas caractériser de catégorie**), un poids moyen sur EXIGENCE et EXCLUSION, et minimum sur DÉPENDANCE. Dans tous les cas il est impératif, pour caractériser une catégorie, que la propriété d'OBLIGATION (y compris l'unicité qui est implicite) soit satisfaite.

**Syntagmes non Easy.** — On s'en sert pour définir des propriétés pour la description des relations entre formes qui ne font partie des mêmes groupes Easy, ou entre groupes (Easy). Pour les repérer on note les syntagmes non Easy en les faisant précéder d'un \*. Les syntagmes utilisés comme constituants dans d'autres syntagmes ont le même jeu de traits que leur tête (p.ex. le NV a les mêmes traits que son V).

**Propriétés distantes ?.** — Il faudrait pouvoir donner des propriétés qui mettent en relation un constituant immédiat et un constituant non immédiat (parce que le 2<sup>e</sup> C.I. est un constituant non-Easy). P.ex., si j'ai un syntagme S1 de niveau n qui a deux constituants immédiats S2 et \*S3 de niveau n-1, dont l'un (\*S3) a deux constituants immédiats S4 et S5 de niveau n-2, je devrais pouvoir définir une relation, dans les propriétés de S1, entre S2 (C.I. de S1) et S4 (C.I. de \*S3 qui est C.I. de S1). Notation conventionnelle utilisée ici pour ces cas :

$$(*S3 : S4) \rightsquigarrow_{Relation} S2.$$

### 3.1 Groupes Syntaxiques

#### 3.1.1 Noyau verbal (NV)

**Constitution d'après Peas 1.6.** — Toujours un et un seul verbe quelconque (**V.....**) (y compris auxiliaire, modal, etc.).

NB : Quand on a un adverbe entre l'auxiliaire et le verbe, on a donc 3 groupes en tout (dont deux NV). En cas de NV non unaire, seules les formes qui précèdent le V y sont adjointes (pour les formes, cf. ci-dessous).

En option : un (plusieurs) pronom(s) clitique(s), y compris le pronom sujet (**Pp...[nad].**), précédant le verbe, ou alors le suivant et étant de la forme *-pronom* (i.e. avec un tiret, comme dans *a-t-on*). D'une manière générale ça fait : tous les pronoms clitiques qui précèdent, plus les nominatifs qui suivent ET qui satisfont une dépendance (accord en personne et en nombre).

En option : adverbe négatif *ne* (**Rpn**) précédant le verbe. Les autres adverbes négatifs consti-

tuent des GR, sauf s'ils figurent entre le *ne* et le verbe (à l'infinitif uniquement). Dans le guide, il est dit que seuls des adverbes négatifs peuvent figurer entre un *ne* et un verbe infinitif, or ce n'est pas exact. Donc, puisqu'il est dit que *ne* fait toujours partie du NV, alors on englobe dans le NV tout ce qu'il y a entre le *ne* et le V, du moment que c'est un adverbe. Pour les constituants d'autres natures qui s'inséreraient (plus complexes), on n'a pas de proposition. En bref : *ne* + le verbe qui suit, ou *ne* + Rpd + le verbe qui suit (dans l'ordre - normalement ça n'arrive qu'avec les infinitifs).

Tout ce qui est constitué d'un verbe + les éléments ci-dessus est un NV (i.e. quel que soit le mode du verbe), sauf certains participes, qui dans ce cas devraient être étiquetés comme étant des adjectifs. (cf. B.1.II vs. B.2 sur le guide). → Pas de moyen automatique de faire une distinction de forme efficace entre ce qui est considéré comme participe NV vs. participe GA. La distinction sera donc faite en fonction de l'étiquetage : les participes étiquetés Adj formeront des GA, ceux étiquetés Vb formeront des NV, et on aura une chance sur deux de tomber sur l'étiquette juste.

NB : Les particules euphoniques (*-t*, *l'*), d'après ce qu'on voit dans les exemples de Peas, sont étiquetées comme une seule unité avec le clitique qui les suit – donc on n'en tient pas compte dans les règles.

NB : Il n'est pas fait allusion aux pronoms réfléchis, mais on va considérer qu'ils sont à traiter comme des clitics.

## Grammaire de Propriétés du NV. — cf. figure 1.

OBLIGATION	Toujours un et un seul V dans un NV : <b>Oblig = V.....</b> <i>Rappel : Les éléments cités comme "obligatoires" sont implicitement uniques.</i>
FACULTATIVITÉ	Pronoms clitics (Pp non obliques) : <b>Facult = Pp...n.</b> <b>Facult = Pp...j.</b> <b>Facult = Pp...d.</b> Pronoms réfléchis : <b>Facult = Px.....</b> Adverbes : <b>Facult = R..</b>
UNICITÉ	Le <i>ne</i> est unique : <b>Unic = Rpn</b>
EXIGENCE	Les adverbes qui ne sont pas des <i>ne</i> exigent un V inf et un <i>ne</i> : <b>Rpñ ⇒ { V.n..... ∧ Rpn }</b> <i>Note : Rpñ signifie R [ scat≠p ∧ type≠n ]</i>
EXCLUSION	Le V qui n'est pas infinitif exclut tout autre adverbe que le <i>ne</i> : <b>V.ñ..... ⇏ Rpñ</b> <i>Rappel : Les propriétés d'exclusion sont réversibles (si a exclut b alors b exclut a).</i>
LINÉARITÉ	Les pronoms clitics non nominatifs et les réfléchis précèdent le verbe : <b>{ Pp...ñ. ∨ Px..... } &lt; V.....</b> Les adverbes précèdent les pronoms clitics accusatifs et datifs, les réfléchis, et les verbes : <b>R.. &lt; { Pp...ñ. ∨ Px..... ∨ V..... }</b> L'adverbe <i>ne</i> précède les autres adverbes (pour le NV inf) : <b>Rpn &lt; Rpñ</b>
DÉPENDANCE	Relation Suj-V : <b>Pp...n. ∼<sub>Suj-V</sub> V.....</b> Relation Cod-V : <b>P...j. ∼<sub>Cod-V</sub> V.....</b> Relation Cpl-V : <b>P...d. ∼<sub>Cpl-V</sub> V.....</b>

FIG. 1 – Description du NV.

Notes & remarques :

– Tel quel, on ne traite pas ce qui peut se trouver dans un NV, entre un *ne* et un verbe



(infinitif), et qui n'est pas un adverbe.

### 3.1.2 Non EASY : Syntagme verbal (\*SV)

**Constitution.** — Le \*SV contient le NV principal (**NVm.....**), les éventuels NV qui y sont rattachés (auxiliaires etc. : **NV[ae].....**) et les compléments de verbe (et circonstants) divers.

NB : Quand on a une forme auxiliaire + verbe, c'est le Vm (ou NVm) qui est principal, alors que quand on a une forme modal + verbe c'est le modal qui est considéré comme le verbe principal.

**Grammaire de Propriétés du \*SV.** — cf. figure 2.

Notes et remarques :

- Subordonnées infinitives (NV.n et \*SV.n) – COD-V ou CPL-V : suivant que l'on peut pronominaliser l'infinitif (Cod) ou non (Cpl). Je n'ai pas trouvé de moyen de faire la distinction entre les deux, c'est à discuter. Pour l'instant on pose que l'on génère les deux, comme ça on trouve forcément la bonne.
- GN et \*SN — COD-V ou MOD-V : même remarque que ci-dessus.

OBLIGATION	Toujours un NV principal dans un *SV, ou un *SV (coord ou juxtap) : <b>Oblig</b> = { NVm..... ∨ *SV..... }
FACULTATIVITÉ	D'autres NV (auxiliaires, principaux à l'infinitif ou au participe) : <b>Facult</b> = NVa..... <b>Facult</b> = NVe..... <b>Facult</b> = NVmn..... <b>Facult</b> = NVmp..... Compléments directs : <b>Facult</b> = GN <b>Facult</b> = *SN <b>Facult</b> = GA <b>Facult</b> = *SA <b>Facult</b> = GR <b>Facult</b> = *SR Compléments indirects : <b>Facult</b> = GP <b>Facult</b> = PV <b>Facult</b> = *SP Propositions subordonnées complément : <b>Facult</b> = *SVmn..... <b>Facult</b> = *SVmp..... <b>Facult</b> = *SUB Pour coord et juxtap : <b>Facult</b> = *SV..... <b>Facult</b> = Cc.
UNICITÉ	Chaque forme verbale : <b>Unic</b> = NVa..... <b>Unic</b> = NVe..... <b>Unic</b> = NVmn..... <b>Unic</b> = NVmp..... Un certain nombre de compléments : <b>Unic</b> = GA <b>Unic</b> = *SA <b>Unic</b> = GR <b>Unic</b> = *SR
EXIGENCE	Pour coord et juxtap. : <b>*SV.....</b> ⇒ { <b>*SV.....</b> ∨ Cc. } Cc. ⇒ <b>*SV.....</b>
LINEARITÉ	Les NV (tous) précèdent tout, sauf les syntagmes adjectivaux et adverbiaux : <b>NV.....</b> < { <b>GN</b> ∨ <b>*SN</b> ∨ <b>GP</b> ∨ <b>PV</b> ∨ <b>*SP</b> ∨ <b>*SV</b> ∨ <b>*SUB</b> } Les NV aux. précèdent aussi tout le reste : { <b>NVa.....</b> ∨ <b>NVe.....</b> } < { <b>NVm.....</b> ∨ <b>GA</b> ∨ <b>*SA</b> ∨ <b>GR</b> ∨ <b>*SR</b> } Pour coord. : { <b>*SV.....</b> ∧ Cc. } < <b>*SV.....</b>
DÉPENDANCE	Relation Aux-V : <b>NVa.....</b> ∼ <sup>Aux_V</sup> <b>NV.ps....</b> <i>NB : que ce soit le vb principal ou l'auxiliaire être.</i> <b>NVe.....</b> ∼ <sup>Aux_V</sup> <b>NVmp.....</b> Relation Cod-V : { <b>GN</b> ∨ <b>NV.n.....</b> } ∼ <sup>Cod_V</sup> <b>NVm.....</b> Relations Cod-V à distance : { ( <b>*SUB</b> : <b>NVm.....</b> ) ∨ ( <b>*SN</b> : <b>GN</b> ) ∨ ( <b>*SV.n.....</b> : <b>NV.n.....</b> ) } ∼ <sup>Cod_V</sup> <b>NVm.....</b> Relation Cpl-V : { <b>GP</b> ∨ <b>PV</b> ∨ <b>NV.n.....</b> } ∼ <sup>Cpl_V</sup> <b>NVm.....</b> Relations Cpl-V à distance : { ( <b>*SP</b> : <b>PV</b> ) ∨ ( <b>*SP</b> : <b>GP</b> ) ∨ ( <b>*SV.n.....</b> : <b>NV.n.....</b> ) } ∼ <sup>Cpl_V</sup> <b>NVm.....</b> Relation Mod-V : { <b>GR</b> ∨ <b>GN</b> } ∼ <sup>Mod_V</sup> <b>NVm.....</b> Relations Mod-V à distance : { ( <b>*SR</b> : <b>GR</b> ) ∨ ( <b>*SN</b> : <b>GN</b> ) ∨ ( <b>*SUB</b> : <b>NV</b> ) } ∼ <sup>Mod_V</sup> <b>NVm.....</b> Relation Atb-So : <b>GA</b> ∼ <sup>Atb_So(-,-,o)</sup> { <b>GN</b> ∨ <b>NV.n.....</b> } Relations Atb-So à distance : { <b>GA</b> ∨ ( <b>*SA</b> : <b>GA</b> ) } ∼ <sup>Atb_So(-,-,s)</sup> ( <b>NV.....</b> : Pp..n.) { <b>GA</b> ∨ ( <b>*SA</b> : <b>GA</b> ) } ∼ <sup>Atb_So(-,-,o)</sup> { <b>GN</b> ∨ ( <b>*SN</b> : <b>GN</b> ) ∨ <b>NV.n.....</b> ∨ ( <b>*SV.n.....</b> : <b>NV.n.....</b> ) } Relations Coord et Juxt à distance : Cc. ∼ <sup>Coord</sup> { ( <b>*SV.....</b> : <b>NV.....</b> ) ∧ ( <b>*SV.....</b> : <b>NV.....</b> ) } <i>NB : Il faut noter les 3 éléments dans cet ordre pour la relation de Coord.</i> ( <b>*SV.....</b> : <b>NV.....</b> ) ∼ <sup>Juxt</sup> ( <b>*SV.....</b> : <b>NV.....</b> )

FIG. 2 – Description du \*SV.

### 3.1.3 Groupe Nominal (GN)

Constitution d'après Peas 1.6. — Obligatoires :

- un et un seul nom commun (**N[ck]....**)<sup>291</sup>,
- ou bien un (ou plusieurs) nom(s) propre(s) (plusieurs si juxtaposés – mais pas un nom commun et un nom propre) (**Np....**),
- ou bien un et un seul pronom non clitique (**Pp...o.** et **Pp'....**),
- ou bien d'un adjectif “transcatégorisé” (**A....**).

Les noms (quels qu'ils soient) peuvent être précédés d'un déterminant (**D.....**), et/ou d'un ou plusieurs adjectifs (**A....**) avec ses (leurs) éventuels modificateurs (modificateurs à voir, ne sont pas indiqués dans le guide à la section GA – on pense à **A....** et/ou **R..** – grosso modo, tout ce qui est entre le déterminant et le nom). Parmi ce qu'on peut trouver dans un GN entre le Det et le N : une forme constituée de (Adv +) Adj + Coord + (Adv +) Adj.

**Grammaire de Propriétés du GN.** — cf. figure 3. N.B. : Lors de la réalisation de cette grammaire du GN, on a modifié le dictionnaire, en y remplaçant les étiquettes **Nd....** (ex. “nom propre avec déterminant” – noms d'habitants, de pays et sigles divers) soit par **Nc....** (noms d'habitants), soit par **Np....** (noms de pays et sigles).

OBLIGATION	Un nom, ou un pronom ni clitique ni réfléchi, ou un adjectif “transcatégorisé” : <b>Oblig={N..... ∨ Pp...o. ∨ Pd..... ∨ Pi..... ∨ Pr..... ∨ Ps..... ∨ Pt..... ∨ A.... }</b>
FACULTATIVITÉ	Adjonctions : <b>Facult = D.....</b> Expansions : <b>Facult = Np....</b> <b>Facult = A....</b> <b>Facult = Rpi'</b> <b>Facult = Cc.</b>
EXIGENCE	L'adjectif exige au moins un déterminant (si “transcatégorisé”), un nom ou un adjectif : Exemples : <i>la bleue, brave chien, brave petit</i> <b>A....</b> ⇒ { <b>D.....</b> ∨ <b>N.....</b> ∨ <b>A....</b> } Le Cc. exige deux adjectifs : <b>Cc.</b> ⇒ { <b>A....</b> ∧ <b>A....</b> } L'adverbe (modifieur de l'Adj.) exige un adjectif : <b>Rpi'</b> ⇒ <b>A....</b> <i>Rmq : enlever cette propriété pour prendre en compte des formes tq des super vacances.</i>
EXCLUSION	Le nom commun exclut tout autre nom : <b>Np....</b> ↯ <b>N.....</b> <i>NB : essayer sans ça pour voir si c'est plus couvrant pour les cas complexes.</i>
LINÉARITÉ	Le déterminant précède tout (nom, adjectif, adverbe et/ou pronom) sauf le coord. : <b>D.....</b> < { <b>N.....</b> ∨ <b>P.....</b> ∨ <b>R..</b> ∨ <b>A....</b> } L'adjectif et l'adverbe modifieur précèdent le nom : { <b>R..</b> ∨ <b>A....</b> } < <b>N.....</b> Coordination : { <b>A....</b> ∧ <b>Cc.</b> } < <b>A....</b>
DÉPENDANCE	Le nom (ou l'adjectif “transcatégorisé”) donnent le <b>genre</b> et le <b>nombre</b> du dét. : <b>D.....</b> ∼ { <b>N.....</b> ∨ <b>A....</b> } Relation Mod-N : { <b>A....</b> ∨ <b>Dk.....</b> } ∼ <sup>Mod-N</sup> <b>N.....</b> <b>Np....</b> ∼ <sup>Mod-N</sup> <b>Np....</b> Relation Mod-A : { <b>R..</b> ∨ <b>A....</b> ∨ <b>Dk.....</b> } ∼ <sup>Mod-A</sup> <b>A....</b> Relation Coord : <b>Cc.</b> ∼ <sup>Coord</sup> { <b>A....</b> ∧ <b>A....</b> }

FIG. 3 – Description du GN.

Notes et remarques :

- On ne note pas que le **Nc....** exige un Det parce qu'un certain nombre de GN seront uniquement constitués d'un **N.....** (comme dans *une tarte MAISON*). Par effet de bord, également, on a bloqué les formes *une tarte maison* (unicité du nom commun).
- A noter : on doit pouvoir étiqueter comme un seul GN : *de dix à quinze garçons*.

