



Linx

Revue des linguistes de l'université Paris X Nanterre

39 | 1998

Modèles linguistiques : convergences, divergences

Héritage de propriétés dans les grammaires d'arbres polychromes

Property Inheritance in Polychrome Tree Grammars

Marcel Cori et Jean-Marie Marandin



Édition électronique

URL : <http://journals.openedition.org/linx/871>

DOI : 10.4000/linx.871

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Édition imprimée

Date de publication : 15 décembre 1998

Pagination : 13-41

ISSN : 0246-8743

Référence électronique

Marcel Cori et Jean-Marie Marandin, « Héritage de propriétés dans les grammaires d'arbres polychromes », *Linx* [En ligne], 39 | 1998, mis en ligne le 26 juillet 2012, consulté le 30 avril 2019. URL : <http://journals.openedition.org/linx/871> ; DOI : 10.4000/linx.871

Ce document a été généré automatiquement le 30 avril 2019.

Département de Sciences du langage, Université Paris Ouest

Héritage de propriétés dans les grammaires d'arbres polychromes

Property Inheritance in Polychrome Tree Grammars

Marcel Cori et Jean-Marie Marandin

- 1 L'occupation d'une position peut varier en fonction de facteurs configurationnels ou lexicaux. A ce propos, Zwicky (1994) développe l'exemple suivant en anglais : un GV fini ou une phrase finie ne peuvent pas être sujet dans une phrase canonique, comme le montre (7) ci-dessous :

| | | |
|-----|----|--|
| (7) | a. | * <i>Build an igloo tires me out quickly</i> |
| | b. | * <i>We build an igloo astonished Terry</i> |

Mais, ils peuvent occuper la position occupée par le sujet dans une phrase clivée inversée (8) :

| | | |
|-----|----|--|
| (8) | a. | <i>Build an igloo was what we did</i> |
| | b. | <i>We build an igloo was what we did</i> |

Zwicky en tire la conclusion que la relation d'occupation est sensible à la construction. La position <S,1> de l'anglais accueille soit un GN (dans le cas général) soit un GN ou une phrase finie dans le cas particulier de la phrase clivée¹⁰.

- 2 Une propriété peut être associée à une position de façon absolue (par exemple : <S,1> accueille un GN), ou de façon relative à une construction (<S,1> accueille une phrase finie dans telle construction). Nous distinguerons donc les propriétés absolues d'une position (les propriétés qui se retrouvent dans toutes les constructions dans laquelle la position apparaît) et les propriétés relatives à une construction.

2.2. Aspects formels

- 3 La caractéristique centrale des GAP est le maintien des conditions classiques sur les arbres (Wall 1972) : racine unique, exclusivité et non-croisement. Du point de vue formel, les arbres polychromes sont des arbres syntagmatiques classiques.

2.2.1. Les couleurs

- 4 Nous ajoutons une dimension (métaphoriquement appelée *couleur*) : les couleurs permettent de distinguer plusieurs types de relations de dominance. C'est «l'outil» formel qui permet de distinguer entre positions sans devoir faire appel aux catégories et aux fonctions. Par exemple, on parle de position $\langle S,1 \rangle$ et non pas de position $\langle S,GN \rangle$ ou de position sujet.
- 5 La notion de couleur est une notion «topologique»¹¹. Dans la grammaire que nous élaborons pour le français, on distingue trois types de relations : être pivot (repère pour la localisation des positions), être immédiatement proche du pivot, être proche. Par exemple, dans l'arbre (9) ci-dessous, $\langle X,3 \rangle$ constitue la position pivot, $\langle X,2 \rangle$ et $\langle X,4 \rangle$ constituent les deux positions immédiatement proches à gauche et à droite du repère, $\langle X,1 \rangle$ et $\langle X,5 \rangle$ les deux positions proches à gauche et à droite du repère :

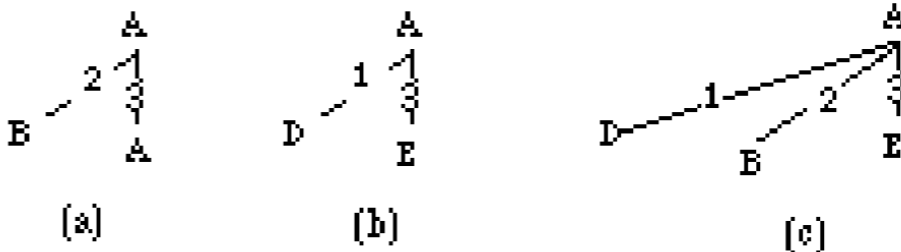


- 6 (9)

La notion de pivot est très proche de celle de «centre topologique» dans les cadres qui admettent la notion de champ (Heltoft 1992). Quant à la distinction entre position adjacente et position proche, elle ressemble à celle qu'opère Dowty 1997.

2.2.2. Composition/compactage

- 7 L'opération de composition des GAP est originale en ce qu'elle est sensible à la couleur. Lorsqu'on compose un arbre de racine A (cf. (10.b) ci-dessous) dans la position pivot d'un arbre élémentaire qui est aussi de racine A (10.a), les deux arbres fusionnent en «s'aplatissant» (10.c) ; c'est ce que nous appelons compactage. Dans le compactage, les branches de l'arbre résultant sont ordonnées selon leur couleur, ce qui peut entraîner un effet d'intrication. Nous illustrons ce cas sous (10) :



8 (10)

C'est cette définition de la composition qui permet de dépasser une des limites bien connues des grammaires syntagmatiques : leur incapacité à traiter des phénomènes de discontinuité. Nous renvoyons à Cori & Marandin (1995) sur ce point.

2.2.4. Les traits

9 Nous représentons les propriétés des positions à l'aide de traits (dans le format «attribut-valeur») associés aux sommets. Ce choix autorise la généralisation de l'opération d'unification qui nous permet de rendre compte des phénomènes qui ne peuvent pas être représentés comme un partage de valeurs. Cette généralisation prend la forme de relations, que l'on appelle *relations d'interdépendance*. Nous ne développons pas cet aspect ici¹².

2.2.5 Problèmes

10 Bien que visant une représentation totalement autonome des positions, la version actuelle du formalisme des GAP n'est pas complètement émancipée de la notion de catégorie. En effet, les positions décrites par les arbres élémentaires d'une GAP, formant une base d'arbres, sont nécessairement associées à une catégorie spécifiée. De plus, elle ne permet pas de faire émerger les principes d'organisation généraux : les propriétés absolues des positions (cf. §2.1.4 ci-dessus) ou des configurations doivent être associées à chaque arbre de la grammaire au même titre que les propriétés relatives ; il est impossible de les factoriser. Ce qui donne une représentation redondante et inadéquate d'un point descriptif. L'introduction d'une hiérarchie entre les arbres de la grammaire permet de remédier à ces deux insuffisances.

3. Un exemple : la hiérarchie des types de phrases

11 Un type de phrase (ou de syntagme) – ce que l'on appellera une construction – présente des propriétés générales et des propriétés particulières (du point de vue d'une langue particulière ou d'un point de vue translinguistique). Nous prenons pour exemple dans ce paragraphe l'inversion locative en français ; l'inversion locative est illustrée en (11) ci-dessous :

| | |
|------|--|
| (11) | a. <i>Sur la place centrale se dresse la cathédrale Saint Paul</i> |
| | c. <i>Dans le château dormait un régiment de spahis</i> |
| | b. <i>Des cheminées sortait une épaisse fumée noire</i> |

Au §3.1, nous présentons l'analyse en introduisant les faits caractéristiques : l'inversion locative est une construction hybride qui présente les propriétés d'une construction à trou et celles de la phrase bipoissonnelle¹³. Aux §3.2 et §3.3, nous construisons le fragment de hiérarchie des constructions phrastiques du français où prend place l'inversion locative.

3.1. Rappel de l'analyse linguistique

3.1.1 L'inversion locative est une construction à trou

- 12 On parle de construction à trou lorsqu'un complément du verbe n'est pas réalisé sous forme d'un syntagme dans la phrase et qu'il peut être mis en relation avec un terme à l'extérieur de ce domaine. La théorie syntaxique contemporaine a étudié cette relation sous le nom de relation «filler/gap» et lui reconnaît les trois propriétés suivantes : (i) le filler peut être arbitrairement éloigné de la phrase qui contient le trou (distance), (ii) le filler présente un certain nombre de traits que devrait présenter le complément manquant (connexité) et (iii) la relation est sensible aux contraintes d'îles.

3.1.1.1. Le groupe prépositionnel frontal de l'inversion locative (GP dorénavant) peut être analysé comme un filler dans les phrases de (11). En effet, on observe les trois propriétés caractéristiques d'une construction à trou.

