

Sciences pour  
la communication

# Initiation au Programme Minimaliste

Eléments de syntaxe comparative

Genoveva Puskás

Peter Lang

Cet ouvrage est une introduction à la théorie syntaxique basée sur le *programme minimaliste*, une version récente de la grammaire générative. Il se veut accessible tant aux linguistes débutants qu'à ceux, déjà familiarisés avec la grammaire générative, qui souhaitent mettre leurs connaissances à jour.

Les chapitres explorent de manière progressive divers aspects de l'analyse de la phrase comme la structure des constituants, de la phrase déclarative simple, des clauses enchâssées, ainsi que les constructions plus complexes telles les phrases passives, les phrases négatives ou interrogatives. Les notions théoriques sont présentées de façon graduelle, dans un langage clair et accessible. L'accent est mis sur la présentation et l'analyse d'exemples en français, mais dans les cas pertinents, la discussion s'enrichit de comparaisons avec des langues très diverses, comme le hongrois, le gungbe, le japonais ou le basque.

Ce livre est un des rares ouvrages sur la théorie générative à être écrit en français; il vise à combler une lacune dans le vaste domaine des ouvrages introductifs sur la syntaxe.

Genoveva Puskas est professeure de linguistique à l'Université de Genève. Elle y enseigne son domaine de spécialisation, la syntaxe générative, ainsi que la sémantique, la pragmatique et les variétés de l'anglais. Sa recherche porte sur la syntaxe du subjonctif, la quantification et l'interface syntaxe-sémantique des marqueurs de structure informationnelle.

# Initiation au Programme Minimaliste



## Sciences pour la communication

Vol. 103

### **Comité scientifique**

- D. Apothéloz, Université de Nancy 2  
J.-P. Bronckart, Université de Genève  
P. Chilton, Université de Lancaster  
W. De Mulder, Université d'Anvers  
J.-P. Desclés, Université Paris-Sorbonne  
F.H. van Eemeren, Université d'Amsterdam  
V. Escandell-Vidal, UNED, Madrid  
F. Gadet, Université de Paris Ouest Nanterre La Défense  
J.-M. Marandin, CNRS et Université Paris-Diderot  
F. Martineau, Université d'Ottawa  
M. Milton Campos, Université de Montréal  
J. Rouault, Université Stendhal (Grenoble 3)

Les ouvrages publiés dans cette collection ont été sélectionnés  
par les soins du comité éditorial, après révision par les pairs.

Collection publiée sous la direction de  
Marie-José Béguelin, Alain Berrendonner,  
Denis Miéville et Louis de Saussure

Genoveva Puskás

# Initiation au Programme Minimaliste

Eléments de syntaxe comparative



PETER LANG

Bern • Berlin • Bruxelles • Frankfurt am Main • New York • Oxford • Wien

**Information bibliographique publiée par «Die Deutsche Nationalbibliothek»**  
«Die Deutsche Nationalbibliothek» répertorie cette publication dans la  
«Deutsche Nationalbibliografie»; les données bibliographiques détaillées sont disponibles  
sur Internet sous <http://dnb.d-nb.de>.

Graphisme: Atelier 4b, Sandra Meyer

ISBN 978-3-0343-1388-9 br.      ISBN 978-3-0352-0198-7 eBook  
ISSN 0933-6079 br.      ISSN 2235-7505 eBook

© Peter Lang SA, Editions scientifiques internationales, Berne 2013  
Hochfeldstrasse 32, CH-3012 Berne, Suisse  
[info@peterlang.com](mailto:info@peterlang.com), [www.peterlang.com](http://www.peterlang.com)

Tous droits réservés.

Cette publication est protégée dans sa totalité par copyright.  
Toute utilisation en dehors des strictes limites de la loi sur le copyright est interdite  
et punissable sans le consentement explicite de la maison d'édition.  
Ceci s'applique en particulier pour les reproductions, traductions, microfilms,  
ainsi que le stockage et le traitement sous forme électronique.

