

Syntaxe théorique et formelle

Volume 1 : modélisation, unités,
structures

Sylvain Kahane

Kim Gerdes

Textbooks in Language Sciences 9



Textbooks in Language Sciences

Editors : Stefan Müller, Martin Haspelmath

Editorial Board : Claude Hagège, Marianne Mithun, Anatol Stefanowitsch, Foong Ha Yap

In this series :

1. Müller, Stefan. Grammatical theory : From transformational grammar to constraint-based approaches.
2. Schäfer, Roland. Einführung in die grammatische Beschreibung des Deutschen.
3. Freitas, Maria João & Ana Lúcia Santos (eds.). Aquisição de língua materna e não materna : Questões gerais e dados do português.
4. Roussarie, Laurent. Sémantique formelle : Introduction à la grammaire de Montague.
5. Kroeger, Paul. Analyzing meaning : An introduction to semantics and pragmatics.
6. Ferreira, Marcelo. Curso de semântica formal.
7. Stefanowitsch, Anatol. Corpus linguistics : A guide to the methodology.

Syntaxe théorique et formelle

Volume 1 : modélisation, unités,
structures

Sylvain Kahane

Kim Gerdes

Kahane, Sylvain & Kim Gerdes. 2021. *Syntaxe théorique et formelle : Volume 1 : modélisation, unités, structures* (Textbooks in Language Sciences 9). Berlin : Language Science Press.

This title can be downloaded at :

<http://langsci-press.org/catalog/book/241>

© 2021, Sylvain Kahane & Kim Gerdes

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0) :

<http://creativecommons.org/licenses/by/4.0/> 

ISBN : **no digital ISBN**

no print ISBNs!

ISSN : 2364-6209

no DOI

Source code available from www.github.com/langsci/241

Collaborative reading : paperhive.org/documents/remote?type=langsci&id=241

Cover and concept of design : Ulrike Harbort

Fonts : Libertinus, Arimo, DejaVu Sans Mono, Source Han Serif, Amiri, XITS

Math

Typesetting software : Xe_{La}TeX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität



Berlin

Table des matières

I	Modéliser la langue	1
II	Les unités de la langue : Les trois composantes du signe linguistiques	3
III	Structures syntaxiques et topologique : Comment les unités se combinent	5
1	La syntaxe profonde : Entre syntaxe et sémantique	7
1.1	Sémantique, syntaxe profonde, syntaxe de surface	7
1.2	Actant et modifieur	10
1.3	Les unités (potentielles) de la syntaxe profonde	12
1.3.1	Les unités de la syntaxe profonde	12
1.3.2	Les unités de la syntaxe de surface qui ne sont pas des unités de la syntaxe profonde	13
1.3.3	Les unités potentielles de la syntaxe profonde	14
1.4	Structure prédicative des sémantèmes	25
1.4.1	Les verbes	26
1.4.2	Les noms	26
1.4.3	Les modifieurs	28
1.4.4	Les sémantèmes grammaticaux	29
1.5	Lexique syntaxique	30
1.6	Contrôle	41
1.7	Montée	44
1.8	Un dernier cas de distorsion structurelle : l'insertion modificative	47
1.9	Les différents types de relations syntaxiques profondes	49
	Références bibliographiques	61

Table des encadrés

1.1	Historique de la notion de syntaxe profonde	9
1.2	Structure communicative et syntaxe profonde	23
1.3	Lexique syntaxique et interface sémantique-syntaxe	31
1.4	Ancrage et portée des quantifieurs	38

Première partie

Modéliser la langue

Deuxième partie

Les unités de la langue : Les trois composantes du signe linguistiques

Troisième partie

**Structures syntaxiques et
topologique :
Comment les unités se combinent**

1 La syntaxe profonde : Entre syntaxe et sémantique

1.1 Sémantique, syntaxe profonde, syntaxe de surface

La SYNTAXE PROFONDE étudie le lien entre le niveau syntaxique et le sens. C'est le contrepoint de la topologie qui s'intéresse au lien entre le niveau syntaxique et le texte.

La structure syntaxique proprement dite, qui décrit comment se combinent les syntaxèmes, est aussi appelée la STRUCTURE SYNTAXIQUE DE SURFACE, par contraste avec la syntaxe profonde. La structure sémantique, et plus précisément la structure prédicative (voir la section ?? *Partir d'un sens*), décrit les relations prédicat-argument entre les sémantèmes. La syntaxe profonde s'intéresse à la correspondance entre la structure sémantique et la structure syntaxique de surface, c'est-à-dire à l'INTERFACE SÉMANTIQUE-SYNTAXE. Cette correspondance est décrite au travers d'une structure qu'on appelle la structure syntaxique profonde.

Définition 1.1 : structure syntaxique profonde, relation syntaxique profonde

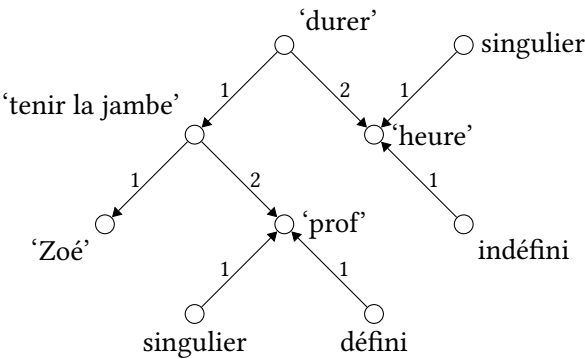
La STRUCTURE SYNTAXIQUE PROFONDE d'un énoncé est une structure qui indique comment les *sémantèmes* de cet énoncé se sont *combinés*. Les RELATIONS SYNTAXIQUES PROFONDES entre les sémantèmes indiquent à la fois la nature de la relation sémantique et de la relation syntaxique entre eux.

Donnons un premier exemple de structure syntaxique profonde, en la contrastant avec la structure syntaxique de surface et la structure sémantique (figure 1.1).

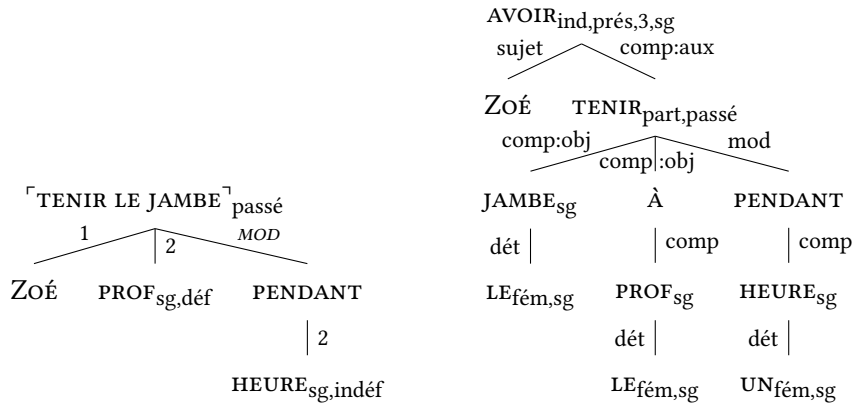
- (1) *Zoé a tenu la jambe à la prof pendant une heure.*

Dans l'exemple (1), phrasemetenir la jambe est un phrasème (voir la section ?? sur le *Phrasème*), c'est-à-dire un sémantème complexe, qui forme dont un

seul nœud dans la structure syntaxique profonde, mais correspond à plusieurs syntaxèmes et donc plusieurs nœuds dans la structure syntaxique de surface. Les conventions utilisées dans la structure syntaxique profonde seront explicitées dans la suite. On notera tout de suite que les articles, qui sont des lexèmes très grammaticalisés, sont considérés comme la réalisation par défaut d'un sémantème de définitude. Rappelons que les signifiés des sémantèmes lexicaux ou lexies sont indiqués entre guillemets simples ('prof', 'tenir la jambe', ...). Les signifiés des sémantèmes grammaticaux ou grammies peuvent être désignés par des termes (singulier, passé, ...) ou par des paraphrases ('un', 'avant maintenant', ...).



a. Structure sémantique



b. Structure syntaxique profonde

c. Structure syntaxique de surface

FIGURE 1.1 – Structures pour (1)

On peut voir la structure syntaxique profonde essentiellement comme une *projection de la structure prédicative sur la structure syntaxique de surface* et donc

comme une structure syntaxique de surface dont la granularité serait celle des sémantèmes. Néanmoins les relations syntaxiques profondes indiquent à la fois les connexions syntaxiques et les relations prédicat-argument, qui peuvent, dans certains cas, ne pas se superposer aux connexions syntaxiques. On peut aussi voir la structure syntaxique profonde, à l'inverse, comme une projection de la structure syntaxique sur la structure prédictive, c'est-à-dire comme une *structure sémantique hiérarchisée*. La structure syntaxique profonde se distingue néanmoins de la structure sémantique par la nature des unités en jeu : si la structure sémantique représente a priori un sens et donc la combinaison des signifiés des sémantèmes, la structure syntaxique profonde représente la combinaison des sémantèmes proprement dit, c'est-à-dire d'unités lexicales et grammaticales. Nous allons préciser ce point dans la section suivante.



Encadré 1.1 : Historique de la notion de syntaxe profonde

La distinction entre syntaxe profonde et syntaxe de surface telle que nous la concevons est due aux travaux d'Igor Mel'čuk dans le cadre de la théorie Sens-Texte. Pour Mel'čuk, la structure syntaxique profonde est une structure intermédiaire entre la structure sémantique et la structure syntaxique de surface. Dans son cadre théorique, le passage du sens au texte est modélisé par un premier ensemble de règles qui transforme la structure sémantique, qui comprend des relations prédicat-argument entre signifiants lexicaux, en une structure syntaxique profonde arborescente, qui est ensuite transformée en la structure syntaxique de surface. Plutôt qu'une structure intermédiaire, nous préférons voir la structure syntaxique profonde comme un témoin de la correspondance entre la structure sémantique et la structure syntaxique de surface (voir l'encadré 1.3 sur *Lexique syntaxique et interface sémantique-syntaxe*).

L'idée d'une structure syntaxique profonde, appelée STRUCTURE TECTOGRAMMATICALE (voir l'encadré 1.3 pour l'origine du terme), est également présente dans les travaux des Pragoïs réunis autour de Petr Sgall, qui est l'un des premiers, si ce n'est le premier linguiste (voir son article de 1967), à défendre l'idée d'un modèle stratifié des langues, avec différents niveaux de représentation en correspondance les uns avec les autres. On retrouve également un niveau de représentation profond dans un des mo-

dèles post-généralistes, la *Lexical Functional Grammar* (LFG) de Ronald Kaplan et Joan Bresnan (1981) : ici une structure syntaxique en constituants, la *c-structure*, encode la syntaxe de surface et est opposée à une structure de dépendance dite STRUCTURE FONCTIONNELLE, la *f-structure*, qui s'apparente à une structure syntaxique profonde.

L'opposition terminologique entre structure profonde (*deep structure*) et structure de surface (*surface structure*) a également été utilisée par Noam Chomsky (1965) dans le cadre de la grammaire générative-transformationnelle. L'usage est différent : la structure profonde n'est pas réellement un niveau de représentation différent de la structure de surface, mais une structure syntaxique sous-jacente à la structure de surface, la structure de surface étant obtenue par l'application (éventuelle) de transformations sur la structure profonde. Alors que la structure profonde de Mel'čuk est clairement une structure qui manipulent des sémantèmes et pas des syntaxèmes, la structure profonde de Chomsky manipulent les mêmes unités que sa structure de surface. De plus, chez Mel'čuk, la structure syntaxique profonde est un arbre de dépendance non ordonnée, l'ordre linéaire n'étant introduit qu'au moment de l'interface entre la syntaxe de surface et le texte, tandis que chez Chomsky, la structure profonde et la structure de surface sont des structures de constituants ordonnées. Il s'ensuit des discussions théoriques, qui nous semblent sans fondement, sur l'*ordre de base* des constructions syntaxiques, l'ordre de base étant l'ordre dans lequel les éléments se trouvent dans la structure profonde avant que des transformations les déplacent vers leur position en surface (voir l'encadré ?? *Mouvement et ordre de base*).

1.2 Actant et modifieur

Les *relations syntaxiques de surface* comme les *relations sémantiques* sont *asymétriques* : les relations syntaxiques de surface lient un gouverneur à un dépendant, tandis que les relations sémantiques lient un prédicat à un argument. Cette *double asymétrie* entraîne qu'il existe deux grands types de RELATIONS SYNTAXIQUES PROFONDES (voir les figures ?? et ?? du chapitre ?? *Produire un énoncé*).

Définition 1.2 : relation actancielle, actant

La relation entre deux sémantèmes est dite ACTANCIELLE quand l'un des sémantèmes est *à la fois le dépendant syntaxique et l'argument sémantique* de l'autre sémantème. Le sémantème dépendant est appelé un ACTANT du sémantème gouverneur.

Définition 1.3 : relation modificative, modifieur

La relation entre deux sémantèmes est dite MODIFICATIVE quand l'un des sémantèmes est *à la fois le gouverneur syntaxique et l'argument sémantique* de l'autre sémantème. Le sémantème dépendant est appelé un MODIFIEUR du sémantème gouverneur.

Le terme *modifieur* est également utilisé pour désigner les sémantèmes qui ont la capacité de modifier un autre sémantème. Nous dirons ainsi que les adjectifs, les adverbes, les prépositions et les conjonctions de subordination sont des modifieurs.

Les arguments d'un sémantème sont numérotés dans l'ORDRE D'OBLICITÉ croissante. L'ordre d'oblicité et son inverse, l'ORDRE DE SAILLANCE, seront définis dans le chapitre 18 (vol. 2). Disons juste que le sujet est la relation la plus saillante, suivie du complément d'objet indirect, puis du complément d'objet direct. Le sujet, lorsqu'il est un argument du verbe, est donc le premier argument et par conséquent le *premier actant*. Pour les modifieurs, le gouverneur syntaxique est considéré comme le premier argument. Ceci est justifié par le fait que lorsqu'un modifieur est verbalisé, son gouverneur devient le sujet de la construction (*une maison blanche, la maison est blanche*). Le deuxième argument d'un modifieur est appelé le deuxième actant (même s'il n'est pas à proprement parler le deuxième des actants).

Les relations modificatives sont étiquetées MOD dans les structures syntaxiques profondes. Les relations actanciennes sont numérotées dans l'ordre d'oblicité croissante, comme les arguments. Les actants portent les mêmes numéros que les arguments, sauf lorsqu'il y a une redistribution. C'est le cas avec le passif, où le deuxième argument devient le sujet et donc le premier actant (voir la figure 1.6).

Nous verrons dans la suite (et notamment dans la section 1.6 et les suivantes)

qu'il existe des cas où les relations syntaxiques de surface et les relations sémantiques ne se superposent pas, ce qui nous amènera à considérer deux autres types de relations syntaxiques profondes, que nous appellerons par abus de langage, les RELATIONS SYNTAXIQUES PURES et les RELATIONS SÉMANTIQUES PURES (bien qu'il s'agisse de relations syntaxiques profondes).

Les structures syntaxiques profondes contiennent également des relations de coréférence et des relations d'ancrage, dont nous discuterons dans la section 1.3 et l'encadré 1.4.

Lorsque nous étudierons les listes paradigmatiques (chapitre 20 (vol. 2)) et la macrosyntaxe (chapitre 21 (vol. 2)), nous introduirons encore d'autres relations syntaxiques profondes.