(i). Le GP peut être éloigné du champ verbal. En (12), la relation filler/gap s'établit à travers plusieurs GV à l'infinitif comportant des verbes à montée (*sembler*) ou à contrôle (*essayer, vouloir*) du sujet et, en (13), elle s'établit à travers la barrière phrastique¹⁴ :

| | |
|------|---|
| (12) | a. <i>Sur la mer essayaient de s'élever des vapeurs d'un noir de suie</i> |
| | b. <i>Du brouillard semblait vouloir émerger un château étrange</i> |

| | |
|------|---|
| (13) | a. <i>Sur cette place, Paul croyait que se trouvait un café</i> |
| | b. <i>De cette enquête, Marie ne savait pas que sortirait la vérité</i> |

(ii). Le GP de l'inversion locative respecte les contraintes de sous-catégorisation ou de sélection du verbe tête lorsqu'il correspond à un argument de ce verbe :

| | | |
|------|-----|---|
| (14) | a. | * <i>Du brouillard se dresse une cathédrale</i> |
| | a'. | * <i>Une cathédrale se dresse du brouillard</i> |
| | b. | * <i>Dans le mur dépassent deux poutres</i> |
| | b'. | * <i>Deux poutres dépassent dans le mur</i> |

(iii). Le GP de l'inversion locative se comporte comme le GP topicalisé en ce qui concerne les contraintes d'îles. En particulier, il peut apparaître en dehors d'une interrogative tensée en français (cf. *A Paul, que pouvait-il répondre ?*) (Godard 1980) :

| | |
|------|---|
| (15) | a. <i>Sur la place, pourquoi se saoulaient tous ces soldats ?</i> |
| | b. <i>De cette famille, pourquoi sortirait un génie ?</i> |

3.1.1.2. On observe, par ailleurs, que l'inversion locative est bien formée dans la relative (16.a) ou dans la complétive (16.b) tout comme la topicalisation de GP (17) :

| | |
|------|---|
| (16) | a. <i>Il aime la majesté avec laquelle sur la place se dresse la cathédrale</i> |
| | b. <i>Paul croyait que sur la place se trouvait un café russe</i> |

| | |
|------|--|
| (17) | a. <i>Il a bien décrit la fougue avec laquelle, aux attaques, le ministre des finances savait répondre</i> |
| | b. <i>Paul savait pourtant qu'aux attaques verbales, il faut répondre par la plus grande douceur</i> |

3.1.1.3. Étant donné ces propriétés, l'inversion locative apparaît comme un cas de topicalisation de GP. Mais, plusieurs traits amènent à la distinguer de la topicalisation «ordinaire» de GP.

Tout d'abord, un certain nombre de GP ne peuvent jamais être topicalisés : il s'agit des GP à prépositions locatives descriptives (Bonami, cp). Par contre, ils donnent lieu à des inversions locatives bien formées :

| | |
|------|---|
| (18) | a. * <i>A côté de l'épicerie, une boulangerie se trouve</i> |
| | a'. <i>A côté de l'épicerie se trouve une boulangerie</i> |

Par ailleurs, les propriétés globales de la phrase présentant une inversion locative sont distinctes de celles de la phrase présentant une topicalisation de GP ordinaire.

3.1.2. L'inversion locative est une construction phrastique

- 13 La phrase présentant une inversion locative a les mêmes propriétés globales que la phrase canonique du français. De ce point de vue, elle se distingue clairement de la phrase présentant une topicalisation de GP.

3.1.2.1. On observe que la phrase présentant une inversion locative est identique à la phrase canonique en ce qui concerne (i) l'intonation et (ii) le conditionnement de l'information (*information packaging*).

(i). Elle est associée à une intonation qui est identique à celle de la phrase canonique. Le GP initial n'est pas détaché du reste de la phrase par une intonation montante ; c'est particulièrement le cas lorsque l'inversion locative est réalisée dans une phrase enchâssée (il faut contraster de ce point de vue (16) et (17) ci-dessus).

(ii). La phrase canonique du français n'est pas associée à un conditionnement informationnel particulier : elle peut être associée à un conditionnement où elle constitue toute entière le focus (*all focus*) ou bien à un conditionnement où certains constituants seulement appartiennent au focus (en particulier le groupe verbal dans le conditionnement appelé *predicate focus*). Il en est de même pour la phrase présentant une inversion locative comme l'indiquent les tests usuels :

| | |
|------|---------------------------------|
| (18) | a. <i>[Que se passe-t-il ?]</i> |
|------|---------------------------------|

| | | |
|--|----|--|
| | | <i>De la cheminée sort une fumée inquiétante</i> |
| | b. | <i>[Dans la chambre de Paul, qu'y a-t-il d'affiché sur les murs ?]</i> |
| | | <i>Sur le mur du fond est affichée une photo de Marie et sur la porte un plan de Paris</i> |

De ce point de vue, elle se distingue de la phrase à topicalisation de GP ordinaire¹⁵ :

| | | |
|------|----|--|
| (19) | a. | <i>[Que se passe-t-il ?]</i> |
| | | <i># A Paul, Marie donne une gifle</i> |
| | b. | <i>[A qui Marie a-t-elle donné une récompense ?]</i> |
| | | <i># A Paul, Marie a donné une récompense</i> |

3.1.2.2. Si les propriétés globales de la phrase présentant une inversion locative sont identiques à celle de la phrase canonique, on peut former l'hypothèse que la phrase à inversion locative présente la même configuration de positions que la phrase canonique ; autrement dit, le GP occupe dans l'inversion locative la position que le GN sujet occupe dans la phrase canonique.

- ¹⁴ Un indice distributionnel soutient cette hypothèse : le sujet préverbal peut être réalisé comme une relative sans antécédent (cf. *Qui m'aime me suive*) ; l'inversion locative présente la même possibilité¹⁶ :

| | | |
|------|----|--|
| (20) | a. | <i>Où dorment les héros se trouve le bonheur éternel</i> |
| | b. | <i>Où ils ont passé se forme une race nouvelle</i> |

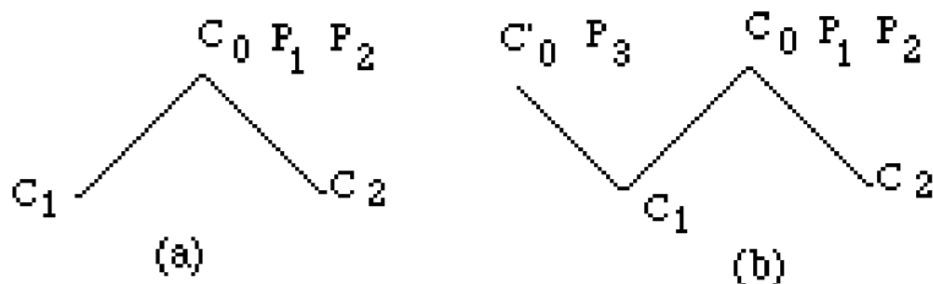
Cet emploi de la relative en où n'est acceptable que dans le cadre d'une inversion locative : il est illicite en position disloquée (*?? Où la maladie se développait, les médecins étaient réduits à l'impuissance*) ou bien à l'intérieur du champ verbal (** La maladie se développe où il n'y a pas de vaccin*).

3.1.3. Synthèse de l'analyse

- ¹⁵ On peut résumer l'analyse de la manière suivante : le GP de l'inversion locative a les propriétés d'un filler (un terme qui correspond à un complément non réalisé dans le GV), mais, à la différence d'un filler ordinaire (GP topicalisé, QUE¹⁷), il occupe une position à l'intérieur du domaine phrastique. Cette analyse permet de réconcilier les deux facettes de la phrase à inversion locative. D'une part, le contexte de réalisation à droite du sujet est identique à celui que l'on observe dans la relative ou l'interrogative : le sujet est réalisé à droite lorsqu'un complément n'est pas réalisé in situ (autrement dit, les contextes à extraction ou à trou)¹⁸. D'autre part, les propriétés globales de la phrase à inversion locative peuvent être identifiées à celles d'une phrase qui diffère d'une phrase canonique par le fait que la position <S,1> est occupée par un GP.

3.2. L'idée de hiérarchie. Approche informelle

- 16 Nous présentons de façon informelle le dispositif qui permet de représenter le fait que plusieurs constructions partagent un même ensemble de propriétés. Si les constructions C_1 , C_2 partagent une ou plusieurs propriétés, on postule une construction C_0 associée à ces propriétés ; C_1 et C_2 en héritent dès lors qu'elles sont mises en relation avec C_0 par la relation «être plus spécifiée que» :



Dans le schéma (21.a), les propriétés P_1 et P_2 sont héritées par C_1 et C_2 ; C_1 et C_2 ont nécessairement des propriétés spécifiques qui les distinguent. Les propriétés peuvent appartenir à plusieurs dimensions (syntaxe, sémantique, conditionnement de l'information, etc.). Une même construction peut hériter de plusieurs constructions moins spécifiées : dans le schéma (21.b), C_1 hérite de C_0 et de C'_0 . C'est cette possibilité de l'héritage multiple que nous exploitons pour représenter l'analyse de l'inversion locative : elle est une construction à trou du point de vue de son organisation syntaxique et une phrase canonique du point de vue de l'organisation de l'information.

3.3. Représentation de l'analyse de l'inversion locative

- 17 Nous montrons dans ce paragraphe comment représenter le caractère «hybride» de l'inversion locative en l'insérant dans une hiérarchie où elle hérite à la fois des propriétés des constructions à trou et de celles de la phrase bpositionnelle (cf. (23) ci-dessous).