<sup>291</sup>**Nk....** est un nom cardinal : *celui qui a une paire de DEUX ouvre la partie*.

- *Autour de* et *près de* sont des déterminants (et caractérisent un GN) quand ils sont suivis par un adjectif numéral (*autour de dix personnes*), et des prépositions (caractérisant un GP) dans le cas contraire (*autour de la table*). Vérifier comment la grammaire (et en amont, le dictionnaire) traite(nt) cela et trouver un moyen de traiter de cette façon spécifiquement pour la campagne. (cf. Peas, section D.1.XII)

### 3.1.4 Non EASY : Syntagme nominal (\*SN)

**Constitution.** — C'est un GN avec ses éventuels compléments postposés : GN, GA, \*SA, GP, PV, \*SP, \*SUB, \*SV inf (NV inf), \*SV ppe (NV ppe).

**Grammaire de Propriétés du \*SN.** — cf. figure 4.

OBLIGATION	Un GN ou un *SN (coord ou juxtap) : Oblig = { GN..... ∨ *SN }
FACULTATIVITÉ	Facult = GN..... Facult = GA Facult = *SA Facult = GR Facult = *SR Facult = GP Facult = PV Facult = *SP Facult = *SUB Facult = *SV.n..... Facult = NV.n..... Facult = *SV.p..... Facult = NV.p..... Coord ou juxtap : Facult = Cc. Facult = *SN
EXIGENCE	Coordination et juxtaposition : Cc. ⇒ { GN..... ∨ *SN..... } *SN..... ⇒ { *SN..... ∨ Cc. }
LINÉARITÉ	Le GN précède tout sauf la conj. de coord. : GN..... < { GA ∨ *SA ∨ GR ∨ *SR ∨ GP ∨ PV ∨ *SP ∨ *SUB ∨ *SV..... ∨ NV..... } Le GA, le GP et le PV précèdent les subordonnées : { GA ∨ *SA ∨ GR ∨ *SR ∨ GP ∨ PV ∨ *SP } < { *SUB ∨ *SV..... ∨ NV..... }
DÉPENDANCE	Relation Mod-N : { GA ∨ GP ∨ GR } ∼ <sup>Mod_N</sup> GN..... Relations Mod-N à distance : { (*SA :GA) ∨ (*SP :GP) ∨ (*SN..... :NVm.....) ∨ (*SR :GR) } ∼ <sup>Mod_N</sup> GN..... (*SUB : *SV..... :NVm.....) ∼ <sup>Mod_N</sup> GN..... Relation Mod-A : GP ∼ <sup>Mod_A</sup> GN..... Celle-ci risque de bcp surgénérer pour pas grand-chose. On pourrait limiter en la remplaçant par GP ∼ <sup>Mod_A</sup> (GN..... :A....) Coordination et juxtaposition : Cc. ∼ <sup>Coord</sup> { GN..... ∧ GN..... } GN..... ∼ <sup>Juxt</sup> GN..... Coordination et juxtaposition à distance : Cc. ∼ <sup>Coord</sup> { (*SN..... :GN.....) ∧ (*SN..... :GN.....) } (*SN..... :GN.....) ∼ <sup>Juxt</sup> (*SN..... :GN.....)

FIG. 4 – Description du \*SN.

Notes et remarques :

- Tel quel, on ne note pas les coordinations entre compléments du GN dans le \*SN.

### 3.1.5 Groupe Prépositionnel (GP)

**Constitution d'après Peas 1.6.** — Obligatoire : une et une seule préposition (**S..+.....**, i.e. amalgamée ou non avec un déterminant), suivie de :

- le **GN.....** qu'elle introduit (cf. formation du **GN**), ou
- un adverbe (*de trop*) (**R..**).

Ou alors des formes de pro-GP, comme les formes *dont*, *où*.

**Grammaire de Propriétés du GP.** — *Première possibilité* — analyse à plusieurs niveaux : Plus simple pour fabriquer les groupes et les relations (parce que le GN a déjà eu un certain nombre de relations, il ne reste plus que les spécifiques à la prep+det amalgamés - et encore je ne sais même pas si ces formes existent). cf. figure 5.

OBLIGATION	Une et une seule préposition, ou un pronom relatif (pour <i>dont</i> et <i>où</i> ) : <b>Oblig</b> = { <b>S..+.....</b> , <b>Pr...o.</b> } <i>Note : Les Pr...o. sont à vérifier sur tests.</i>
FACULTATIVITÉ	Un GN, ou un adverbe : <b>Facult</b> = <b>GN.....</b> <b>Facult</b> = <b>Rpt'</b>
UNICITÉ	Tous les éléments sont uniques : <b>Unic</b> = <b>GN.....</b> <b>Unic</b> = <b>Rpt'</b>
EXIGENCE	La préposition exige un GN ou un adverbe : <b>S..+.....</b> $\Rightarrow$ { <b>GN.....</b> $\vee$ <b>Rpt'</b> }
EXCLUSION	Le <b>Pr...o.</b> exclut tout autre élément : <b>Pr...o.</b> $\nRightarrow$ { <b>GN.....</b> $\vee$ <b>Rpt'</b> } Les deux éléments facultatifs s'excluent l'un l'autre : <b>GN.....</b> $\nRightarrow$ <b>Rpt'</b>
LINÉARITÉ	La préposition est toujours au début du GP : <b>S..+.....</b> $\prec$ { <b>GN.....</b> $\vee$ <b>Rpt'</b> }
DÉPENDANCE	Le déterminant amalgamé dans la prep s'accorde avec le GN en genre et en nombre : <b>S..+D.....</b> $\leadsto$ <b>GN.....</b> Relation Mod-N à distance : <b>S..+Dk.....</b> $\leadsto^{Mod\_N}$ ( <b>GN :N.....</b> ) Relation Mod-A à distance : <b>S..+Dk.....</b> $\leadsto^{Mod\_A}$ ( <b>GN :A.....</b> )

FIG. 5 – Description du GP.

*Seconde possibilité* — analyse plate : Plus compliqué pour fabriquer les groupes et les relations, mais peut-être plus simple pour formater les sorties (on n'a pas à "effacer" des GN). cf. figure 6.

Notes et remarques :

- Le seul moyen que je vois a priori pour différencier les *dont* qui sont des Pro-GP non subordonnants de ceux qui sont des subordonnants, c'est de prendre en compte leur contexte, et ça, je sais pas trop comment le faire. Pour l'instant ça n'est donc pas géré.
- Si la préposition est éloignée du GN qu'elle introduit (*avec, dans sa poche, un petit discours*), alors on ne note que le GN (la prep ne reçoit pas d'étiquette de constituant, le GP sera pour Easy une relation et pas un constituant — on reconstituera le syntagme dans un \*SP). Si une même prep est factorisée pour deux GNs (*dans les couloirs et la cour*), alors le premier groupe est noté GP et le deuxième GN. La relation de dépendance entre le GN et la prep sera aussi une relation à donner.
- On ne dit pas comment traiter les formes comme *je suis POUR* : prépositions simples ? ou GP ? ou GA ? A priori plutôt les laisser comme des Preps simples, et envisager une relation

OBLIGATION	Une et une seule préposition, ou un pronom relatif oblique (pour <i>dont</i> et <i>où</i> ) : <b>Oblig</b> = { <b>S..+.....</b> , <b>Pr...o.</b> }
FACULTATIVITÉ	Tous les éléments constitutifs d'un GN : <b>Facult</b> = <b>N.....</b> <b>Facult</b> = <b>Pp...o.</b> <b>Facult</b> = <b>Pd.....</b> <b>Facult</b> = <b>Pi.....</b> <b>Facult</b> = <b>Pr.....</b> <b>Facult</b> = <b>Ps.....</b> <b>Facult</b> = <b>Pt.....</b> <b>Facult</b> = <b>A.....</b> <b>Facult</b> = <b>D.....</b> <b>Facult</b> = <b>Rp<i>ix</i>'</b> <b>Facult</b> = <b>Cc.</b>
UNICITÉ	Tous les éléments sont uniques sauf l'Adv. ( <i>de très près</i> ) et l'Adj. : <b>Unic</b> = <b>N.....</b> <b>Unic</b> = <b>Pp...o.</b> <b>Unic</b> = <b>Pd.....</b> <b>Unic</b> = <b>Pi.....</b> <b>Unic</b> = <b>Pr.....</b> <b>Unic</b> = <b>Ps.....</b> <b>Unic</b> = <b>Pt.....</b> <b>Unic</b> = <b>D.....</b> <b>Unic</b> = <b>Cc.</b>
EXIGENCE	L'adjectif exige au moins un déterminant (si "transcatégorisé"), un nom ou un adjectif : <b>A....</b> $\Rightarrow$ { <b>D.....</b> $\vee$ <b>S..+D.....</b> $\vee$ <b>N.....</b> $\vee$ <b>A....</b> } Le Cc. exige deux adjectifs ou deux noms : <b>Cc.</b> $\Rightarrow$ { <b>A....</b> $\vee$ <b>N.....</b> } { <b>Cc.</b> $\wedge$ <b>A....</b> } $\Rightarrow$ <b>A....</b> { <b>Cc.</b> $\wedge$ <b>N.....</b> } $\Rightarrow$ <b>N.....</b> La préposition non amalgamée exige un nom, ou un pronom, ou un adverbe : <b>S..+D.....</b> $\Rightarrow$ { <b>N.....</b> $\vee$ <b>P.....</b> $\vee$ <b>Rp<i>ix</i>'</b> } La préposition amalgamée exige un nom : <b>S..+D.....</b> $\Rightarrow$ <b>N.....</b>
EXCLUSION	Le Pr...o. exclut tout autre élément : <b>Pr...o.</b> $\nRightarrow$ { <b>N.....</b> $\vee$ <b>P.....</b> $\vee$ <b>A....</b> $\vee$ <b>D.....</b> $\vee$ <b>Rp<i>ix</i>'</b> } Le nom commun exclut tout autre nom : <b>Np.....</b> $\nRightarrow$ <b>N.....</b>
LINÉARITÉ	La préposition est toujours au début du GP : <b>S..+.....</b> $\prec$ { <b>N.....</b> $\vee$ <b>P.....</b> $\vee$ <b>A....</b> $\vee$ <b>Rp<i>ix</i>'</b> } <b>S..+D.....</b> $\prec$ <b>D.....</b> Le déterminant précède tout (nom, adjectif, adverbe et/ou pronom) sauf le coord. : <b>D.....</b> $\prec$ { <b>N.....</b> $\vee$ <b>P.....</b> $\vee$ <b>R..</b> $\vee$ <b>A....</b> } L'adjectif et l'adverbe modifieur précèdent le nom : { <b>R..</b> $\vee$ <b>A....</b> } $\prec$ <b>N.....</b> L'adverbe modifieur précède l'adjectif qu'il modifie (à vérifier sur tests) : <b>R..</b> $\prec$ <b>A....</b> Coordination : { <b>A....</b> $\wedge$ <b>Cc.</b> } $\prec$ <b>A....</b> { <b>N.....</b> $\wedge$ <b>Cc.</b> } $\prec$ <b>N.....</b>
DÉPENDANCE	Le déterminant amalgamé dans la prep s'accorde avec le N en genre et en nombre : { <b>S..+D.....</b> $\vee$ <b>D.....</b> } $\rightsquigarrow$ { <b>N.....</b> $\vee$ <b>A....</b> } Relation Mod-N : { <b>S..+Dk.....</b> $\vee$ <b>A....</b> $\vee$ <b>Dk.....</b> } $\rightsquigarrow^{Mod\_N}$ <b>N.....</b> <b>Np.....</b> $\rightsquigarrow^{Mod\_N}$ <b>Np.....</b> Relation Mod-A : { <b>S..+Dk.....</b> $\vee$ <b>Dk.....</b> $\vee$ <b>R..</b> $\vee$ <b>A....</b> } $\rightsquigarrow^{Mod\_A}$ <b>A....</b> Relation Coord : <b>Cc.</b> $\rightsquigarrow^{Coord}$ { <b>A....</b> $\wedge$ <b>A....</b> } <b>Cc.</b> $\rightsquigarrow^{Coord}$ { <b>N.....</b> $\wedge$ <b>N.....</b> }

FIG. 6 – Description du GP, seconde possibilité.

de dépendance entre cette prep et ce qu'elle introduit (qui peut la précéder dans l'énoncé : *Un café, je suis pour*). En l'état ça n'est pas traité.

### 3.1.6 Groupe verbal introduit par une préposition (PV)

**Constitution d'après Peas 1.6.** — Obligatoire : un et un seul noyau verbal dont le verbe n'est pas conjugué (infinitif ou participe) (**V.n....** ou **V.p....**) + une préposition qui l'introduit (**S..+D.....**).

En plus : tout ce qui se trouve entre les deux (adverbes divers : **R..**).

**Grammaire de Propriétés du PV.** — *Première possibilité : analyse à plusieurs niveaux* cf. figure 7.

OBLIGATION	La préposition est obligatoire : <b>Oblig = S..+P.....</b>
FACULTATIVITÉ	Le groupe verbal inf. ou ppe, et les éventuels éléments figurant entre la Prep et lui : <b>Facult = NV.n....</b> <b>Facult = NV.p....</b> <b>Facult = R..</b>
UNICITÉ	Le groupe verbal est unique : <b>Unic = NV.....</b>
EXIGENCE	La Prep (obligatoire) exige un groupe verbal (facult.) : <b>S..+P..... ⇒ NV.....</b> L'adverbe et les preps exigent un groupe verbal : <b>R.. ⇒ NV.....</b>
LINÉARITÉ	La Prep est toujours devant tout le reste : <b>S..+P..... &lt; { R.. ∨ NV..... }</b> L'adverbe est entre la prep et le NV : <b>R.. &lt; NV.....</b>

FIG. 7 – Description du PV.

*Seconde possibilité : analyse plate* cf. figure 8.

OBLIGATION	La préposition est obligatoire : <b>Oblig = S..+P.....</b>
FACULTATIVITÉ	Le verbe inf. ou ppe, les pronoms et adverbes du NV : <b>Facult = V.n....</b> <b>Facult = V.p....</b> <b>Facult = Pp...j.</b> <b>Facult = Pp...d.</b> <b>Facult = Px....</b> <b>Facult = R..</b>
UNICITÉ	<b>Unic = V.....</b>
EXIGENCE	La Prep (obligatoire) exige un V (facult.) : <b>{ S..+P..... ∨ R.. ∨ P..... } ⇒ V.....</b> Les adverbes qui ne sont pas des <i>ne</i> exigent un V inf : <b>Rpt' ⇒ V.n....</b>
EXCLUSION	Le V qui n'est pas inf. (donc qui est ppe) exclut tout autre adverbe que le <i>ne</i> : <b>V.p.... ≠ Rpt'</b>
LINÉARITÉ	La Prep est toujours devant tout le reste : <b>S..+P..... &lt; { R.. ∨ V..... ∨ P..... }</b> Les adverbes sont juste après la Prep : <b>R.. &lt; { V..... ∨ P..... }</b> L'adverbe <i>ne</i> précède les autres adverbes (pour le NV inf) : <b>Rpn &lt; Rpt'</b> Les pronoms précèdent le verbe : <b>P..... &lt; V.....</b>
DÉPENDANCE	Relation Cod-V : <b>Pp...j. <math>\leadsto</math> <sup>Cod</sup> V V.....</b> <b>Px...j. <math>\leadsto</math> <sup>Cod</sup> V V.....</b> Relation Cpl-V : <b>Pp...d. <math>\leadsto</math> <sup>Cpl</sup> V V.....</b> <b>Px...d. <math>\leadsto</math> <sup>Cpl</sup> V V.....</b>

FIG. 8 – Description du PV, seconde possibilité.

### 3.1.7 Non EASY : Syntagme prépositionnel (\*SP)

**Constitution.** — Il sert à relier les SPreps dont la préposition est à distance, et à former des gros SPreps avec es éventuels compléments du \*SN qu'ils contiennent.