Imprimé en Suisse

*Ce livre est dédié à la mémoire  
de mon frère Claudius, grand amoureux de la langue,  
et de mon père Imre, dont la curiosité intellectuelle et le talent pour les  
langues ont nourri mon goût pour la linguistique.*



## Préface

Moi, je crois que la grammaire, c'est une voie d'accès à la beauté [...] Quand on fait de la grammaire, on a accès à une autre dimension de la beauté de la langue. Faire de la grammaire, c'est la décortiquer, regarder comment elle est faite, la voir toute nue, en quelque sorte. Et c'est là que c'est merveilleux : parce qu'on se dit : « comme c'est bien fait, qu'est-ce que c'est bien fichu ! ».

Muriel Barbery, *L'élégance du hérisson*.

Ce livre est né des « notes de cours » que j'ai accumulées au cours de plus de dix ans d'enseignement d'introduction à la syntaxe. Lorsque j'ai enseigné ce cours pour la première fois, j'ai très naïvement exposé à mes étudiants des théories recueillies dans diverses introductions. Mal m'en a pris, puisque j'ai réalisé qu'un cours d'introduction amenait des questions – les étudiants se sont chargés de me le montrer – qu'il fallait sans cesse réexaminer, reprendre et auxquelles les réponses sont difficiles à trouver. La dizaine de versions qui ont suivi ont certainement contribué à affuter ma propre approche quant au sujet. Je suis donc extrêmement reconnaissante aux générations d'étudiants qui se sont succédés sur les bancs de notre bâtiment historique des Bastions d'avoir sans cesse remis les problèmes sur le tapis, d'avoir questionné chacun des points, même ceux – et surtout ceux – qui paraissaient les plus évidents. Bien qu'ils soient trop nombreux pour être nommés ici individuellement, je les remercie un par un.

En transformant des notes de cours en ouvrage, je suis consciente de figer, en quelque sorte, la matière. Cela a impliqué des choix particuliers, qui sont liés à l'état de la théorie au moment de la rédaction, et à la possibilité de l'appliquer à un niveau introductif. Chaque lecteur pourra – et même devra – se livrer au même exercice de questionnement que celui que mes étudiants ont pratiqué, et prendre ce cours comme un guide à la théorie syntaxique plutôt que comme une référence absolue. Pour cela, il sera encouragé à comparer les données discutées avec celles de sa langue maternelle, dans une approche comparative. Les données qui illustrent ce cours,

outre celles que j'ai proposées moi-même, sont partiellement tirées d'ouvrages spécifiques. Pour le reste, elles m'ont été aimablement fournies par différentes personnes. Je tiens tout particulièrement à remercier Anamaria Bentea et Gabriela Soare pour le roumain, Shensen Luo pour le chinois, Luka Nerima pour le slovène, Lorenza Russo pour l'italien, Tomislav Socanac pour le croate et Lonneke van der Plas pour le néerlandais.

Dans ce processus de rédaction, j'ai eu la grande chance de pouvoir bénéficier de conseils avisés de la part de mes collègues. Je voudrais remercier plus particulièrement les personnes qui ont participé au programme d'introduction, à savoir Lucienne Rasetti, Lena Baunaz, Marc-Ariel Friedman, Paola Minen, Gabriela Soare et Anamaria Bentea. Chacun de leurs commentaires a contribué à affiner l'approche introductive et la cohérence de la matière.

Je voudrais aussi remercier chaleureusement les collaborateurs qui ont relu le manuscrit, Lena Baunaz, Tabea Ihsane et Tomislav Socanac. Leur lecture attentive et sans concession m'a permis de retravailler le texte pour en éliminer les incohérences et les imprécisions. J'espère avoir été à la hauteur de leurs exigences dans ce domaine. J'ai aussi eu la grande chance de bénéficier de la relecture minutieuse et critique de la toute dernière version par Sibylle Blanc, étudiante qui a souffert de mes hésitations de dernière minute. Ses commentaires pertinents, ses questions toujours justifiées et ses suggestions avisées m'ont permis d'identifier des problèmes qu'avec l'expérience on a tendance à négliger. Qu'elle en soit mille fois remerciée. Comme le veut la tradition, je tiens à préciser qu'aucun de ces relecteurs n'est en rien responsable des erreurs qui pourraient avoir subsisté.