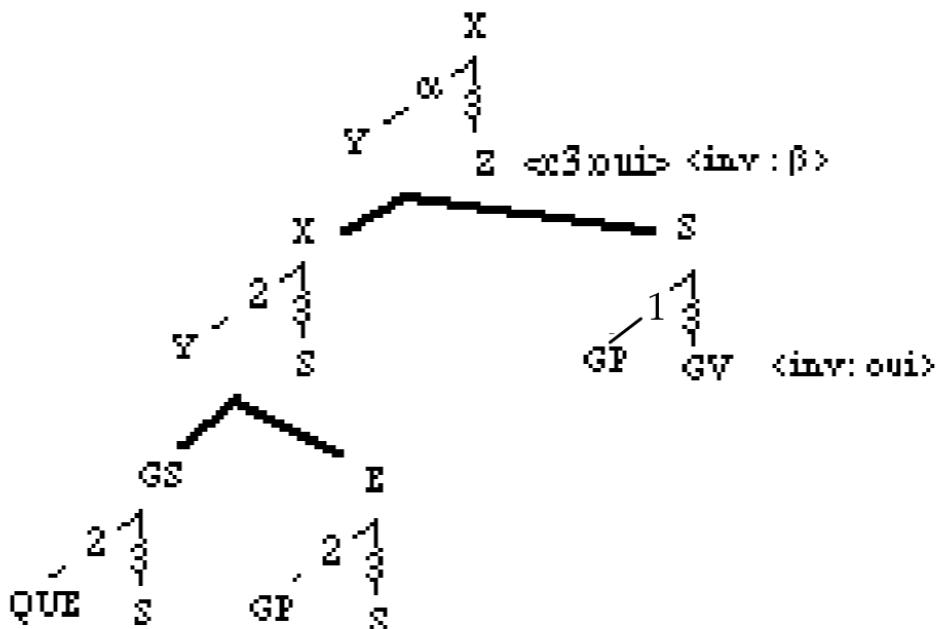
3.3.1. Les constructions à trou

3.3.1.1. Afin de rendre lisible nos représentations, nous précisons une particularité de la GAP que nous élaborons : nous ne traitons pas de manière unifiée les «trous». Nous postulons trois traits c_1 , c_2 et c_3 qui correspondent respectivement à trois types de complément verbal : le sujet, l'objet direct et l'objet indirect¹⁹. Ces traits sont appropriés sur V, GV, S et GS. Le trait $\langle c_3 : \text{non} \rangle$, appliqué à la racine d'un arbre, indique que cet arbre ne peut pas être combiné avec un arbre contenant un complément de type c_3 ; le trait $\langle c_3 : \text{oui} \rangle$ indique que l'on peut mettre en relation la configuration décrite par l'arbre avec un complément de type c_3 ailleurs dans la configuration englobante. Ce traitement n'a aucune incidence sur l'exemple que nous construisons ici ; le lecteur peut lire le trait c_3 comme un trait slash dans ce qui suit.

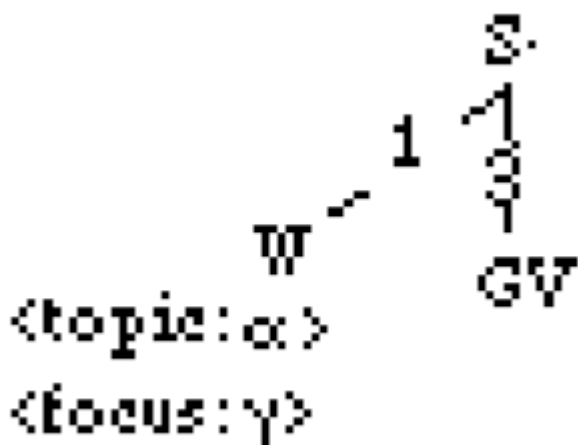
- 18 La position des fillers a été un problème souvent escamoté dans les cadres syntaxiques contemporains si on excepte Chomsky & Lasnik (1977) ou Delaveau (1991) qui, sur la base de (16) pour le français et des équivalents en anglais, posent deux positions qu'ils décrivent par les deux règles : (i) $S'' \rightarrow \text{COMP } S'$ et (ii) $S' \rightarrow X'' S$. Notre analyse de

l'inversion locative revient à admettre une troisième position : la position qu'occupe le sujet dans la phrase canonique. Cette possibilité n'est réalisée que dans le cadre de l'inversion locative. Pour résumer, un filler c3 peut occuper une des trois positions suivantes²⁰ : <E,2> (*A Paul, Marie a donné une récompense*), <GS,2> (*[l'enfant] auquel Marie a donné une récompense*) et <S,1> (*Sur la place se dresse une cathédrale*).

3.3.1.2. La relation filler/gap qui nous intéresse peut se déployer dans trois configurations qui se distinguent par leur catégorie et par le positionnement du filler par rapport au pivot. On figure la hiérarchie des constructions à trous c3 par le diagramme suivant :



- 19 L'inversion locative diffère de la phrase canonique de deux manières : le sujet est obligatoirement postverbal et le GP occupe la position <S,1>. Nous postulons que l'inversion locative et la phrase canonique héritent d'une même construction représentée en (23) :



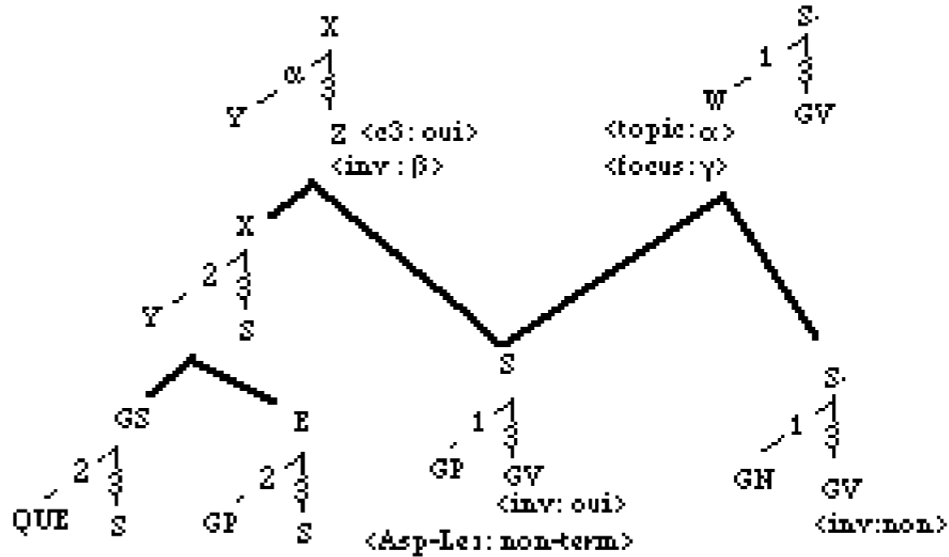
20 (23)

La construction (23), qui définit le domaine phrastique bipositionnel, a une existence autonome du point de vue du conditionnement de l'information : (23) n'est pas associée à un conditionnement particulier²¹. La position <S,1> est remarquable de deux points de vue

(par rapport aux autres positions qui accueillent des GN dans S) : elle accepte tous les types de GN et elle est la seule (dans S) à pouvoir accueillir un topic²².

3.3.3. Hiérarchie

- 21 On peut figurer la position de l'inversion locative dans la hiérarchie des constructions en français par le diagramme suivant :

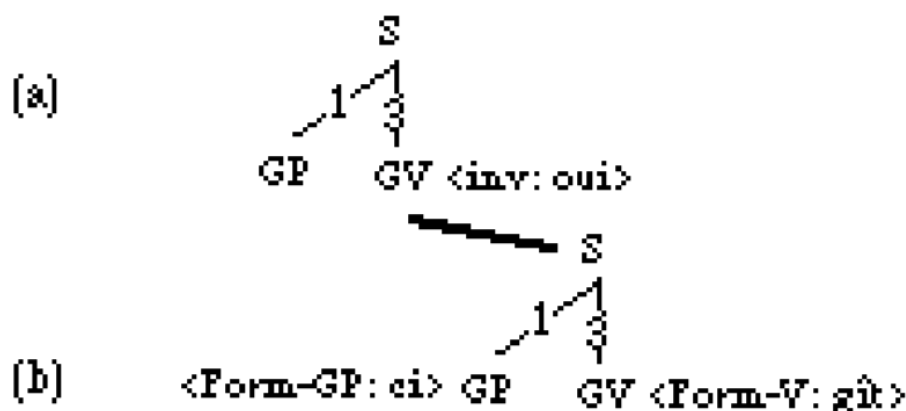


- 22 (24)

L'inversion locative présente de plus des propriétés qui lui sont propres et qui la caractérisent globalement. Nous en avons noté une dans (24) à l'aide du trait Asp(ect)-Lex(ical) : le fait qu'elle ne peut accueillir que des verbes dénotant une éventualité terminative (*atelic*)²³. Ces propriétés ne sont pas héritées, par contre elles sont héritables par toute construction qui hérite de l'inversion locative.

3.3.4. Extension de la hiérarchie

- 23 La spécification des propriétés d'occupation peut faire appel à toutes les propriétés lexicales des termes. En particulier, elle peut faire appel à une catégorie sémantique ou à une forme lexicale. Cette possibilité permet de traiter les expressions toutes faites (expressions idiomatiques). Par exemple, l'expression (pragmatiquement figée) *ci gît GN* peut-être analysée comme une instantiation particulière de l'inversion locative du français. On représente l'analyse par le diagramme (25) qui ajoute une feuille plus spécifiée à la hiérarchie figurée en (24) :



24 (25)

Les phrases quiinstancient cette expression héritent des propriétés spécifiques de l'inversion locative²⁴. L'analyse complète de l'expression «ci gît GN» doit mentionner des propriétés qui lui sont spécifiques ; dans ce cas, des conditions d'emploi particulières (l'inscription sur une pierre tombale).

25 Une expression toute faite est analysée comme une construction régulière dans laquelle les termes sont lexicalement fixés. L'analyse que nous proposons rejoint celle qui est proposée dans les grammaires de constructions²⁵ et, plus généralement, dans toutes les approches de style modulaire du phénomène de l'idiomaticité (Simatos, 1986) : les expressions toutes faites ne sont pas des entités lexicales, mais des entités syntaxiquement définies présentant de fortes restrictions lexicales sur les termes qu'elles admettent²⁶.