**Grammaire de Propriétés du \*SP.** — cf. figure 9.

OBLIGATION	Un GP ou un PV ou une Prep : <b>Oblig</b> = { S..+D..... ∨ GN ∨ PV }
FACULTATIVITÉ	<b>Facult</b> = GP <b>Facult</b> = GN <b>Facult</b> = *SN <b>Facult</b> = GA <b>Facult</b> = *SA <b>Facult</b> = GR <b>Facult</b> = *SR <b>Facult</b> = *SUB <b>Facult</b> = NV.n.... <b>Facult</b> = *SV.n.... <b>Facult</b> = NV.p.... <b>Facult</b> = *SV.p.... <b>Facult</b> = Cc.
EXIGENCE	Compléments du GP et du PV : { GP ∨ PV } ⇒ { GP ∨ GN ∨ *SN ∨ GA ∨ *SA ∨ GR ∨ *SR ∨ *SUB } Exigence de la Prep : S..+D..... ⇒ { GN ∨ *SN ∨ NV ∨ *SV ∨ GA ∨ *SA ∨ GR ∨ *SR ∨ *SUB } Autres : *SUB ⇒ { GP ∨ PV ∨ GN ∨ *SN } Cc. ⇒ { GN ∨ *SN } { Cc. ∧ GN } ⇒ GN { Cc. ∧ *SN } ⇒ *SN
LINÉARITÉ	La préposition est toujours au début du *SP : S..+D..... < { GN ∨ *SN ∨ NV ∨ *SV ∨ GA ∨ *SA ∨ GR ∨ *SR ∨ *SUB } La subordonnée est en finale : { GN ∨ *SN ∨ NV ∨ *SV ∨ GA ∨ *SA ∨ GR ∨ *SR } < *SUB Coordination : { Cc. ∧ GN } < GN { Cc. ∧ *SN } < *SN
DÉPENDANCE	Relation Comp : S..+..... ∼ <sup>Comp</sup> { GN ∨ NVm.... } S..+..... ∼ <sup>Comp</sup> { (*SN :GN) ∨ (*SV :NV) } Relation Mod-R (va surgénérer à tous les coups) : GR ∼ <sup>Mod-R</sup> S..+..... (*SR :GR) ∼ <sup>Mod-R</sup> S..+..... Relation Coord : Cc. ∼ <sup>Coord</sup> { GN ∧ GN } Cc. ∼ <sup>Coord</sup> { (*SN :GN) ∧ (*SN :GN) }

FIG. 9 – Description du \*SP.

### 3.1.8 Groupe Adjectival (GA)

**Constitution d'après Peas 1.6.** — Un adjectif (non épithète-antéposé sinon il est dans le GN) **A....**, ou un participe passé employé comme adjectif (cf. ce qui a été dit à ce propos ci-dessus : normalement ils seront étiquetés comme Adj – ou alors trouver un moyen de faire la différence).

**Grammaire de Propriétés du GA.** — cf. figure 10.

OBLIGATION	Un et un seul adjectif : <b>Oblig</b> = A....
------------	--

FIG. 10 – Description du GA.

### 3.1.9 Non EASY : Syntagme adjectival (\*SA)

**Constitution.** — Un GA avec ses éventuelles expansions : GR (\*SR), coordinations, GP, PV (\*SP, postposés).



Grammaire de Propriétés du **\*SA**. — cf. figure 11.

OBLIGATION	Un GA : <b>Oblig= GA</b>
FACULTATIVITÉ	Expansions préposées : <b>Facult = GR</b> <b>Facult = *SR</b> Expansions postposées : <b>Facult = GP</b> <b>Facult = PV</b> <b>Facult = *SP</b> <b>Facult = *SUB</b> <b>Facult = *SV.p....</b> <b>Facult = NV.p....</b> Coordination et juxtaposition : <b>Facult = GA</b> <b>Facult = Cc.</b>
EXIGENCE	Coordination : $Cc. \Rightarrow \{ GA \vee *SA \}$ $\{ Cc. \wedge GA \} \Rightarrow GA$ $\{ Cc. \wedge *SA \} \Rightarrow *SA$
LINÉARITÉ	Le GR (*SR) précède le GA : $\{ GR \vee *SR \} \prec GA$ Tout le reste le suit : $GA \prec \{ GP \vee PV \vee *SP \vee *SUB \vee *SV \vee NV \}$ Coordination : $\{ Cc. \wedge GA \} \prec GA$ $\{ Cc. \wedge *SA \} \prec *SA$
DÉPENDANCE	Relation Mod-A : $\{ GP \vee NV \vee PV \} \sim^{Mod\_A} GA$ $\{ (*SUB : NV) \vee (*SP : GP) \vee (*SV : NV) \} \sim^{Mod\_A} GA$ Relation Coord : $Cc. \rightsquigarrow^{Coord} \{ GA \wedge GA \}$ $Cc. \rightsquigarrow^{Coord} \{ (*SA : GA) \wedge (*SA : GA) \}$

FIG. 11 — Description du **\*SA**.

### 3.1.10 Groupe Adverbial (GR)

**Constitution d'après Peas 1.6.** — Toujours un et un seul adverbe, sauf *ne* : **Rpn**. Pas de GR à tête adjectivale : dans *voter utile*, *utile* est un GA.

Grammaire de Propriétés du **GR**. — cf. figure 12.

OBLIGATION	Un et un seul adverbe, sauf <i>ne</i> : <b>Oblig={Rpn}</b>
------------	---

FIG. 12 — Description du GR

### 3.1.11 Non EASY : Syntagme Adverbial (\*SR)

**Constitution.** — Deux GR dont l'un modifie l'autre, ou deux GR coordonnés.

Grammaire de Propriétés du **\*SR**. — cf. figure 13.

OBLIGATION	Un GR : <b>Oblig = GR</b>
FACULTATIVITÉ	Un autre GR : <b>Facult = GR</b> Coordination : <b>Facult = Cc.</b>
EXIGENCE	Coordination : <b>Cc. <math>\Rightarrow</math> GR</b>
LINÉARITÉ	Coordination : <b>{ Cc. <math>\wedge</math> GR } <math>\prec</math> GR</b>
DÉPENDANCE	Relation Mod-R : <b>GR <math>\rightsquigarrow^{Mod\_R}</math> GR</b> Relation Coord : <b>Cc. <math>\rightsquigarrow^{Coord}</math> { GR <math>\wedge</math> GR }</b>

FIG. 13 – Description du \*SR.

### 3.1.12 Non EASY : Phrase (\*PHR)

**Constitution.** — Un NV<sub>m</sub> ou un \*SV<sub>m</sub> avec un GN (\*SN) sujet.

**Grammaire de Propriétés de la \*PHR.** — cf. figure 14.

OBLIGATION	Un GN ou un *SN : <b>Oblig = { GN <math>\vee</math> *SN }</b>
FACULTATIVITÉ	Un groupe verbal : <b>Facult = NV<sub>m</sub>....</b> <b>Facult = *SV<sub>m</sub>....</b>
EXIGENCE	<b>{ GN <math>\vee</math> *SN } <math>\Rightarrow</math> { *SN <math>\vee</math> NV }</b>
LINÉARITÉ	<b>{ GN <math>\vee</math> *SN } <math>\prec</math> { *SN <math>\vee</math> NV }</b>
DÉPENDANCE	Relation Suj-V : <b>GN <math>\rightsquigarrow^{Suj\_V}</math> NV</b> <b>{ GN <math>\vee</math> (*SN :GN) } <math>\rightsquigarrow^{Suj\_V}</math> { (*SV :NV) <math>\vee</math> NV }</b>

FIG. 14 – Description du \*PHR.

### 3.1.13 Non EASY : Subordonnée (\*SUB)

**Constitution.** — Un \*SV ou une \*PHR précédés d'un subordonnant (**Cs.** ou **Pr....**).

**Grammaire de Propriétés de la \*SUB.** — cf. figure 15.

## 3.2 Relations

### 3.2.1 Sujet - Verbe : SUJ\_V

Un ou plusieurs sujets pour un seul verbe conjugué (cela signifie que s'il y a un auxiliaire, c'est lui qui entre dans la relation Suj-V –temps composés–).

NB : erreur dans l'exemple 32 : le verbes sont notés NV et non F. Puisque ça va à l'encontre de ce qui est dit plus tôt, on n'en tient pas compte.

OBLIGATION	Un subordonnant : <b>Oblig</b> = { <b>Cs.</b> ∨ <b>Pr....</b> }
FACULTATIVITÉ	Un groupe verbal ou une phrase : <b>Facult</b> = <b>NVm....</b> <b>Facult</b> = <b>*SVm....</b> <b>Facult</b> = <b>*PHR</b>
EXIGENCE	{ <b>Cs.</b> ∨ <b>Pr....</b> } $\Rightarrow$ { <b>NV</b> ∨ <b>*SV</b> ∨ <b>*PHR</b> }
LINÉARITÉ	{ <b>Cs.</b> ∨ <b>Pr....</b> } $\prec$ { <b>NV</b> ∨ <b>*SV</b> ∨ <b>*PHR</b> }
DÉPENDANCE	Relation Suj-V : <b>Pr...n.</b> $\leadsto^{Suj-V}$ <b>NVm....</b> <b>Pr...n.</b> $\leadsto^{Suj-V}$ { (*SN :NVm....) ∨ (*PHR :NVm....) } Relation Cod-V : <b>Pr...j.</b> $\leadsto^{Cod-V}$ <b>NVm....</b> <b>Pr...j.</b> $\leadsto^{Cod-V}$ { (*SN :NVm....) ∨ (*PHR :NVm....) } Relation Cpl-V : <b>Pr...d.</b> $\leadsto^{Cpl-V}$ <b>NVm....</b> <b>Pr...d.</b> $\leadsto^{Cpl-V}$ { (*SN :NVm....) ∨ (*PHR :NVm....) } Relation Comp : <b>Cs.</b> $\leadsto^{Comp}$ <b>NVm....</b> <b>Cs.</b> $\leadsto^{Comp}$ { (*SV :NVm....) ∨ (*PHR :NVm....) }

FIG. 15 – Description du \*SUB.

Constitution possible de la relation : cf. figure 16.

Syntagme porteur	SUJ	V
NV	P...n.	Vm....
*PHR	GN	NVm....
*SUB	Pr...n.	NVm....
*SUB	GN	NV.p...

FIG. 16 – Description du SUJ-V.

### 3.2.2 Auxiliaire - Verbe : AUX\_V

Relation entre un auxiliaire et le verbe auquel il se rapporte. Uniquement valable pour les verbes auxiliaires (*être* et *avoir*), mais pas pour les verbes modaux (p.ex. *pouvoir*) qui sont considérés comme des verbes principaux (NB : dans le dico, les modaux ne sont pas repérés comme tels, et sont notés comme des principaux). Le passif contient une relation auxiliaire-verbe.

NB : On note la relation entre les formes et non entre les NV parce que c'est ce qui est dit dans le paragraphe et ce qui est noté dans les exemples ; cependant dans le tableau récapitulatif il n'est plus question de relations entre formes mais entre NV. On ne tient pas compte de ceci.

cf. figure 17.

Syntagme porteur	AUX	V
*SV	Va....	Vmp....
*SV	Va....	Veps...
*SV	Ve....	Vmp....

FIG. 17 – Description du AUX-V.

### 3.2.3 Objet direct - Verbe : COD\_V

Relation entre les groupes objet direct et les verbes auxquels ils se rapportent (on prend le verbe principal dans le cas des temps composés, ou alors –au choix– le NV dont il fait partie).

NB : Il y a vraisemblablement une erreur dans l'énoncé 32 : un PV considéré comme COD d'un Verbe (alors que d'après le protocole ça devrait être un Cpl-V).

cf. figure 18.

Syntagme porteur	COD	V
NV	P...j.	Vm.....
*SV	GN	Vm.....
*SV	V.n....	Vm.....
*SV	*SUB :NVm.....	Vm.....
*SUB	Pr...j.	Vm.....

FIG. 18 – Description du COD-V

### 3.2.4 Complément - Verbe : CPL\_V

Caractérise les relations entre un GP ou PV et le verbe principal auquel il se rapporte. (Y compris les compléments d'agent du passif.)

cf. figure 19.

Syntagme porteur	CPL	V
NV	P...d.	Vm.....
*SV	PV	Vm.....
*SV	GP	Vm.....
*SV	V.n...	Vm.....
*SV	Pr...d.	Vm.....

FIG. 19 – Description du CPL-V

### 3.2.5 Modifieur - Verbe : MOD\_V

Caractérise les compléments de verbe tels que les GR (y compris “pas”), les GN circonstanciels, les propositions circonstancielles. cf. figure 20.

Syntagme porteur	MOD	V
*SV	GR	Vm.....
*SV	GN (circonstanciels)	Vm.....
*SV	*SUB :NVm.....	Vm.....

FIG. 20 – Description du MOD-V

### 3.2.6 Complémenteur : COMP

Principalement pour les subordonnées conjonctives. (?) Permet de relier la préposition et le GN correspondant quand ils sont éloignés (et ne forment donc pas un GP) – idem pour les PV. cf. figure 21.

Syntagme porteur	COMP	V
*GP	S..+.....	GN
*GP	S..+.....	NVm.....
*SUB	Cs.	Vm.....

FIG. 21 – Description du COMP.

### 3.2.7 Attribut - Sujet/Objet : ATB\_SO

Entre l'attribut et l'un des verbes l'introduisant (sur liste). L'attribut peut être un SA ou autre chose. Ici on ne va traiter que les cas avec le SA parce que je ne sais pas comment traite le reste puisque ça demande un accès lexical. cf. figure 22.

Syntagme porteur	ATB	de...	Suj./Obj.
Phrase	(*SV :GA)	GN	s
*SV	GA	GN	o

FIG. 22 – Description du ATB-SO

### 3.2.8 Modifieur - Nom : MOD\_N

cf. figure 23.

### 3.2.9 Modifieur - Adjectif : MOD\_A

Rattachés à la forme de l'adjectif auxquels ils se rapportent (c'est ce qui est dit dans le texte, mais ça ne correspond pas à ce qui est donné dans les exemples). Ne pas oublier ceux qui modifient un adjectif dans un GN sans nom (cf. D.1.III). cf. figure 24.

Syntagme porteur	MOD	N
GN	A....	N.....
*SN	GA	GN
*SN	SP	GN
*SN	(*SUB :NVm.....)	GN
*SN	GR	GN
GN	Dk.....	N
GP	Dk.....	N
GN	Np....	Np....
GP	Np....	Np....

FIG. 23 – Description du MOD-N.

Syntagme porteur	MOD	A
*SA	GP	GA
*SA	*SUB :NVm.....	GA
*SN	GP	GN ou (GN :A....)

FIG. 24 – Description du MOD-A.

Exemple : la plus belle (GN) de la collection (GP) – le GP est un modifieur de belle.

### 3.2.10 Modifieur - Adverbe : MOD\_R

cf. figure 25.

Syntagme porteur	MOD	Adverbe
*SR	GR	GR

FIG. 25 – Description du MOD-R.

### 3.2.11 Modifieur - Préposition : MOD\_P

cf. figure 26.

Syntagme porteur	MOD	Prep
*GP	GR	GP ou (GP :S..+.....)

FIG. 26 – Description du MOD-P.

Exemple pour la 1 : il arrive peu (GR) avant le début (GP) du cours.

### 3.2.12 Coordination : COORD

Relation à trois éléments : 1 le coordonnant (forme), 2 et 3 les éléments coordonnés (les groupes si ce sont des groupes EASY, sinon les NV qu'ils contiennent).

cf. figure 27.

Syntagme porteur	Cc.	X1	X2
X	Coord	X	X

FIG. 27 – Description du COORD.

### 3.2.13 Apposition : APP

Relation de coréférence. Non traitée ici. cf. figure 28.

Syntagme porteur	GN apposé	GN auquel il s'appose
*SN	GN	GN

FIG. 28 – Description du APP

### 3.2.14 Juxtaposition : JUXT

Même problème que pour la coordination. Non traitée ici. cf. figure 29.

Syntagme porteur	X1	X2
*X	X	X

FIG. 29 – Description du JUXT

## 4 Interprétation des résultats

Les résultats de la campagne ne sont pas encore disponibles <sup>292</sup>; il serait donc prématuré de parler d'une évaluation complète de nos résultats. Cependant à la lumière des travaux de développement effectués au cours de la participation à la campagne et des comparaisons entre les différentes sorties obtenues en fonction des techniques employées, on a pu mettre en évidence un certain nombre de caractéristiques (avantages et limites) propres à chacune des approches adoptées, pour le traitement d'un même corpus à l'aide d'une même grammaire.

<sup>292</sup>Ici encore, il s'agit d'une citation intégrale de [Balfourier *et al.*, 2005].

Il est bien évident que, nos parseurs se basant sur des entrées étiquetées et cette étape préalable n'étant pas fiable à 100%<sup>293</sup>, les erreurs provenant de la phase d'étiquetage sont autant de causes d'erreurs systématiques de parsing (bien que dans ce cas ce ne soit pas le parsing lui-même qui soit à mettre en cause). Cela étant dit, même à partir d'entrées correctement étiquetées, nous avons pu lors de nos étapes de test successives mettre en évidence un certain nombre de différences caractéristiques de traitements entre les parseurs. Compte-tenu du fait que les indications d'annotation ne sont pas censées laisser place à l'ambiguïté, cela signifie que dans ces cas seule l'une des réponses est à considérer comme étant celle attendue. Nous allons donc maintenant présenter quelques-unes de ces différences de traitements, qui peuvent provenir soit des techniques de parsing (méthodes d'introduction des "groupes" ou des "relations"), soit de la grammaire elle-même.