Finalement, je remercie Louis de Saussure pour son soutien dans la publication de cet ouvrage, ainsi que Thierry Waser pour sa patience et son amabilité dans le processus de mise en place du manuscrit.

La rédaction – sur une période ma foi assez longue – de ce manuscrit n'aurait pas pu être achevée sans le soutien indéfectible de Luka Nerima, mon époux, et de mes filles Mateja et Anja. Je leur exprime ici toute ma gratitude et mon amour.

# Table des matières

Préface.....	VII
Table des matières.....	IX
Introduction.....	1
La grammaire générative.....	1
1. Le modèle générativiste (Principes et Paramètres).....	1
2. Le programme Minimaliste.....	5
3. La syntaxe comparative.....	6
4. But de ce cours.....	8
5. Les données et les références.....	10
Chapitre 1.....	13
Structure argumentale.....	13
1.1 Sous-catégorisation.....	14
1.2 Les rôles thématiques.....	19
Chapitre 2.....	25
Catégories de mots.....	25
2.1 Catégories lexicales.....	26
2.1.1 Les noms.....	26
2.1.2 Les adjectifs.....	30
2.1.3 Les adverbes.....	32
2.1.4 Les verbes.....	33
2.2 Catégories fonctionnelles.....	35
2.2.1 Les déterminants.....	35
2.2.2 Les auxiliaires.....	37
2.2.3 Les conjonctions.....	38
2.2.4 Les prépositions.....	39
2.2.5 Note sur les marqueurs d'aspect et de mode.....	41
Chapitre 3.....	45
Structure phrasale.....	45
3.1 Constituants.....	45
3.2 Ambiguïtés.....	50
3.3 Le constituant verbal.....	51

Chapitre 4 .....	55
Représentations syntaxiques .....	55
4.1 La représentation de l'organisation hiérarchique .....	55
4.2 Les syntagmes .....	57
4.2.1 Le constituant verbal .....	57
4.2.2 Le constituant nominal .....	61
4.2.3 Adjectifs .....	65
4.2.4 Les adverbes .....	67
4.2.5 Les prépositions .....	67
4.3 Les projections fonctionnelles .....	69
4.3.1 Le déterminant .....	69
4.4 La notion de projection .....	71
Chapitre 5 .....	75
La phrase tendue .....	75
5.1 La tête de la phrase .....	75
5.2 Les temps simples .....	77
5.3 Le mouvement de tête .....	85
5.4 Le mouvement du sujet .....	92
5.5 Les auxiliaires .....	97
5.5.1 L'auxiliaire et le verbe ne forment pas une unité syntaxique ..	97
5.5.2 L'auxiliaire peut avoir une fonction temporelle .....	98
5.6 La copule .....	102
5.7 Les modaux .....	104
Chapitre 6 .....	109
La phrase complexe .....	109
6.1 La clause subordonnée .....	109
6.2 Les complétives tendues .....	112
6.3 Les complétives en position de sujet .....	115
Chapitre 7 .....	121
Les subordonnées complétives infinitives .....	121
7.1 Les verbes à « contrôle par le sujet » .....	121
7.2 Le problème du cas .....	129
7.2.1 Cas structurel .....	131
7.2.2 Cas inhérent .....	139
7.3 Les verbes « ECM » .....	142
7.4 Les clauses réduites .....	146

Chapitre 8 .....	151
Le mouvement <i>wh</i> .....	151
8.1 Questions <i>wh</i> .....	151
8.2 Le mouvement <i>wh</i> .....	152
8.3 La motivation pour le mouvement .....	156
8.4 Les stratégies <i>wh</i> .....	162
8.5 Les questions fermées (questions « <i>yes-no</i> ») .....	167
Chapitre 9 .....	173
Contraintes sur le mouvement <i>wh</i> .....	173
9.1 Les interrogatives enchâssées .....	173
9.2 Le mouvement « long » .....	176
9.3 La notion d'îlot .....	180
9.3.1 Îlot interrogatif ( <i>wh</i> -island) .....	180
9.3.2 DP-complexe .....	182
9.3.3 Sujet phrastique .....	183
9.3.4 Îlots Ajouts .....	184
9.4 Une analyse du mouvement <i>wh</i> .....	185
9.5 Problèmes résiduels .....	192
Chapitre 10 .....	195
Les relatives .....	195
10.1 Le pronom relatif .....	195
10.2 Les pronoms relatifs sujet et objet .....	200
Chapitre 11 .....	207
Le mouvement A .....	207
11.1 Les verbes à montée .....	208
11.2 Le passif .....	215
11.3 Les verbes inaccusatifs .....	222
11.4 Localité du mouvement A .....	224
11.5 L'accord .....	227
11.6 Tableau récapitulatif des verbes .....	232
Chapitre 12 .....	233
Les constructions à double objet .....	233
12.1 Les constructions causatives .....	235
12.2 Les constructions à double complément .....	241
12.3 Les verbes à contrôle par l'objet .....	244
12.4 Les verbes transitifs .....	248