1.3 Les unités (potentielles) de la syntaxe profonde

L'objectif de la syntaxe profonde est d'étudier les combinaisons entre sémantèmes, c'est-à-dire les unités lexicales et grammaticales qui ont une contribution sémantique. Les unités de base de la structure syntaxique profonde sont donc avant tout les sémantèmes. Mais plusieurs questions se posent et nous allons donc passer en revue les unités qui sont nécessairement dans la structure, celles qui n'y apparaissent pas explicitement et celles qui pourraient y apparaître.

1.3.1 Les unités de la syntaxe profonde

1.3.1.1 Les sémantèmes lexicaux

Ce sont les *unités lexicales* ou *lexies*. Les lexies peuvent correspondre, du côté syntaxique, à un lexème ou à un phrasème composé de plusieurs lexèmes, comme «TENIR LA JAMBE» dans l'exemple (1). Elles peuvent éventuellement contenir des grammèmes, comme la lexie TRAVAUX (*Il y a des **travaux** dans ma rue.*), qui contient un pluriel inhérent.

Du côté sémantique, les lexies ont un signifié bien déterminé et sont donc des unités non ambiguës, qui sont associées à des acceptions précises de lexèmes. (Dans nos représentations, nous n'indiquons pas quelle acception de chaque unité lexicale est considérée, car cela nécessite d'avoir un lexique de référence.)

1.3.1.2 Les sémantèmes grammaticaux

Ce sont les *unités grammaticales* ou *grammies*. Une grammie peut correspondre du côté syntaxique à un grammème, comme l'imparfait, ou à une combinaison de

1.3 Les unités (potentielles) de la syntaxe profonde

grammèmes et de lexèmes, l'un des grammèmes se combinant avec un lexème ne faisant pas partie de la grammaire. Ce dernier cas peut être illustré par l'accompli, formé en français d'un auxiliaire, AVOIR ou ÊTRE, et d'un grammème de participe passé (voir la figure 1.2).

(2) *J'ai peur d'avoir répondu trop vite.*

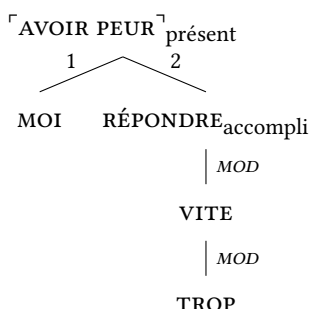


FIGURE 1.2 – Structure syntaxique profonde de (2)

(Il existe une autre acception de *avoir peur* qui est une collocation, où PEUR est modifiable et *avoir* commute avec FAIRE : *j'ai une peur bleue des araignées*. Mais le sens figuré utilisé en (2) est bien un phrasème « AVOIR PEUR ».)

1.3.2 Les unités de la syntaxe de surface qui ne sont pas des unités de la syntaxe profonde

1.3.2.1 Les lexèmes polysémiques

Nous avons défini les syntaxèmes sur des critères purement syntaxiques. Un lexème, c'est-à-dire un syntaxème lexical, peut tout à fait être polysémique et correspondre à plusieurs sémantèmes. Dans ce cas, c'est une acception précise du lexème, correspondant à un sens particulier, qui figure dans la structure syntaxique profonde.

1.3.2.2 Les lexèmes qui font partie d'un phrasème

Dans ce cas, le lexème n'apparaît pas en tant que tel : c'est le phrasème qui sera une unité minimale de la structure syntaxique profonde.

1.3.2.3 Les régimes

Les syntaxèmes qui marquent la relation syntaxique entre deux sémantèmes n'apparaissent pas explicitement dans la structure syntaxique profonde. Ils ne correspondent pas à un choix séparé du locuteur, mais sont imposés par le régime du gouverneur. C'est le cas de la préposition *À* dans l'exemple (1), qui est imposé par le régime de «TENIR LA JAMBE». C'est aussi le cas des syntaxèmes flexionnels de cas, comme le nominatif porté par les pronoms personnel sujet en français (cf. *je* = *MOI*_{nominatif}, dans l'exemple (2)).

1.3.2.4 Les syntaxèmes d'accord

Les syntaxèmes flexionnels qui marquent l'accord, comme l'accord en genre des adjectifs du français (*maison blanche*), n'ont pas de contribution sémantique. Ces syntaxèmes servent généralement à marquer des relations syntaxiques. Le cas de l'accord en nombre entre le nom et l'article (*les chevaux*) est plus complexe, car il y a bien un sémantème de pluriel, qui correspond à deux syntaxèmes. Nous positionnons le sémantème sur le nom, puisque c'est sur le nom que porte sémantiquement le nombre (même si en français, le nombre est morphologiquement marqué avant tout sur l'article).

1.3.3 Les unités potentielles de la syntaxe profonde

1.3.3.1 Les collocatifs

Les collocatifs sont des sémantèmes, mais leur choix est contraint par la base de la collocation et leur sens dans le contexte de la collocation est généralement différent de leur sens habituel. Dans l'exemple (3), FAIRE et BLEU sont des collocatifs de PEUR, qui expriment respectivement des sens de causation ('Zoé cause que j'ai peur') et d'intensification ('Ma peur est intense'), que nous représentons dans la représentation sémantique de la phrase par les signifiés génériques 'causer' et 'intense'. À partir de là deux choix sont possibles : on peut introduire des lexies FAIRE et BLEU particulières, utilisées avec les sens 'causer' et 'intense' dans le contexte de PEUR. Ou bien, comme le propose Igor Mel'čuk, considérer que FAIRE et BLEU sont des lexèmes qui réalisent en surface les valeurs d'un « sémantème » plus abstrait, qu'il appelle des FONCTIONS LEXICALES (voir lencadré ?? sur les *Fonctions lexicales*) et que nous nommons Caus et Magn dans la figure 1.3.

- (3) Zoé m'a fait une peur bleue.

1.3 Les unités (potentielles) de la syntaxe profonde

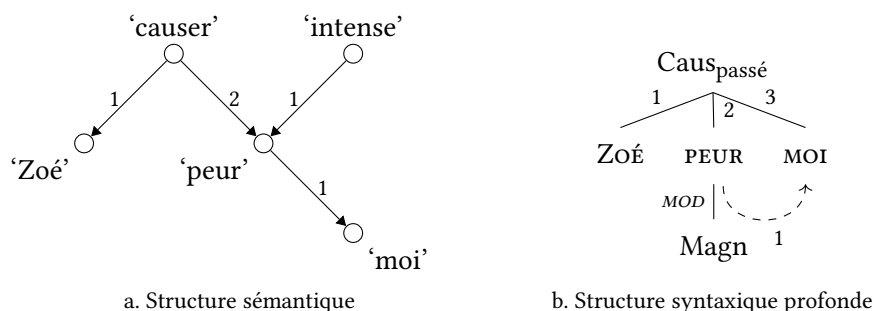


FIGURE 1.3 – Structures pour les fonctions lexicales de (3)

1.3.3.2 Les translatifs purs

LES TRANSLATIFS sont des syntaxèmes dont la fonction est de permettre à un lexème d'une catégorie donnée d'occuper une position syntaxique dont les éléments prototypiques appartiennent à une autre catégorie (voir le chapitre 18 (vol. 2) sur les *Catégories microsyntaxiques*). Ainsi dans l'exemple (4)a, la copule ÊTRE permet à l'adjectif CONTENT de se comporter comme un prédicat verbal et d'occuper la position de complément du verbe PENSER. Une autre construction est possible, (4)b, sans copule. La synonymie entre les deux constructions montre l'absence de contribution sémantique de la copule. Un translatif sans réelle contribution sémantique est dit PUR.

- (4) a. *Ali trouve que Zoé est sympa.*
 b. *Ali trouve Zoé sympa.*

Malgré l'absence de contribution sémantique des translatifs purs, nous décidons de les faire figurer dans la structure syntaxique profonde, car on peut considérer que le fait de ne pas utiliser, dans une position donnée, une lexie de la catégorie attendue est un choix du locuteur (souvent contraint par l'absence d'une possible réalisation dans la catégorie attendue du sens à lexicaliser) et que ce choix induit une lexicalisation particulière dans la position considérée. De plus, dans un cas comme celui de CONTENT dans (4)a, le fait que l'adjectif soit combiné avec un translatif en verbe entraîne la présence d'une grammaire de temps, le présent dans cet exemple, dont le choix est en partie libre. Il existe plusieurs possibilités pour modéliser la copule dans la structure syntaxique profonde de (4)a. Dans la figure 1.4a, nous représentons la copule comme un opérateur V de verbalisation, tandis que, en b, nous lui attribuons une véritable position dans la structure (ce qui nous rapproche davantage de la structure syntaxique de surface).

Dans ce deuxième cas, nous utilisons l'étiquette *Pred*, proposée par Mel'čuk. La flèche hachée représente une dépendance sémantique qui n'est pas réalisée par une dépendance syntaxique entre les mêmes éléments. Nous y reviendrons dans la section 1.6 sur le *Contrôle*.

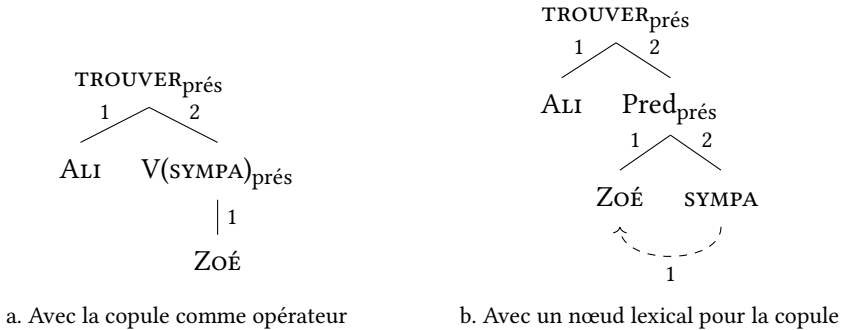


FIGURE 1.4 – Deux représentations pour la structure syntaxique profonde de (4)a

Notons que la conjonction de subordination *QUE* est également un translatif (de verbe en substantif). Nous aurions donc pu aussi l'introduire dans les représentations de la figure 1.4. Nous ne l'avons pas fait, car on considère que la conjonction de subordination *QUE* fait partie du régime de *TROUVER*.

Notons également que les translatifs peuvent en plus être des collocatifs : tel est le cas des verbes supports qui permettent à des noms prédicatifs d'occuper des positions verbales (*poser une question*, *faire une sieste*, *pousser un cri*, etc.) (voir l'encadré ?? sur les *Verbes supports et unités grammaticales*).

1.3.3.3 Les sémantèmes constructionnels

Il existe des syntaxèmes qui n'expriment pas des sens proprement dits, mais qui ont à voir avec la structure communicative, la façon dont on présente l'information (voir l'encadré 1.2). Nous considérons qu'il s'agit de sémantèmes d'un type particulier, que nous appelons les *SÉMANTÈMES CONSTRUCTIONNELS*. Nous distinguons ceux comme le clivage, qui sont réalisés par des lexèmes et des grammèmes distincts de la forme verbale et que nous traitons comme des lexies, et ceux comme le passif, qui sont réalisés par un grammème sur le verbe et que nous traitons comme des grammies.

Nous avons déjà parlé du clivage dans la section ?? sur *Les tests de constituency* et dont nous reparlerons dans le chapitre 20 (vol. 2). Le clivage, réalisé par *c'est X qui/que Y*, s'applique à une proposition *Y* dont il promet l'un

1.3 Les unités (potentielles) de la syntaxe profonde

des éléments X. Il possède donc deux actants : l'élément promu X est le premier actant et la proposition Y privée de cet élément est le deuxième actant.

- (5) a. *C'est Zoé qui viendra.*
 b. *C'est à Zoé que j'ai parlé.*

Nous modélisons le clivage comme une lexie que nous appelons « clivage ». Les structures syntaxiques profondes de nos deux exemples sont données dans la figure 1.5 (voir la section 1.6 sur le *Contrôle* pour la flèche hachée).

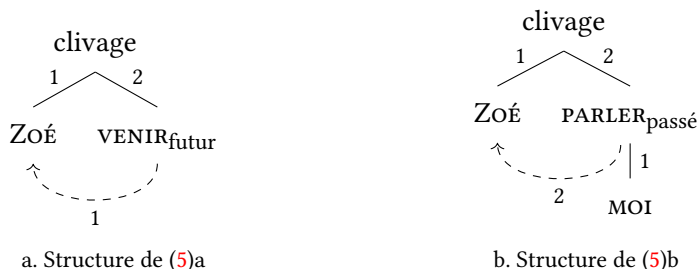


FIGURE 1.5 – Structures syntaxiques profondes pour le clivage

Le passif est l'exemple le plus connu de redistribution : il a pour effet de promouvoir l'objet d'un verbe transitif dans la position sujet et d'effacer ou de rétrograder le sujet du verbe. En français, il est réalisé par un grammème de participe passé sur le verbe, généralement combiné avec la copule ÊTRE. En conséquence de cette redistribution, le deuxième argument du verbe devient le premier actant, tandis que le premier argument est rétrogradé dans un rôle que nous notons ∞ . (Nous utilisons cet étiquette pour indiquer que le COMPLÉMENT RETROGRADÉ, appelé COMPLÉMENT D'AGENT par la grammaire traditionnelle, occupe toujours une position plus oblique que les autres actants. Voir la section suivante pour la numérotation des actants.)

- (6) a. *une fille poursuivie par un chien*
 b. *Zoé est poursuivie par un chien.*

Notons que dans le cas du participe passif dépendant d'un nom, le passif, en plus d'opérer une redistribution, joue aussi le rôle de translatif de verbe en adjectif (figure 1.6a) (nous n'indiquons pas cette translation, puisqu'elle est contenue dans la grammaire du passif). Le passif en ÊTRE peut alors être vu comme une double translation, du verbe en adjectif de la grammaire du passif, puis d'adjectif

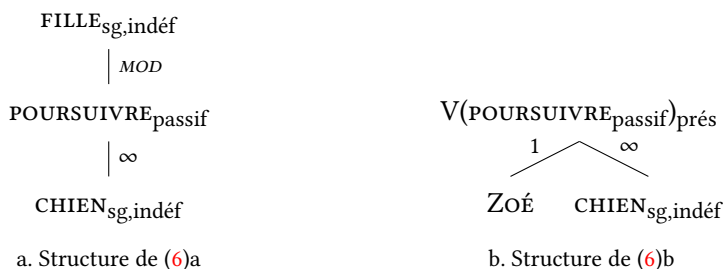


FIGURE 1.6 – Structures syntaxiques profondes pour le passif

en verbe par la copule. C’est ainsi que nous le modélisons (en français) (figure 1.6b).

Nous n’indiquons pas quand le verbe est à l’actif, considérant qu’il s’agit de la construction de base du verbe et qu’il n’y a donc pas de redistribution en jeu.

1.3.3.4 Les pronoms

Certains pronoms résultent du dédoublement d’un nœud sémantique, comme le pronom *elle* en (7)a. Le pronom n’est pas un sémantème standard, puisqu’il n’a pas de signifié distinct qui apparaisse dans la structure sémantique, comme on le voit dans la structure sémantique de la figure 1.7 commune aux deux exemples en (7).