**Constituants.** — Comme on l'a vu précédemment, le fichier que les parseurs prennent en input est un texte étiqueté. A chaque token correspond une liste d'étiquettes possibles, et parmi elles une (sous-)liste des propositions retenues par le désambiguïseur. Pour des raisons de simplicité (et de probabilité), les deux premiers parseurs ont retenu comme technique de ne considérer que le premier élément de cette liste pour leurs analyses, alors que le troisième prend en compte toutes les possibilités proposées. Il s'est avéré que dans certains cas cette dernière technique ait permis de "rattrapper" une imprécision récurrente du désambiguïseur, et ait ainsi permis de produire une analyse en constituants juste là où les deux autres étaient nécessairement erronées, par exemple dans les cas fréquents d'ambiguïté entre un déterminant et un amalgame préposition + déterminant qui ont la même forme (*des, du,...*), et pour lesquels le désambiguïseur ne choisissait pas toujours la meilleure possibilité. Cela avait pour conséquence que les deux parseurs se contentant de la première possibilité retenue par le désambiguïseur introduisaient systématiquement, en cas d'erreur, des GN à la place de GP et *vice versa*, alors que le troisième pouvait vérifier la cohérence de son choix et opter pour l'étiquette qu'il évaluait comme fournissant l'analyse la plus satisfaisante.

**Relations.** — Seuls deux des trois parseurs ont produit des relations, le premier et le troisième. Les deux techniques d'introduction des relations relèvent de deux approches totalement différentes :

- Comme il le fait pour les groupes, le premier parseur établit des relations en fonction de leur *constituance* (construction des tables regroupant les candidats possibles) et de leur *linéarité* (introduction des relations pour tous les cas où l'ordre des éléments est celui recherché).
- Le troisième parseur a intégré les relations comme étant des contraintes de *dépendance* caractéristiques, mettant en relation les deux ou trois éléments concernés au sein de syntagmes de niveaux supérieurs aux groupes Easy.

Il en résulte que là où le premier parseur génère non seulement toutes les relations attendues mais aussi un nombre conséquent de relations superflues, le troisième introduit souvent moins de relations, cependant chacune de celles-ci dépend directement de l'exactitude des groupes qui les contiennent et a par conséquent une plus forte probabilité d'être juste. Prenons l'exemple du traitement de l'énoncé suivant :

---

<sup>293</sup>Pas plus que la transcription elle-même : on peut trouver autant de coquilles dans les textes de sources écrites que dans les transcriptions de corpus oraux, coquilles qui augmentent d'autant la probabilité d'erreur d'étiquetage.



- (178) Tout en adoptant le principe de l'adhésion de ces pays, le Conseil européen a précisé que ceux-ci devraient répondre à certains critères et que la capacité de l'Union à accueillir de nouveaux membres devrait également être prise en compte.

Pour cet énoncé, notre premier parseur a introduit 20 relations, dont 7 justes. Notre troisième parseur a introduit 22 relations, dont 11 justes. Le détail est donné en figure 30<sup>294</sup>.

Relation	Parseur 1		Parseur 2	
	Nombre	dont justes (précision)	Nombre	dont justes (précision)
Mod-N	9	4 (44 %)	4	3 (75 %)
Suj-V	3	1 (33 %)	1	1 (100 %)
Cod-V	2	0 (0 %)	2	2 (100 %)
Cpl-V	1	0 (0 %)	4	2 (50 %)
Mod-V	1	0 (0 %)	1	1 (100 %)
Aux-V	1	1 (100 %)	2	2 (100 %)
Comp	2	0 (0 %)	8	0 (0 %)
Coord	1	1 (100 %)	0	0
Total	20	7 (35 %)	22	11 (50 %)

FIG. 30 – Détail des relations pour les parseurs 1 et 3 sur l'énoncé de l'exemple (178).

On voit que même si le nombre total de relations introduites n'est pas très différent (20 dans un cas, 22 dans l'autre), par contre la pertinence de ces relations diffère d'un parseur à l'autre, puisque dans la moitié des cas le parseur 3 ne fournit que des bonnes relations (pour Suj-V, Cod-V, Mod-V et Aux-V), et au moins 50 % de justes dans la moitié des cas restants (Pour Cpl-V et Mod-N). Par contre toutes ses propositions de Comp sont erronées, et il n'a pas trouvé la relation de Coord que le premier parseur a su construire.

**Grammaire.** — Le Protocole d'Evaluation propose une description des Groupes Nominaux indiquant, globalement, que fait partie du GN tout ce qui est compris entre le déterminant et le nom (ou l'objet qui occupe sa place). Cela ne comprend pas, donc, tous les possibles constituants du syntagme nominal (au sens classique) qui figurent après le nom. Le traitement de ce point a été facilement représentable dans notre GP. Cependant, on peut lire plus loin dans le guide PEAS que cette description a une limite : elle n'est plus valable pour les “*éléments en langue étrangère, (l)es formules, (l)es équations mathématiques ou chimiques*”, ni pour les “*références bibliographiques au sein de textes (comme dans les articles)*” ([Gendner & Vilnat, 2004], section D.1.X). Pour ces cas précis, il est dit qu’ “*ils peuvent être regroupés dans des constituants*” (*ibid.*). Les trois exemples présentés (et leurs analyses respectives) dans PEAS sont les suivants :

- (179) a. La patiente présentait <GN> un placenta prævia </GN>.  
 b. L' amour est plein de quiétude et gardé de sentinelles <GP> à toutes les portes des sens </GP>, et <GP> in cunctis sensibus custoditus </GP>.  
 c. Le cas souvent étudié ( <GN> Hamburger 99 </GN> ) est revu dans cet article.

L'exemple (179a) montre un cas où un mot étranger (*prævia*) a reçu un traitement différent de la norme, du fait de sa nature de “mot étranger” : s'il avait été considéré comme un mot “normal” il n'aurait pas été intégré au GN *un placenta*, mais aurait été l'objet d'un GA unaire

<sup>294</sup>Bien évidemment ces pourcentages ne sont calculés que sur l'énoncé donné en exemple et ne sauraient en aucun cas être représentatifs des résultats généraux des parseurs ; il s'agit juste ici d'illustrer les différences de traitements et non d'évaluer les résultats.

(puisque postposé au nom auquel il se rapporte). L'exemple (179b) montre le parallèle fait entre le GP *à toutes les portes des sens* et le groupe suivant, qui est une expression latine, *in cunctis sensibus custoditus*, qui se voit affecter l'étiquette de GP parce que coordonné au GP qui le précède<sup>295</sup>. Enfin, l'exemple (179c) montre le cas exceptionnel de la citation (référence) où un objet qui non seulement n'est pas un Nom propre (une date en l'occurrence) mais qui est également postposé, peut faire partie d'un GN lui-même constitué d'un Nom propre (ce qui théoriquement est impossible, en vertu de la description du GN donnée en section B.2).

Les cas tels que celui de l'exemple (179c) ont pu être décrits sans problème dans la grammaire. En revanche, les exceptions illustrées par les exemples (179a) et (179b) ont été impossibles à exprimer. En effet, les groupes que la grammaire permet d'introduire ne contiennent pas d'information qui permette de savoir si un groupe donné est constitué d'éléments de langue étrangère, de formules ou d'équations mathématiques ou chimiques. Cette information, si elle pouvait figurer, proviendrait de l'étiquette des constituants et non de l'analyse, puisqu'en termes strictement syntaxiques, ces "natures" de groupes ne sont pas pertinentes en soi : leur analyse demeure la même que pour tous les autres groupes. Or ni l'étiquetage fourni pour les besoins de la campagne, ni notre couple étiqueteur-désambiguïseur, ne permet cela : soit le lexique contient le mot à étiqueter (cas des expressions mathématiques) et nos étiquettes ne donnent pas d'information de ce type, soit le lexique ne contient pas le mot à étiqueter (cas des mots de langue étrangère<sup>296</sup>), et dans ce cas la tâche consistera à évaluer quelle est la catégorie la plus probable du mot inconnu en fonction de son contexte, mais rien ne pourra nous permettre d'affirmer qu'il s'agit d'un mot de langue étrangère.

Dans ces cas donc, nous avons fait le choix de nous référer uniquement aux étiquettes fournies aux analyseurs, et aux propriétés de la grammaire. Tous les mots donc, qu'ils soient d'origine étrangère ou non, qu'ils soient des formules mathématiques ou non, ont été traités selon les définitions des groupes fournies en section B, même si cela constitue une limite de l'adéquation des résultats de nos parseurs avec l'annotation de référence.

**Protocole.** — A l'étude détaillée du corpus test, nous avons pu mettre en évidence certains cas dont le traitement demandait un choix, lequel n'était pas spécifié dans PEAS. C'est le cas notamment du traitement des bribes et des amorces : il n'est pas spécifié dans le guide si dans un cas de disfluente, l'on doit considérer chaque occurrence de la répétition (du *reparandum* au *repair*) comme faisant partie du groupe (ce qui donne l'annotation de (180a)), ou alors si l'on ne doit faire figurer dans le groupe que l'occurrence du *repair* (exemple (180b)) :

- (180) a. <NV> il il se tachait </NV> sa sa <NV> il ne ne buvait </NV> que des Blancs  
 b. il <NV> il se tachait </NV> sa sa il ne <NV> ne buvait </NV> que des Blancs

Dans un cas comme celui-ci, nous avons donc du faire un choix parmi les possibilités, ne sachant pas lequel de ces choix avait été fait par les annotateurs lors de l'établissement de la référence.

<sup>295</sup>Où alors du fait de l'analyse syntaxique du groupe latin, mais le parsing du latin n'étant pas l'objet de la grammaire ni de la campagne, et le latin n'étant pas la seule langue possible dans ce cas, nous avons par principe abandonné cette possibilité.

<sup>296</sup>Sauf certains latinismes et anglicismes courants. Il serait d'ailleurs intéressant de définir précisément ce que l'on entend par "élément de langue étrangère" dans le cadre du Protocole, pour savoir si des entrées telles que *a priori*, *cool* ou bien *(e)mail*, *ersatz* en font partie ou non.

En l'occurrence, nous avons choisi de faire figurer toutes les occurrences des répétitions dans les groupes, comme dans l'exemple (180a). Mais si la décision des annotateurs a été différente, alors dans tous ces cas notre annotation sera considérée comme erronée alors qu'il ne s'agit précisément que d'une question de convention et non de justesse d'analyse.



# Table des figures

1	Représentation modulaire des unités en LFG. . . . .	16
2	Représentation modulaire des objets linguistiques en TAG. . . . .	17
3	Représentation modulaire des objets linguistiques en HPSG. . . . .	17
4	Représentation procédurale des objets linguistiques en DS. . . . .	18
5	Représentation procédurale des objets linguistiques dans une Grammaire Minimaliste. . . . .	18
6	Représentation modulaire des analyses en TST. . . . .	19
7	Représentation modulaire des analyses en GB dans FIPS. . . . .	20
8	Architecture générale du formalisme. . . . .	26
9	Architecture de la grammaire. . . . .	26
10	Forme de base de la représentation d'une construction dans la grammaire. . . . .	29
11	Exemple de construction : Le Syntagme Nominal déterminé. . . . .	30
12	Exemple d'identifiant dans le Syntagme Nominal déterminé. . . . .	31
13	Exemple d'informations intrinsèques dans le Syntagme Nominal déterminé. . . . .	33
14	Exemple d'informations extrinsèques dans le Syntagme Nominal déterminé. . . . .	35
15	Architecture la spécification des propriétés. . . . .	37
16	Forme de base d'une définition de type de propriété dans la spécification. . . . .	38
17	Exemple de définition de type de propriété : la Précédence. . . . .	38
18	Définition de type de propriété : forme et identité d'une propriété de Précédence. . . . .	41
19	Définition de type de propriété : Conditions d'une propriété de Précédence. . . . .	44
20	Architecture générale du processus d'analyse. . . . .	46
21	Forme de base d'une construction introduite lors d'une analyse. . . . .	48
22	Structuration des informations intrinsèques dans la grammaire. . . . .	52
23	spécification du type Majeur. . . . .	59
24	spécification du type Mineur. . . . .	59
25	spécification du type Exigence. . . . .	62
26	spécification du type Exclusion. . . . .	62
27	spécification du type Unicité. . . . .	63
28	spécification du type Précédence. . . . .	64
29	spécification du type Adjacence. . . . .	65
30	spécification du type Simultanéité. . . . .	66
31	spécification du type Accord. . . . .	67
32	Un extrait de $\mathcal{GUP}_{sem}$ générant le graphe sémantique de (17). . . . .	73
33	Spécification des types Exigence et Exclusion pour GUST. . . . .	75
34	Extrait de $\mathcal{GP}_{sem}$ correspondant à l'extrait de $\mathcal{GUP}_{sem}$ . . . . .	76
35	Représentation GP du graphe sémantique généré par $\mathcal{GP}_{sem}$ . . . . .	76
36	Un extrait de $\mathcal{GUP}_{synt}$ générant le graphe syntaxique de (17). . . . .	77

37	La structure initiale en $\mathcal{GP}_{synt}$ . . . . .	78
38	Spécification du type Majeur pour GUST. . . . .	78
39	Les règles lexicales dans $\mathcal{GP}_{synt}$ . . . . .	79
40	Les règles sagittales dans $\mathcal{GP}_{synt}$ . . . . .	79
41	Les règles grammaticales dans $\mathcal{GP}_{synt}$ . . . . .	80
42	Spécification des types Mineur et Unicité pour GUST. . . . .	80
43	Représentation syntaxique générée par $\mathcal{GP}_{synt}$ . . . . .	81
44	Extrait de grammaire de correspondance syntaxe-sémantique en GUP. . . . .	81
45	Le graphe sémantique de (17) avant, pendant et après l'interfaçage syntaxe-sémantique. . . . .	82
46	Le graphe sémantique GP doublement polarisé. . . . .	83
47	La grammaire de correspondance $\mathcal{GP}_{synt-sem}$ : les entrées Pierre et pomme. . . . .	84
48	La grammaire de correspondance $\mathcal{GP}_{synt-sem}$ : l'entrée deux. . . . .	85
49	La grammaire de correspondance $\mathcal{GP}_{synt-sem}$ : l'entrée mange. . . . .	86
50	La sortie de l'analyse du graphe sémantique par $\mathcal{GP}_{synt-sem}$ . . . . .	87
51	Résumé des équivalences entre représentations GUP et GP. . . . .	88
52	Le nœud racine en DS. . . . .	92
53	La description d'arbre à l'issue de la première étape d'analyse, en GP. . . . .	93
54	Spécification du type Exigence pour DS. . . . .	93
55	Description d'arbre à l'issue de la deuxième étape, en DS. . . . .	94
56	La description d'arbre à l'issue de la deuxième étape d'analyse, en GP. . . . .	95
57	Une construction figurant la règle d'introduction, en GP. . . . .	95
58	Introduction de nouveaux noeuds en DS. . . . .	96
59	Description de l'arbre à l'issue de la troisième étape, en GP. . . . .	97
60	La construction figurant la règle générale de prédiction, en GP. . . . .	97
61	La construction figurant la prédiction d'un foncteur, en GP. . . . .	98
62	La construction figurant la prédiction d'un argument, en GP. . . . .	98
63	L'entrée lexicale <i>Jean</i> en DS. . . . .	99
64	Application des informations de <i>Jean</i> à l'arbre en DS. . . . .	99
65	Description de l'arbre à l'issue de la quatrième étape, en GP. . . . .	100
66	L'entrée lexicale <i>Jean</i> en GP. . . . .	100
67	Complétion de l'arbre en DS. . . . .	101
68	Description de l'arbre à l'issue de la cinquième étape, en GP. . . . .	102
69	La construction figurant la règle de complétion, en GP. . . . .	102
70	L'entrée lexicale <i>admire</i> en DS. . . . .	102
71	Description d'arbre à l'issue de l'étape 6, en DS. . . . .	103
72	L'entrée lexicale <i>admire</i> en GP. . . . .	103
73	Description de l'arbre à l'issue de la sixième étape, en GP. . . . .	104
74	Description d'arbre à l'issue de la septième étape, en DS. . . . .	105
75	Description de l'arbre à l'issue de la septième étape, en GP. . . . .	106
76	L'entrée lexicale <i>Pierre</i> en DS. . . . .	106
77	Description d'arbre à l'issue de la huitième étape, en DS. . . . .	107
78	Description de l'arbre à l'issue de la huitième étape, en GP. . . . .	108
79	Description d'arbre à l'issue de la neuvième étape, en DS. . . . .	108
80	Description de l'arbre à l'issue de la neuvième étape, en GP. . . . .	109
81	Représentation du sujet en position canonique selon [Crabbé, 2005]. . . . .	113
82	Spécification du type Dominance Immédiate pour MG. . . . .	114
83	Spécification du type Précédence Linéaire pour MG. . . . .	114