12.5	Tableau récapitulatif 2.....	250
Chapitre 13.....		253
La négation.....		253
13.1	Négation phrastique.....	253
13.2	NegP.....	256
Chapitre 14.....		265
Le CP.....		265
14.1	V2.....	266
14.2	Le CP éclaté.....	271
14.2.1	Focus.....	271
14.2.2	Topic.....	276
14.2.3	CP éclaté.....	278
Chapitre 15.....		283
La théorie du liage.....		283
15.1	Les pronoms.....	284
15.2	La théorie du liage (Binding Theory).....	285
15.3	Le liage et les domaines.....	287
15.4	Problèmes résiduels.....	291
Glossaire.....		295
Références bibliographiques.....		309
Index.....		315

## Introduction

# La grammaire générative

La linguistique est l'étude scientifique du langage. Elle vise à mettre au jour les propriétés universelles des langues et à découvrir de quels principes généraux les propriétés universelles découlent. Elle vise également à déterminer les principes additionnels propres à chaque langue individuelle qui constituent, conjointement avec les principes universels, la grammaire de chaque langue.

(Hirschbühler & Labelle 1998 : 1)

## 1. Le modèle générativiste (Principes et Paramètres)

Ce cours d'introduction à la syntaxe comparative s'inscrit dans le cadre du modèle dit de **grammaire générative** (*generative grammar*). Le courant est né à la fin des années 1950, sous l'impulsion de Noam Chomsky. La publication de son ouvrage *Syntactic Structures* en 1957 est considérée comme le point de départ de la théorie générative.

Les prémisses adoptées dans le modèle générativiste, et qui, depuis le début, servent d'hypothèse de travail, sont les suivantes :

- (i) le langage est unique à l'espèce humaine ; l'humain est génétiquement prédéterminé pour apprendre une/des langue(s).
- (ii) l'humain naît donc avec une capacité langagière, mais bien plus, avec un système « pré-encodé », composé d'un certain nombre de principes universels. Seul l'humain est biologiquement programmé pour apprendre une langue humaine.
- (iii) les principes universels font partie de l'héritage biologique de l'homme. Ce sont des principes linguistiques de base, qui trouvent des réalisations dans toutes les langues du monde. Ces principes, ainsi que leur organisation, correspondent à un état initial, appelé Etat<sub>0</sub> (*State<sub>0</sub>*) de la connaissance linguistique. La théorie qui s'occupe à décrire et à modé-

liser cet état initial est la **Grammaire Universelle**. Parmi les principes universels, on peut citer, entre autres, que toute langue a des catégories de type *verbe* et de type *nom*, toutes les langues utilisent la récursion et l'enchâssement, toutes les langues ont des consonnes et des voyelles, toute phrase a un sujet, etc.