4. Les GAP hiérarchisées (GAPH)

26 On trouvera dans cette partie une définition informelle des principales notions nécessaires à l'introduction de l'organisation hiérarchique entre les arbres dans la grammaire. Le lecteur se reportera à l'annexe (§6) où les définitions formelles sont présentées de façon systématique.

4.1. Les arbres sous-spécifiés

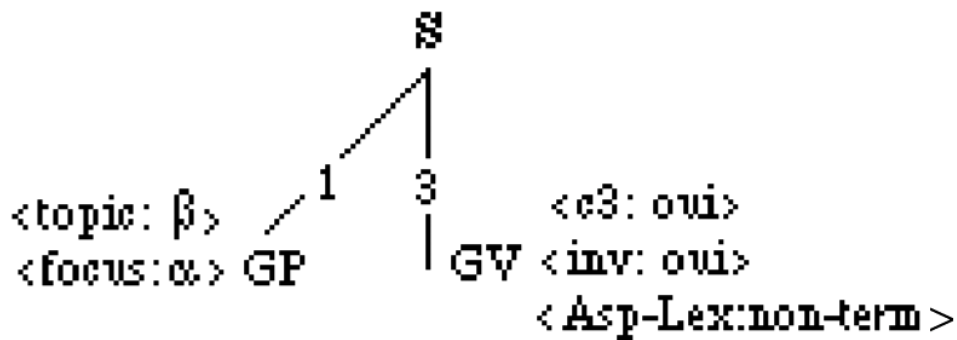
27 L'établissement de la hiérarchie repose sur la possibilité de définir des constructions sous-spécifiées. Par exemple, dans la construction (23), seule l'occupation catégorielle d'une position est laissée sous-spécifiée ; en (22), l'arbre à la racine de la hiérarchie est sous-spécifié sauf pour un trait de la position pivot. Certains arbres appartenant à la grammaire sont donc sous-spécifiés pour l'une ou l'autre de leurs dimensions : la couleur des branches, la catégorie étiquetant les sommets ou bien un trait. C'est ce qui motive l'introduction de variables de couleur, de catégorie et de trait dans la nouvelle définition des arbres polychromes (cf. §6.1.2).

4.2. La relation d'ordre

- 28 On définit une relation d'ordre entre les arbres : l'arbre A_2 subsume l'arbre A_1 . L'arbre A_2 subsume A_1 lorsque A_1 est plus spécifié que A_2 (cf. §6.1.3). Par exemple, la construction (25.b) ci-dessus est plus spécifiée que la construction (25.a) parce que la forme lexicale du terme occupant est fixée.
- 29 C'est cette relation d'ordre qui permet le gain d'adéquation descriptive du formalisme des GAP : elle permet, en effet, d'abstraire les propriétés communes à plusieurs constructions et d'isoler les propriétés qui distinguent les constructions. Les constructions les moins spécifiées représentent celles qui sont les plus générales. C'est ainsi que l'on peut construire des familles de ressemblances entre constructions sans sacrifier le détail de chacune d'entre elles. Comme on l'a vu avec l'exemple de l'inversion locative, la ressemblance peut privilégier la dimension syntaxique (cf. les constructions à trou c3 par exemple, cf. (22) ci-dessus) ou bien la dimension informationnelle (cf. les constructions héritant de (23) ci-dessus). La ressemblance est toujours portée par une configuration de positions, mais la dimension de ressemblance n'est pas fixée a priori.

4.3. Grammaire

- 30 Une grammaire GAPH est constituée d'arbres élémentaires ; à la différence d'une GAP non hiérarchisée, elle est structurée selon une hiérarchie que représente un graphe sans circuit ; on se reportera au §6.2.
- 31 On notera que l'on peut obtenir à n'importe quel degré de la hiérarchie un arbre, qu'on appelle arbre déduit, qui présente la somme des traits particuliers et des traits hérités (cf. §6.2.3). Par exemple, l'arbre déduit de l'inversion locative est le suivant :



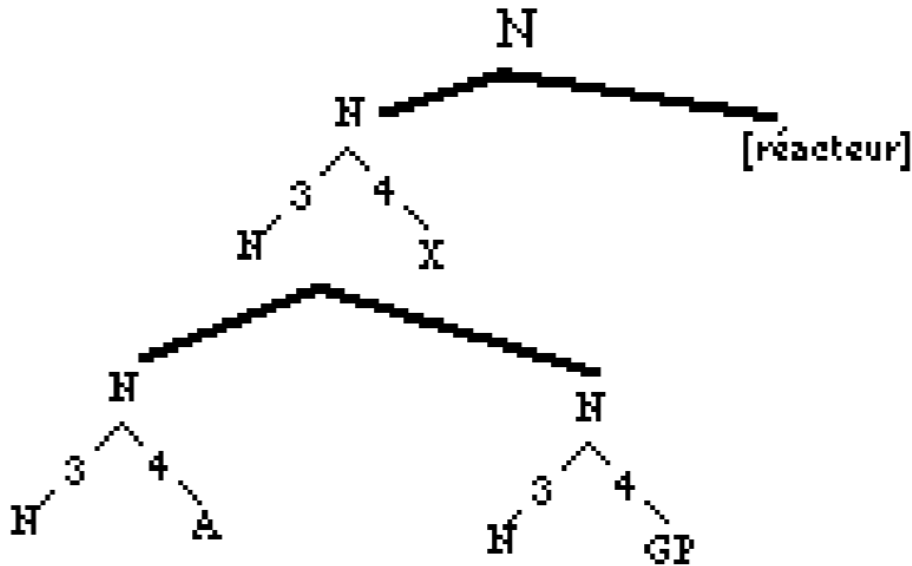
32 >

(26)

Une représentation telle que (26) en regroupant propriétés générales et particulières permet d'isoler dans la grammaire du français un tour particulier, alors qu'une représentation comme (23) ci-dessus permet d'appréhender une forme générale de construction du français. Ce degré variable de généralité caractérise la notion de construction en général²⁷.

- 34 Les études en HPSG ont montré la fécondité de l'approche hiérarchique pour modéliser le lexique conçu comme un ensemble de formes (cf. Sag (à par.) et les références qui y sont données). On peut postuler, à côté de la hiérarchie grammaticale, une hiérarchie lexicale ;

les deux hiérarchies peuvent se croiser en certains points. On peut modéliser grâce à ces points de croisement les phénomènes d'hybridité entre lexique et syntaxe. C'est, par exemple, le cas de certaines unités nominales polylexicales qui peuvent être analysées comme des unités syntaxiques (ressemblance avec le GN) et des unités lexicales (en particulier pour les propriétés référentielles) (David 1993). On figure par le diagramme (27) un fragment de hiérarchie montrant les réalisations de N : N peut être une unité simple (*réacteur*), une unité «N+GP» (*réacteur à neutrons*) ou «N+A» (*réacteur nucléaire*) :



35 (27)

4.4. Composition d'arbres

- 36 L'ajout de la hiérarchie constitue un enrichissement de la structure de la base d'arbres et n'entraîne pas de modification de l'opération de composition prise en elle-même et telle que nous l'avons définie dans le formalisme de départ.
- 37 Une contrainte sur les arbres à composer doit toutefois être ajoutée dans une GAP (cf. §6.3.1) : lorsque l'on compose un arbre élémentaire A avec un arbre complexe B, on doit vérifier qu'il n'existe pas un arbre A' plus spécifié que A au regard des attentes de B. Prenons l'exemple de (25.b) (l'arbre décrivant l'expression *ci gît*) : il doit être composé avec un arbre de racine GV muni du trait <inv : oui> et <V-form : gît>. Si on le composait avec un arbre ne spécifiant pas la forme du verbe, on pourrait obtenir un énoncé mal formé comme : * *ci dort Paul* .

4.5. Expressivité de la représentation

- 38 Nous avons insisté dans l'illustration développée au §3 sur le fait que le dispositif hiérarchique permet d'avoir une approche analytique de la syntaxe en dégageant des structures plus ou moins générales. Nous montrons dans ce paragraphe qu'il permet également la représentation unifiée des notions de la théorie syntaxique, en particulier des fonctions grammaticales. Nous terminons donc notre illustration en donnant une représentation unifiée de la notion de sujet grammatical en français.

4.5.1. Définition du sujet grammatical

- 39 On admet que le sujet est le complément distingué du verbe : il présente des propriétés que ne présentent pas les autres compléments. Ses propriétés sont variables selon les langues «à sujet»; en français, on peut tenir son caractère obligatoire et l'accord en nombre et en personne pour les traits de distinction les plus évidents. Par ailleurs, on peut admettre, à la suite de la plupart des approches syntagmatiques, qu'il correspond à l'argument distingué du verbe.