- (7) a. Zoé pense qu’elle viendra.
b. Zoé pense venir.

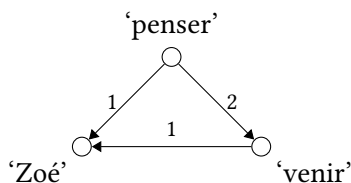


FIGURE 1.7 – Structure sémantique avec un cycle commune à (7)a et b

La représentation que nous proposons pour (7)a est d’indiquer qu’il y a un sémantème « pro » coréférent avec ZOÉ. La coréférence est indiquée par une double flèche en pointillé (voir la figure ??a). Ce lien indique que le sémantème « pro » provient du dédoublement du sémantème ‘Zoé’ et il permet d’assurer

l'accord de « pro » avec ZOÉ au niveau syntaxique de surface Dans le cas de (7)b, où il n'y a pas de pronom, nous indiquons que la dépendance entre le verbe subordonné et son premier argument est uniquement sémantique (par une flèche hachée) (figure ??b). Nous reviendrons sur cette construction dans la section 1.6 sur le *Contrôle*.

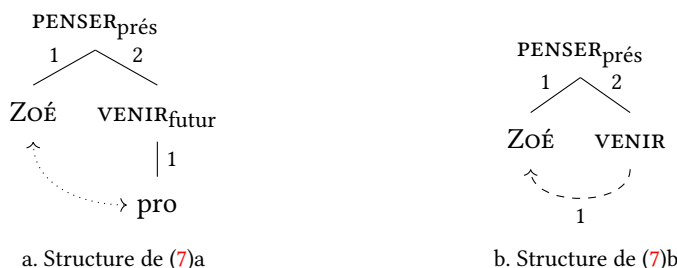


FIGURE 1.8 – Structures syntaxiques profondes dans le cas d'un cycle

D'autres représentations plus proches de la sémantique ont été proposées : Mel'čuk propose une représentation commune pour les deux phrases de (7), avec deux nœuds ZOÉ lié par un lien de coréférence. Une alternative à cette représentation est de garder un seul nœud ZOÉ et d'avoir deux gouverneurs syntaxiques pour ce nœud, ce qui donne une structure de dag (*directed acyclic graph*, voir l'encadré ?? sur *Graphe et arbre*).

1.3.3.5 La finitude-mode

On appelle FINITUDE-MODE la catégorie comprenant les grammèmes indicatif, subjonctif, impératif, infinitif, participe présent et participe passé. A l'exception de l'impératif, ces grammèmes n'ont généralement pas de contribution sémantique. Par exemple, l'indicatif ou l'infinitif sur le verbe VENIR dans les exemples (7) est imposé par le verbe PENSER qui le régit et nous ne le faisons donc pas figurer dans les structures syntaxiques profondes de ces exemples. Cette décision est tout de même discutable, car même si l'indicatif et l'infinitif n'ont pas ici de contribution sémantique, le choix de l'indicatif plutôt que l'infinitif a des conséquences sur la présence d'un sémantème de temps et sur la réalisation d'un pronom.

La question se pose aussi pour l'indicatif sur le verbe principal. Nous considérons que la réalisation d'un verbe à l'indicatif n'est pas réellement un choix du locuteur et ne correspond pas à l'expression d'un sens particulier. Ce n'est

pas tout à fait vrai, puisque le choix de l'indicatif (*Tu fais ce que tu veux.*) s'oppose à celui de l'impératif (*Fais ce que tu veux!*) ou du subjonctif (*Qu'il fasse ce qu'il veut!*) et indique qu'il s'agit d'une assertion ou d'une question et pas d'une injonction.

Il existe des cas où l'infinitif possède réellement une contribution sémantique et n'est pas imposé par le régime d'un verbe. C'est par exemple le cas dans l'exemple (8)a : la grammie infinitif de *fumer* en position sujet réalise une valeur générique du premier argument de FUMER, que l'on peut aussi exprimer en français avec le pronom ON, comme le montre la paraphrase avec (8)b.

- (8) a. *Fumer est dangereux pour la santé.*
b. *Quand on fume, on met sa santé en danger.*

Nous donnons dans la figure 1.9 les représentations sémantique et syntaxique profonde de (8). Nous représentons le sens générique par une étiquette « générique ». Ce sens est exprimé par grammie infinitif dans la structure syntaxique profonde. Notons que l'infinitif n'apparaît pas dans la structure syntaxique profonde quand il ne s'agit pas d'un sémantème comme ici (voir l'analyse des exemples (25) et (28)).

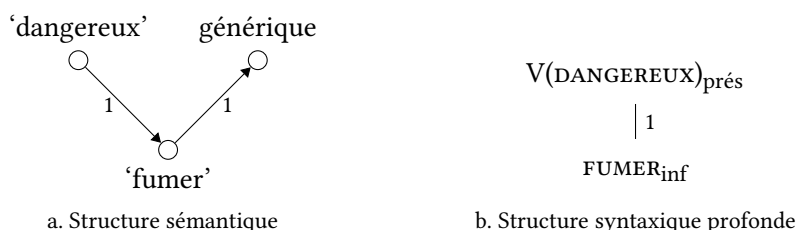


FIGURE 1.9 – Structures pour (8)

Concernant le grammème de participe passé, nous avons déjà donné des exemples et vu qu'il pouvait faire partie d'une grammie complexe exprimant le passé (AVOIR_{présent} + V_{part-passé}) ou l'accompli (AVOIR + V_{part-passé}) ou qu'il pouvait réaliser la grammie du passif (voir l'analyse de (6)).

Le grammème de participe présent, lui, est utilisé dans deux emplois en français : comme un translatif pur de verbe en adjectif, comme en (9)a, ou comme un translatif de verbe en adverbe dans la grammie complexe EN + V_{part-présent}, comme en (9)b (voir la figure 1.10).

- (9) a. *un chemin suivant la rivière*
b. *Ali est allé à la poste en suivant la rivière.*

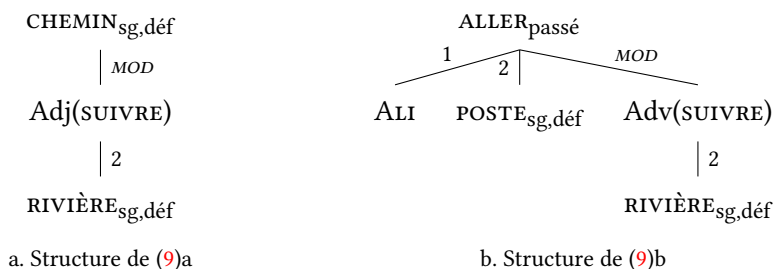


FIGURE 1.10 – Structures syntaxiques profondes pour le participe présent

1.3.3.6 Les sémantèmes cachés

Nous appelons SÉMANTEMES CACHÉS des sens qui naissent d'une configuration particulière sans être réellement réalisés par un syntaxème. Notre premier exemple est illustré par une construction particulière du russe relevée par Mel'čuk 1988 : 141. En russe, le numéral se place normalement avant le nom. Il est néanmoins possible de placer le numéral après le nom, mais cela change le sens : le numéral est alors interprété comme une valeur approximative. C'est donc l'ordre des mots qui est signifiant.

- (10) a. *Ja polučil desjat' rublej.*
'J'ai reçu dix roubles.'
b. *Ja polučil rublej desjat'.*
'J'ai reçu environ dix roubles.'

Nous indiquons cette valeur par un sémantème « approx », qui se combine avec le numéral (voir figure 1.11). Plus généralement, nous représentons les sémantèmes cachés comme des sémantèmes opérationnels s'appliquant à un autre sémantème.

La *dislocation* est un cas de sémantème constructionnel, comme le clivage, qui, à la différence du clivage, n'a pas vraiment de marqueur lexical : la dislocation met en jeu un pronom qui reprend l'élément disloqué, mais le pronom n'est pas en soi le signifiant d'une dislocation. Nous considérons donc qu'il s'agit d'un sémantème caché, que nous notons « dislocation » (voir la figure 1.12).

- (11) *Zoé, j'ai l'intention de lui parler.*

Un dernier exemple de sémantème caché est celui des conversions massif-comptable. Un nom comme VIN ou SABLE est dit MASSIF, car on ne compte pas

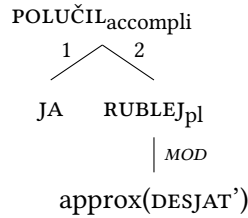


FIGURE 1.11 – Structure syntaxique profonde pour le sémantème configurationnel de (10)b

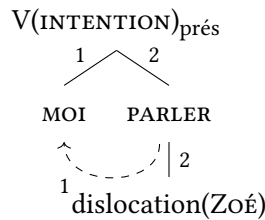


FIGURE 1.12 – Structure syntaxique profonde pour la dislocation de (11)

le vin ou le sable et que du vin séparé en deux donne toujours du vin. Dans l'exemple (12)a, nous considérons que VIN ne se combine pas avec une grammie de nombre, car le grammème de singulier ne résulte pas d'un choix et ne s'oppose pas à un grammème de pluriel (voir la figure 1.13a). Pour quantifier du vin ou du sable, on doit ajouter un « classifieur » : *deux bouteilles de vin, trois kilos de sable*. On peut néanmoins combiner les massifs directement avec des numéraux, mais alors le nom X est interprété comme dénotant un « type de X ». Nous considérons donc que dans l'exemple (12)b se cache un sémantème opérationnel, « type » (figure 1.13b). On notera aussi, dans les exemples (12), le contraste entre l'indéfini réalisé par DU pour les massifs et par UN pour les comptables.

- (12) a. Zoé a bu du vin.
 b. Zoé a goûté un bon vin.

1.3 Les unités (potentielles) de la syntaxe profonde

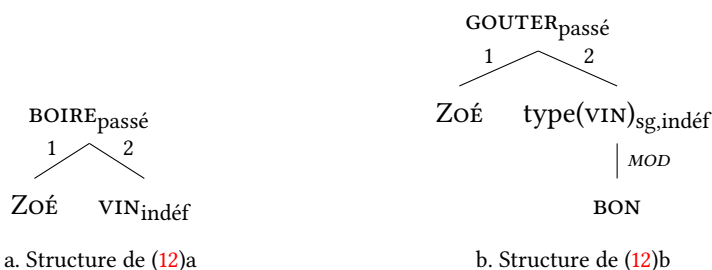


FIGURE 1.13 – Structures syntaxiques profondes pour la conversion massif-comptable



Encadré 1.2 : Structure communicative et syntaxe profonde

La STRUCTURE COMMUNICATIVE, encore appelée *information packaging* en anglais, est une composante de la représentation sémantique qui se superpose à la structure prédicative pour indiquer comment l'information doit être communiquée (voir l'encadré ?? sur *Les composantes du sens*). La principale composante de la structure communicative est la partition thème-rhème : le RHÈME est *ce qu'on dit*, l'information qui est réellement communiquée, tandis que le THÈME désigne *ce dont on parle*, ce à propos de quoi le rhème communique une information.

Nous avons vu que, dans les langues dites à ordre libre, l'ordre des mots peut être utilisé pour encoder la structure communicative (avec le support de la prosodie). Par contre, dans les langues à ordre plus strict, comme le français, il existe des constructions dédiées pour exprimer la structure communicative. Nous en avons introduit deux dans la section 1.3 sur *Les unités (potentielles) de la syntaxe profonde* : le clivage et la dislocation. Le clivage est l'expression d'un RHÈME FOCALISÉ, c'est-à-dire un rhème que l'on souhaite contraster avec les informations concurrentes. À l'inverse, la dislocation gauche marque un THÈME FOCALISÉ, indiquant que c'est à propos de cet élément et pas d'un autre que l'information est communiquée.

- (13) a. *C'est à Zoé que je parle* (et ce n'est pas à quelqu'un d'autre).
 b. *Zoé, je ne lui parle pas* (les autres, je leur parle).

Nous avons fait le choix d'indiquer explicitement les constructions qui sont déclenchées, puisqu'elles peuvent mettre en jeu des unités lexicales, comme le clivage, et faire intervenir une réorganisation de la structure syntaxique de surface. L'exemple (14) montre, par exemple, que si l'argument de l'adjectif EXCELLENT est disloqué, alors il n'est plus nécessaire de traduire l'adjectif en verbe pour réaliser cet argument. Voir la figure 1.14 qui donne les structures syntaxiques profondes correspondantes.

- (14) a. *Excellent, ce café!*
 b. *Ce café est excellent.*

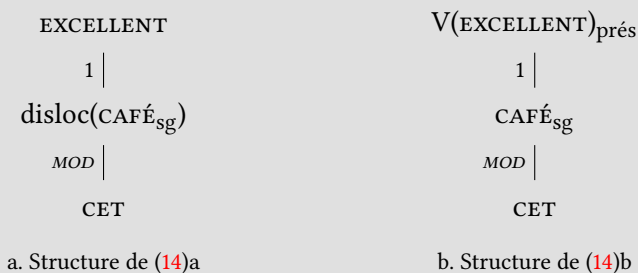


FIGURE 1.14 – Structures syntaxiques profondes avec et sans dislocation

A la différence du français, dans d'autres langues, comme les langues slaves, la structure communicative ne déclenche pas de modifications de la structure syntaxique proprement dite, mais va être réalisée par des variations dans l'ordre des mots (voir l'encadré ?? sur *Les langues dites à ordre libre*). On peut considérer que, dans ce cas, la structure communicative n'est pas consommée par l'interface sémantique-syntaxe et qu'elle devra être prise en compte directement par le module topologique. On peut donc décider de ne pas la mentionner dans la structure syntaxique profonde. C'est ce qu'on fera dans un modèle distribué où les différents niveaux de représentation peuvent communiquer entre eux et où la grammaire topologique peut accéder à des informations de niveau sémantique. Dans un modèle stratifié, la structure communicative devra être recopiée

aux différents niveaux de représentation jusqu'à ce qu'elle soit consommée et elle apparaîtra donc dans les représentations syntaxiques profonde et de surface.

1.4 Structure prédicative des sémantèmes

Les sens linguistiques fonctionnent comme des prédicats qui prennent d'autres sens comme arguments. Lorsque deux sémantèmes sont combinés, nous constatons que l'un des deux est l'argument de l'autre. Cette propriété, que nous ne pouvons pas démontrer, nous permet de postuler la structure prédicative des différents sémantèmes.

Définition 1.4 : valence sémantique, structure prédicative, régime

La STRUCTURE PRÉDICATIVE du sémantème est l'*ensemble des positions argumentales* qu'il ouvre. La VALENCE SÉMANTIQUE d'un sémantème est le *nombre d'arguments sémantiques* que possède le sémantème. Le RÉGIME du sémantème est l'*ensemble des contraintes syntaxiques de surface* s'appliquant sur ses arguments.

Nous adoptons une définition très sémantique des arguments. Les arguments correspondent à des éléments essentiels dans la *définition du sens* d'un sémantème. Par exemple, un verbe comme VENDRE, comparé à DONNER, est considéré comme quadrivalent : quelqu'un donne quelque chose à quelqu'un, mais pour vendre il faut en plus recevoir un montant en échange. Le verbe LOUER est quant à lui est pentavalent, puisque, contrairement à la vente qui est une cession pleine, la location se fait pour une durée déterminée.