84	Représentation de l'équivalent de la figure 81 en GP. . . . .	114
85	Représentation du Sujet Relativisé et de l'Objet Canonique dans la MG de [Crabbé, 2005]. . . . .	115
86	Extrait de la hiérarchie des Arguments Verbaux dans la MG de [Crabbé, 2005]. . . . .	115
87	Représentation de l'équivalent de la figure 85 en GP. . . . .	116
88	Combinaison de fragments de descriptions selon l'approche de [Crabbé, 2005]. . . . .	117
89	Représentation de l'équivalent de la figure 88 en GP. . . . .	117
90	Interprétation des alternatives comme une hiérarchie d'héritage en GP. . . . .	118
91	Représentation des "méta-constructions" avec double héritage. . . . .	119
92	Interprétation des alternatives de combinaisons d'alternatives en GP. . . . .	119
1	Représentation dans la hiérarchie de l'héritage sans la notation "au choix". . . . .	131
2	Représentation dans la hiérarchie de l'héritage avec la notation "au choix". . . . .	131
3	Représentation dans une construction de l'héritage <i>au choix</i> . . . . .	132
4	Représentation dans une construction de l'héritage conjoint. . . . .	132
5	Hiérarchie générale de la grammaire : racine et premiers niveaux. . . . .	136
6	La Construction Grammaticale (CG), racine de la grammaire. . . . .	137
7	La construction syntagmatique générale, Syntagme (Synt). . . . .	138
8	Le X construit (X). . . . .	140
9	Le Syntagme X (SX). . . . .	140
10	La construction paradigmatique générale, Paradigme (Par). . . . .	141
11	Hiérarchie des Noms construits dans la grammaire. . . . .	146
12	Le Nom construit (N). . . . .	147
13	Exemple de hiérarchisation des compléments du Nom. . . . .	148
14	Le Syntagme Nominal (SN). . . . .	149
15	Le Syntagme Nominal exocentrique (SN-x). . . . .	150
16	Le Syntagme Nominal endocentrique (SN-n). . . . .	150
17	Le Syntagme Nominal déterminé (SN-d). . . . .	152
18	Le syntagme nominal non déterminé (SN-nd). . . . .	153
19	Hiérarchie des Verbes construits (V) dans la grammaire. . . . .	158
20	Le Verbe construit (V). . . . .	159
21	Le Verbe à négation (V-n). . . . .	162
22	Le Verbe modal (V-m). . . . .	163
23	Le Verbe au participe passé avec auxiliaire <i>avoir</i> (V-a). . . . .	165
24	Le Verbe au participe passé avec auxiliaire <i>être</i> (V-e). . . . .	166
25	Le Verbe participe passé avec auxiliaires <i>avoir</i> et <i>être</i> (V-ae). . . . .	167
26	Le Verbe infinitif à modal (V-im). . . . .	168
27	Le Verbe infinitif à verbe à montée (V-imm). . . . .	169
28	Le Verbe infinitif avec verbe causatif (V-imc). . . . .	170
29	Le Verbe avec clitique nominatif (V-cn). . . . .	171
30	Le Verbe avec clitique nominatif inversé (V-cni). . . . .	173
31	Le Verbe avec clitique accusatif (V-ca). . . . .	174
32	Le Verbe à clitique accusatif attribut du sujet (V-cas). . . . .	175
33	Le Verbe avec clitique accusatif objet direct (V-cod). . . . .	175
34	Le Verbe avec clitique datif (V-cd). . . . .	176
35	Le Verbe à clitiques nominatif et accusatif objet direct (V-cnod). . . . .	177
36	Le Verbe à clitiques nominatif et accusatif attributif (V-cnas). . . . .	177
37	Le Verbe à sujet et complément indirect clitiques (V-cnd). . . . .	178

38	Le Verbe à clitiques accusatif et datif (V-cad).	179
39	Le Verbe à clitiques nominatif, accusatif et datif (V-cnad).	180
40	Hiérarchie des Syntagmes Verbaux (SV) dans la grammaire.	181
41	Le Syntagme Verbal général (SV).	181
42	Le Syntagme Verbal exocentrique (SV-x).	182
43	Exemple de hiérarchisation des circonstanciels dans un SV endocentrique.	183
44	Le Syntagme Verbal endocentrique (SV-n).	183
45	Le Syntagme Verbal à complément direct (SV-1d).	184
46	Le Syntagme Verbal à attribut du sujet (SV-1as).	185
47	Le Syntagme Verbal transitif direct (SV-1od).	185
48	Le Syntagme Verbal transitif indirect (SV-1oi).	186
49	Le Syntagme Verbal transitif direct avec attribut de l'objet (SV-2ao).	188
50	Le Syntagme Verbal bitransitif (SV-2odi).	188
51	Le Syntagme Verbal tritransitif (SV-3).	189
52	Hiérarchie des constructions Propositionnelles dans la grammaire.	195
53	La Proposition (P).	195
54	La Proposition endocentrique (P-n).	196
55	La Proposition exocentrique (P-x).	196
56	La Proposition exocentrique spécifiée (P-xs).	198
57	La Proposition exocentrique non-spécifiée (P-xns).	198
58	La Proposition exocentrique nominale (P-xn).	199
59	La Proposition exocentrique verbale (P-xv).	199
60	Hiérarchie des constructions de subordination dans la grammaire.	204
61	Le Subordonnant (Sub).	204
62	Le Subordonnant conjonctif (Sub-c).	205
63	Le Subordonnant complétif (Sub-cq).	205
64	Le Subordonnant circonstanciel (Sub-cc).	206
65	Le Subordonnant relatif (Sub-r).	207
66	Le Conjoint (Conj).	208
67	Le Conjoint de subordination (Conj-s).	209
68	Le Conjoint de subordination conjonctif (Conj-sc).	209
69	Le Conjoint de subordination complétif (Conj-scq).	210
70	Le Conjoint de subordination circonstanciel (Conj-scc).	210
71	Le Conjoint de subordination relatif (Conj-sr).	210
72	Le Conjoint de subordination relatif nominatif (Conj-srn).	211
73	Le Conjoint de subordination relatif accusatif (Conj-sra).	211
74	Le Conjoint de subordination relatif datif (Conj-srd).	212
75	Le Conjoint de subordination relatif datif objet indirect (Conj-srdi).	212
76	Le Conjoint de subordination relatif datif circonstanciel (Conj-srdc).	212
77	Hiérarchie des Adjectifs construits dans la grammaire.	215
78	L'Adjectif construit (Adj).	216
79	L'Adjectif qualificatif (Adj-q).	217
80	L'Adjectif numéral (Adj-n).	218
81	Hiérarchie des Entassements Paradigmatiques dans la grammaire.	223
82	Relations dans un énoncé sans EP.	226
83	Relations entre occurrences.	226
84	Relations entre places syntagmatiques.	227



85	Les entassements paradigmatiques comme occurrences complètes de syntagmes. . . . .	228
86	Les EP comme accumulation de places syntagmatiques, 1. . . . .	229
87	Les EP comme accumulation de places syntagmatiques, 2. . . . .	230
88	Les EP comme accumulation de places syntagmatiques, 3. . . . .	231
89	Relations dans la séquence maximale dans (129). . . . .	231
90	L'Entassement Paradigmatique (EP). . . . .	232
91	La Coordination (EP-c). . . . .	235
92	La Coordination asyndétique (EP-ca). . . . .	237
93	Le Coordonnant construit (Coord). . . . .	239
94	Deux structurations possibles de la Coordination. . . . .	239
95	Le Conjoint de coordination (Conj-c). . . . .	240
96	La Coordination syndétique (EP-cs). . . . .	241
97	La Disfluence (EP-d). . . . .	245
1	Construction lexicale <i>Jean</i> dans (157). . . . .	260
2	Construction lexicale <i>admire</i> dans (157). . . . .	260
3	Construction lexicale <i>la</i> dans (157). . . . .	260
4	Construction lexicale <i>maison</i> dans (157). . . . .	261
5	Mise en relation des unités, niveau 1. . . . .	261
6	Interprétation du système de contraintes, niveau 1. . . . .	262
7	Niveau 1 : Le SN-nd <i>Jean</i> . . . . .	263
8	Niveau 1 : Le SN-d <i>la maison</i> . . . . .	264
9	Mise en relation des unités, niveau 2. . . . .	264
10	Interprétation du système de contraintes, niveau 2. . . . .	265
11	Le SV-cd <i>admire la maison</i> . . . . .	265
12	Mise en relation des unités, niveau 3. . . . .	266
13	Interprétation du système de contraintes, niveau 3. . . . .	266
14	La P <i>Jean admire la maison</i> . . . . .	267
15	Exemple de structure de l'analyse d'une construction verbale simple. . . . .	270
16	Exemple de structure de l'analyse d'une construction verbale composée. . . . .	270
17	Exemple de structure de l'analyse d'une construction verbale à contrôle. . . . .	271
18	Exemple de structure de l'analyse d'une construction verbale à modal. . . . .	271
19	L'entrée lexicale correspondant à <i>Marie</i> dans (159). . . . .	272
20	L'entrée lexicale correspondant à <i>a</i> dans (159). . . . .	272
21	L'entrée lexicale correspondant à <i>mangé</i> dans (159). . . . .	273
22	L'entrée lexicale correspondant à <i>une</i> dans (159). . . . .	273
23	L'entrée lexicale correspondant à <i>pomme</i> dans (159). . . . .	273
24	Introduction du SN <i>Marie</i> dans (158). . . . .	274
25	Introduction du SN <i>une pomme</i> dans (158). . . . .	274
26	Introduction du V <i>a mangé</i> dans (159). . . . .	275
27	Introduction du SV <i>mange une pomme</i> dans (159). . . . .	276
28	Introduction de la P <i>Marie mange une pomme</i> dans (158). . . . .	277
29	L'entrée lexicale correspondant à <i>de</i> dans (160). . . . .	278
30	L'entrée lexicale correspondant à <i>manger</i> dans (160). . . . .	278
31	Introduction du SPrep <i>de manger une pomme</i> dans (160). . . . .	279
32	Introduction du SV <i>essaie de manger une pomme</i> dans (160). . . . .	280
33	L'entrée lexicale correspondant à <i>peut</i> dans (161). . . . .	280

34	Introduction du V <i>peut manger</i> dans (161). . . . .	281
35	Exemple de structure de l'analyse d'une proposition à sujet plein. . . . .	282
36	Exemple de structure de l'analyse d'une proposition à sujet clitique. . . . .	282
37	Introduction du V <i>elle aime</i> dans (164). . . . .	283
38	Introduction de la P <i>elle aime le théâtre</i> dans (164). . . . .	284
39	Exemple de structure de l'analyse d'un énoncé présentant des compléments de verbe clitiques. . . . .	285
40	Introduction du V <i>le lui donne</i> dans (166). . . . .	286
41	Exemple de structure de l'analyse d'un énoncé présentant des compléments de verbe plein. . . . .	287
42	Exemple de structure de l'analyse d'un énoncé présentant un complément de verbe clitique et un plein. . . . .	287
43	Exemple de structure de l'analyse d'un énoncé présentant un sujet et un complément de verbe clitiques. . . . .	287
44	Structure de l'analyse d'un énoncé présentant un auxiliaire et un complément de verbe clitique. . . . .	288
45	Introduction du V <i>l'a donné</i> dans (169). . . . .	288
46	Structure de l'analyse d'un énoncé présentant un modal et un complément de verbe clitique. . . . .	289
47	Structure de l'analyse d'un énoncé présentant un modal et un complément de verbe clitique. . . . .	290
48	Structure de l'analyse d'un énoncé présentant un modal et un complément de verbe clitique. . . . .	290
49	Structure de l'analyse d'un énoncé présentant un modal et un complément de verbe clitique. . . . .	291
50	Le SN <i>Pierre</i> dans (174). . . . .	294
51	Le Verbe <i>est</i> dans (174). . . . .	294
52	Le SN <i>président</i> dans (174). . . . .	294
53	Le Coordonnant <i>et</i> dans (174). . . . .	294
54	L'Adjectif construit <i>fier de l'être</i> dans (174). . . . .	295
55	Le Conjoint <i>et fier de l'être</i> dans (174). . . . .	295
56	La Coordination syndétique <i>président et fier de l'être</i> dans (174). . . . .	296
57	Le SV attributif <i>est président et fier de l'être</i> dans (174). . . . .	297
58	Une structure arborescente représentant notre analyse de (174). . . . .	298
59	Le SN-d <i>le cinéma</i> dans (175). . . . .	298
60	La Coordination syndétique <i>Pierre et Marie</i> dans (175). . . . .	299
61	La Coordination asyndétique <i>le cinéma le théâtre</i> dans (175). . . . .	300
62	Le SV transitif direct <i>aime le cinéma le théâtre</i> dans (175). . . . .	301
63	La Proposition <i>Pierre aime le cinéma et Marie le théâtre</i> . . . . .	302
64	Une structure arborescente syntagmatique représentant notre analyse de (175). . . . .	302
65	Une structure arborescente syntaxique représentant notre analyse de (175). . . . .	303
66	Une structure arborescente syntaxique à deux dimensions représentant notre analyse de (175). . . . .	303
1	Exemple de construction pour <i>il y a</i> . . . . .	315
2	Exemple de construction pour <i>il y en a</i> . . . . .	316
3	Exemple de construction pour <i>maisons</i> . . . . .	317
4	Exemple de construction morphématique pour la base <i>maison</i> . . . . .	319
5	Exemple de construction morphématique pour un affixe pluriel. . . . .	319
6	Relativité de la satisfaction de la propriété d'Exclusion <i>b<del>ne</del>e</i> à la zone considérée. . . . .	323
1	Description du NV. . . . .	332
2	Description du *SV. . . . .	334
3	Description du GN. . . . .	335
4	Description du *SN. . . . .	336
5	Description du GP. . . . .	337

---

6	Description du GP, seconde possibilité. . . . .	338
7	Description du PV. . . . .	339
8	Description du PV, seconde possibilité. . . . .	339
9	Description du *SP. . . . .	340
10	Description du GA. . . . .	340
11	Description du *SA. . . . .	341
12	Description du GR . . . . .	341
13	Description du *SR. . . . .	342
14	Description du *PHR. . . . .	342
15	Description du *SUB. . . . .	343
16	Description du SUJ-V. . . . .	343
17	Description du AUX-V. . . . .	343
18	Description du COD-V . . . . .	344
19	Description du CPL-V . . . . .	344
20	Description du MOD-V . . . . .	345
21	Description du COMP. . . . .	345
22	Description du ATB-SO . . . . .	345
23	Description du MOD-N. . . . .	346
24	Description du MOD-A. . . . .	346
25	Description du MOD-R. . . . .	346
26	Description du MOD-P. . . . .	346
27	Description du COORD. . . . .	347
28	Description du APP . . . . .	347
29	Description du JUXT . . . . .	347
30	Détail des relations pour les parseurs 1 et 3 sur l'énoncé de l'exemple (178). . . . .	349



# Bibliographie

- [Abeillé, 1991] Anne Abeillé (1991). *Une grammaire lexicalisée d'arbres adjoints pour le français*. Thèse de doctorat, Université Denis Diderot, Paris 7.
- [Abeillé, 1993] Anne Abeillé (1993). *Les nouvelles syntaxes – Grammaires d'unification et analyse du français*. Armand Colin, Paris.
- [Abeillé, 2002] Anne Abeillé (2002). *Une grammaire électronique du français*. CNRS Editions, Paris.
- [Abeillé, 2003] Anne Abeillé (2003). A lexicalist and construction-based approach to coordinations. In Stefan Mueller, ed., *Proceedings of the HPSG03 Conference*, pages 5–24, Michigan State University, East Lansing.
- [Abeillé & Candito, 2000] Anne Abeillé & Marie-Hélène Candito (2000). FTAG : A lexicalized tree adjoining grammar for french. In A. Abeillé & O. Rambow, eds., *Tree Adjoining Grammars : Formalisms, Linguistic Analysis and Processing*, pages 305–330. CSLI.
- [Abeillé et al., 2000] Anne Abeillé, Marie-Hélène Candito & Alexandra Kinyon (2000). FTAG : Developing and maintaining a wide-coverage grammar for french. In E. Hinrichs, D. Meurers & S. Wintner, eds., *Proceedings of the Workshop on Linguistic Theory and Grammatical Implementation*, pages 21–32, Birmingham, Royaume-Uni.
- [Abeillé & Godard, 1999] Anne Abeillé & Danièle Godard (1999). La position de l'adjectif épithète en français : le poids des mots. In *Recherches linguistiques de Vincennes*.
- [Abercrombie, 1983] David Abercrombie (1983). *Studies in phonetics and linguistics*. Oxford University Press, Londres.
- [Amblard, 2006] Maxime Amblard (2006). Treating clitics with minimalist grammars. In *Proceedings of Formal Grammars*, Malaga, Espagne.
- [Antoine et al., 2003] Jean-Yves Antoine, Jérôme Goulian & Jeanne Villaneau (2003). Quand le TAL robuste s'attaque au langage parlé : analyse incrémentale pour la compréhension de la parole spontanée. In B. Bel & I. Marlien, eds., *Actes de TALN 2003*, Batz-sur-Mer, France.
- [Arnauld & Lancelot, 1830] Arnauld & Lancelot (1830). *Grammaire générale et raisonnée*. Paulet, Paris. Edition de 1969.
- [Baccus, 2002] Nathalie Baccus (2002). *Grammaire française*. Flammarion, Paris.
- [Baldrige & Kruijff, 2003] Jason Baldrige & Geert-Jan Kruijff (2003). Multi-modal combinatory categorial grammar. In *Proceedings of the 10th Annual Meeting of the EACL*, Budapest, Hongrie.
- [Balfourier et al., 2005] Jean-Marie Balfourier, Philippe Blache, Marie-Laure Guénot & Tristan Vanrullen (2005). Comparaison de trois analyseurs symboliques dans une tâche d'annotation syntaxique. In *Actes de TALN 2005 - Workshop EASY*, Dourdan, France.