- (iv) Le langage humain est réalisé sous différentes formes par les langues. Bien que les principes universels sous-tendent toutes les langues, ces dernières présentent aussi des différences. Cependant, les variations ne sont pas infinies, mais très restreintes au contraire. Elles sont l'effet de variations paramétriques sur la base des principes universels. Les langues du monde ne varient donc pas de façon infinie et/ou aléatoire, mais sont le résultat d'une combinaison de paramètres différents. Par exemple, le fait que toute phrase a un sujet, et ceci dans toutes les langues, est un principe universel. En revanche, la présence réalisée du sujet dans une phrase est déterminée par un paramètre, le paramètre du sujet nul. Ainsi, bien que l'on considère que l'anglais et l'italien sont fondamentalement semblables puisque les deux langues ont des sujets, il existe une différence « de surface », dans la mesure où le sujet doit être exprimé en anglais, mais peut être omis (dans le cas des pronoms) en italien. On expliquera la différence entre anglais et italien par la fixation d'une valeur différente pour le sujet nul (anglais= non, italien= oui).
- (v) La fixation de ces paramètres se fait au cours de l'acquisition d'une langue. Pour ce faire, l'apprenant (le nourrisson) utilise les données qu'il perçoit dans son environnement, et s'en sert pour fixer les paramètres de sa langue (naturellement, l'acquisition se fait de manière inconsciente – le nourrisson ne décide pas consciemment d'utiliser les données qu'il perçoit pour se construire une grammaire). Par données, nous entendons ici des données linguistiques, à savoir des mots, des phrases, des échantillons de discours dans une ou plusieurs langues dont l'enfant est entouré et qu'il perçoit. Ainsi, de l'état initial  $E_0$ , il va graduellement atteindre un état stable  $Etats$ , qui correspond à la connaissance d'une langue  $L$ , celle à laquelle il a eu accès. Notons qu'un environnement bilingue peut donner lieu au développement de deux systèmes de connaissance linguistiques (bilinguisme). La théorie qui

s'occupe à décrire et à modéliser l'ensemble de ces connaissances est la **grammaire générative**.

Parmi les arguments souvent cités en faveur d'une approche innéiste (à savoir, qui prend comme hypothèse de base que non seulement la capacité d'apprendre une langue, mais les structures linguistiques sont innées), il nous faut mentionner :

*L'acquisition du langage* : l'observation que tous les enfants du monde acquièrent une langue (leur langue maternelle) de la même manière, en passant par les mêmes étapes et plus ou moins au même rythme. Ceci est présenté comme un argument déterminant, puisque quelle que soit la qualité des données à la portée de l'enfant, celui-ci est capable de « construire » sa langue et d'en avoir la maîtrise (on parle de compétence linguistique) dans un laps de temps relativement court. En effet, l'environnement linguistique d'un nourrisson est de qualité très variable. Loin d'avoir accès à des phrases complètes et bien formées, il est le plus souvent soumis à un ensemble hétérogène de bruits, d'exclamations, de bribes de phrase et de langage simplifié que les adultes ont tendance à utiliser en présence d'un bébé. Le fait que malgré la relative pauvreté de stimulus que compose cet ensemble hétéroclite de données linguistiques, l'enfant soit capable d'élaborer une connaissance linguistique complète parle en faveur de l'hypothèse que les principes de base de cette connaissance sont déjà là à la naissance, et sont donc génétiquement prédéterminés.

*L'étude comparative des langues* : plus les études se multiplient, plus on observe que les langues partagent des caractéristiques liées à des principes d'organisation et de fonctionnement universels. Bien qu'en surface, les langues présentent de grandes disparités, ces dernières s'avèrent être essentiellement de nature lexicale. Vu le nombre de langues dans le monde, les structures, c'est-à-dire l'organisation des éléments de la phrase, présentent des similitudes auxquelles on ne s'attendrait pas si chaque langue était un système indépendant et unique.

Il va sans dire que la variation entre langues n'est pas seulement une affaire de structures déterminées par des paramètres fixant des variations de principes universels. Tout locuteur est conscient que la connaissance d'une langue comprend aussi la connaissance d'un système phonologique, ainsi

que d'un lexique. Bien que ce ne soit pas le sujet des investigations premières de la grammaire générative, les chercheurs défendent l'hypothèse que le système phonologique est aussi soumis à des universaux (voir, entre autres, Hyman (1985), McCarthy and Prince (1986), Hayes (1989)). Quant à ce qui paraît, de prime abord, être la grande différence entre les langues, à savoir le lexique, Chomsky suggère qu'il s'agit en fait d'un épiphénomène. En effet, bien que les concepts soient universaux et pré-linguistiques, l'association entre concept et séquence sonore relève de l'arbitraire, comme initialement introduit par Ferdinand de Saussure, auquel Chomsky se réfère :