Les arguments peuvent aussi être caractérisés par des contraintes de réalisation particulières au niveau syntaxique de surface. Par exemple, le montant pour un verbe comme VENDRE est réalisé par un complément direct (*elle l'a vendu 100 euros*) qui peut en plus commuter avec l'adjectif CHER (*elle l'a vendu cher*). Ces propriétés montrent qu'il ne s'agit pas d'un modifieur. Les modifieurs verbaux peuvent être combinés avec la plupart des verbes, ce qui n'est évidemment pas le cas d'un tel complément.

Nous allons étudier la structure prédicative des différents sémantèmes en procédant par parties du discours en commençant par les verbes, les noms, puis les modificateurs.

1.4.1 Les verbes

La valence des verbes est la plus étudiée. Nous en avons déjà donné quelques exemples. Il existe des *verbes avalents*, comme les verbes météorologiques PLEUVOIR ou NEIGER, des *verbes monovalents* comme DORMIR ou COURIR, des *verbes bivalents* comme MANGER ou PENSER. (Evidemment, tous ces verbes sont polysémiques et quand nous disons que tel verbe a telle valence, nous faisons référence à l'acception la plus courante de ce verbe. Le verbe COURIR, par exemple, est monovalent dans *courir pour rester en forme*, mais bivalent dans *courir les filles* ou *courir après son bus*.)

Parmi les compléments locatifs, on distingue les modificateurs qui indiquent le lieu et le moment du procès d'arguments qui indiquent une destination. Par exemple, ALLER est bivalent (quelqu'un va quelque part), METTRE est trivalent (quelqu'un met quelque chose quelque part). Certains cas sont délicats à trancher : par exemple, le complément *à la bibliothèque* est un modificateur s'il indique le lieu où je travaille aujourd'hui, mais il est un argument si (15) est utilisé pour dire 'je suis un employé de la bibliothèque'.

(15) *Je travaille à la bibliothèque.*

1.4.2 Les noms

Les noms posent un problème délicat. Un nom comme SŒUR est bivalent : il exprime la relation entre deux personnes et le sens de 'sœur' ne peut être défini sans faire intervenir ces deux personnes. Néanmoins son premier argument ne peut être exprimé que lorsque le nom est translaté en verbe comme en (16)a. Lorsque SŒUR occupe une position nominale, comme en (16)b, seul son deuxième argument est exprimable. On dit, dans ce cas, que le nom INTÈGRE son premier argument, le nom désigne le premier argument lui-même.

- (16) a. *Zoé est la sœur de Luce.*
b. *La sœur de Luce dort.*

Dans la représentation syntaxique profonde de (16)a, nous indiquons explicitement le fait que SŒUR est translaté en V (voir la figure 1.16a). C'est seulement dans ce cas que le sémantème SŒUR peut réaliser son argument.

1.4 Structure prédicative des sémantèmes



FIGURE 1.15 – Structures sémantiques d'un nom intégrant son premier actant

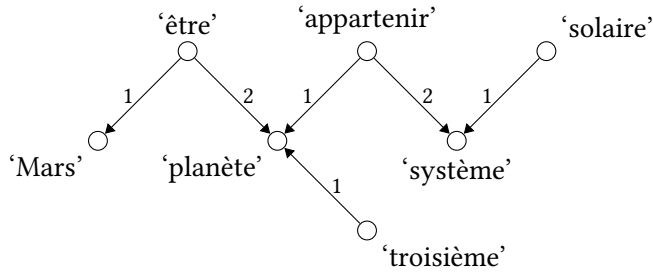


FIGURE 1.16 – Structures syntaxiques profondes d'un nom intégrant son premier actant

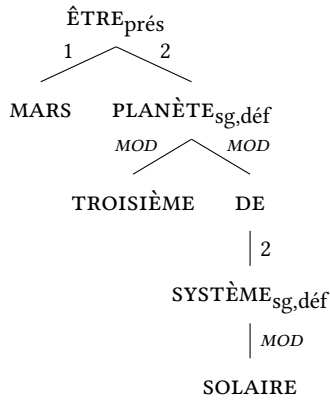
On peut considérer, comme l'on fait les logiciens depuis au moins [Frege \(1892\)](#), que tous les noms intègrent un premier actant, qui ne peut être réalisé que lorsque le nom est utilisé comme attribut du sujet. Cependant la plupart des noms ne sont quasiment jamais utilisés comme attribut du sujet. De plus, lorsque le nom est défini, il s'agit souvent d'une PROPOSITION ÉQUATIVE, où le verbe ÊTRE indique l'identité de deux choses, comme en (17). Dans ce cas, nous considérons que le verbe ÊTRE est un sémantème indiquant l'équation entre ses deux actants (voir la figure 1.17 ; on notera que DE est traité comme un sémantème exprimant le sens 'appartenir').

(17) *Mars est la troisième planète du système solaire.*

Il existe aussi des noms qui sont fondamentalement prédicatifs. Par exemple, un nom comme QUESTION est parallèle au verbe QUESTIONNER : *X questionne Y à propos de Z, la question de X à Y à propos de Z*. Il peut être utilisé dans des constructions à verbes support : *X pose une question Y, Y répond à la question de X*. Nous considérons donc que le nom QUESTION est un nom trivalent et nous numérotions les actants du nom QUESTION comme ceux du verbe QUESTIONNER.



a. Structure sémantique



b. Structure syntaxique profonde

FIGURE 1.17 – Structures pour (17)

1.4.3 Les modifieurs

Les adjectifs, les adverbes, les prépositions ou les conjonctions de subordination sont intrinsèquement des MODIFIEURS, c'est-à-dire des sémantèmes, qui dépendent syntaxiquement de leur premier argument. Des adjectifs comme ROUGE ou BEAU désigne des propriétés d'une entité et ne peuvent être définis sans faire intervenir cette entité. Des adverbes comme VITE ou FACILEMENT désigne des propriétés d'un procès et ne peuvent être définis sans faire intervenir ce procès. Une conjonction de subordination comme PARCE QUE est bivalente : elle indique une relation de cause à effet entre deux faits. Des prépositions comme SUR ou CHEZ exprime également la relation entre deux éléments (*quelque chose est sur quelque chose*; *quelque chose est chez quelqu'un*) et sont donc bivalentes (voir la figure 1.18a et b). Une préposition comme AVANT est même trivalente : *son anniversaire est deux jours avant Noël*. Les prépositions peuvent intégrer leur premier

actant comme les noms, lorsqu'elles sont l'actant d'un verbe de mouvement (voir la figure 1.18c).

- (18) a. *Luce est chez Zoé.*
 b. *Luce dort chez Zoé.*
 c. *Luce va chez Zoé.*

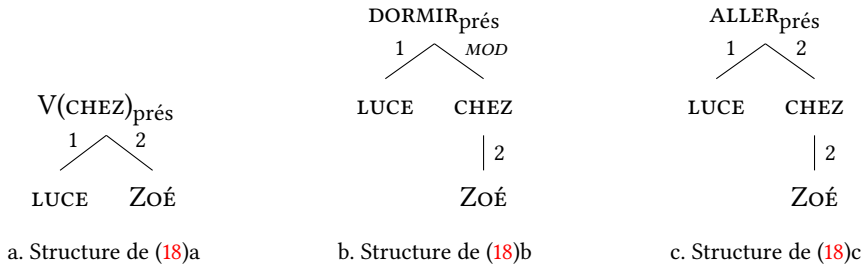


FIGURE 1.18 – Structures syntaxiques profondes pour la préposition CHEZ

Les adjectifs ou les adverbes peuvent aussi occuper une position actancielle, comme dans (19)a et c. Ils intègrent alors leur premier argument : on comparera les deux emplois de FRANÇAIS dans la figure 1.19.

- (19) a. *la production française de lait*
 b. *une tomate française*
 c. *Luce s'est mal comportée.*

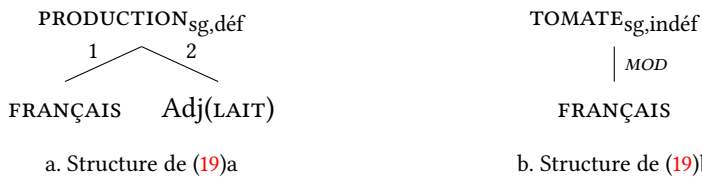


FIGURE 1.19 – Structures syntaxiques profondes pour un adjectif

1.4.4 Les sémantèmes grammaticaux

Les grammies sont toujours des prédicats unaires, qui prennent une lexie comme unique argument. La grammie exprime une propriété associée à la lexie : par

exemple, un temps passé exprime que le procès décrit par le verbe a lieu avant maintenant, un nombre pluriel exprime qu'il y a plus d'une entité dénotée par le nom. Dans la représentation sémantique, on peut d'ailleurs faire figurer aussi bien le terme associé à la grammaire (passé ou pluriel) que la glose lexicale ('avant maintenant' ou 'plus d'un').

D'autres exemples sont donnés dans les exercices. Les conjonctions de coordination seront étudiées dans le chapitre 20 (vol. 2) et les marqueurs de discours et les interjections dans le chapitre 21 (vol. 2). Dans le chapitre 17 (vol. 2), nous discuterons des parties du discours et nous verrons qu'il existe des langues avec une organisation différente des catégories et notamment des langues, comme le nahuatl, où les sémantèmes « nominaux » sont toujours prédicatif.

1.5 Lexique syntaxique

Un LEXIQUE est une liste d'entrées lexicales associées à des informations. Pour les sémantèmes, on considère deux lexiques.

Définition 1.5 : lexique sémantique, lexique syntaxique

Un LEXIQUE SÉMANTIQUE associe à chaque lexie une description de son sens, tandis qu'un LEXIQUE SYNTAXIQUE associe à chaque lexie une description de son syntactique, c'est-à-dire de sa combinatoire syntaxique.

Les dictionnaires monolingues traditionnelles, avec des définitions associées à chaque lexie, constituent des lexiques sémantiques. (Les définitions ne sont pas la seule façon de représenter le sens lexical, mais cela nous emmènerait trop loin de notre sujet de développer ce point.) Nous nous intéressons ici au lexique syntaxique. Chaque lexie y est associée à une description de la réalisation de ses arguments. Une telle description est traditionnellement appelée un TABLEAU DE RÉGIME.

Définition 1.6 : tableau de régime

Un TABLEAU DE RÉGIME de la lexie L indique la correspondance entre les arguments sémantiques de L et leur réalisation en syntaxe de surface. Il indique pour chaque argument de la lexie L quelle relation syntaxique le lie à L, à quelle catégorie il peut appartenir et quel régime L lui impose.

Dans la grammaire générative, les tableaux de régime sont appelés des CADRES DE SOUS-CATÉGORISATION (angl. *sub-categorization frame*). (Le terme vient du fait que chaque régime définit une sous-catégorie, au sens où les verbes transitifs forment une sous-catégorie de la catégorie des verbes.)

Nous donnons dans la figure 1.20 quelques exemples de tableaux de régime. Chaque colonne correspond à un actant pour lequel on indique sa fonction et son régime, à moins qu'il soit gouverneur. À chaque fonction correspond des propriétés par défaut : par exemple le sujet est obligatoire et les compléments sont facultatifs ; on indique donc explicitement lorsqu'un complément est obligatoire. Le cas du verbe à montée SEMBLER est expliqué dans la section 1.7 sur la *Montée*.

Nous verrons dans le chapitre 18 (vol. 2) sur les *Relations syntaxiques* que chaque relation syntaxique est associée à un faisceau de propriétés. Par exemple, la relation sujet en français suppose un accord du verbe, un placement particulier, l'absence de préposition, mais un marquage casuel des pronoms personnels (*je, tu, il, on, etc.*), ainsi que des propriétés de contrôle ou de redistribution que nous ne décrivons pas ici. On peut donc voir l'utilisation des relations syntaxiques dans le tableau de régime comme un moyen de décrire de façon synthétique une partie des propriétés syntaxiques d'un argument.

**Encadré 1.3 : Lexique syntaxique et interface sémantique-syntaxe**

Avec un lexique syntaxique, il devient possible à partir d'une structure syntaxique profonde de reconstituer la structure syntaxique de surface. En fait, chaque tableau de régime constitue la description d'une structure

1 La syntaxe profonde : Entre syntaxe et sémantique

「TENIR LA JAMBE」: *X tient la jambe à Y*

- nature : V

X = 1	Y = 2
<ul style="list-style-type: none"> • fonction : sujet • régime : N 	<ul style="list-style-type: none"> • fonction : complément d'objet indirect • régime : à N, Pro_{datif} • obligatoire

SEMBLER : *X semble Y*

- nature : V

X = +1	Y = 2
<ul style="list-style-type: none"> • fonction : sujet • régime : N • montée : 1^{er} actant de Y 	<ul style="list-style-type: none"> • fonction : complément • régime : V_{inf} • obligatoire

QUESTION : *question de X à Y sur Z*

- nature : N

X = 1	Y = 2	Z = 3
<ul style="list-style-type: none"> • fonction : complément • régime : de N, Pro_{poss} 	<ul style="list-style-type: none"> • fonction : complément • régime : à N 	<ul style="list-style-type: none"> • fonction : complément • régime : sur N, à propos de N

CHEZ : [X] *chez Y*

nature : Prép

X = 1	Y = 2
<ul style="list-style-type: none"> • gouverneur • nature : V, N 	<ul style="list-style-type: none"> • fonction : complément • régime : N • obligatoire

HEUREUX : [X] *heureux de Y*

nature : Adj, généralement postposé

X = 1	Y = 2
<ul style="list-style-type: none"> • gouverneur • nature : N 	<ul style="list-style-type: none"> • fonction : complément • régime : de V_{inf}, que V_{subj}

FIGURE 1.20 – Tableaux de régime de 「TENIR LA JAMBE」, SEMBLER, QUESTION, CHEZ, HEUREUX

élémentaire mettant en correspondance un fragment de structure sémantique avec un fragment de structure syntaxique de surface. En combinant ces fragments de structure, on peut construire en même temps la structure prédicative et la structure syntaxique de surface. (Pour obtenir une structure syntaxique de surface complète, il faudra aussi des règles de grammaires associées aux sémantèmes grammaticaux, notamment ceux qui opèrent des redistributions comme le passif.) On peut alors interpréter la structure syntaxique profonde comme le témoin de ces combinaisons, comme une structure indiquant quelle lexie s'est combinée avec quelle autre et dans quelle position argumentale.

L'idée d'interpréter le témoin de la production ou de l'analyse d'un énoncé comme une structure linguistique remonte aux premières heures des grammaires formelles. Dans les grammaires catégorielles, le calcul associé à une suite de mots prouvant que cette suite est bien une phrase peut être interprété comme la structure syntaxique de cette phrase (voir l'encadré ?? *Calcul symbolique et grammaires catégorielles*). Une idée similaire a été exploitée par Chomsky dans ses premiers travaux : les grammaires de réécriture de Chomsky (1957) génèrent des suites de mots. Le processus qui permet de produire une phrase par l'application successive de règles de réécriture est appelé une DÉRIVATION. Le témoin de cette dérivation est une structure hiérarchique qui décrit quelle règle doit s'appliquer après quelle règle et que Chomsky appelle l'ARBRE DE DÉRIVATION de la phrase et qu'il interprète comme la structure syntaxique de la phrase. Autrement dit, dans ces modèles, la structure syntaxique est le « témoin » du fait que la suite de mots est une phrase.