4.5.2. Une autre position pour le sujet grammatical

- 40 Le français présente un autre cas «d'inversion du sujet» qui n'est pas liée à une construction à trou ; ce cas est illustré en (28) :

| | | |
|------|----|--|
| (28) | a. | <i>S'efforceront de rendre un devoir supplémentaire tous les élèves qui n'ont pas la moyenne</i> |
| | b. | <i>Recevra un reçu chaque héritier qui en fera la demande</i> |

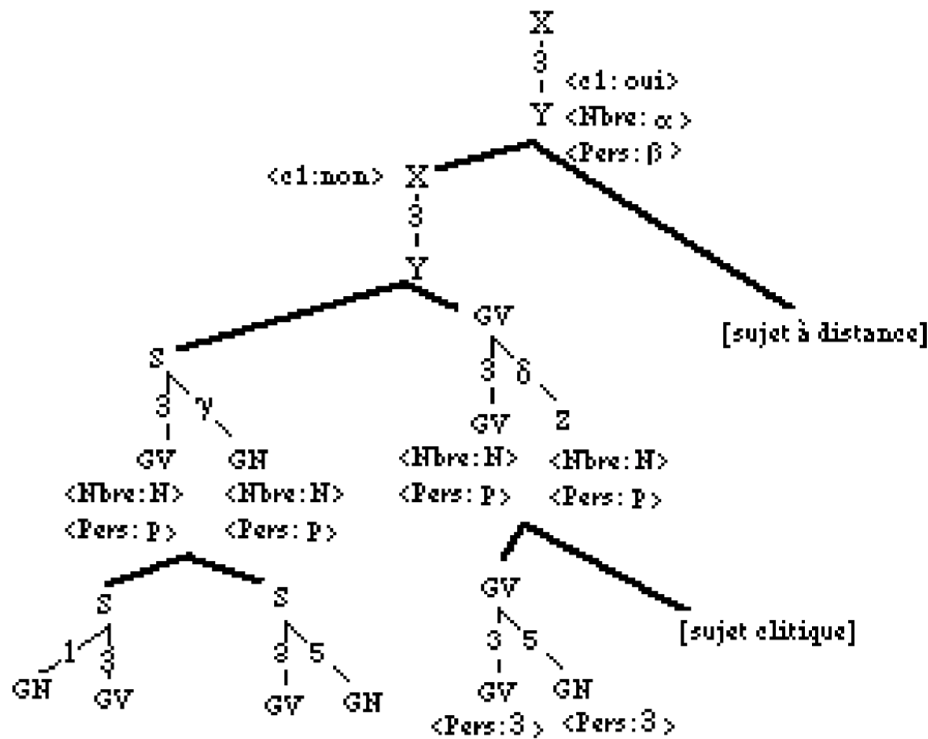
Ces phrases ont une propriété tout à fait distinctive : la réalisation des compléments dans le champ verbal n'est pas affecté par la réalisation postverbale du sujet. En (29.a), la réalisation postverbale du sujet bloque la réalisation d'un GN objet référentiel (à la différence de (28.b)) et en (29.b), la réalisation postverbale du sujet n'est pas possible avec un verbe à contrôle du sujet sans extraction (Bonami & Godard, à par.) (à la différence de (28.a)) :

| | | |
|------|----|---|
| (29) | a. | <i>* La salle où rencontreront le proviseur les élèves en grève</i> |
| | b. | <i>* Le jour où s'efforceront de rendre un devoir supplémentaire les élèves en retard, je reprendrai espoir</i> |

C'est un des indices pour analyser les phrases de (28) comme «une permutation GN-GV» : le sujet occupe une position <S,5> dans les phrases de (28).

4.5.3. Représentation syntaxique du sujet grammatical

Si la définition est unitaire, par contre les positions de réalisation du sujet sont diverses. On en a recensé trois jusqu'à présent en ce qui concerne les constituants : <S,1> (*L'homme est venu*), <GV,5> (*Sur la place se dresse une cathédrale*) et <S,5> (*Passeront à mon bureau tous les élèves qui n'ont pas la moyenne*)²⁸. Nous proposons la hiérarchie suivante qui capte la diversité de réalisation du sujet :



41 (29)

Le diagramme (29) représente une analyse qui est diamétralement opposée à la définition configurationnelle classique donnée dans le cadre de *Aspects*. Selon *Aspects*, la fonction sujet est réduite à un trait distinctif : être le GN le plus élevé dans le domaine phrastique. Le trait définitoire du sujet est configurationnel et est établi en structure profonde. La diversité phénoménologique est obtenue en admettant une notion de transformation²⁹.

42 Dans (29), le contenu configurationnel de la définition du sujet est faible (cf. l'arbre en racine du fragment de hiérarchie) : le trait <cl : oui> note qu'un sujet doit être réalisé dans une configuration. Surtout, (29) met bien en évidence un fait qui n'a pas assez retenu l'attention jusqu'à présent : il y a un grand nombre de positions possibles pour le sujet à la différence du GN objet direct et du GP objet indirect qui sont confinés dans le groupe verbal. Ce caractère doit être compté au titre des traits caractérisant le sujet grammatical en français.

5. Conclusion

43 L'utilisation d'un formalisme à base d'arbres pour représenter les unités élémentaires des langues a fait l'objet de critiques radicales. En réponse à ces critiques, nous avons déjà montré qu'il suffisait de modifier l'opération de composition pour pallier la surgénération hiérarchique (contra Milner 1984) et pour traiter les phénomènes de discontinuité syntagmatique (contra, entre autres, McCawley 1982). Nous venons de montrer que la hiérarchisation de la grammaire permet de contourner la limitation descriptive des grammaires syntagmatiques en général : elle permet à la fois de faire émerger des formes générales et de sauver l'extrême détail des organisations de langue.

44 La théorie syntaxique qui s'exprime dans les GAP postule un imperceptible spécifique : la position. Dans les GAP, chaque position constitue un individu ; dans les GAPH,

l'introduction de la forme hiérarchique permet de travailler avec des classes de positions. Elle permet, par ailleurs, de distinguer clairement les propriétés absolues et relatives des positions. Mais, surtout, la hiérarchie permet de sauver le détail des langues sans entraîner la théorie syntaxique dans le pur et simple enregistrement de petits faits.

6. Annexe : les définitions formelles des GAPH

6.1 Arbres polychromes sous-spécifiés

6.1.1 Arbres polychromes sans variable

- 45 On se donne un nombre entier strictement positif p (le nombre de couleurs), un nombre k inférieur ou égal à p (la couleur noyau), un ensemble fini Cat de catégories, un ensemble T d'attributs et un ensemble fini E de valeurs de traits. On note $Col = \{1, \dots, p\}$ l'ensemble des couleurs.
- 46 Un arbre polychrome sans variable est défini par un 6-uplet $A = \langle X, L, d, P, \tau, t \rangle$ où X est l'ensemble des sommets, L est une application de X dans Cat , d est un sous-ensemble de $X \times X \times Col$, P est un ordre strict partiel de $X \times X$, la relation de précédence, τ est une application de X dans l'ensemble des parties de T qui définit pour chaque sommet l'ensemble des attributs qui lui sont associés ; t enfin associe à chaque sommet x une application $t(x) : \tau(x) \rightarrow E$ qui définit les valeurs associées à chaque attribut pour le sommet. On écrit $t(x, y)$ plutôt que $t(x)(y)$.
- 47 Nous renvoyons le lecteur à Cori & Marandin 1994 pour le détail des conditions qui doivent être vérifiées. Elles expriment les conditions de Wall, précisent qu'il y a un ordre entre les couleurs pour les fils d'un même sommet et que la couleur pivot est toujours présente de manière unique parmi les fils d'un sommet donné.

6.1.2 Introduction de variables

- 48 On se donne un ensemble (dénombrable) Φ de variables (de catégories, de couleurs ou de traits). Un arbre polychrome est maintenant défini par un 7-uplet $A = \langle X, L, d, P, \tau, t, R \rangle$ où X, P et τ jouent le même rôle que ci-dessus, L est une application de X dans $Cat \cup \Phi$, d est un sous-ensemble de $X \times X \times (Col \cup \Phi)$, t associe à chaque sommet x une application $t(x) : \tau(x) \rightarrow E \cup \Phi$ ^[2]
- 49 Les ensembles $\Phi_{Cat}(A) = \Phi \cap L(X)$, $\Phi_{Col}(A) = \{f \in \Phi ; \exists \langle x, y, f \rangle \in d\}$ et $\Phi_E(A) = \{f \in \Phi ; \exists x \in X \exists u \in \tau(x) t(x, u) = f\}$ doivent être disjoints. Soit $\Phi(A) = \{f_1, \dots, f_m\}$ le sous-ensemble (fini) de Φ obtenu par la réunion de ces trois ensembles.
- 50 R est une relation m -aire non vide telle que pour tout m -uplet $\langle u_1, \dots, u_m \rangle \in R$:
- 51 - si $f_i \in \Phi_{Cat}(A)$ alors $u_i \in Cat$;
- 52 - si $f_i \in \Phi_{Col}(A)$ alors $u_i \in Col - \{k\}$ ³⁰ ;
- 53 - si $f_i \in \Phi_E(A)$ alors $u_i \in E$;

et la structure obtenue en remplaçant dans A chaque variable f_i par la valeur u_i correspondante est un arbre polychrome sans variable.

R est la relation d'interdépendance, qui englobe à présent les variables de catégorie, de couleur et de traits.

On notera que chaque variable de Φ qui sert à la définition d'un arbre est particulièrement dédiée à représenter une catégorie, une couleur ou une valeur de trait. Cette spécialisation doit être préservée dans les substitutions qui portent sur de telles variables. A une variable représentant une catégorie, une substitution ne peut associer qu'une catégorie ou une variable de catégorie, etc.