- [Bally, 1965] Charles Bally (1965). *Linguistique générale et Linguistique française*. Leroux, Paris.
- [Bar-Hillel, 1964] Yehoshua Bar-Hillel (1964). *Language and information*. Addison Wesley, Reading.
- [Baschung, 1998] Karine Baschung (1998). Le contrôle revisité : y a-t-il une différence entre verbes ‘à contrôle’ et verbes ‘à montée’ ? *Journal of french language studies*, 8(1) :1–27.
- [Bellengier & Priego-Valverde, 2004] Emmanuel Bellengier & Béatrice Priego-Valverde (2004). Modélisation de la modulation. In *Actes de TALN 2004*, Fès, Maroc.
- [Bentley, 2003] Delia Bentley (2003). Sur la force d’une approche non-dérivationnelle de l’analyse linguistique : quelques données de l’italo-roman. *Cahiers du Crisco*, 13 :51–75.
- [Berrendonner, 2002] Alain Berrendonner (2002). Les deux syntaxes. *Verbum*, XXIV(1-2) :23–35.
- [Berrendonner et al., 1983] Alain Berrendonner, Michel Le Guern & Gilbert Puech (1983). *Principes de grammaire polylectale*. Presses Universitaires de Lyon, Lyon.
- [Bertrand & Priego-Valverde, 2005] Roxanne Bertrand & Béatrice Priego-Valverde (2005). Le corpus d’interactions dialogiques – présentation et perspectives. Rapport technique, LPL – CNRS / Université de Provence.
- [Bès & Blache, 1999] Gabriel Bès & Philippe Blache (1999). Propriétés et analyse d’un langage. In *Proceedings of TALN 1999*, Cargèse, France.
- [Bès et al., 1999] Gabriel Bès, Philippe Blache & Caroline Hagège (1999). The 5P paradigm. Rapport de recherche, GRIL / LPL.
- [Bilger, 1982] Mireille Bilger (1982). Contribution à l’analyse en grilles. *Recherches sur le Français Parlé*, 4 :195–215.
- [Bilger, 1997] Mireille Bilger (1997). Pour une nouvelle approche des phénomènes de coordination. In Ramon Lorenzo, ed., *Actes du XIXe Congrès International de Linguistique et Philologie Romanes*, pages 925–932, Coruna, Espagne. Université de Santiago de Compostela.
- [Bilger, 2000] Mireille Bilger (2000). *Corpus, méthodologies et applications linguistiques*. Champion, Paris.
- [Bird & Liberman, 1999] Steven Bird & Mark Liberman (1999). A formal framework for linguistic annotation. Rapport technique MS-CIS-99-01, Dept of Computer and Information Science, University of Pennsylvania.
- [Blache, 2000a] Philippe Blache (2000a). *Contraintes et théories linguistiques : des Grammaires d’Unification aux Grammaires de Propriétés*. Habilitation à diriger des recherches, Université Denis Diderot, Paris 7.
- [Blache, 2000b] Philippe Blache (2000b). Le rôle des contraintes dans les théories linguistiques et leur intérêt pour l’analyse automatique : les Grammaires de Propriétés. In *Actes de TALN 2000*, Lausanne, Suisse.
- [Blache, 2001] Philippe Blache (2001). *Les Grammaires de Propriétés – des contraintes pour le traitement automatique des langues naturelles*. Hermès, Paris.
- [Blache, 2002] Philippe Blache (2002). Analyse des dépendances à distance à l’aide de graphes de contraintes – le cas des disloquées. *T.A.L.*, 43(3) :109–128.

- 
- [Blache, 2003a] Philippe Blache (2003a). Meta-level constraints for linguistic domain interaction. In *Proceedings of International Workshop on Parsing Technologies (IWPT-03)*, Nancy, France.
- [Blache, 2003b] Philippe Blache (2003b). Vers une théorie cognitive de la langue basée sur les contraintes. In *Proceedings of TALN 2003*, pages 321–326, Batz-sur-Mer, France.
- [Blache, 2004a] Philippe Blache (2004a). Constraints : an operational framework for construction grammars. In *Proceedings of the 3rd International Conference on Construction Grammars (ICCG3)*, Marseille, France.
- [Blache, 2004b] Philippe Blache (2004b). Densité d'information syntaxique et gradient de grammaticalité. In B. Bel & I. Marlien, eds., *Actes de TALN 2004*, Fès, Maroc.
- [Blache, 2005] Philippe Blache (2005). Property grammars : A fully constraint-based theory. In H. Christiansen, P. Skadhauge & J. Villadsen, eds., *Constraint Satisfaction and Language Processing*. Springer.
- [Blache & Azulay, 2002] Philippe Blache & David-Olivier Azulay (2002). Parsing ill-formed inputs with constraints graphs. In A. Gelbukh, ed., *Intelligent Text Processing and Computational Linguistics*. Springer.
- [Blache & Balfourier, 2001] Philippe Blache & Jean-Marie Balfourier (2001). Parsing with property grammars. In *Proceedings of SICS-2001*.
- [Blache & Guénot, 2002] Philippe Blache & Marie-Laure Guénot (2002). Flexible corpus annotation with property grammars. In K. Simov, ed., *Proceedings of Treebanks and Linguistic Theories*, Sozopol, Bulgarie.
- [Blache *et al.*, 2003] Philippe Blache, Marie-Laure Guénot & Tristan Vanrullen (2003). A corpus-based technique for grammar development. *UCREL Technical Paper, Proceedings of the Corpus Linguistics 2003 Conference*, 16 :124–131.
- [Blache *et al.*, 2006] Philippe Blache, Barbara Hemforth & Stéphane Rauzy (2006). Acceptability prediction by means of grammaticality quantification. In *Proceedings of the COLING-ACL Conference*, Sydney, Australie.
- [Blache & Hirst, 2001] Philippe Blache & Daniel Hirst (2001). Aligning prosody and syntax in property grammars. In *Proceedings of Eurospeech-2001*.
- [Blache & Meunier, 2004] Philippe Blache & Christine Meunier (2004). Domaines et propriétés : une descriptions de la répartition de l'information linguistique. In *Actes des 4e Journées d'Etudes Linguistiques*.
- [Blache & Morin, 1998] Philippe Blache & Jean-Yves Morin (1998). TAO et théories linguistiques : institutions grammaticales. In A. Clas, S. Mejri & T. Baccouche, eds., *La mémoire des mots, Actes des Ve Journées scientifiques du Réseau Thématique "Lexicologie, Terminologie, Traduction"*, pages 161–176, Tunis, Tunisie. AUPELF-UREF.
- [Blache & Morin, 2003] Philippe Blache & Jean-Yves Morin (2003). Une grille d'évaluation pour les analyseurs syntaxiques. In *Actes de TALN 2003 - Workshop Evaluation des Analyseurs Syntaxiques*, Batz-sur-Mer, France.
- [Blackburn & Meyer-Viol, 1994] Patrick Blackburn & Wilfried Meyer-Viol (1994). Linguistics, logic and finite trees. *Bulletin of Interest Group of Pure and Applied Logics*, 2 :2–39.
- [Blackburn *et al.*, 1996] Patrick Blackburn, Wilfried Meyer-Viol & M. de Rijke (1996). A proof system for finite trees. *Computer Science Logic, Lecture notes in computer science*, 1092 :86–106.

- [Blake, 1990] B. Blake (1990). *Relational Grammar*. Routledge, Londres.
- [Blanche-Benveniste, 1975] Claire Blanche-Benveniste (1975). *Recherches en vue d'une théorie de la grammaire française*. Champion, Paris.
- [Blanche-Benveniste, 1981] Claire Blanche-Benveniste (1981). La complémentation verbale : valence, rection et associés. *Recherches sur le Français Parlé*, 3 :57–98.
- [Blanche-Benveniste, 1987] Claire Blanche-Benveniste (1987). Syntaxe, choix du lexique et lieux de bafouillage. *DRLAV*, 36-37 :123–157.
- [Blanche-Benveniste, 1991] Claire Blanche-Benveniste (1991). Divers types de relative en français parlé. *TA Informations*, 21(2) :16–25.
- [Blanche-Benveniste, 1995] Claire Blanche-Benveniste (1995). Le semblable et le dissemblable en syntaxe. *Recherches sur le Français Parlé*, 13 :7–32.
- [Blanche-Benveniste, 1997] Claire Blanche-Benveniste (1997). A propos de 'qu'est-ce que c'est' et de 'c'est quoi'. *Recherches sur le Français Parlé*, 14 :127–146.
- [Blanche-Benveniste, 2000] Claire Blanche-Benveniste (2000). *Approches de la langue parlée en français*. Ophrys, Paris.
- [Blanche-Benveniste, 2002] Claire Blanche-Benveniste (2002). Phrase et construction verbale. *Verbum*, XXIV(1-2) :7–22.
- [Blanche-Benveniste et al., 1990] Claire Blanche-Benveniste, Mireille Bilger, Christine Rouget & Karel Van Den Eynde (1990). *Le français parlé : Etudes grammaticales*. Sciences du langage. CNRS Editions, Paris.
- [Blanche-Benveniste et al., 1979] Claire Blanche-Benveniste, B. Borel, Henri-José Deulofeu, Jacques Durand, Alain Giacomi, Claude Loufrani, B. Meziane & N. Pazery (1979). Des grilles pour le français parlé. *Recherches sur le Français Parlé*, 2 :163–205.
- [Blanche-Benveniste & Jeanjean, 1986] Claire Blanche-Benveniste & Colette Jeanjean (1986). *Le français parlé. Edition et transcription*. Didier, Paris.
- [Bloomfield, 1961] Leonard Bloomfield (1961). *Language*. Holt, New York.
- [Bresnan, 2001] Joan Bresnan (2001). *Lexical-Functional Syntax*. Blackwell, Oxford.
- [Bresnan & Kaplan, 1982] Joan Bresnan & Ronald Kaplan (1982). *The mental representation of grammatical relations*. MIT Press, Cambridge.
- [Bresnan & Zaenen, 1990] Joan Bresnan & Annie Zaenen (1990). Deep unaccusativity in LFG. In K. Dziwirek et al., ed., *Grammatical Relations. A Cross-Theoretical Perspective*, pages 45–57. CSLI Publications, Stanford.
- [Bril & Rebuschi, 2005] I. Bril & G. Rebuschi, eds. (2005). *Actes du Colloque Typologie et modélisation de la coordination et de la subordination*, Paris.
- [Burroni, 1993] A. Burroni (1993). Higher-dimensional word problems with applications to equational logic. *Theoretical Computer Sciences*, 115 :43–62.
- [Candito, 1999] Marie-Hélène Candito (1999). *Organisation modulaire et paramétrable de grammaires électroniques lexicalisées*. Thèse de doctorat, Université Denis Diderot, Paris 7.
- [Candito & Kahane, 1998] Marie-Hélène Candito & Sylvain Kahane (1998). Can the TAG derivation tree represent a semantic graph? an answer in the light of meaning-text theory. *IRCS Technical Report Series : Proceedings of the Fourth International Workshop on Tree-Adjoining Grammars and Related Frameworks (TAG+4)*, 98(12).



- 
- [Cann, 1993] Ronnie Cann (1993). *Formal semantics*. Cambridge University Press.
- [Cann *et al.*, 2005] Ronnie Cann, Ruth Kempson & Lutz Marten (2005). *The dynamics of language*. Elsevier.
- [Cappeau & Deulofeu, 2001] Paul Cappeau & Henri-José Deulofeu (2001). Partition et topicalisation : *il y en a* “stabilisateur” de sujets et de topiques indéfinis. *Cahiers de praxématique*, 37.
- [Cardinaletti & Starke, 2000] A. Cardinaletti & M. Starke (2000). Overview : the grammar (and acquisition) of clitics. In S. M. Powers & C. Hamann, eds., *The acquisition of scrambling and cliticization*, pages 165–186. Kluwer Academic Publishers, Dordrecht.
- [Chevalier *et al.*, 1997] Jean-Claude Chevalier, Claire Blanche-Benveniste, Michel Arrivé & Jean Peytard (1997). *Grammaire du français contemporain*. Larousse, Paris.
- [Chomsky, 1957] Noam Chomsky (1957). *Syntactic Structures*. Mouton, la Hague.
- [Chomsky, 1961] Noam Chomsky (1961). Some methodological remarks on generative grammar. *Word*, 17 :219–239.
- [Chomsky, 1965] Noam Chomsky (1965). *Aspects of the theory of syntax*. MIT Press, Cambridge.
- [Chomsky, 1981] Noam Chomsky (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- [Chomsky, 1993] Noam Chomsky (1993). *Lectures on Government and Binding : the Pisa Lectures*. Mouton de Gruyter.
- [Chomsky, 1995] Noam Chomsky (1995). *The Minimalist Program*. MIT Press.
- [Clément & Kinyon, 2003a] Lionel Clément & Alexandra Kinyon (2003a). Generating LFG with a MetaGrammar. In *Proceedings of Lexical-Functional Grammars (LFG-03)*. CSLI Publications.
- [Clément & Kinyon, 2003b] Lionel Clément & Alexandra Kinyon (2003b). Generating parallel multilingual LFG-TAG grammars from a metagrammar. In *Proceedings of ACL-03*.
- [Colmerauer, 1975] A. Colmerauer (1975). Les grammaires de métamorphose. Rapport technique, Université d’Aix-Marseille.
- [Copestake & Flickinger, 2000] Ann Copestake & Dan Flickinger (2000). An open-source grammar development environment and broad-coverage english grammar using HPSG. In *Proceedings of the 2nd Conference on Language resources and evaluation (LREC)*, Athènes, Grèce.
- [Corbin, 1987] Danielle Corbin (1987). *Morphologie dérivationnelle et structuration du lexique*. M. Niemeyer, Tübingen. Thèse d’Etat, Université Denis Diderot, Paris 7.
- [Cotton & Bird, 2002] Scott Cotton & Steven Bird (2002). An integrated framework for treebanks and multilayer annotations. In *Proceedings of the Third International Conference on Language Resources and Evaluation*.
- [Crabbé, 2005a] Benoît Crabbé (2005a). Grammatical development with XMG. In *Proceedings of Logical Aspects of Computational Linguistics*. Springer.
- [Crabbé, 2005b] Benoît Crabbé (2005b). *Représentation de grammaires d’arbres fortement lexicalisées : application à la grammaire d’arbres adjoints*. Thèse de doctorat, Université Nancy 2, Nancy, France.