L'idée de voir la construction d'une structure par la combinaison de structures élémentaires remonte aux grammaires d'arbres, déjà évoquée dans l'encadré 4.4 sur les *Modèles génératif, équatif et transductif*, et à la plus célèbre d'entre elles, la *Tree Adjoining Grammar* (Grammaire d'adjonction d'arbre), plus simplement appelée TAG, portée par les travaux du linguiste indo-américain Aravind Joshi, à partir de l'article de 1975. Chaque unité lexicale est associée à une structure élémentaire, un petit bout d'arbre, qui permet de décrire la combinatoire de l'unité lexicale. Une suite de mots est alors une phrase si les structures élémentaires associées aux unités lexicales (il peut s'agir de plusieurs mots, si l'unité lexicale est un phrasème, mais aussi si son régime comprend des mots) peuvent

se combiner pour former une structure bien formée, qui est appelée la STRUCTURE DÉRIVÉE et qui est interprétée comme la structure syntaxique. Il existe deux opérations de combinaison des arbres élémentaires, la SUBSTITUTION et l'ADJONCTION, qui correspondent aux relations actancielles et modifieurs de la structure syntaxique profonde. La STRUCTURE DE DÉRIVATION, qui est le témoin de la dérivation et enregistre comment les structures élémentaires se sont combinées, peut alors être vue comme la structure syntaxique profonde.

On retrouve déjà la distinction entre structure dérivée et structure de dérivation dans un article du mathématicien Haskell Curry de 1961. Celui-ci nomme *structure phénogrammaticale* la structure dérivée, celle qui est construite et que l'on peut observer, et *structure tectogrammaticale* la structure sous-jacente, qui indique comment la construction a eu lieu. Les termes ont été repris par Petr Sgall dans le cadre du modèle pragois (voir l'encadré 1.1 sur l'*Historique de la notion de syntaxe profonde*).

Les grammaires d'arbres ont également été utilisées en GRAMMAIRE DE DÉPENDANCE pour produire des arbres de dépendance. Nous présentons dans la figure 1.22 un fragment de grammaire pour l'interface sémantique-syntaxe permettant de produire la phrase (1). Dans ce formalisme, initialement proposé par Alexis Nasr en 1995, puis développé dans les travaux de Sylvain Kahane, les sémantèmes sont associés à des structures élémentaires qui sont des fragments d'arbres syntaxiques de surface. Les positions syntaxiques de cet arbre sont associées à des polarités blanches ou noires : les polarités noires indiquent des positions instanciées par le sémantème, qu'il s'agisse des syntaxèmes qui composent le sémantème ou ceux qui font partie du régime qu'il impose à d'autres ; les polarités blanches indiquent des positions argumentales qui devront être instanciées par la combinaison avec la position noire d'une autre structure élémentaire. Les positions des lexèmes sont indiquées par des ronds et celles des grammèmes par des losanges. Le losange blanc de la structure élémentaire de PROF indique qu'il doit recevoir un grammème de définitude qui lui sera donné par son déterminant. La figure 1.21 montre la combinaison de ces deux structures élémentaires. Le résultat est une structure saturée dont toutes les polarités sont noires. (La grammaire a été simplifiée, nous n'avons pas introduit les grammèmes de nombre.)

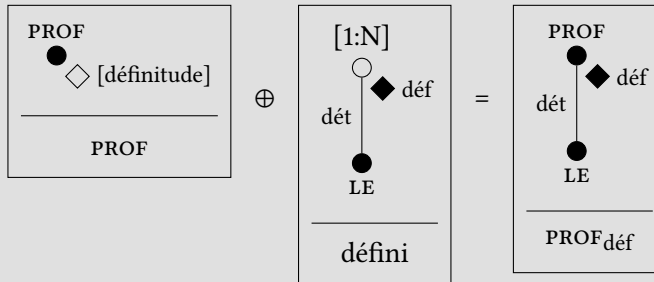


FIGURE 1.21 – Combinaison des structures élémentaires d’une lexie et d’une grammie

La figure 1.22 donne l’ensemble des règles nécessaires à la production de l’arbre syntaxique de surface de la phrase (1). Le losange blanc de la structure élémentaire de «TENIR LA JAMBE» indique qu’il doit recevoir un grammème de mode-temps. Celui-ci sera instancié par la grammie de passé, qui comporte un auxiliaire. Un mécanisme que nous ne détaillerons pas assure la montée du sujet sur l’auxiliaire, qui sera donc un pur dépendant syntaxique de AVOIR tout en restant l’argument sémantique de «TENIR LA JAMBE». Les positions argumentales de chaque sémantème sont numérotées ; ainsi, l’arbre élémentaire de PENDANT comporte deux positions argumentales : le premier argument doit être un verbe qui est gouverneur syntaxique, tandis que le deuxième argument est un nom qui est complément, ce qu’indiquent les étiquettes [1:V] et [2:N]. Enfin, la règle « déclaration », qui initie le processus de dérivation, demande à ce que la racine de l’arbre syntaxique de surface soit un verbe à l’indicatif.

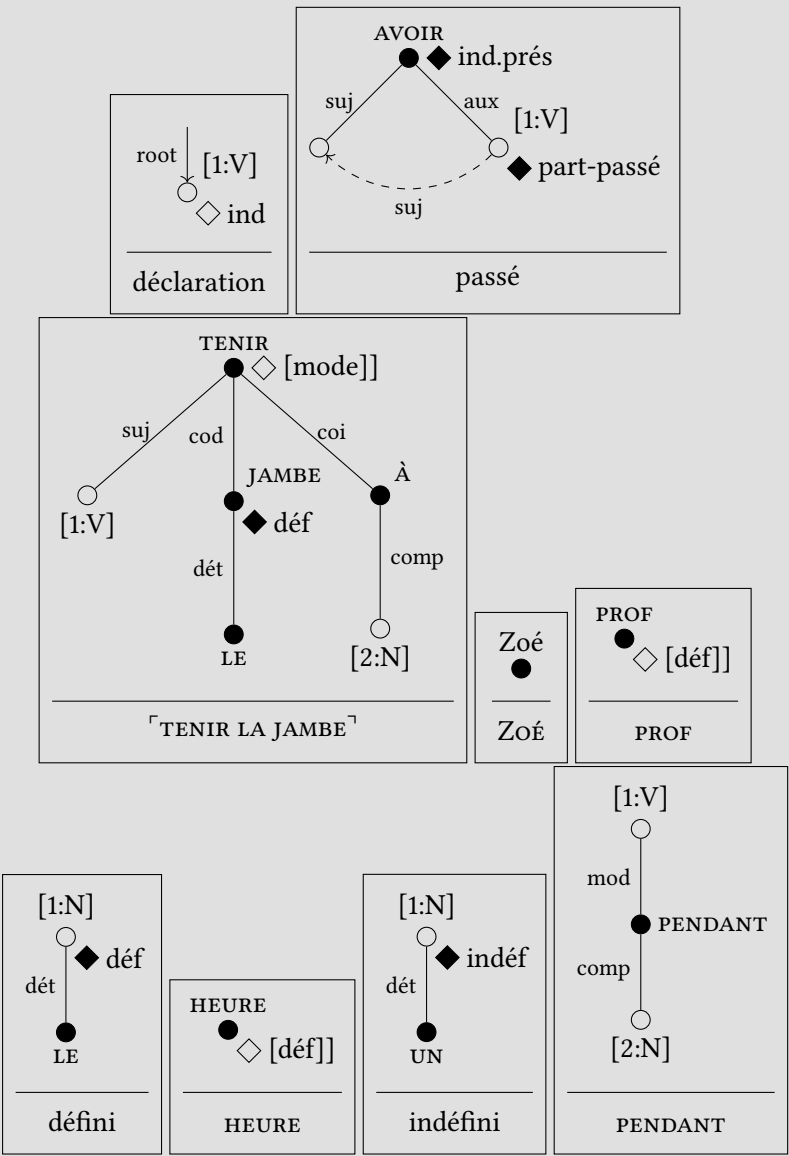


FIGURE 1.22 – Structures élémentaires pour une grammaire de dépendance

La structure dérivée qui résulte de la combinaison des 9 structures élé-

mentaires de la figure 1.22 est donnée dans la figure 1.23. On notera que toutes les polarités blanches ont été saturées par une polarité noire, ce qui indique que la dérivation peut s'arrêter.

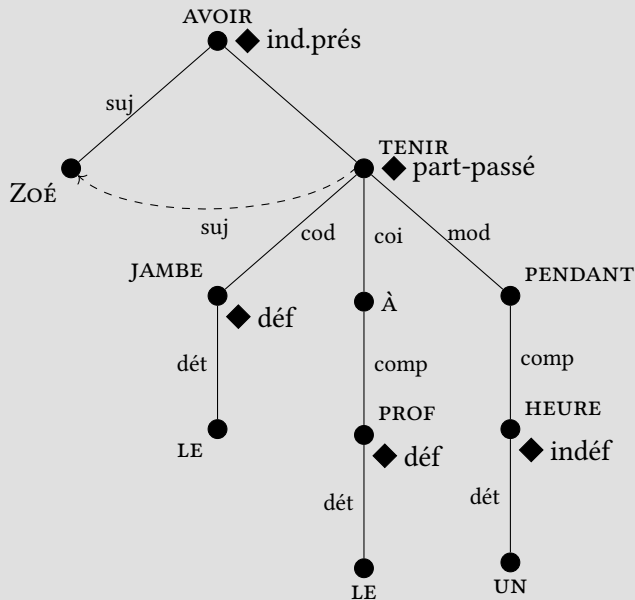


FIGURE 1.23 – Structure dérivée de (1) par la grammaire

Cette structure dérivée correspond à la structure syntaxique surface de (1), que nous avons donné dans la figure 1.1c. (Il manque les grammèmes de nombres que nous n'avons pas introduits pour simplifier. Notons aussi qu'il reste à appliquer des règles d'accord, qui concernent la bonne formation de la structure syntaxique de surface et pas directement l'interface sémantique-syntaxe.) La façon dont les structures élémentaires se sont combinées les unes avec les autres est décrit par la structure de dérivation de la figure 1.24. Chaque structure élémentaire est représentée par le nom du sémantème auquel elle correspond. Les flèches sont orientées du sémantème prédicat vers son argument, qui vient saturer une des positions blanches du prédicat. Les flèches vers le bas, où le prédicat gouverne syntaxiquement son argument, sont des relations actancielles, tandis que les flèches vers le haut, où le prédicat dépend de son argument, sont des

relations modificatives. On retrouve donc bien la structure syntaxique profonde de (1), donnée initialement dans la figure 1.1b.

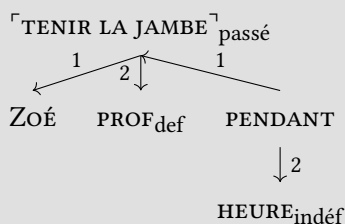


FIGURE 1.24 – Structure de dérivation de (1) par la grammaire



Encadré 1.4 : Ancrage et portée des quantifieurs

Les quantifieurs posent un problème intéressant, auquel les logiciens se sont beaucoup intéressés, depuis les travaux fondateurs de Gottlob Frege (1892). Le problème peut être illustré par la paire d'énoncés suivante.

- (20) a. *Tous les étudiants ont résolu un exercice.*
 b. *Un exercice a été résolu par tous les étudiants.*

Ces deux énoncés n'ont pas la même interprétation a priori. Dans le premier, l'interprétation privilégiée est que chaque étudiant a résolu un exercice, sans que ce soit a priori le même exercice, tandis que dans le deuxième, le même exercice a été résolu par l'ensemble des étudiants. Cette différence est généralement modélisée en termes de portée. On dit qu'un élément B est dans la PORTÉE d'un élément A, lorsque l'interprétation de B est fonction de A. En (20)a, on considère que 'un' est dans la portée de 'tous' lorsqu'on considère que l'exercice résolu est propre à chaque étudiant. Dans ce type de modélisation, les quantifieurs 'tous' et 'un' (appelés respectivement le quantifieur universel et le quantifieur existentiel) se voient associer deux arguments : un premier argument qui est le nom auquel ils sont combinés syntaxiquement et qu'on appelle leur

RESTRICTION et un deuxième argument qu'on appelle leur portée et qui comprend une prédication. On obtient alors les description sémantique suivantes pour les énoncés (20)a et b, données sous forme textuelle en (21)a et b, puis sous la forme d'une formule où chaque prédicat est suivi de la liste de ses arguments en (20)a'et b'.

- (21) a. Pour tout étudiant x , il existe un exercice y tel que x a résolu y .
 $\text{tout}(x, \text{étudiant}(x), \text{un}(y, \text{exercice}(y), \text{résoudre}(x, y)))$
- b. Il existe un exercice y tel que, pour tout étudiant x , x a résolu y .
 $\text{un}(y, \text{exercice}(y), \text{tout}(x, \text{étudiant}(x), \text{résoudre}(x, y)))$

Bien que cette modélisation soit aujourd'hui largement dominante, elle nous semble à la fois inutilement compliquée et non totalement satisfaisante. Pour comprendre pourquoi la notion de portée ne fonctionne pas complètement, considérons l'exemple suivant.

- (22) *Ali et Zoé ont acheté une pizza et bu une bière.*

Cet énoncé montre à nouveau que les groupes indéfinis (ici *une pizza* et *une bière*) peuvent avoir plusieurs interprétations : une interprétation possible est que Ali et Zoé ont acheté en tout une pizza et deux bières, qu'ils ont partagé la pizza et bu chacun une bière. Cette double « interprétation » de l'article UN peut difficilement être résolue en termes de portée dans ce cas, puisqu'il n'y a pas de quantifieur, mais simplement un groupe sujet coordonné et que les deux groupes indéfinis sont dans la même « portée ». On parle plutôt dans ce cas d'interprétation collective ou distributive du groupe coordonné *Ali et Zoé*. Mais comme on le voit ce groupe peut avoir en même temps une interprétation COLLECTIVE vis-à-vis de la première prédication (*acheter une pizza*) et DISTRIBUTIVE vis-à-vis de la deuxième prédication (*boire une bière*).

Nous proposons une autre modélisation qui résout les différents problèmes. Cette modélisation est basée sur un *concept inverse de la notion de portée*, que nous appelons l'ANCRAGE. Le problème est, selon nous, la question du calcul du référent des indéfinis. Les groupes substantifs indéfinis vont construire leur référent en s'ancrant dans un monde : il peut s'agir du monde construit par le discours précédent (qu'on appelle l'univers du discours) ou d'un monde ouvert par un autre élément de l'énoncé.

Dans l'exemple (22), les groupes indéfinis *une pizza* et *une bière* peuvent soit s'ancrer dans l'univers du discours et alors il y a un seul élément, soit s'ancrer sur *Ali et Zoé* et alors il y a deux éléments. Dans le cas de (20)a, *un exercice* peut s'ancrer sur *tous les étudiants* et alors un exercice pour chaque étudiant est considéré, ou bien *un exercice* s'ancre sur l'univers du discours et il y a un seul exercice en tout. Dans le cas de (20)b, *un exercice* occupe une position plus saillante que *tous les étudiants* et peut donc moins facilement s'ancrer sur ce dernier. (Cela n'est pas impossible, comme le montre un exemple tel que *Un garde du corps accompagnait chaque représentant*.)