Ceci [variation dans la perception et l'articulation] mis à part, la variation dans le langage semble résider dans le lexique. Un aspect en est 'l'arbitraire Saussurien', le lien arbitraire entre concepts et sons : le programme génétique ne détermine pas si le concept *arbre* est associé avec les sons « tree » (en anglais) ou « Baum » (en allemand). L'association entre concept et son peut être acquise sur la base d'évidence minimale, et la variation n'est donc pas surprenante. Mais les sons possibles sont étroitement contraints, et les concepts pourraient être fixés virtuellement. Il est difficile d'imaginer autre chose, étant donné la rapidité d'acquisition qui est d'environ un mot par heure des âges de deux à huit ans, avec des éléments lexicaux typiquement acquis après une seule exposition, dans des conditions hautement ambiguës, mais compris dans leur délicate et extraordinaire complexité, bien au-delà de ce qui est donné dans les dictionnaires les plus complets qui, comme toute grammaire traditionnelle (même la plus complète) ne peuvent donner que des indications par ailleurs suffisantes pour des individus qui connaissent les réponses dans la mesure où elles sont largement innées. (Chomsky 2000 :120) [ma traduction].

L'étude des principes universaux et des variations paramétriques est donc l'objet principal de la grammaire générative. De ce fait, une approche comparative prend tout son sens, puisque les différences entre langues sont postulées comme étant le fait d'une différence dans la fixation de certains paramètres, et que les similitudes sont la manifestation des principes universels.

## 2. Le programme Minimaliste

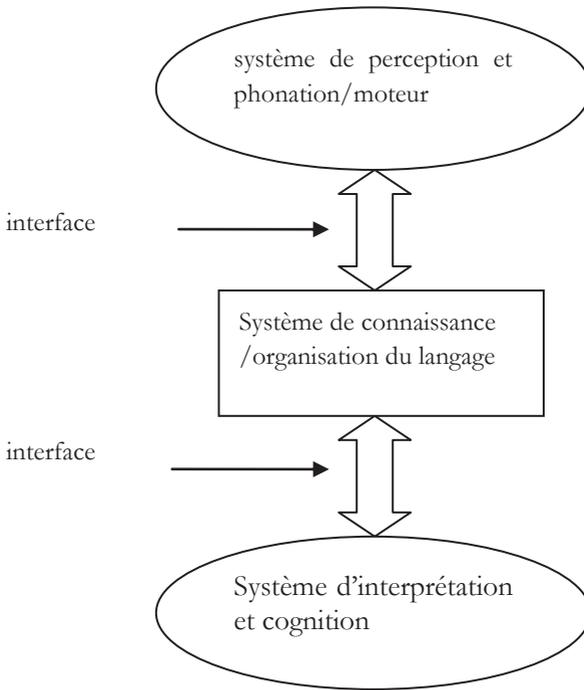
Au cours des années 1990, le courant générativiste a vu apparaître un certain nombre d'innovations importantes. Il ne s'agit pas, selon Chomsky (voir Chomsky 1995), d'une rupture radicale dans la théorie, mais d'une réorientation de la recherche, motivée par la redondance que la recherche en linguistique a petit-à-petit introduite dans le système descriptif développé précédemment et connu sous le nom de **Principes et Paramètres**. Le programme minimaliste, dont les principes sont exposés dans l'ouvrage *Minimalist Program*, par Noam Chomsky, n'est pas une théorie mais un programme de recherche. Celui-ci vise à optimiser l'adéquation entre **modèle descriptif** (qui cherche à décrire du mieux possible les variations et les spécificités des diverses langues dans le monde) et **modèle explicatif** (qui cherche à rendre compte non seulement des variations mais aussi, et surtout, des principes de bases qui permettent ces variations). Dans ce sens, beaucoup des propositions élaborées dans le cadre de ce programme éliminent des formalisations complexes et redondantes, traquent la multiplication des niveaux de représentation ainsi que les conceptualisations biaisées. Il est ici impossible de faire un compte-rendu détaillé des objectifs du programme minimaliste, mais le lecteur curieux est invité à lire l'introduction de l'ouvrage éponyme, qui donne une idée claire des buts poursuivis par l'auteur.