6.1.3 Relation d'ordre entre arbres

54 Soient deux arbres $A_1 = \langle X_1, L_1, d_1, P_1, \tau_1, t_1, R_1 \rangle$ et $A_2 = \langle X_2, L_2, d_2, P_2, \tau_2, t_2, R_2 \rangle$ tels que $\Phi(A_1) = \{f_1, \dots, f_m\}$ et $\Phi(A_2) = \{h_1, \dots, h_q\}$. On écrit $A_1 \sqsubseteq A_2$, et on dit que A_2 *subsume* A_1 , si et seulement si il existe une application injective $g : X_2 \rightarrow X_1$ et une substitution

$$\sigma : \Phi(A_2) \cup \text{Cat} \cup \text{Col} \cup E \rightarrow \Phi(A_1) \cup \text{Cat} \cup \text{Col} \cup E$$

qui vérifie les conditions suivantes :

55 (i) $\leftrightarrow x \in X_2 \quad L_1(g(x)) = \sigma(L_2(x))$

Des catégories qui étaient variables dans A_2 sont constantes dans A_1 , ou des catégories variables dans A_2 restent variables dans A_1 mais doivent être égales ;

56 (ii) $\leftrightarrow x, y \in X_2 \quad \langle x, y, \alpha \rangle \in d_2 \Rightarrow \langle g(x), g(y), \sigma(\alpha) \rangle \in d_1$

Un sommet a autant ou davantage de fils dans A_1 que dans A_2 . Ceci se justifie par le fait que lorsque dans un arbre on indique les fils d'un sommet donné, on n'interdit pas que ce sommet ait d'autres fils. De plus, des couleurs qui étaient variables dans A_2 sont constantes dans A_1 , ou des couleurs variables dans A_2 restent variables dans A_1 mais doivent être égales ;

57 (iii) $\leftrightarrow x, y \in X_2 \quad \langle x, y \rangle \in P_2 \Rightarrow \langle g(x), g(y) \rangle \in P_1$

58 (iv) $\leftrightarrow x \in X_2 \quad \tau_2(x) \subset \tau_1(g(x))$

Il peut y avoir en A_1 des attributs qu'il n'y a pas en A_2 ;

59 (v) $\leftrightarrow x \in X_2 \quad \leftrightarrow u \in \tau_2(x) \quad t_1(g(x), u) = \sigma(t_2(x, u))$

Des valeurs de traits qui étaient variables dans A_2 sont constantes dans A_1 , ou des valeurs qui étaient variables dans A_2 restent variables dans A_1 mais doivent être égales ;

60 (vi) Pour tout m -uplet $\langle u_1, \dots, u_m \rangle \in R_1$, il doit exister $\langle v_1, \dots, v_q \rangle \in R_2$ tel que pour tout $i \in \{1, \dots, q\}$, si $\sigma(h_i) \notin \Phi(A_1)$ alors $v_i = \sigma(h_i)$; sinon, $\sigma(h_i) = f_j$ et $v_i = u_j$

La relation d'interdépendance de A_2 comporte plus de n -uplets (ramenés aux variables communes) que la relation d'interdépendance de A_1 .

6.2. Grammaires hiérarchisées

6.2.1. Arbres élémentaires

61 Un arbre polychrome sans variable est dit *élémentaire* si et seulement si il ne comporte qu'un niveau (seule la racine a des fils) et si toutes les couleurs en sont distinctes. Autrement dit :

62 (i) $\langle x, y, i \rangle \in d$ et $\langle u, v, j \rangle \in d \Rightarrow x = u$

63 (ii) $\langle x, y, i \rangle \in d$ et $\langle x, z, i \rangle \in d \Rightarrow y = z$

Un arbre polychrome avec variables est dit *élémentaire* si et seulement si chaque n -uplet de sa relation d'interdépendance définit un arbre polychrome élémentaire.

6.2.2 Graphe d'arbres

64 Une grammaire hiérarchisée sera donnée par un ensemble d'arbres élémentaires (auquel on pourra joindre le lexique, c'est-à-dire un ensemble d'arbres ponctuels) qui sera structuré selon un graphe sans circuit. Les descendants d'un arbre élémentaire donné A dans ce graphe devront tous être des arbres plus spécifiques que A .

65 Afin d'éviter de répéter des traits identiques à plusieurs niveaux de la hiérarchie, ce qui serait redondant, on les indique au niveau le plus haut. C'est pourquoi une grammaire est définie en deux étapes. On donne d'abord un graphe dans lequel les traits sont indiqués au niveau adéquat (la grammaire hiérarchisée effective). A partir de ce premier graphe, on construit un second graphe, le *graphe déduit*, qui contient tous les traits et pour lequel la relation de subsomption est effectivement vérifiée entre un sommet et tous ses descendants. Cette construction n'est possible que s'il n'y a pas d'incompatibilités dans la grammaire.

6.2.2.1. Une *grammaire hiérarchisée* est définie par un couple $\langle B, \Gamma \rangle$, B étant un ensemble d'arbres élémentaires (ou ponctuels) et Γ étant une application qui, à un arbre A de B , associe un ensemble d'arbres $\Gamma(A)$ de B de telle manière que le graphe ainsi construit n'admette pas de circuit.

6.2.2.2. Une grammaire hiérarchisée est *bien hiérarchisée* si et seulement si, pour tout arbre polychrome $A = \langle X, L, d, P, \tau, t, R \rangle$ de la grammaire dont tous les ascendants sont A_1, \dots, A_n (avec $A_i = \langle X_i, L_i, d_i, P_i, \tau_i, t_i, R_i \rangle$), il existe n applications injectives $g_i : X_i \rightarrow X$ et une substitution

66 $\sigma : (\cup_{i \in \{1, \dots, n\}} \Phi(A_i)) \cup \text{Cat} \cup \text{Col} \cup E \rightarrow \Phi \cup \text{Cat} \cup \text{Col} \cup E$
vérifiant :

67 (i) $\leftrightarrow i \leftrightarrow x \in X_i \quad L(g_i(x)) = \sigma(L_i(x))$
(compatibilité de catégories) ;

68 (ii) $\leftrightarrow i \leftrightarrow x, y \in X_i \quad \langle x, y, \alpha \rangle \in d_i \Rightarrow \langle g_i(x), g_i(y), \sigma(\alpha) \rangle \in d$
(compatibilité de domination) ;

69 (iii) $\leftrightarrow i \leftrightarrow x, y \in X_i \quad \langle x, y \rangle \in P_i \Rightarrow \langle g_i(x), g_i(y) \rangle \in P$
(compatibilité de précédence) ;

70 (iv.1) $\leftrightarrow i \leftrightarrow x \in X \leftrightarrow y \in X_i \quad g_i(y) = x \Rightarrow$
 $\leftrightarrow u \in \tau(x) \cap \tau_i(y) \quad t(x, u) = \sigma(t_i(y, u))$

71 (iv.2) $\leftrightarrow i \leftrightarrow j \leftrightarrow y \in X_i \leftrightarrow z \in X_j \quad g_i(y) = g_j(z) \Rightarrow$
 $\leftrightarrow u \in \tau_i(y) \cap \tau_j(z) \quad \sigma(t_i(y, u)) = \sigma(t_j(z, u))$
(compatibilité entre les traits) ;

(v) Pour tout r -uplet $\langle u_1, \dots, u_r \rangle \in R$ et pour tout $i \in \{1, \dots, n\}$, il doit exister $\langle v_1, \dots, v_q \rangle \in R_i$ tel que pour tout $j \in \{1, \dots, q\}$, si $\sigma(h_j) \notin \Phi$ alors $v_j = \sigma(h_j)$; sinon, $\sigma(h_j) = f_1$ et $v_j = u_1$ (en supposant que $\Phi(A_i) = \{h_1, \dots, h_q\}$). (Compatibilité entre les relations d'interdépendance).

6.2.3 Graphe déduit

72 Soit $A = \langle X, L, d, P, \tau, t, R \rangle$ l'arbre d'une grammaire bien hiérarchisée, soient A_1, \dots, A_n (avec $A_i = \langle X_i, L_i, d_i, P_i, \tau_i, t_i, R_i \rangle$) tous les ascendants de A dans le graphe, soient g_1, \dots, g_n les

applications injectives et σ la substitution qui assurent que le graphe est bien hiérarchisé, prises de telle façon que σ soit la plus générale possible.

- 73 On définit l'arbre déduit $D(A) = \langle X, L, d, P, \tau', t', R' \rangle$ de A dans le graphe comme suit :
- 74 (i) $\leftrightarrow x \in X \quad \tau'(x) = \tau(x) \cup \{u \in T; \exists i \in \mathbb{N} \exists y \in X_i \quad g_i(y) = x \text{ et } u \in \tau_i(y)\}$
- 75 (ii.1) $\leftrightarrow i \leftrightarrow y \in X_i \leftrightarrow u \in \tau_i(y) \quad t'(g_i(y), u) = \sigma(t_i(y, u))$
- 76 (ii.2) $\leftrightarrow x \in X \quad \leftrightarrow u \in \tau(x) \quad t'(x, u) = t(x, u)$
- 77 (iii) $\Phi(D(A)) = \Phi(A) \cup [(\cup_{i \in \{1, \dots, n\}} \sigma(\Phi(A_i))) \cap \Phi]$
- 78 On pose $\Phi(A) = \{f_1, \dots, f_m\}$ et $\Phi(D(A)) = \{f_1, \dots, f_m, f_{m+1}, \dots, f_r\}$ et on écrit que $\langle u_1, \dots, u_r \rangle \in R'$ ssi $\langle u_1, \dots, u_m \rangle \in R$ et pour tout i il existe $\langle v_1, \dots, v_s \rangle \in R'$ tels que et pour tout j , si $\sigma(h_j) \notin \Phi$ alors $v_j = \sigma(h_j)$; sinon, $\sigma(h_j) = f_i$ et $v_j = u_i$ (en supposant que $\Phi(A_i) = \{h_1, \dots, h_q\}$).
- 79 On démontre que $D(A)$ est subsumé par tous ses ascendants A_i , ainsi que par les arbres déduits $D(A_i)$ des ascendants.
- 80 Le graphe obtenu en remplaçant chacun des arbres d'un graphe bien hiérarchisé par son arbre déduit est le *graphe déduit*.