L'ancrage ne fait pas à notre avis partie de la structure prédicative, mais de la structure référentielle, qui se superpose à la structure prédicative. Nous représentons dans la figure 1.25 la structure référentielle des deux énoncés de (20). Nous indiquons le référent des groupes substantifs par des boîtes rectangulaires. Le référent d'un groupe pluriel, comme *tous les étudiants* ou *Ali et Zoé*, introduit une variable qui parcourt l'ensemble dénoté et sur lequel un indéfini peut s'ancrer. Nous indiquons l'ancrage par une flèche en pointillé. L'univers du discours est noté Ω .

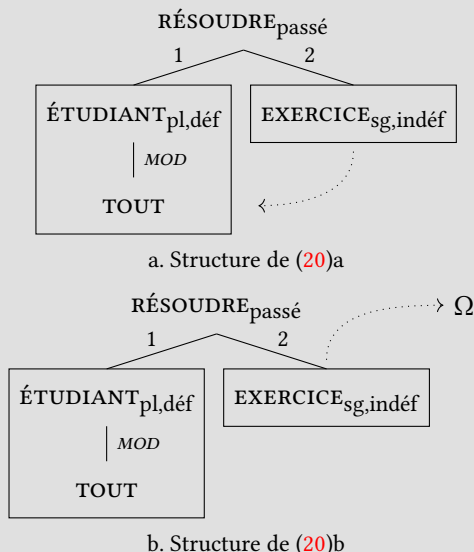


FIGURE 1.25 – Structure référentielle avec ancrage

Le cas de la portée de la négation relève à notre avis d'un autre aspect de la représentation sémantique, qui est la structure communicative (voir l'encadré 1.2). Dans un énoncé tel que (23)a, plusieurs interprétations sont possibles selon que la négation « porte » sur *donner*, *Zoé*, *ce livre*, à *Luce* ou sur un empan plus large de texte.

- (23) a. *Zoé ne veut pas donner ce livre à Luce.*
 b. *C'est Zoé qui ne veut pas donner ce livre à Luce.*
 c. *C'est ce livre que Zoé ne veut pas donner à Luce.*

On peut forcer de telles interprétations en clivant l'élément en question comme le montre les exemples (23)b et c. On voit alors que la négation porte sur le rhème, c'est-à-dire sur l'information qui est communiquée, notamment lorsqu'elle est extraite par un clivage. Le problème de l'ancrage est différent, puisque que ce soit *une pizza* ou *Ali et Zoé* qui est communiqué, les deux ancrages de *une pizza* restent possibles, comme on peut le constater dans l'exemple suivant.

- (24) a. *C'est une pizza que Ali et Zoé ont acheté.*
 b. *C'est Ali et Zoé qui ont acheté une pizza.*

1.6 Contrôle

La notion de CONTRÔLE est une notion assez générale : il y a contrôle dès qu'un élément de l'énoncé impose des contraintes à un autre élément. Les éléments qui sont connectés syntaxiquement se contrôlent l'un l'autre, puisque chacun restreint le paradigme de commutation de l'autre. Par exemple, dans *maison blanche*, la forme adjectivale *blanche* impose à son gouverneur d'être un nom féminin au singulier, tandis que la forme nominale *maison* impose à son dépendant adjectival de s'accorder en genre et en nombre. Nous nous intéressons ici à une forme de contrôle particulière entre des éléments sans lien syntaxique de surface.

Définition 1.7 : construction à contrôle, verbe à contrôle

On appelle CONSTRUCTION À CONTRÔLE une construction où un élément X est l'argument de deux éléments Y et Z en même temps, dont un avec lequel il n'a pas de lien syntaxique. Lorsque Y est un verbe recteur qui subordonne un verbe infinitif Z, Y est appelé un VERBE À CONTRÔLE quand Y contrôle Z en lui imposant le contrôle d'un des actants X de Y.

Ce phénomène est illustré par les exemples classiques suivants.

- (25) a. *Zoé promet à Ali de venir.*
b. *Zoé permet à Ali de venir.*
c. *Zoé propose à Ali de venir.*

Dans chacun de ces exemples, le verbe VENIR, qui est à l'infinitif, ne peut pas réaliser son premier actant comme sujet. (Nous parlons bien du premier actant potentiel et pas du premier argument, car, si le verbe est passivé, c'est son deuxième argument qui devient premier actant et sera contrôlé par le verbe recteur, comme dans *Zoé permet à Ali d'être accompagné par un ami.*). Il y a néanmoins un élément de la phrase qui réalise le premier actant de VENIR et le choix de cet élément est contraint par le verbe régissant VENIR : PROMETTRE impose son sujet comme premier actant de VENIR ('il est promis que Zoé vienne'), PERMETTRE impose son complément ('il est permis que Ali vienne'), PROPOSER impose une réalisation plus lâche par son sujet et/ou son complément ('il est proposé que Zoé ou Ali vienne, l'un ou l'autre, ou les deux').

Nous représentons la relation syntaxique profonde entre le verbe infinitif et son premier actant par une flèche hachurée indiquant qu'il s'agit d'une RELATION SÉMANTIQUE PURE, qui n'a pas de contrepartie en syntaxe de surface (voir la figure ?? où sont proposées les structures syntaxiques profondes des exemples (25)a et b).

Les constructions à contrôle se produisent quand la structure sémantique contient un cycle (non orienté). Ce cycle doit être réalisé en syntaxe de surface par un arbre, c'est-à-dire par une structure acyclique. Il faut donc couper le cycle quelque part.

Reprenons l'exemple (7) :

- (26) a. *Zoé pense qu'elle viendra.*
b. *Zoé pense venir.*

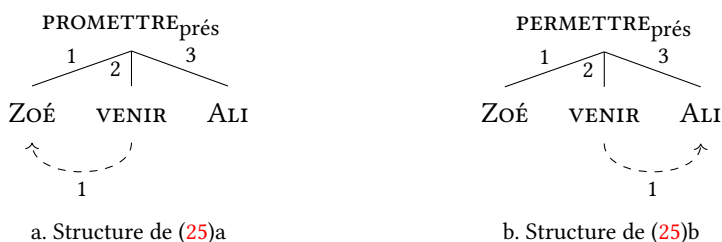


FIGURE 1.26 – Structures syntaxiques profondes avec un verbe à contrôle

La structure sémantique de ces paraphrases, qui a été donnée dans la figure 1.7, contient un cycle. Il y a deux façons de couper ce cycle. La première façon est de *couper au niveau d'un nœud sémantique*, comme 'Zoé' : le nœud est alors dédoublé dans la structure syntaxique et on obtient un lien de coréférence entre les deux réalisations syntaxiques de 'Zoé' (voir la figure ??a). La deuxième façon est de *couper au niveau de la dépendance sémantique* et c'est ce qui donne la construction à contrôle (voir la figure 1.7b). Nous illustrons ces deux solutions dans la figure 1.27.

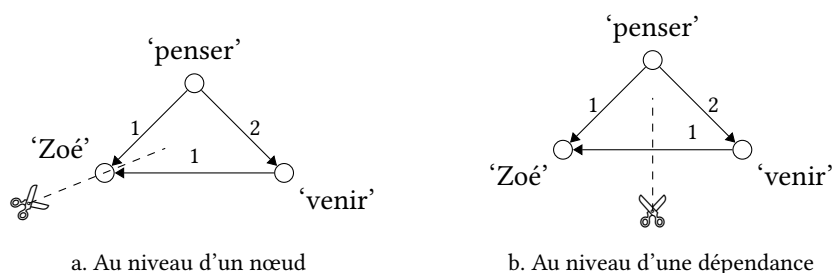


FIGURE 1.27 – Deux façons de couper un cycle dans une représentation sémantique

Les *constructions à verbes support* sont aussi des constructions à contrôle. Nous reprenons deux exemples donnés dans l'encadré ?? sur *Verbes supports et unités grammaticales*.

- (27) a. *Marie se prend une gifle.*
 b. *Pierre donne une gifle à Marie.*

Le nom GIFLE est un nom prédicatif qui possède trois arguments : *la gifle de Pierre à Marie* (à propos de Z). Dans une construction à verbe support, le nom

prédicatif contrôle les actants du verbe support (dont la contribution sémantique est quasi nulle). On peut voir le verbe support comme une réification de la relation sémantique, c'est-à-dire un élément lexical qui réalise la relation sémantique et lui donne ainsi un poids sémantique plus important.

Nous proposons dans la figure ?? la structure syntaxique profonde des exemples (27). Le verbe support 'SE PRENDRE' est modélisé par une fonction lexicale $Oper_2$, terme introduit par Igor Mel'cuk signifiant qu'il s'agit d'un opérateur qui *réifie* le 2^e argument du nom prédicatif. Le verbe support DONNER a deux actants contrôlés par le nom GIFLE : il est noté $Oper_{12}$ car il réalise comme actants les arguments 1 et 2 du nom GIFLE.

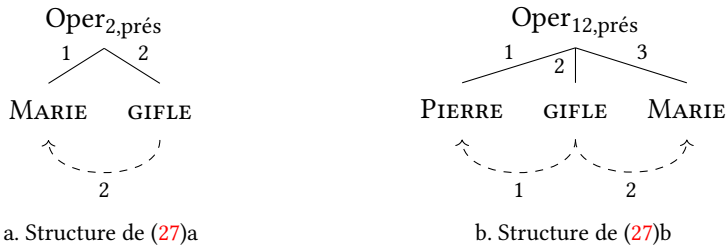


FIGURE 1.28 – Structures syntaxiques profondes de constructions à verbe support

1.7 Montée

La montée est une construction qui ressemble au contrôle, mais s'en distingue. Cette différence est bien illustrée par la paire suivante.

- (28) a. Zoé veut dormir.
b. Zoé semble dormir.

Le deuxième énoncé se distingue sémantiquement du premier par le fait que le verbe SEMBLER ne contrôle pas son propre sujet. Alors qu'on comprend que Zoé veut quelque chose en (28)a, on ne peut pas considérer que Zoé semble quelque chose. Cette différence sémantique s'illustre par plusieurs contrastes.

- (1) Le verbe à contrôle peut s'utiliser sans le verbe subordonné dans une réponse partielle, mais pas le verbe à montée :

- (29) a. *Est-ce que Zoé dort ? Non, mais elle veut.*
b. * *Est-ce que Zoé dort ? Elle semble.*

(2) Le clivage ou le pseudo-clivage du verbe subordonné est possible avec un verbe à contrôle, mais pas avec un verbe à montée :

- (30) a. *Ce que Zoé veut, c'est dormir.*
 b. * *Ce que Zoé semble, c'est dormir.*

(3) A l'inverse, le verbe à montée se combine à un verbe impersonnel, mais pas le verbe à contrôle :

- (31) a. * *Il veut pleuvoir.*
 b. *Il semble pleuvoir.*

(4) Le verbe à monter peut également se combiner avec un phrasème verbal qui inclut son sujet comme « LA MOUTARDE MONTER AU NEZ » (*la moutarde me monte au nez* signifie 'la colère grandit en moi') :

- (32) a. * *La moutarde veut lui monter au nez.*
 b. *La moutarde semble lui monter au nez.*

Il apparaît donc que le sujet du verbe à montée SEMBLER n'est pas son argument sémantique. Ceci est confirmé par la paraphrase entre (28)b et les phrases suivantes :

- (33) a. *Il semble que Zoé dorme.*
 b. *Zoé dort, semble-t-il.*

On considère alors que le verbe SEMBLER possède un unique actant qui est son complément verbal et que son sujet est réalisé par la « montée » (angl. *raising*) du premier actant de son complément.

Définition 1.8 : construction à montée, verbe à montée

On appelle CONSTRUCTION À MONTÉE une construction où un élément X n'est pas contrôlé par son gouverneur syntaxique Y, mais par un dépendant Z de Y. On dit alors que X est monté sur Y. Lorsque Y est un verbe recteur qui subordonne un verbe infinitif Z, Y est appelé un VERBE À MONTÉE.

Il existe aussi des verbes qui permette la montée dans la position objet, comme le verbe TROUVER. La paraphrase entre les deux phrases de (34) (qui reprennent

celle de (4)) montre que ce verbe possède deux arguments et qu'il y a bien montée en position de complément d'objet de TROUVER en (34)a. (La structure syntaxique de (34)b a été donnée dans la figure 1.4.)

- (34) a. *Ali trouve Zoé sympa.*
b. *Ali trouve que Zoé est sympa.*

La figure 1.29 propose une représentation de la structure syntaxique profonde des exemples (28)b et (34)a. Le dépendant non contrôlé par le verbe est un QUASI-ACTANT, qui se voit imposer un régime par le verbe, même s'il n'est pas un argument. Nous indiquons les quasi-actants par une étiquette actancielle à laquelle nous ajoutons un signe +.



FIGURE 1.29 – Structures syntaxiques profondes de constructions avec un verbe à montée

Les constructions avec auxiliaires peuvent être vues comme des cas particuliers de constructions à montée : en effet, l'auxiliaire exprime lexicalement une grammie qui a le verbe auxilié comme actant, tandis que le sujet de l'auxiliaire résulte de la montée du premier actant du verbe. Ceci est illustré en (35)a, où l'auxiliaire AVOIR du passé a pour sujet ZOÉ et pour unique actant le verbe COURIR, dont il spécifie le moment du procès.

- (35) a. *Zoé a couru.*
b. *Zoé fait rire Ali.*
c. *Zoé fait manger du tofu à Ali.*
d. *Zoé le fait rire.*
e. *Zoé le lui fait manger.*

La construction causative en (35)b et c illustre un cas possible de montée en position de complément. Nous préférons dire « cas possible », sans être plus affirmatif, car la très grande cohésion de la construction FAIRE V_{inf} ne permet pas

1.8 Un dernier cas de distorsion structurelle : l'insertion modificative

de décider clairement de quel verbe, FAIRE ou V, dépend le sujet rétrogradé de V. Comme le montre les exemples (35)d et e, le sujet rétrogradé se cliticise sur FAIRE, mais c'est aussi le cas du complément d'objet de MANGER en (35)e. De plus, la fonction du sujet rétrogradé de V dépend de la valence de V : il s'agit d'un complément d'objet direct quand V n'a pas de complément, comme RIRE, et d'un complément d'objet indirect quand V a déjà un complément d'objet direct, comme MANGER. Pour ces raisons, nous considérons qu'il s'agit d'une construction avec auxiliaire : le causatif est traité en syntaxe profonde comme une grammaie (voir la figure 1.30), à l'instar d'autres redistributions comme le passif (voir la figure 1.6). La relation syntaxique profonde du sujet rétrogradé reçoit l'étiquette ∞ déjà introduite dans la discussion sur le passif de la section 1.3.3 sur *Les unités potentielles de la syntaxe profonde*.