Dans ce volume, le modèle théorique choisi se base sur certaines propositions issues du programme minimaliste. Mais comme toute théorie en plein essor, le minimalisme n'a pas nécessairement atteint une stabilité dans tous ses domaines d'application. C'est pourquoi certaines formulations de cet ouvrage sont résolument « minimalistes », alors que dans d'autres domaines, le choix du formalisme sera plus neutre, se basant sur des développements moins récents de la théorie. Nous avons cependant veillé à exposer une théorie cohérente et moderne de la grammaire générative, et le lecteur assidu ne rencontrera aucune difficulté à lire et à comprendre des articles récents dont l'orientation serait plus radicalement minimaliste.

### 3. La syntaxe comparative

Le lecteur l'aura sans doute compris, au cœur de la théorie générative se trouve la syntaxe. Bien que la grammaire générative, en tant que théorie de la connaissance du langage, se doive de tenir compte de la connaissance phonologique, sémantique et même pragmatique du locuteur, l'essentiel du travail de recherche de Chomsky et des « générativistes » s'est concentré sur la syntaxe, c'est-à-dire sur l'étude du système qui permet d'organiser les différents composés du langage. En effet, le modèle générativiste part de l'observation suivante. La connaissance du langage interagit avec d'autres modules de connaissance au moyen de deux portes d'accès, deux interfaces. L'une interagit avec le **système articulatoire-perceptuel** et l'autre avec le **système conceptuel-intentionnel**. Le premier fait le pont entre systèmes de perception (audition, vision dans le cas de la langue des signes) et de production (phonation, moteur dans le cas du signe) et système de connaissance linguistique, alors que le second donne accès aux systèmes d'interprétation et systèmes cognitifs non-linguistiques. Le schéma en (1) ci-dessous illustre ces connections :

(1)



Seules ces deux « portes » communiquent avec les autres modules cognitifs. Cependant, les systèmes des sons et des interprétations ne sont, à priori, pas appariés (autrement dit, nous n'avons pas, dans un grenier de notre cerveau, des expressions complexes toutes faites qui correspondent à chacune des séquences sonores associées à leur sens que nous produisons). Le mécanisme qui permet d'associer ces systèmes, qui permet de les organiser et de créer à l'infini des énoncés linguistiques interprétables est le module décrit comme la syntaxe. Autrement dit, la syntaxe permet d'organiser les pensées en formes communicables et interprétables. C'est la raison pour laquelle elle a un poids prépondérant dans le modèle générativiste.

Comme nous l'avons vu plus haut, les recherches récentes ont mis l'accent sur la syntaxe comparative. En effet, l'étude et la comparaison de différentes langues montre que bien que les différences (prévisibles) entre langues soient évidentes, les variations (modulo le fameux arbitraire dans

l'association entre un sens et une séquence sonore) sont en fait très restreintes. L'étude d'un grand nombre de langues fait donc ressortir ce que les langues ont en commun. Le but de la grammaire générative étant d'identifier ce qui appartient aux principes universels et ce qui relève de la paramétrisation, une étude comparative est l'approche la plus cohérente pour atteindre ce but.

#### 4. But de ce cours

Ce cours donne les outils nécessaires à la compréhension de la théorie générative. Il se veut essentiellement pratique, et se propose de :

- présenter une approche à la langue qui permette de comprendre les mécanismes de la phrase et de ses composants, le but de la linguistique théorique étant l'explicitation et la formalisation des connaissances du locuteur.
- donner une introduction – forcément non-exhaustive – de la théorie syntaxique utilisant un outillage minimaliste. Il comporte donc des choix, comme celui de traiter certains aspects, mais d'en laisser d'autres, plus complexes, de côté.
- présenter une introduction à la syntaxe comparative. Il se réfère donc à d'autres langues, pour en montrer la diversité mais aussi les ressemblances.
- donner accès à des références qui permettent au lecteur, s'il le souhaite, d'étudier certains sujets de manière plus approfondie.