- [Crabbé & Duchier, 2004] Benoît Crabbé & Denys Duchier (2004). Metagrammar redux. In *International Workshop on Constraint Solving and Language Processing (CSLP04)*, Copenhagen.
- [Creissels, 2004] Denis Creissels (2004). Cours de syntaxe générale. Université Lyon 2.
- [Croft & Cruse, 2003] W. Croft & D. Cruse (2003). *Cognitive linguistics*. Cambridge University Press.
- [Dahl & Blache, 2004] Verónica Dahl & Philippe Blache (2004). Implantation des grammaires de propriétés en CHR. In Fred Mesnard, ed., *Programmation en logique avec contraintes, Proceedings of JFPLC 2004*, pages 149–166. Hermès.
- [Damourette & Pichon, 1968] Jacques Damourette & Edouard Pichon (1968). *Essai de grammaire de la langue française — Tome 2, 1911–1930*. D’Artrey, Paris.
- [Damourette & Pichon, 1987] Jacques Damourette & Edouard Pichon (1987). *Essai de grammaire de la langue française — Tome 1, 1911–1927*. D’Artrey, Paris.
- [De Fornel & Marandin, 1996] Michel De Fornel & Jean-Marie Marandin (1996). L’analyse grammaticale des auto-réparations. *Le Gré des Langues*, 10 :8–68.
- [Debaisieux & Deulofeu, 2003] Jeanne-Marie Debaisieux & Henri-José Deulofeu (2003). Fonctionnement microsyntaxique de modifieur et fonctionnement macrosyntaxique en parataxe des constructions introduites par *que* et *parce que* en français parlé, avec extension au cas de *perché* et *che* en italien parlé. In F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy, eds., *Il Parlato Italiano. Atti del Convegno Nazionale*, Naples, Italie. D’auria.
- [Debusmann *et al.*, 2004] Ralph Debusmann, Denys Duchier & Geert-Jan Kruijff (2004). Extensible dependency grammar : A new methodology. In *Recent advances in Dependency Grammar, COLING 2004*.
- [Deulofeu, 1982] Henri-José Deulofeu (1982). A propos des préjugés logicistes dans l’analyse grammaticale : le cas des prépositions ; hypothèses linguistiques et sociolinguistiques. *Recherches sur le Français Parlé*, 4 :45–68.
- [Deulofeu, 1999] Henri-José Deulofeu (1999). Problèmes méthodologiques de l’analyse morphosyntaxique de *que* en français contemporain. *Recherches sur le Français Parlé*, 15.
- [Deulofeu, 2003] Henri-José Deulofeu (2003). L’approche macrosyntaxique en syntaxe : un nouveau modèle de rasoir d’Occam contre les notions inutiles. *Scolia*, 16.
- [Deulofeu, 2004] Henri-José Deulofeu (2004). Lontan degli occhi lontan dal cuore : Les énoncés non verbaux sont des énoncés comme les autres. In *La rime et la raison, Hommage à Benoît de Cornulier*. Champion, Paris.
- [Deulofeu, 2005] Henri-José Deulofeu (2005). The limits between syntax and discourse in the analysis of “peripheral” constituents : Towards a solution in a “parallel architecture” model of linguistic structure. In *Workshop IDP-05*, Aix en Provence.
- [Deulofeu, 2006] Henri-José Deulofeu (2006). Pour une linguistique du rattachement. In *Les linguistiques du détachement*, pages 26–27, Nancy, France.
- [Dik, 1989] Simon Dik (1989). *The theory of Functional Grammar*, volume 1. Foris, Dordrecht.
- [Dowding *et al.*, 1999] J. Dowding, J. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore & D. Moran (1999). Gemini : A natural language system for spoken language understanding. In *Proceedings of ARPA Workshop on Human Language Technology*, pages 21–24.

- 
- [Dubois *et al.*, 1994] J. Dubois, L. Gespin, M. Giacomo, C. Marcellesi, J.-B. Marcellesi & J.-P. Mével (1994). *Dictionnaire de linguistique et des sciences du langage*. Larousse, Paris.
- [Dubois & Lagane, 1973] Jean Dubois & René Lagane (1973). *La nouvelle grammaire du français*. Larousse, Paris.
- [Duchier *et al.*, 2005] Denys Duchier, Joseph Le Roux & Yannick Parmentier (2005). XMG : un compilateur de méta-grammaires extensible. In *Actes de TALN 2005*, Dourdan, France.
- [Estratat, 2006] Mathieu Estratat (2006). *Vers les grammaires des configuration*. Thèse de doctorat, Université de la Méditerranée, Aix-Marseille III.
- [Estratat & Hencoque, 2004] Mathieu Estratat & Laurent Hencoque (2004). Parsing languages with a configurator. In *Proceedings of ECAI 2004*, pages 591–595.
- [Flaux & Van de Velde, 2000] Nelly Flaux & Danièle Van de Velde (2000). *Les noms en français : Esquisse de classement*. Ophrys, Paris.
- [Flickinger, 1987] Dan Flickinger (1987). *Lexical rules in the hierarchical lexicon*. PhD thesis, Stanford University.
- [Fuchs & Le Goffic, 1992] Catherine Fuchs & Pierre Le Goffic (1992). *Les linguistiques contemporaines – Repères théoriques*. Hachette, Paris. (première édition 1975).
- [Gabbay, 1996] Dov Gabbay (1996). *Labelled deductive systems*. Oxford University Press, Oxford.
- [Gadet, 1987] F. Gadet (1987). *Saussure, une science du langage*. Presses Universitaires de France, Paris.
- [Gaiffe *et al.*, 2001] Bertrand Gaiffe, Benoît Crabbé & Azim Roussanaly (2001). A new metagrammar compiler. In *Proceedings of TAG+6*, Venise, Italie.
- [Gardent & Kallmeyer, 2003] Claire Gardent & Laura Kallmeyer (2003). Semantic construction in feature-based tree-adjoining grammar. In *Proceedings of the 10th Conference of the European Chapter of the ACL*.
- [Gazdar *et al.*, 1985] Gerald Gazdar, Ewan Klein, Geoffrey Pullum & Ivan Sag (1985). *Generalized Phrase Structure Grammar*. Blackwell, Oxford.
- [Gendner & Vilnat, 2004] Véronique Gendner & Anne Vilnat (2004). *Les annotations syntaxiques de référence PEAS, version 1.6*. Révisions par : Laura Monceaux, Patrick Paroubek, Isabelle Robba.
- [Gerdes & Kahane, 2004] Kim Gerdes & Sylvain Kahane (2004). L’amas verbal au coeur d’une modélisation topologique du français. In K. Gerdes & C. Muller, eds., *Actes des Journées de la syntaxe*, Bordeaux, France.
- [Gerdes & Muller, 2004] K. Gerdes & C. Muller, eds. (2004). *Actes des Journées de la syntaxe : Ordre des mots dans la phrase française, positions et topologie*, Bordeaux, France.
- [Givón, 1984] Tom Givón (1984). *Syntax*, volume 1. Benjamins.
- [Godard, 1988] Danièle Godard (1988). *La syntaxe des relatives en français*. Editions du CNRS, Paris.
- [Godfrey *et al.*, 1992] J. Godfrey, E. Holliman & J. McDaniel (1992). Switchboard : A telephone speech corpus for research and development. In *Proceedings of the IEEE*, pages 517–520.

- [Goldberg, 1995] Adele Goldberg, ed. (1995). *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago University Press.
- [Gross, 1967] Maurice Gross (1967). Sur une règle de cacophonie. *Langages*, 7 :105–119.
- [Gross, 1975] Maurice Gross (1975). *Méthodes en syntaxe*. Hermann.
- [Gross, 1986] Maurice Gross (1986). *Grammaire transformationnelle du français – tome 2 : Syntaxe du nom*. Cantilène, Paris.
- [Guénot, 2001] Marie-Laure Guénot (2001). *Essai de formalisation de l'analyse en constituants immédiats en vue de la réalisation d'un analyseur syntaxique automatique de la langue française*. Mémoire de maîtrise, Université de Provence, Aix-Marseille I.
- [Guénot, 2002] Marie-Laure Guénot (2002). *Vers une grammaire électronique du français parlé : Prolégomènes*. Mémoire de DEA, Université de Provence, Aix-Marseille I.
- [Guénot, 2005a] Marie-Laure Guénot (2005a). Des constructions à l'interface entre lexique et grammaire. In Sylvain Kahane, ed., *Journée de l'Atala Interfaces Lexique-Grammaire et lexiques syntaxiques et sémantiques*, Paris.
- [Guénot, 2005b] Marie-Laure Guénot (2005b). Parsing de l'oral : traiter les disfluences. In *Actes de TALN 2005*, pages 323–332, Dourdan, France.
- [Guénot, 2006] Marie-Laure Guénot (2006). La coordination considérée comme un entassement paradigmatique : description, représentation et intégration. *Cahiers du Cental*, 2(1) :178–187.
- [Guénot & Bellengier, 2004] Marie-Laure Guénot & Emmanuel Bellengier (2004). Quelques principes pour une grammaire multimodale du français. In B. Bel & I. Marlien, eds., *Actes de RECITAL 2004*, pages 51–60, Fès, Maroc.
- [Guénot & Blache, 2005] Marie-Laure Guénot & Philippe Blache (2005). A descriptive and formal perspective for grammar development. In R. Kempson & G. Morrill, eds., *Workshop on Foundations of Natural-Language Grammars*, Edimbourg, Royaume-Uni.
- [Guénot & Vanrullen, 2003] Marie-Laure Guénot & Tristan Vanrullen (2003). Un outil de développement des grammaires de propriétés. In *Actes de RECITAL 2003*, pages 521–526, Batz-sur-Mer, France.
- [Guillaume, 1971] Gustave Guillaume (1971). *Leçons de linguistique — 1948-1949, série A : Structure sémiologique et structure psychique de la langue française 1*. Klincksieck, Paris.
- [Guillaume, 1990] Gustave Guillaume (1990). *Leçons de linguistique — 1944-1945, série B : Sémantèmes, morphèmes et systèmes*. Presses Universitaires de Lille, Lille.
- [Halliday, 1985] Mickael Halliday (1985). *An Introduction to functional grammar*. Arnold.
- [Harris, 1968] Zellig Harris (1968). *Mathematical structures of English*. Wiley, New York. (trad. fr. Structures mathématiques du langage, Dunod, Paris).
- [Henry & Pallaud, 2003] Sandrine Henry & Berthille Pallaud (2003). Word fragments and repeats in spontaneous spoken french. In R. Eklund, ed., *Proceedings of DiSS'03*, pages 77–80, Göteborg University.
- [Hockett, 1958] Charles F. Hockett (1958). *A course in modern linguistics*. Macmillan, New York.
- [Huddleston & Pullum, 2002] Rodney Huddleston & Geoffrey Pullum, eds. (2002). *The Cambridge Grammar of English Language*. Cambridge University Press.

- 
- [Hudson, 1984] Richard Hudson (1984). *Word Grammar*. Blackwell, Oxford.
- [Jacobson, 1990] Pauline Jacobson (1990). Raising as functional composition. *Linguistics and Philosophy*, 13 :423–475.
- [Johnson & Lappin, 1999] D. Johnson & Stuart Lappin (1999). *Local Constraints vs. Economy*. CSLI Publications.
- [Johnson *et al.*, 2004] Mark Johnson, Eugene Charniak & Matthew Lease (2004). An improved model for recognizing disfluencies in conversational speech. In *Rich Transcription 2004 Fall workshop*.
- [Joshi *et al.*, 1975] Aravind Joshi, Leon Levy & M. Takahashi (1975). Tree adjunct grammars. *Journal of the Computer and System Sciences*, 10 :136–163.
- [Joshi & Schabès, 1992] Aravind Joshi & Yves Schabès (1992). Definability and recognizability of sets of trees. In M. Nivat & M. Podelski, eds., *Tree-Adjoining Grammars and Lexicalized Grammars*, pages 31–81, Princetown. Elsevier.
- [Joshi & Schabès, 1997] Aravind Joshi & Yves Schabès (1997). Tree adjoining grammars. In G. Rozenberg & A. Salomaa, eds., *Handbook of Formal Languages*. Springer, Berlin.
- [Joshi, 1987] Aravind K. Joshi (1987). Introduction to tree adjoining grammar. In A. Manaster Ramer, ed., *The mathematics of language*, pages 87–114. Benjamins, Amsterdam.
- [Kahane, 2001] Sylvain Kahane (2001). Grammaires de dépendance formelles et théorie sens-texte. In *Proceedings of TALN 2001 — Tome 2 : Tutoriels*, pages 17–76, Paris.
- [Kahane, 2002] Sylvain Kahane (2002). *Grammaire d’unification sens-texte : Vers un modèle mathématique articulé de la langue naturelle*. Document de synthèse de l’habilitation à diriger les recherches, Université Denis Diderot, Paris VII.
- [Kahane, 2004] Sylvain Kahane (2004). Grammaires d’unification polarisées. In *Actes de TALN 2004*, Fès, Maroc.
- [Kahane, 2006] Sylvain Kahane (2006). On the status of phrases in head-driven phrase-structure grammar : Illustration by a totally lexical treatment of extraction. In A. Polguère (ed.). Benjamins. A paraître.
- [Kahane & Lareau, 2005] Sylvain Kahane & François Lareau (2005). Grammaire d’unification sens-texte : Modularité et polarisation. In *Actes de TALN 2005*, pages 23–32, Dourdan, France.
- [Kallmeyer, 2002] Laura Kallmeyer (2002). Using an enriched TAG derivation structure as basis for semantics. In *Proceedings of the Sixth International Workshop on Tree-Adjoining Grammar and Related Framework (TAG+6)*.
- [Kamp & Reyle, 1993] Hans Kamp & Uwe Reyle (1993). *From discourse to logic - Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation*. Kluwer Academic Publishers, Dordrecht.
- [Kaplan & Bresnan, 1982] Ronald Kaplan & Joan Bresnan (1982). Lexical-functional grammar : A formal system for grammatical representation. In Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*, pages 173–281. The MIT Press, Cambridge.
- [Karlsson, 1990] F. Karlsson (1990). Constraint grammar as a framework for parsing running text. In *Proceedings of COLING’90*.
- [Kay, 1979] Martin Kay (1979). Functional grammars. In *Proceedings of 5th annual meeting of the Berkeley Linguistics Society*, pages 142–158, Berkeley.

- [Kay, 1995] Paul Kay (1995). Construction grammar. In J. Verschueren, J.-O. Östman & J. Blommaert, eds., *Handbook of Pragmatics*, pages 171–177. Benjamins, Amsterdam / Philadelphia.
- [Kay & Fillmore, 1999] Paul Kay & Charles Fillmore (1999). Grammatical constructions and linguistic generalizations : The *what's X doing Y?* construction. *Language*, 75(1) :1–33.
- [Kempson *et al.*, 2001] Ruth Kempson, Wilfried Meyer-Viol & Dov Gabbay (2001). *Dynamic Syntax – The flow of language understanding*. Blackwell, Oxford.
- [Kesper & Mönnich, 2003] S. Kesper & U. Mönnich (2003). Graph properties of hpsg feature structures. In *Proceedings of Formal Grammars*, pages 115–124, Vienne, Autriche.
- [Kleiber, 2003] Georges Kleiber (2003). Faut-il dire adieu à la phrase? *L'information grammaticale*, 98 :17–22.
- [Koch, 1995] Peter Koch (1995). Subordination, intégration syntaxique et “oralité”. *Etudes Romanes : La subordination dans les langues romanes*, 34.
- [Kordoni, 2003] Valia Kordoni (2003). *Morphosyntax-semantics interface in lexicalist theories*. Vienne, Autriche. ESSLLI Lecture Notes.
- [Kracht, 2005a] Marcus Kracht (2005a). *The Emergence of Syntactic Structure*. Ms, UCLA.
- [Kracht, 2005b] Marcus Kracht (2005b). The emergence of syntactic structure. In *Proceedings of workshop on Foundations of Natural-Language Grammars*, Edimbourg, Royaume-Uni.
- [Kruijff, 2004] Geert-Jan Kruijff (2004). *Multi-modal combinatory categorial grammar*. Nancy, France. ESSLLI Lecture Notes.
- [Laenzlinger & Wehrli, 1991] Christopher Laenzlinger & Eric Wehrli, eds. (1991). *Syntax and the Lexicon*, volume 26. Academic Press.
- [Laenzlinger & Wehrli, 1993] Christopher Laenzlinger & Eric Wehrli (1993). Théorie linguistique et traitement automatique du langage. In P. Bouillon & A. Clas, eds., *La traductique*, pages 199–211. Presses de l'Université de Montréal.
- [Lalande, 1972] André Lalande, ed. (1972). *Vocabulaire technique et critique de la philosophie*. Presses Universitaires de France, Paris.
- [Langacker, 1987] R. Langacker (1987). *Foundations of Cognitive Grammar*, volume 1. Stanford University Press.
- [Le Goffic, 1993] Pierre Le Goffic (1993). *Grammaire de la phrase française*. Hachette Supérieur, Paris.
- [Lefeuve, 1999] Florence Lefeuve (1999). *La phrase averbale en français*. L'Harmattan, Paris.
- [Levelt, 1983] William Levelt (1983). Monitoring and self-repair in speech. *Cognition*, 14 :41–104.
- [Levin, 1986] Lori Levin (1986). *Operations on lexical forms : Unaccusative rules in Germanic languages*. PhD thesis, MIT.
- [Loufrani, 1981] Claude Loufrani (1981). Locuteur collectif ou locuteur tout court. *Recherches sur le Français Parlé*, 3 :215–243.
- [Loufrani, 1984] Claude Loufrani (1984). Le locuteur collectif. typologie de configurations discursives. *Recherches sur le Français Parlé*, 6 :169–193.
- [Martinet, 1979] André Martinet (1979). *Grammaire fonctionnelle du français*. Didier, Paris.