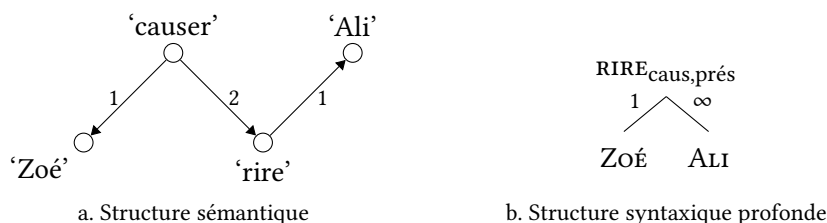


FIGURE 1.30 – Structures pour la construction causative (35)b

1.8 Un dernier cas de distorsion structurelle : l'insertion modificative

On appelle DISTORSIONS SÉMANTIQUE-SYNTAXE tous les cas où la structure sémantique et la structure syntaxique de surface ne se superposent pas. Il y a distorsion sémantique-syntaxe dès que les sémantèmes et les syntaxèmes ne se correspondent pas un à un : c'est le cas s'il y a des syntaxèmes vides qui marquent l'accord ou le régime ou s'il y a des sémantèmes complexes, composés de plusieurs syntaxèmes (voir la section 1.3). On distingue ces distorsions dues à la non-correspondance entre unités des DISTORSIONS STRUCTURELLES à proprement parler, lorsque *les relations prédicatives et les dépendances syntaxiques de surface ne se superposent pas*.

Nous avons déjà vu deux cas de distorsions structurelles, que sont le contrôle et la montée. Nous allons en étudier un troisième, que nous appelons l'insertion modificative. Ce cas de distorsion est illustré par l'exemple (36), déjà discuté dans l'encadré ?? sur les *Distorsions sémantique-syntaxe*.

(36) *Félix boit un verre de vin.*

Dans cet exemple, la tête sémantique et la tête syntaxique de l'unité *un verre de vin* sont distinctes : VIN est la tête sémantique, puisque 'vin' est l'argument sémantique de 'boire' (Félix boit du vin), tandis que VERRE est la tête syntaxique. En fait, 'un verre' indique la quantité de vin qui a été bu et prédique donc sur 'vin'. Mais au lieu que VIN soit réalisé comme un modifieur de VERRE et VERRE comme l'actant de BOIRE, VERRE vient gouverner VIN et s'insérer entre BOIRE et VIN.

Définition 1.9 : insertion modificative, modifieur inséré

Nous appelons INSERTION MODIFICATIVE une construction où un élément X gouverne syntaxiquement Y mais Y est la tête sémantique de l'unité qu'ils forment ensemble. Autrement, si Z est le gouverneur syntaxique de cette unité, Z a X pour dépendant syntaxique, mais c'est Y qu'il contrôle. L'élément X est qui est venu s'insérer entre Z et Y est appelé un MODIFIEUR INSÉRÉ.

L'insertion modificative a été bien modélisée dans le cadre de TAG (voir l'encadré 1.3 sur *Lexique syntaxique et interface sémantique-syntaxe*), où elle a été nommée l'*adjonction prédicative* (contrastée avec l'adjonction modificative des modifieurs ordinaires). Nous ne retenons pas ce terme, car nous considérons que tout modifieur est un prédicat sémantique qui s'adjoit sur son argument. La particularité des modifieurs insérés n'est pas de prédiquer, mais de *gouverner syntaxiquement leur premier argument* (tout en fonctionnant sémantiquement comme des modifieurs).

Nous proposons pour l'insertion modificative la représentation de la figure 1.31b. La figure 1.31a montre le processus d'insertion. L'insertion de VERRE entre BOIRE et VIN entraîne que le lien actanciel 2 entre BOIRE et VIN est dédoublé en lien purement syntaxique entre BOIRE et VERRE, que nous étiquettons (2), et un lien entre BOIRE et VIN, qui est maintenant un lien purement sémantique.

L'insertion modificative concerne aussi les déterminants et en tout premier lieu les dits « déterminants complexes » comme en (37)a (voir l'encadré ?? sur les « *Déterminants complexes* »).

- (37) a. *Félix a lu **plus de la moitié** du livre.*
b. *Félix veut acheter **ce** livre.*

1.9 Les différents types de relations syntaxiques profondes

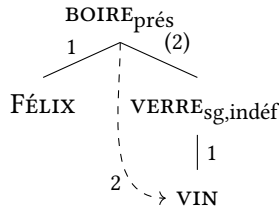


FIGURE 1.31 – Insertion modificative de (36)

L'exemple (37)a montre de plus que les modifieurs insérés peuvent s'insérer de manière récursive, puisque dans cet exemple MOITIÉ vient s'insérer sur LIVRE, puis l'adverbe PLUS s'insère sur MOITIÉ, comme on le voit dans la structure syntaxique profonde de la figure ??a. Si l'on considère que les déterminants sont la tête du groupe substantif, comme nous l'avons défendu dans la section ?? *Déterminant comme tête* ?, alors il faut également considérer que le déterminant CET s'est inséré dans l'exemple (37)b. Cette analyse est proposée dans la figure ??b et est à contraster avec l'analyse de la figure 1.14, où le nom est considéré comme la tête du groupe substantif et CET est traité comme un modifieur ordinaire.

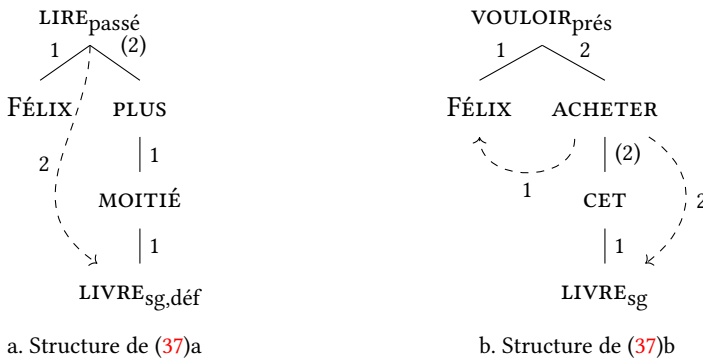


FIGURE 1.32 – Structures syntaxiques profondes de déterminants insérés

1.9 Les différents types de relations syntaxiques profondes

En guise de conclusion de ce chapitre, nous souhaitons récapituler les différents types de relations considérées au niveau syntaxique profond. Nous avons au final 6 types de relations syntaxiques profondes à proprement parler :

1 La syntaxe profonde : Entre syntaxe et sémantique

- les relations actanciennes, où une relation sémantique et une relation syntaxique se superposent et sont orientées de la même façon; elles sont étiquetées 1, 2, 3, etc. ou encore ∞ pour un actant rétrogradé;
- les relations modificatives, où une relation sémantique et une relation syntaxique se superposent, mais ont des orientations inverses; elles sont étiquetées *MOD*;
- les relations lexie-grammie, qui sont aussi des cas où une relation sémantique et une relation syntaxique se superposent, mais que nous annotons différemment en raison du statut particulier de la connexion syntaxique entre lexies et grammies;
- les relations entre un opérateur et son argument, que nous utilisons pour la translation syntaxique et pour les sémantèmes cachés;
- les relations syntaxiques pures, qui ne se superposent pas à une relation sémantique; elles sont étiquetées +1, +2, etc. lorsqu'il s'agit de quasi-actants résultant d'une montée ou (1), (2), etc. lorsqu'il de positions actanciennes où un modifieur s'est inséré;
- les relations sémantiques pures qui ne se superposent pas une relation syntaxique; elles sont représentées par des flèches hachurées et sont numérotées comme les relations actanciennes.

On peut encore ajouter à cette liste deux relations qui appartiennent à la structure référentielle :

- les relations de coréférence, qui indique la scission d'un nœud sémantique; elles sont représentées par une double flèche en pointillés;
- les relations d'ancrage, qui indique l'encrage du référent d'un indéfini sur un autre élément de la structure ou sur l'univers du discours; elles sont représentées par un flèche en pointillés.

D'autres exemples de structures syntaxiques profondes seront donnés dans le chapitre 19 (vol. 2) sur l'*Extraction* et le chapitre 20 (vol. 2) sur les *Listes paradigmatiques*.

Le volume 1 de *Syntaxe théorique et formelle* se termine ici. Le volume 2 rentrera plus en détail dans la structure syntaxique de surface (sans oublier les structures topologique et syntaxique profonde) en s'intéressant d'abord à la *nanosyntaxe* et à la notion de mot et de catégorie flexionnel, puis à la *microsyntaxe* et aux catégories lexicales ou parties du discours, aux relation syntaxique et à deux phénomènes complexes, l'extraction et la coordination, et enfin à la *macrosyntaxe* et à la notion controversée de phrase.



Exercices

Exercice 1. (modifieurs vs actants) Pour les compléments de nom suivants, déterminer s'il s'agit d'un actant ou d'un modifieur.

- a. *la réponse de Zoé*
- b. *le portrait de Zoé*
- c. *la main de Zoé*
- d. *la trousse de Zoé*
- e. *le chat de Zoé*
- f. *une huître de Bretagne*
- g. *le phare du Cap Fréhel*

Exercice 2. Reprenons notre exemple de base du chapitre ?? :

Beaucoup de gens aimeraient passer Noël en Laponie.

Donner la structure sémantique et la structure syntaxique profonde de cet exemple.

Exercice 3. Pour les exemples suivants, déterminez quels sont les sémantèmes, puis proposez une représentation syntaxique profonde. On notera que ces exemples contiennent un sémantème caché.

- a. *Zoé, faut qu'on y aille !*
- b. *Il y a de l'œuf sur ma chemise.*

Exercice 4. On considère la construction suivante :

Je ne comprends rien à ce problème.

- 1. Montrer que les éléments qui peuvent commuter avec *rien* forme un paradigme très réduit que vous décrirez et tenterez de caractériser
- 2. Quel est la contribution sémantique de *rien* dans cette construction ?

3. Bien qu'il soit réalisé en position d'objet direct, pourquoi peut-on considérer qu'il s'agit d'un modifieur ?

Exercice 5. (noms temporels) Quel problème posent les noms temporels pour la modélisation ? Comment modéliser les énoncés suivants en syntaxe profonde ?

- a. *Je viendrai la semaine prochaine.*
- b. *Il a dormi deux heures.*

Exercice 6. Nous nous intéressons à la construction suivante.

Luc casse les œufs dans un bol.

1. Montrer que le complément locatif *dans un bol* n'est pas un modifieur indiquant les circonstance du procès.
2. Comment modéliser cette construction ?

Exercice 7. Nous nous intéressons à la construction suivante, appelée *tough-movement* par les générativistes. (Le terme comporte un jeu de mots puisque *tough movement* évoque à la fois un problème épineux qui concerne des adjectifs comme *tough*, qui signifie 'dur, coriace, épineux'.)

- a. *un livre difficile à lire*
 - b. *Ce livre est difficile à lire.*
1. En utilisant la paraphrase avec « *Lire ce livre est difficile* », montrer qu'il s'agit potentiellement d'une construction à montée.
 2. Donner la structure sémantique et la structure syntaxique profonde de ces exemples.

Exercice 8. La synonymie entre les deux phrases suivantes est un peu étrange si on regarde les choses de près.

- a. *On doit dire la vérité.*
- b. *On ne doit pas mentir.*

Montrer qu'il y a un apparent phénomène de montée de la négation en jeu.



Lectures additionnelles

Comme nous l'avons dit, la notion de structure syntaxique profonde doit beaucoup aux travaux d'Igor Mel'čuk. Celui-ci a théorisé la notion et a aussi développé des lexiques sémantiques et syntaxiques pour le russe, puis pour le français lorsqu'il a émigré au Québec en 1977. On consultera tout particulièrement son ouvrage de sémantique en 3 volumes (2012-2015) et les dictionnaires explicatifs et combinatoires du français (1984-1999). Sa modélisation repose de façon essentielle sur la notion d'actant, à laquelle il a consacré de nombreux articles : on retiendra tout particulièrement les deux articles de 2004 (Mel'čuk 2004a ; 2004b).

Le flambeau a été repris par son étudiant, Alain Polguère, dont nous recommandons encore une fois l'ouvrage de sémantique lexicale et lexicologie. Celui-ci développe également un lexique sémantique et syntaxique électronique sous forme de réseau lexical, consultable en ligne sur <https://lexical-systems.atilf.fr/spiderlex/>. Dans les travaux de Mel'čuk et Polguère, l'accent est particulièrement mis sur la combinatoire lexicale restreinte décrite à l'aide des fonctions lexicales, dont nous n'avons donné qu'un très faible aperçu dans ce chapitre.

La formalisation de l'interface sémantique sous la forme d'une correspondance entre un graphe sémantique et un arbre de dépendance syntaxique de surface est développée dans le cadre de la Théorie Sens-Texte. La formalisation sous la forme d'une combinaison de structure élémentaire est développée dans les travaux de Sylvain Kahane, dont on pourra consulter le tutoriel sur la Théorie Sens-Texte et les grammaires formelles de 2001 et le mémoire d'habilitation de 2002 consacré à la formalisation de la Théorie Sens-Texte par une grammaire comme celle de l'encadré 1.3. L'article de 2015 sur *Les trois dimensions d'une modélisation formelle de la langue* présente une comparaison entre une telle grammaire et les grammaires TAG, avec notamment des exemples de distorsions sémantique-syntaxe.

- Kahane, Sylvain. 2001. Grammaires de dépendance formelles et théorie Sens-Texte. In *Actes de la conférence sur le Traitement Automatique des Langues Naturelles (TALN)*, t. 2, 1-63.
- Kahane, Sylvain. 2002. *Grammaire d'Unification Sens-Texte : Vers un modèle mathématique articulé de la langue*. Habilitation à diriger les recherches. Université Paris 7.
- Kahane, Sylvain. 2015. Les trois dimensions d'une modélisation formelle de la langue : syntagmatique, paradigmaticque et sémiotique. *TAL* 56(1). 39-63.
- Mel'čuk, Igor Aleksandrovič, Nadia Arbatchewsky-Jumarie, Lidija Iordanskaja, Suzanne Mantha & Alain Polguère. 1984–1999. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques*. T. I–IV. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, Igor. 2004a. Actants in semantics and syntax I : Actants in semantics. *Linguistics* 42(1). 1-66.
- Mel'čuk, Igor. 2004b. Actants in semantics and syntax II : Actants in syntax. *Linguistics* 42(2). 247-291.
- Mel'čuk, Igor. 2012-2015. *Semantics : From meaning to text*. T. 1-3. Amsterdam/Philadelphia : John Benjamins.

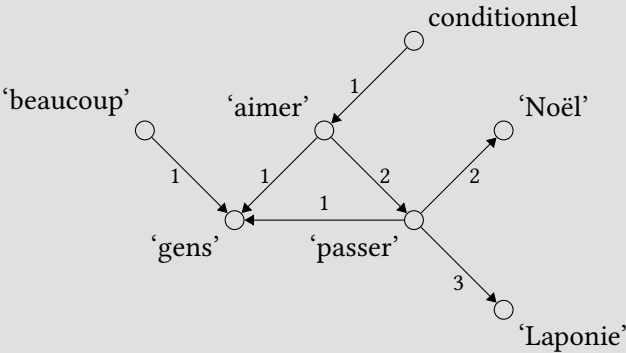


Corrections des exercices

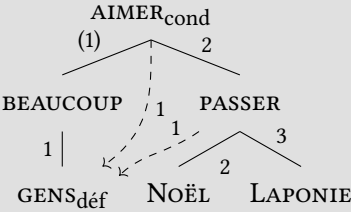
Corrigé 1. La plupart de ces complément de nom sont des actants : RÉPONSE est un nom prédicatif dont ZOÉ est le premier argument ; PORTRAIT est un nom prédicatif à deux arguments (quelqu'un fait le portrait de quelqu'un) et ZOÉ peut être l'un ou l'autre des arguments ; MAIN désigne une partie du corps de quelqu'un ; TROUSSE désigne un artefact, c'est-à-dire un objet fabriqué pour être utilisé et dont l'utilisateur est donc un argument ; CHAT désigne un animal domestique, qui à ce titre possède un maître/utilisateur (un animal, lorsqu'il est domestiqué, peut être vu comme une sorte d'artefact). Il est également défendable de considérer *de Zoé* comme un

modifier dans ces deux derniers exemples. On traitera alors la préposition DE comme la réalisation d'un sémantème « possesseur ». Dans les deux derniers exemples, les compléments locatifs *de Bretagne* et *du Cap Fréhel* peuvent être clairement considérés comme des modificateurs. Nous considérons que la préposition DE est la réalisation d'un sémantème « provenance » dans le premier cas et « location » dans le deuxième.