6.3 Composition d'arbres

6.3.1 Composabilité

- 81 Deux arbres sans variable sont composables quand la catégorie étiquetant une feuille du premier est identique à une catégorie étiquetant la racine du second et qu'il y a compatibilité entre les traits. L'introduction des variables nécessite l'introduction d'une substitution σ (on vérifie si les catégories ou les valeurs de traits peuvent être rendues égales). Il faut par ailleurs s'assurer qu'une éventuelle composition ne forme pas un arbre dont la relation d'interdépendance soit vide (cf. Cori & Marandin 1993). On étend facilement ce qui est dit à propos de la relation d'interdépendance qui ne porte que sur les valeurs de traits à la nouvelle relation d'interdépendance qui porte également sur les catégories et les couleurs.
- 82 Il faut néanmoins introduire ici la nouvelle condition de composabilité ci-dessous qui permet d'assurer qu'on saura dans quel cas de la composition on se trouve, et, que dans le cas où il y a compactage, l'ordre de précedence ne sera pas indéterminé.
- 83 Soient $A' = \langle X', L', d', P', \tau', t', R' \rangle$ et $A = \langle X, L, d, P, \tau, t, R \rangle$ les deux arbres à composer. Soit a la racine de A , b une feuille de A' et c le père de b . Pour que A' et A soient composables selon b il faut que si $\langle c, b, k \rangle \in d'$ alors $L'(c) \in \text{Cat}$ et $L'(b) \in \text{Cat}$ et si $L'(c) = L'(b)$ alors
- 84 (i) $\langle c, x, \mathbb{E} \rangle \in d' \Rightarrow \mathbb{E} \in \text{Col}$
- 85 (ii) $\langle a, y, \mathbb{E} \rangle \in d \Rightarrow \mathbb{E} \in \text{Col}$
- 86 Dans le cas où le sommet b occupe la position noyau dans un arbre élémentaire, il faut que sa catégorie soit connue ainsi que celle de son père. Par ailleurs, lorsqu'il y a compactage, il faut que l'ordre des sommets concernés par ce compactage soit parfaitement déterminé.

6.3.2 Composition

- 87 La définition de la composition de Cori et Marandin 1993, 1994 s'adapte immédiatement à la nouvelle définition des arbres polychromes. On note $A = [A', b, A \gg]$ l'arbre obtenu par la composition de A' par $A \gg$ selon b .

6.3.3 Arbres engendrés par une grammaire bien hiérarchisée

- 88 On considère une grammaire $G = \langle B, \Gamma \rangle$. On définit $T[G]$, l'ensemble des arbres engendrés par la grammaire de la manière suivante :
- 89 (i) si $A \in B$ alors $D(A) \in T[G]$
- 90 (ii) si $A \in T[G]$ et $A_1 \in B$ alors $[A, b, D(A_1)] \in T[G]$ ssi il y a composabilité et s'il n'existe pas dans le graphe un arbre A_2 tel que $D(A_2) \boxtimes D(A_1)$ et $\tau(b) \cap \tau_1'(a) \subset \tau(b) \cap \tau_2'(g(a))$ l'inclusion étant stricte et g étant l'application injective qui définit la subsomption.
- 91 L'ensemble des arbres engendrés par une grammaire représente un ensemble de descriptions : les descriptions des parties d'énoncé et des énoncés bien formés d'une langue donnée.

BIBLIOGRAPHIE

- BONAMI, O. & GODARD, D. (à par.) Inversion du sujet, constituance et ordre des mots, [Lardreau G., ed.], Cahier Jean-Claude Milner, Paris : Verdier.
- BONAMI, O., GODARD, D. & MARANDIN J.-M. (1998) French subject inversion in extraction context, [Bouma G et al., eds.], Proceedings of FHCG '98, U. de Sarrebrück.
- BORSLEY, R. (1996) Modern Phrase Structure Grammar, Oxford : Blackwell Publ.
- BRESNAN, J. (1994) Locative inversion and the architecture of universal grammar, *Language* 70 : 72-131.
- CHOMSKY, N. (1971 [1965]) Aspects de la théorie syntaxique, Paris : Le Seuil.
- CHOMSKY, N. & LASNIK, H. (1977) Filters and Control, *Linguistic Inquiry* 8 : 425-504.
- CINQUE, G. (1997 [1982]) 'Topic' Constructions in some European Languages and 'Connectedness' [Anagnostopoulou E. et al., eds.] *Materials on Left Dislocation*, Amsterdam : J. Benjamins.
- CLECH-DARBON, A., REBUSCHI, G. & RIALLAND, A. (1998) Are there Cleft Sentences in French ?, U. de Paris 3, ms.
- COOPER, R. (1990) Classification-based Phrase Structure Grammar : an Extended Revised Version of HPSG, University of Edinburgh : PhD, np.
- CORI, M. & MARANDIN, J.-M. (1993) Grammaires d'arbres polychromes, *T.A.L.* 34 : 101-132.
- CORI, M. & MARANDIN, J.-M. (1994) Polychrome Tree Grammars (PTG) : a formal approach, [Martin-Vide ed.], *Current Issues in Mathematical Linguistics* : 141-149, Amsterdam : Elsevier.

- CORI, M. & MARANDIN, J.-M. (1995) Trees and the Representation of Disjunct Constituents, ESSLLI Conference on Formal Grammar [Morris G. and Oehrle R., eds], Barcelone, août 1995.
- DAVID, S. (1993) Les unités nominales polylexicales. *Éléments de description et reconnaissance automatique*, Université de Paris 7 : thèse (NR), np.
- DELAVEAU, A. (1991) La structure de la phrase en français : le témoignage des formes disloquées, [Eskénazi A. et al., ed.] *Études de linguistique française réunies à la mémoire de A. Lerond* : 195-212, Université de Paris X-Nanterre.
- DIDERICHSEN, P. (1943) Logische und topische Gliederung des germanischen Satzes, [Bjerrum et al., eds] *Paul Diderischen (...)*, Copenhague : CEG Gads Forlag.
- DOWTY, D. (1997) Towards a minimalist theory of syntactic structure, [Hock & Sitjma, eds] *Discontinuous Constituency*, La Haye : Mouton de Gruyter.
- FAARLUND, J. (1989) Discourse Functions and Syntax, *CLS 25 (Papers from the 25th Annual Meeting of the Chicago Linguistic Society. Part II)* : 30-40, Chicago : CLS.
- FARKAS, D. (1997) Towards a Semantic Typology of Noun Phrases, *Communication à CSSP97*, Université Paris 7, ms.
- GODARD, D. (1980) Les relatives complexes en français, Université de Paris 7 : thèse de 3ème cycle, np.
- HELTOFT, L. (1992) The Topology of Verb Second and SVO Languages. A study in the Sign Functions of Word Order, [Herslund M. ed] *Word order* : 13-64, *Copenhaguen Studies in Language* 15.
- KARTTUNEN, L. & KAY, M. (1985) Parsing a free order language [Dowty D ; et al., eds] *Natural language Parsing*, Cambridge : CUP.
- KAYNE, R. & POLLOCK, J.-Y. (1978) Stylistic Inversion, Successive Cyclicity, and Move NP in French, *Linguistic Inquiry* 9 : 595-621.
- KATHOL, A. & POLLARD, C. (1995) On the Left Periphery of German Subordinate Clause, np.
- LAMBRECHT, K. (1994) Information structure and sentence form. Topic, focus, and the mental representation of discourse referents, Cambridge : Cambridge UP.
- LE BIDOIS, R. (1950) *L'inversion du sujet dans la prose contemporaine*, Paris : Editions d'Artey.
- MARANDIN, J.-M. (1997) Dans le titre se trouve le sujet. L'inversion locative, U. de Paris 7 : *Mémoire d'habilitation*, np.
- MARANDIN, J.-M. (à par.) Sites et constructions dans la théorie de la syntaxe, [Lardreau G., ed.], *Cahier Jean-Claude Milner*, Paris : Verdier.
- MCCAWLEY, J. (1982) Parentheticals and Discontinuous Constituent Structure, *Linguistic Inquiry* 13 : 91-106
- MILNER, J.-C. (1978) Cyclicité successive, comparatives, et Cross-over en français, *Linguistic Inquiry* 9 : 673-693.
- MILNER, J.-C. (1984) De l'inutilité des arbres en linguistique, *T.A. informations* 25 : 5-12.
- MILNER, J.-C. (1989) *Introduction à une science du langage*, Paris : Le Seuil.
- POLLOCK, J.-Y. (1986) Sur la syntaxe de *en* et le paramètre du sujet nul, [Ronat M. et Couquaux D.], *La grammaire modulaire* : 211-248, Paris : Edition de Minuit.

POLLARD, C. & SAG, I.A. (1994) *Head-Driven Phrase Structure Grammar*, Chicago : University of Chicago Press.

SAG, I.A. (à par.) *English Relative Constructions*, *Journal of Linguistics*.