- 
- [Maruyama, 1990] H. Maruyama (1990). Structural disambiguation with constraint propagation. In *Proceedings of COLING-ACL'98 Workshop on Dependency-Based Grammars*.
- [Mathieu, 2003] Yvette Yannick Mathieu (2003). La grammaire de construction. *Linx*, 48 :43–56.
- [McEnery & Wilson, 1996] Tony McEnery & Andrew Wilson (1996). *Corpus Linguistics*. Edinburgh University Press, Edimbourg, Royaume-Uni.
- [Meillet, 1924] Antoine Meillet (1924). *Traité de grammaire comparée des langues indo-européennes*. Champion, Paris.
- [Mel'čuk, 1988] Igor Mel'čuk (1988). *Dependency Syntax : theory and practice*. University of New York Press, Albany.
- [Mel'čuk, 1997] Igor Mel'čuk (1997). *Vers une linguistique Sens-Texte*. Leçon inaugurale, Collège de France.
- [Mertens, 1993] Piet Mertens (1993). Accentuation, intonation et morphosyntaxe. *Travaux de Linguistique*, 26.
- [Miller, 1991] Philip Miller (1991). *Clitics and constituents in phrase-structure grammar*. PhD thesis, University of Utrecht, Hollande.
- [Miller & Sag, 1995] Philip Miller & Ivan Sag (1995). French clitic movement without clitics or movement. CSLI Technical Report.
- [Milner, 2002] Jean-Claude Milner (2002). *Le péricle structural*. Seuil, Paris.
- [Mithun, 2005] Marianne Mithun (2005). On the assumption of the sentence as the basic unit of syntactic structure. In Z. Frajzyngier, ed., *Linguistic Diversity and Language Theories*, pages 169–183. Benjamins, Amsterdam.
- [Moeschler & Auchlin, 1997] J. Moeschler & A. Auchlin (1997). *Introduction à la linguistique contemporaine*. Armand Colin, Paris.
- [Montague, 1974] Richard Montague (1974). English as a formal language. In R. H. Thomason, ed., *Formal philosophy : selected papers of Richard Montague*. Yale University Press.
- [Moortgat, 1989] Michael Moortgat (1989). *Categorial investigations : Logical and linguistic aspects of the Lambek calculus*. Foris, Dordrecht.
- [Moortgat, 1990] Michael Moortgat (1990). *La grammaire catégorielle généralisée : le calcul de Lambek-Gentzen*, chapter 3, pages 127–182. Hermès.
- [Morel, 2003] Marie-Annick Morel (2003). Phrase ? Enoncé ? Paragraphe ? Hyperparagraphe ? Quelles unités intonatives et discursives pour le dialogue oral en français ? *L'information grammaticale*, 98 :39–47.
- [Morin, 1985] Jean-Yves Morin (1985). Théorie syntaxique et théorie du passage : quelques réflexions. *Revue Québécoise de linguistique*, 14(2) :9–48.
- [Morin, 1989] Jean-Yves Morin (1989). *Syntaxe*. Ms. Université de Montréal.
- [Morrill, 1987] Glynn Morrill (1987). Meta-categorial grammar. In G. Morrill, N. J. Haddock & E. Klein, eds., *Unification Grammar and Parsing*, Edimbourg, Royaume-Uni. Centre for Cognitive Science, University of Edinburgh.
- [Mouret, 2003] François Mouret (2003). Conjonction de coordination et constituance en français. In *Actes des 6e rencontres de l'Ecole Doctorale de Paris III*, Paris.

- [Muskens *et al.*, 1997] Reinhard Muskens, J. van Benthem & A. Visser (1997). Dynamics. In J. van Benthem & A. ter Meulen, eds., *Handbook of Logic and Language*, pages 587–648. Elsevier, Amsterdam.
- [Nasr, 1995] Alexis Nasr (1995). A formalism and a parser for lexicalized dependency grammars. In *Proceedings of 4th International Workshop on Parsing Technologies*. State University of New York Press.
- [Nølke, 1994] Henning Nølke (1994). *Linguistique modulaire : de la forme au sens*. Peeters, Bibliothèque de l'Information Grammaticale, Paris.
- [Östman & Fried, 2005] Jan-Ola Östman & Mirjam Fried, eds. (2005). *Constructional Approaches to Language - Construction Grammars : Cognitive grounding and theoretical extensions*, volume VIII. Benjamins.
- [Pallaud, 2006] Berthille Pallaud (2006). Truncations de mots, reprises et interruption syntaxique en français parlé spontané. In *Actes des Journées internationales d'Analyse statistique des Données Textuelles (JADT)*, pages 707–715, Besançon, France.
- [Pallaud & Henry, 2004] Berthille Pallaud & Sandrine Henry (2004). Amorces de mots et répétitions : des hésitations plus que des erreurs en français parlé. In *Actes de JADT 2004*.
- [Paroubek, 2005] Patrick Paroubek (2005). EASY : Campagne d'évaluation des analyseurs syntaxiques. In *Actes de TALN 2005 - Workshop EASY*, Dourdan, France.
- [Partee *et al.*, 1993] Barbara Partee, Alice ten Meulen & Robert Wall (1993). *Mathematical Methods in Linguistics*. Kluwer.
- [Péronnou, 1996] G. Péronnou (1996). Compréhension du dialogue oral : le rôle du lexique dans l'approche par segments conceptuels. In *Actes de Lexique et Communication Parlée*, pages 169–178.
- [Perlmutter, 1970] David Perlmutter (1970). Surface structure constraints in syntax. *Linguistic Inquiry*, 1 :187–255.
- [Perrier, 2002] Guy Perrier (2002). Descriptions d'arbres avec polarités : les grammaires d'interaction. In *Actes de TALN 2002*, Nancy, France.
- [Perrier, 2003] Guy Perrier (2003). *Les grammaires d'interaction*. Habilitation à diriger les recherches, Université Nancy 2.
- [Perrot, 1994] Jean Perrot (1994). Éléments pour une typologie des structures informatives. In *Mémoires de la Société de Linguistique de Paris, tome II : La phrase : Énonciation et information*, pages 13–26. Peeters, Leuven.
- [Polguère, 1998] Alain Polguère (1998). La théorie sens-texte. *Dialangue*, 8-9 :9–30.
- [Pollard & Sag, 1994] Carl Pollard & Ivan Sag (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.
- [Prince & Smolensky, 1993] A. Prince & P. Smolensky (1993). Optimality theory : Constraint interaction in generative grammars. Rapport technique RUCCS TR-2, Rutgers Center for Cognitive Science.
- [Pullum & Scholz, 2003] Geoffrey Pullum & Barbara Scholz (2003). *Model-theoretic Syntax Foundations : Linguistic Aspects*. Vienne, Autriche. ESSLLI Lecture Notes.
- [Retoré, 2005] Christian Retoré (2005). Syntaxe et traitement automatique des langues. Rapport de recherche RR-5459, INRIA.



- 
- [Retoré & Moot, 2004] Christian Retoré & Richard Moot (2004). L'ordre des mots dans les grammaires catégorielles. In K. Gerdes & C. Muller, eds., *Pré-Actes des Journées de la syntaxe : L'ordre des mots dans la phrase française, positions et topologie*, pages 56–59, Bordeaux, France.
- [Retoré & Moot, 2006] Christian Retoré & Richard Moot (2006). Les indices pronominaux du français dans les grammaires catégorielles. *Linguisticae Investigationes*, 29(1) :137–146.
- [Riegel *et al.*, 1999] Martin Riegel, Jean-Christophe Pellat & René Rioul (1999). *Grammaire méthodique du français*. Presses Universitaires de France, Paris.
- [Rogers & Vijay-Shanker, 1994] James Rogers & K. Vijay-Shanker (1994). Obtaining trees from their descriptions : an application to tree-adjoining grammars. *Computational Intelligence*, 10 :401–421.
- [Sag *et al.*, 2003] Ivan Sag, Thomas Wasow & Emily Blender (2003). *Syntactic Theory – A formal introduction*. CSLI Publications, Chicago.
- [Sagot, 2006] Benoît Sagot (2006). *Analyse automatique du français : lexiques, formalismes, analyseurs*. Thèse de doctorat, Université Denis Diderot, Paris 7.
- [Salkoff, 1979] Maurice Salkoff (1979). *Analyse syntaxique du français : grammaire en chaîne*. Benjamins, Amsterdam.
- [Saussure, 1916] Ferdinand De Saussure (1916). *Cours de linguistique générale*. Payot, Paris. réédition de 1990.
- [Schabès & Shieber, 1994] Yves Schabès & Stuart Shieber (1994). An alternative conception of tree-adjoining derivation. *Computational Linguistics*, 20(1) :91–124.
- [Seneff, 1992] S. Seneff (1992). Robust parsing for spoken language systems. In *Proceedings of ICASSP*, pages 189–192.
- [Shriberg, 1994] Elizabeth Shriberg (1994). *Preliminaries to a theory of speech disfluencies*. PhD thesis, University of California, Berkeley.
- [Steedman, 1996] Mark Steedman (1996). *Surface Structure and Interpretation*. MIT Press, Cambridge.
- [Stokoe, 1972] W.-C. Stokoe (1972). *Semiotics and Human Sign Languages*. Mouton, la Hague.
- [Taylor *et al.*, 2004] Ann Taylor, Mitchell Marcus & Beatrice Santorini (2004). *Building and Using Parsed Corpora*, chapter The Penn Treebank : an overview, pages 5–22. Kluwer, Dordrecht.
- [Tesnière, 1953] Lucien Tesnière (1953). *Esquisse d'une syntaxe structurale*. Klincksieck, Paris.
- [Tesnière, 1959] Lucien Tesnière (1959). *Éléments de syntaxe structurale*. Klincksieck, Paris.
- [Thomasset & de la Clergerie, 2005] F. Thomasset & Eric Villemonte de la Clergerie (2005). Comment obtenir plus des métagrammaires. In *Actes de TALN 2005*, Dourdan, France.
- [Touratier, 2002] Christian Touratier (2002). *Morphologie et Morphématique, Analyse en morphèmes*. Publications de l'Université de Provence, Aix en Provence.
- [Touratier, 2004] Christian Touratier (2004). *Syntaxe*. Université de Provence.
- [Touratier, 2005] Christian Touratier (2005). *Analyse et théorie syntaxiques*. Publications de l'Université de Provence, Aix en Provence.

- [Tugwell, 1998] David Tugwell (1998). *Dynamic Syntax*. PhD thesis, University of Edinburgh, Royaume-Uni.
- [Tugwell, 2000] David Tugwell (2000). Towards a dynamic syntax for language modelling. In *Lecture Notes in Artificial Intelligence - Proceedings of 3rd International Workshop on Text, Speech and Dialogue (TSD2000)*, pages 33–38, Brno, République Tchèque. Springer.
- [Tugwell, 2004] David Tugwell (2004). Rethinking the foundations with a dynamic construction approach. In R. Kempson & G. Morrill, eds., *Proceedings of Esslli's workshop on Foundations of Natural Language Grammar*, Edimbourg, Royaume-Uni.
- [Uszkoreit, 1986] Hans Uszkoreit (1986). Categorical unification grammars. In *Proceedings of COLING-86*, pages 187–194, Bonn.
- [van den Eynde & Mertens, 2003] Karel van den Eynde & Piet Mertens (2003). La valence : l'approche pronominale et son application au lexique verbal. *French Language Studies*, 13 :63–104.
- [Van Valin & LaPolla, 1997] Robert Van Valin & R. LaPolla (1997). *Syntax : Structure, Meaning and Function*. Cambridge University Press, Cambridge.
- [Vanrullen, 2004] Tristan Vanrullen (2004). Analyse syntaxique et granularité variable. In *Actes de RECITAL 2004*, Fès, Maroc.
- [Vanrullen, 2005] Tristan Vanrullen (2005). *Vers une analyse syntaxique à granularité variable*. Thèse de doctorat, Université de Provence, Aix-Marseille I.
- [Vanrullen et al., 2003] Tristan Vanrullen, Marie-Laure Guénot & Emmanuel Bellengier (2003). Formal representation of property grammars. In *Proceedings of ESSLLI Student Session*, Vienne, Autriche.
- [Vijay-Shanker, 1987] K. Vijay-Shanker (1987). *A Lexicalized Tree Adjoining Grammar for English*. PhD thesis, University of Pennsylvania.
- [Vijay-Shanker, 1992] K. Vijay-Shanker (1992). Using description of trees in a tree-adjoining grammar. *Computational Linguistics*, 18 :481–517.
- [Vijay-Shanker & Schabès, 1992] K. Vijay-Shanker & Yves Schabès (1992). Structure sharing in lexicalized tree-adjoining grammars. In *Proceedings fo CoLing 92*.
- [Villemonde de la Clergerie, 2005] Eric Villemonde de la Clergerie (2005). DyALog : a tabular logic programming based environment for NLP. In H. Christiansen & J. Villadsen, eds., *Proceedings of International Workshop on Constraint Solving and Language Processing (CSLP04)*, pages 18–33, Sitges, Espagne.
- [Vilnat, 2005] Anne Vilnat (2005). *Fondements du TAL*. Cours de Master Informatique, IUT d'Orsay.
- [Wehrli, 1996] Eric Wehrli (1996). Le traitement des expressions figées dans l'analyseur FIPS. In *Actes de TALN 1996*, pages 210–219.
- [Wehrli, 2004] Eric Wehrli (2004). Un modèle multilingue d'analyse syntaxique. In A. Auchlin, M. Burger, L. Fillittaz, A. Grobet, J. Moeschler, L. Perrin, C. Rossari & L. de Saussure, eds., *Structures et discours, Mélanges offerts à Eddy Roulet*, pages 311–329. Nota Bene.
- [Wilmet, 1998] Marc Wilmet (1998). *Grammaire critique du français*. Hachette - Duculot, Paris.
- [Winograd, 1983] Terry Winograd (1983). *Language as a cognitive processs*, volume 1 : Syntax. Addison Wesley, Reading.

- 
- [Xia, 2001] F. Xia (2001). *Automatic grammar generation from two different perspectives*. PhD thesis, University of Pennsylvania.
- [Ziff, 1964] Paul Ziff (1964). On understanding ‘understanding utterances’. In J. Fodor & J. Katz, eds., *The structure of Language : Readings in the Philosophy of Language*, pages 390–399. Prentice-Hall, Englewood Cliffs, NJ.
- [Zue & Glass, 2000] Victor Zue & James Glass (2000). Conversational interfaces : Advances and challenges. In *Proceedings of the IEEE number 8*, pages 1166–1180.



**Résumé:** Nous proposons un modèle de grammaire basé sur une théorie originale de la langue, et représenté formellement. Dans la première partie nous dressons un bilan des positions théoriques propres aux grammaires formelles coexistantes, afin de faire ressortir quatre éléments dont la conjonction fait de notre proposition une approche nouvelle: le non-générativisme, la non-modularité, la non-lexicalisation et la multi-dimensionnalité. Nous présentons ensuite le formalisme des Grammaires de Propriétés (GP), que nous avons utilisé pour représenter notre grammaire. A la suite de cela nous introduisons notre modèle de grammaire, basé sur les choix théoriques précédents et formalisé en GP, et nous proposons quelques formalisations d'autres modèles afin d'illustrer les possibilités de GP et l'originalité de notre modèle. Dans la seconde partie nous proposons un ensemble de descriptions syntaxiques du français basées sur notre modèle et constituant un noyau de grammaire; nous y présentons notamment les constructions nominales, verbales, adjectivales, propositionnelles, ainsi que les entassements paradigmatiques (coordinations et disfluences). Enfin, dans la troisième partie nous illustrons le fonctionnement notre proposition avec l'analyse de quelques phénomènes syntaxiques, dont notamment le traitement des pronoms clitiques dans les constructions verbales, et celui des coordinations et des disfluences. Ce travail apporte à toute une partie de la linguistique descriptive une validation par son expression formelle, et à la linguistique formelle l'intégration de descriptions syntaxiques jusqu'ici non encore prises en considération. En outre, elle apporte une validation de GP en tant que formalisme linguistique en montrant ce qu'il permet par sa souplesse de représentation.

**Mots-clés:** *Syntaxe - Linguistique descriptive - Linguistique formelle - Linguistique informatique - Développement de grammaire - Grammaires de Propriétés (GP)*

**Abstract:** We propose a grammatical model based on an original natural language theory, and formally expressed. In the first part, we draw up an assessment of theoretical positions that are specific to other formal grammars, in order to bring out four elements whose conjunction makes our proposal a new approach: non-generativism, non-modularity, non-lexicalization and multi-dimensionality. We present then the formalism of Property Grammars (PG), that we used to represent our grammar. Following that, we introduce our grammar model, based on the preceding theoretical choices and formalized into PG, and we propose some formalizations of other models in order to illustrate the possibilities of GP and the originality of our model. In the second part, we propose a set of syntactic descriptions for French, based on our model and constituting a core grammar; we present in particular nominal, verbal, adjectival, and propositional constructions, as well as paradigmatic accumulations (coordinations and disfluences). Lastly, in the third part we illustrate the our proposal's functioning through the analysis of some syntactic phenomena, among which the treatment of clitic pronouns in verbal constructions, and that of coordinations. This work provides a validation to a whole part of descriptive linguistics by its formal representation, and it allows the integration of syntactic descriptions not taken into account up to now into formal linguistics. Moreover, it provides a validation of PG as a linguistic formalism by showing what it allows by its flexibility of representation.

**Keywords:** *Syntax - Descriptive Linguistics - Formal Linguistics - Computational Linguistics - Grammar Development - Property Grammars (PG)*