Corrigé 2. Cet exemple illustre plusieurs phénomènes intéressants. Deux cas de distorsion sémantique-syntaxe : un verbe à contrôle avec AIMER et une insertion modificative avec BEAUCOUP. Notons également que le complément locatif *en Laponie* fait partie de valence de passer (quelqu'un passe du temps quelque part). Enfin le nom GENS est un nom massif pluriel qui ne varie donc pas en nombre.



a. Structure sémantique

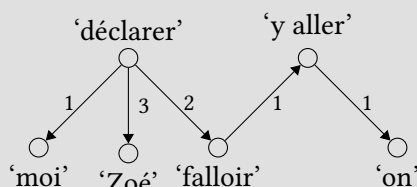


b. Structure syntaxique profonde

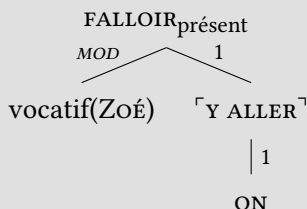
FIGURE 1.33 – Structures pour notre phrase préférée

Corrigé 3.

- a. Le vocatif Zoé contient un sémantème caché. Il ne s'agit pas vraiment d'un dépendant de la construction verbale faut qu'on y aille, mais d'un élément qui se rattache directement à l'illocution et peut être paraphrasé par 'je déclare à Zoé qu'il faut qu'on y aille' (voir la figure ??a). Cet élément de sens n'étant pas lexicalisé, nous le traitons comme un sémantème caché « vocatif ». Par ailleurs, nous traitons 'Y ALLER' comme un phrasème, car dans cette expression Y n'est pas a priori un pronom anaphorique correspondant à une destination précise. Enfin, le subjonctif sur ce verbe est imposé par le verbe FALLOIR et n'est donc pas un sémantème.



a. Structure sémantique



b. Structure syntaxique profonde

FIGURE 1.34 – Structures pour un vocatif

- b. Le nom ŒUF est un nom comptable. Il est employé ici comme un massif, avec le déterminant indéfini des massifs. Nous considérons donc qu'il est combiné avec un sémantème caché « massif », qui signifie 'une matière formée de'. Ce sémantème joue le rôle inverse du sémantème « type », qui produit un nom comptable à partir d'un massif. Par ailleurs, nous traitons *il y a* comme la réalisation d'un phrasème 'IL Y AVOIR' qui « verbalise » la préposition SUR. Enfin,

nous traitons le possessif *ma* comme un modifieur avec un sémantème « possesseur » (voir le corrigé de l'exercice 1).

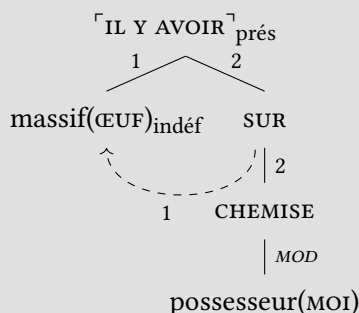


FIGURE 1.35 – Structure syntaxique profonde d'une conversion comptable-massif

Corrigé 4. Nous avons déjà présenté cette unité lexicale étrange dans l'encadré ?? intitulé *Le lexique : un cabinet de curiosités*. Le paradigme de commutation de *rien* comprend uniquement *quelque chose*, *pas grand-chose*, *que dalle* et les formes interrogatives *que* et *qu'est-ce que*. Nous considérons donc que ce paradigme forme un phrasème avec le verbe COMPRENDRE, phrasème que nous proposons d'appeler «COMPRENDRE QUELQUE CHOSE». Ce phrasème doit obligatoirement se combiner avec une négation ou une interrogation et le complément RIEN, qui est la négation de QUELQUE CHOSE, porte donc juste la valeur négative : négation + «COMPRENDRE QUELQUE CHOSE» = *ne rien comprendre*. Cette valeur négative fonctionne comme un modifieur.

Corrigé 5. Les noms temporels indiquent un moment (*la semaine prochaine*) ou une durée (*deux heures*). Dans les exemples donnés, ils sont utilisés comme des modifieurs de verbes. En même temps, ce sont de vrais noms, qui peuvent être utilisés, avec le même sens exactement, dans des positions où on attend un substantif : *deux heures suffiront pour terminer*, *la réunion de la semaine prochaine*. Nous considérons donc qu'il y a des sémantèmes cachés qui se combinent avec les noms temporels et permettent de les utiliser comme modifieurs. Nous nommons ces sé-

mantèmes « moment » et « durée ». On notera que ces sémantèmes sont bi-valents : leur premier argument est le verbe et leur deuxième argument le nom temporel.

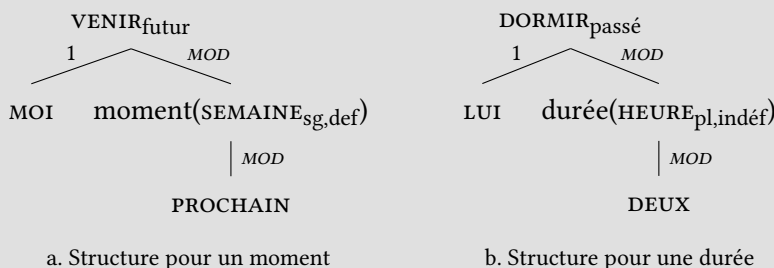


FIGURE 1.36 – Structures syntaxiques profondes de noms temporels modificateurs

Corrigé 6. Lorsque Luc casse des œufs, le procès n’a pas lieu dans un bol : le bol est la destination des œufs. Le verbe *CASSER* a ici le sens habituel ‘casser’, mais la construction de *METTRE* (*Luc met les œufs dans un bol*). On peut donc imaginer plusieurs modélisations possibles. Si l’on traite *dans un bol* comme un modifieur, il y a nécessairement un sémantème caché indiquant qu’il s’agit d’une destination. Ce sémantème, que nous nommons « destination », est une façon d’indiquer que la construction est signifiante. On peut aussi considérer que ce complément est « entré » dans la valence du verbe *CASSER* et qu’on a une acception trivalente de ce verbe, un *CASSER_b*, équivalent au *CASSER_a* + destination. Cette construction, plus courante en anglais qu’en français, a été appelée la *construction résultative*. On pourra notamment consulter le livre d’Adele Goldberg (1995) qui y consacre un chapitre.

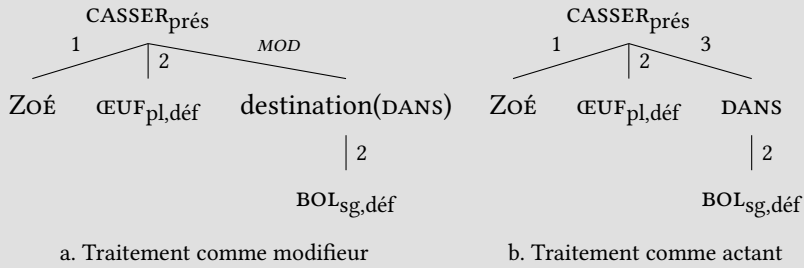


FIGURE 1.37 – Deux structures syntaxiques possibles pour une construction résultative

Corrigé 7. Si l'on accepte la paraphrase entre « *Ce livre est difficile à lire.* » et « *Lire ce livre est difficile* », alors on peut considérer que **DIFFICILE** reste un prédicat à un argument et que c'est le deuxième argument de **LIRE** qui devient son « sujet » (voir la structure sémantique de la figure 1.38c). Comme il existe plusieurs adjectifs qui ont les deux constructions (**FACILE**, **IMPOSSIBLE**, **UTILE** ...), nous considérerons, à la suite des générativistes, qu'il s'agit d'une réorganisation de la valence de l'adjectif (plutôt que de deux acceptions du même adjectif). Dans notre cadre, une telle réorganisation résulte de la combinaison avec un sémantème constructionnel, que nous appellerons « *tough-movement* » (pour ne pas rompre avec la tradition). Il en résulte les structures syntaxiques profondes des figures 1.38a et b.

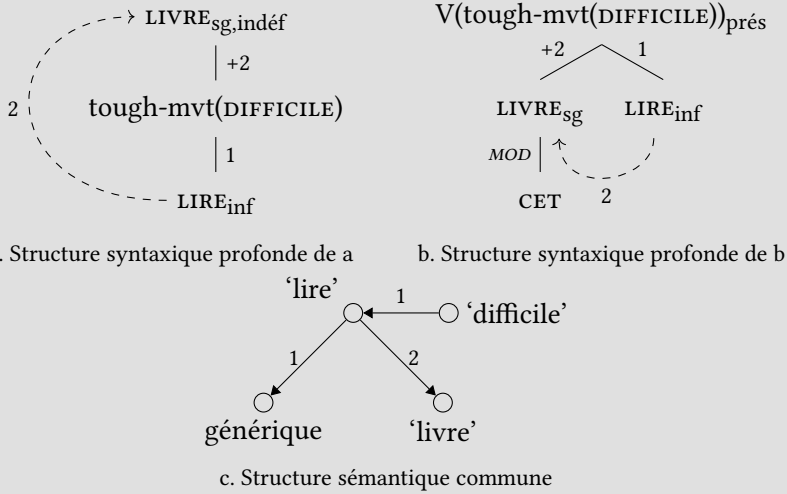


FIGURE 1.38 – Structures pour le tough-movement

Corrigé 8. « *On ne doit pas mentir.* » n'est pas la négation sémantique de « *On doit mentir.* ». Celle-ci serait plutôt exprimée par « *On n'est pas obligé de mentir.* » ou « *On peut ne pas mentir.* ». En fait, « *On ne doit pas mentir.* » est synonyme de « *On doit ne pas mentir.* ». Il y a donc bien un apparent phénomène de montée de la négation : la négation qui porte sémantiquement sur MENTIR se trouve attachée syntaxiquement au verbe recteur DEVOIR. [Tesnière \(1959\)](#) notait déjà ce phénomène auquel il consacre son chapitre 89 intitulé *Anticipation de la négation*. Plutôt qu'un phénomène syntaxique qui verrait véritablement une montée dans l'interface sémantique-syntaxe, on considère plutôt qu'il s'agit d'un phénomène de figement lexical associé au verbe recteur qui combiné à la négation prend un sens particulier (voir par exemple [Forest 1994](#)). On retrouve ce phénomène dans de nombreuses langues. Cet exemple de l'italien montre un exemple intéressant de contraste avec le français :

Il cafe non mi fa dormire.
 'Le café m'empêche de dormir.'
 lit. 'Le café ne me fait pas dormir.'

Références bibliographiques

- Chomsky, Noam. 1957. *Syntactic structures*. Cambridge : MIT Press. [traduction française de M. Bradeau : *Structures syntaxiques*, 1969, Paris : Éditions du Seuil.]
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge : The MIT Press. [traduction française de J.-Cl. Milner : *Aspects de la théorie syntaxique*, 1971, Paris : Éditions du Seuil.]
- Curry, Haskell B. 1961. Some logical aspects of grammatical structure. In Roman Jakobson (éd.), *Structure of language and its mathematical aspects*, t. 12, 56-68. [Proceedings of the Symposium in Applied Mathematics 12]. Providence : American Mathematical Society.
- Forest, Robert. 1994. La négation et les verbes d'adhérence. Pour en finir avec le neg-raising. *Linx* (5). 49-58.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung [Sens et dénotation]. *Zeitschrift für Philosophie und philosophische Kritik* 100. 25-50.
- Goldberg, Adele E. 1995. *Constructions : A construction grammar approach to argument structure*. University of Chicago Press.
- Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10(1). 136-163.
- Kahane, Sylvain. 2001. Grammaires de dépendance formelles et théorie Sens-Texte. In *Actes de la conférence sur le Traitement Automatique des Langues Naturelles (TALN)*, t. 2, 1-63.
- Kahane, Sylvain. 2002. *Grammaire d'Unification Sens-Texte : Vers un modèle mathématique articulé de la langue*. Habilitation à diriger les recherches. Université Paris 7.
- Kahane, Sylvain. 2015. Les trois dimensions d'une modélisation formelle de la langue : syntagmatique, paradigmatique et sémiotique. *TAL* 56(1). 39-63.
- Kaplan, Ronald M. & Joan Bresnan. 1981. *Lexical-Functional Grammar : A formal system for grammatical representation*. Massachusetts Institute of Technology, Center For Cognitive Science.
- Mel'čuk, Igor. 1988. *Dependency syntax : Theory and practice*. Albany : State University of New York Press.

Références bibliographiques

- Mel'čuk, Igor Aleksandrovič, Nadia Arbatchewsky-Jumarie, Lidija Iordanskaja, Suzanne Mantha & Alain Polguère. 1984–1999. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques*. T. I–IV. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, Igor. 2004a. Actants in semantics and syntax I : Actants in semantics. *Linguistics* 42(1). 1-66.
- Mel'čuk, Igor. 2004b. Actants in semantics and syntax II : Actants in syntax. *Linguistics* 42(2). 247-291.
- Mel'čuk, Igor. 2012-2015. *Semantics : From meaning to text*. T. 1-3. Amsterdam/Philadelphia : John Benjamins.
- Nasr, Alexis. 1995. A formalism and a parser for lexicalised dependency grammars. In *Proceedings of the Fourth International Workshop on Parsing Technologies (IWPT)*, 186-195.
- Sgall, Petr. 1967. Functional sentence perspective in a generative description. *Prague studies in mathematical linguistics* 2(203–225).
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris : Klincksieck.

Syntaxe théorique et formelle

Conçu comme une introduction générale à la syntaxe, cet ouvrage présente les notions de base nécessaires à une étude de la combinaison des unités lexicales et grammaticales au sein d'un énoncé. Sans se placer dans un cadre préconçu, l'ouvrage étudie les différentes possibilités pour la représentation des structures syntaxiques, en fonction des principes généraux et des critères particuliers retenus.

Élaboré avec l'objectif de fournir une base pour l'enseignement de la syntaxe à l'université, cet ouvrage souhaite montrer qu'on peut dégager de manière méthodique les propriétés des langues et mettre de l'ordre dans la forêt vierge que constitue chaque langue. Il est divisé en trois parties : comment élaborer le modèle d'une langue, comment déterminer les unités de base de la langue en fonction de leur sens, forme et combinatoire, comment définir et représenter les différents modes d'organisation des unités. Cette dernière partie présente une abondance de diagrammes syntaxiques de diverses natures.

L'ouvrage est découpé en de petites sections, alternant le contenu principal avec des éclairages, des notes historiques, des élaborations plus formelles, des exemples linguistiques dans diverses langues, des propositions de lectures additionnelles et des exercices avec des éléments de correction.