SIMATOS, I. (1986) *Eléments pour une théorie des expressions idiomatiques*, U de Paris 7 : doctorat, np.

VALLDUVI, E. & ENGDahl, E. (1996) *The linguistic realization of information packaging*, *Linguistics* 34 : 459-519.

VALLDUVI, E. & VILKUNA, M. (1997) *On rheme and kontrast* [Culicover & McNally] *The limits of Syntax*, NY : Academic Press.

WALL, R. (1972) *Introduction to Mathematical Linguistics*, Englewood Cliffs : Prentice-Hall.

ZAENEN, A. (1983) *On Syntactic Binding*, *Linguistic Inquiry* 14 : 469-504.

ZWICKY, A. (1994) *Dealing out meaning : Fundamentals of syntactic constructions*, BLS 20.

NOTES

1. Les définitions formelles sont en partie rappelées au §6.1.1 ; nous renvoyons le lecteur à Cori & Marandin 1993 , 1994 pour une présentation de détail.

2. L'ordre dit libre correspond souvent au fait que des termes entretenant des relations à des têtes distinctes appartiennent au même domaine de positionnement.

3. Les configurations dans les GAP ne correspondent pas toujours aux syntagmes dans l'approche syntagmatique classique. Par exemple, prennent place dans la configuration que nous appelons GV les dépendants du verbe (compléments et ajouts) ainsi que les unités interpolées (incises, etc. ; cf. Cori & Marandin 1995).

4. Milner (1989) montre justement que la représentation (1) est ambivalente : elle est susceptible d'une lecture mérologique et d'une lecture positionnelle où la relation d'occupation est envisagée seulement en termes catégoriels. Une des motivations premières des GAP était de lever cette ambivalence.

5. Le diagramme (3) est illustratif ; nous traitons la réalisation du sujet différemment (voir ci-dessous le §4.5) ; la représentation du statut informationnel est librement inspirée de Vallduvi & Engdahl (ibid.) (<Focus :-> = Ground ; le trait <Topic> n'est approprié que pour une position <Focus ;+> ; <Topic :-> = Tail). Cette représentation est provisoire.

6. Pour une présentation plus exhaustive, on se reportera à Marandin (à par.).

7. La c-structure est une représentation ambivalente au même titre que (1) ci-dessus : elle définit les catégories syntagmatiques et localise les termes des fonctions grammaticales (et discursives) dans la phrase. «On c-structure are defined the various structural configurations for syntactic functions (such as the position of the TOP function), [...] In general, c-structure gives the surface positions and forms of the constituents» (Bresnan, 1994 : 104).

8. Dans les versions récentes de HPSG les types de branches sont réduits à deux : tête et non-tête (Sag, à par.).

9. On peut également justifier un niveau de représentation autonome et «pure» des positions en reprenant la distinction entre les positions A (les positions constitutives du domaine phrastique) et les positions A-barre (les positions externes au domaine phrastique) dans le modèle GB. Les premières sont réputées monocatégorielles et monofonctionnelles, les secondes sont polycatégorielles et polyfonctionnelles. Cette distinction revient à poser que les positions A sont spécialisées pour un type de fonctionnement (la réalisation fonctionnelle des dépendants du verbe) et les positions A-barre ne le sont pas. Le fait d'être spécialisé pour l'accueil d'un type de

constituant est un cas particulier du fait de ne pas être spécialisé ; or, précisément, les catégories A-barre ne peuvent pas être définies catégoriellement, ce que marque le recours à des catégories ad hoc comme SPEC.

10. Plus exactement de la phrase à verbe tête *be*. On observe, en effet, que le verbe *be* admet des sujets phrastiques (par exemple : *I am tired is all* (Pollard & Sag, 1994 : 382 fn)). On notera que le français connaît ce type de construction (cf. entre autres Clech-Darbon et al.) mais le site <S,1> ne peut pas accueillir de phrase finie non marquée (les équivalents de (7) et (8) sont agrammaticaux en français). On ne peut donc pas recourir à une explication de nature sémantique pour distinguer entre les deux langues : le facteur discriminant est bien l'occupabilité de la position <S,1>.

11. Au sens de la théorie des champs (Diderichsen 1943, Heltoft 1992, Faarlund 1989). «Topological concepts (fields and positions) are locations or sites for complex and simple constituents, but they do not themselves denote constituents» (Heltoft, 1992 : 18).

12. Il est, par ailleurs, indépendant du formalisme arborescent et de la théorie syntaxique ; voir Bonami dans ce numéro pour un choix différent.

13. On se reportera pour l'analyse de détail à Marandin 1997, Bonami et al. 1998.

14. L'exemple (12.a) est un exemple simplifié de Proust cité dans Le Bidois (1950).

15. La phrase à GP topicalisé est associée à une contrainte de bonne formation portant sur la relation entre le référent de discours (rd dorénavant) introduit par le GP et un rd appartenant à l'univers de discours, qui limite ses possibilités d'occurrence dans le discours ; en particulier, elle ne peut pas apparaître en début absolu de discours. Ce n'est pas le cas de la phrase à inversion locative.

16. L'exemple (20.b) est un exemple de Maurras cité dans Le Bidois.

17. La catégorie QUE regroupe dans une même classe lexicale le complémenteur (*que*) et les unités *qu* (relatif et interrogatif).

18. On retrouve la généralisation sur l'inversion du GN sujet élaborée dans les analyses désormais classiques de Kayne & Pollock (1978), Milner (1978). Mais, on notera que la formulation de la généralisation n'est pas exactement identique : ce qui déclenche la réalisation postverbale du sujet n'est pas l'occupation de COMP, mais directement le fait qu'un dépendant de V n'est pas réalisé lexicalement dans le domaine phrastique. Voir Bonami, Godard & Marandin (à par.) pour un traitement de cette facette de l'inversion locative dans le cadre de HPSG.

19. Nous admettons que les ajouts sont traités comme des objets indirects en français ; ce que montre précisément l'inversion locative.

20. E représente l'énoncé et GS la phase enchâssée.

21. La construction (23) diffère de son équivalent catalan. Selon Valduvi & Engdahl (1996), le domaine phrastique en catalan ne peut être occupé que par des éléments du focus.

22. C'est la position non marquée du domaine phrastique au sens que Zwicky définit pour les constructions : «Characteristically, in a set of etic constructions there is one that is 'un-marked' – that is un-special – by virtue of having (a) the fewest extragrammatical values, (b) the fewest semantic implicatures, (c) the fewest constraints on internal composition, and/or (d) the fewest constraints on external distribution» (Zwicky, *ibid* : 10).

23. Les autres propriétés sont les suivantes : (i) le GN sujet n'est pas interprétable comme un Agent, (2) les ajouts locatifs à droite du verbe sont mal formés et (3) les phrases n'entrent pas dans une relation temporelle avec d'autres phrases.

24. Par exemple, elles ne sont pas susceptibles d'une interprétation agentive (a), elles n'admettent pas de complément locatif dans le champ verbal (b) et ne peuvent pas entrer dans une relation temporelle avec une autre phrase (c) :

25. «Each idiom is an instance of particular constructions (with at least partially idiosyncratic semantics and possibly with lexical stipulations beyond those of the participating constructions)» (Zwicky : 8).

26. Nous laissons ici de côté les propriétés sémantiques spécifiques : l'aréférentialité des GN par exemple.
27. Comme le note Cinque (1997), les constructions «constitute separate units **at some level of analysis** [nous soulignons]».
28. Il faut y ajouter la réalisation clitique (*Il vient*) et la réalisation ex situ (*l'homme que je crois qui est venu*) que nous ne représentons pas ici en détail pour des raisons de place.
29. Elle ne peut être identique dans tous les cas : mouvement de qu-, mouvement de GN (vers la droite dans le contexte d'une extraction), permutation (pour les phrases telles que (28), c'est une transformation stylistique au sens technique du terme). Pour les difficultés à définir la transformation rendant compte du sujet postverbal dans le contexte à trou, voir Marandin 1997 pour la critique de Pollock 1986 et Bonami & Godard pour les analyses récentes.
30. Cette condition empêche qu'une variable prenne pour valeur la couleur pivot ; elle rend obligatoire la présence explicite (indiquée par la valeur *k*) de la couleur pivot pour tout sommet d'un arbre qui a des fils.
-

RÉSUMÉS

L'article présente une extension du formalisme des grammaires d'arbres polychromes (Cori & Marandin 1993) en introduisant une hiérarchie sur les arbres de la grammaire (un arbre polychrome représente une configuration de positions). La hiérarchie permet de structurer les configurations en classes de constructions (une construction est une configuration de positions associée à des propriétés syntaxiques, sémantiques ou informationnelles). L'héritage permet de distinguer les constructions selon leur degré de généralité. On illustre cette extension avec l'analyse de l'inversion locative en français et la représentation du sujet grammatical dans la syntaxe du français.

We present an extension of the Polychrome Tree Grammars formalism (Cori & Marandin 1993) by introducing a hierarchy on the trees of the grammar (a polychrome grammar represents a configuration of positions). Thanks to the hierarchy, configurations can be treated as members of a class of constructions (a construction is a configuration associated with syntactic, semantic or informational properties). Thanks to the inheritance, constructions can be distinguished according to their generality. The analysis of locative inversion and the modelization of the grammatical subject in French are taken as illustrations.

AUTEURS

MARCEL CORI

Université Paris 10, URA 1028 CNRS

JEAN-MARIE MARANDIN

Université Paris 10, URA 1028 CNRS