Mais il reste un cours destiné aux étudiants francophones. C'est pourquoi, là où il n'y a pas de pertinence à la comparaison, seul le français est illustré. Afin de ne pas alourdir le cours, nous avons choisi de ne pas donner de références systématiques à un grand nombre de langues ; plutôt, l'emphase sera mise sur une discussion des points qui sont importants, dans la variation, mais aussi dans les similitudes.

De façon plus pratique, il nous semble important, en abordant ce cours, de garder en tête un certain nombre de définitions fondamentales. Ces dernières sont brièvement exposées ci-dessous :

### 1. *La linguistique théorique*

La linguistique théorique s'intéresse à la **connaissance** du langage, c'est-à-dire à l'ensemble des connaissances qu'un locuteur a, souvent de façon non-explicite, de sa langue maternelle. En utilisant les outils de la linguistique théorique, le linguiste cherche à :

- (i) déterminer l'ensemble des connaissances (ou règles) qui génèrent (=énumèrent explicitement au moyen de règles) les énoncés produits ou acceptés comme grammaticaux par un locuteur.
- (ii) déterminer les propriétés communes à l'ensemble des langues du monde, et expliquer l'existence de ces propriétés en les mettant en rapport avec les principes innés de l'être humain.

L'objectif de la linguistique théorique est donc de développer un **modèle** de la faculté du langage, en se basant sur l'étude de langues particulières. La **grammaire générative** est une grammaire explicite (où grammaire est comprise comme une **théorie** de la connaissance, un ensemble de règles) qui permet de générer toutes les phrases possibles d'une langue, et seulement celles-ci. La grammaire générative d'une langue est une **modélisation** du système de connaissance que possède chaque locuteur de cette langue.

### 2. *Prescriptif et descriptif*

Il existe plusieurs approches à l'étude des langues. Une approche **prescriptive** (ou normative), qui codifie et impose l'usage d'une variété de langue. L'approche prescriptive se retrouve dans les grammaires de langue étrangère, par exemple. Une approche **descriptive**, elle, décrit toutes les variétés, de la façon la plus fidèle et objective possible. Elle sert de base à la formulation des règles et des principes permettant de modéliser la connaissance linguistique. Les **données** utilisées pour cette élaboration sont :

(i) des échantillons de discours ou de texte. Bien que le texte écrit soit incontournable, notamment lorsqu'il s'agit de travailler sur le développement diachronique d'une langue, les échantillons de langage parlé sont extrêmement précieux, dans le sens où ils correspondent souvent à une forme spontanée de production langagière.

(ii) des jugements de **grammaticalité**. Le locuteur (natif) d'une langue est capable de juger de la grammaticalité d'un énoncé : il a des intuitions sur ce qui est grammatical et ce qui ne l'est pas dans sa langue. Ces jugements sont primordiaux pour le linguiste, puisqu'ils sont le reflet de la connaissance linguistique du locuteur, et donc de sa grammaire. Il nous faut faire ici une distinction entre *grammaticalité* et *acceptabilité*. En effet, on considère comme grammaticale une séquence (phrase) qui est bien formée du point de vue de la grammaire (donc de l'ensemble de règles) d'une langue. La notion d'acceptabilité est plus vaste, puisqu'elle inclut aussi la validité sémantique d'un énoncé, ou même l'acceptabilité selon des normes sociales, etc. Ainsi, une phrase peut être grammaticale, sans pour autant être acceptable, comme l'illustre le très célèbre exemple *colourless green ideas sleep furiously* (des idées vertes incolores dorment furieusement) (Chomsky 1957).

## 5. Les données et les références

Finalement, un mot sur les sources qui sont utilisées dans cet ouvrage. Pour illustrer les points théoriques discutés dans ce cours, nous avons recours à des données. Comme mentionné plus haut, la majorité des exemples illustrent le cas du français. Sauf indications, les exemples en français sont des exemples originaux de l'auteur, inspirés par les personnages et les situations tirés de « L'abécédaire musical à l'usage des enfants et des personnes qui téléphonent », recueil de chansons créées par Boris Vian et chantées par Lucienne Vernay et les Quatre Barbus. Lorsque cela est pertinent, des exemples et des discussions sur d'autres langues sont fournis. Ces exemples dans d'autres langues sont soit créés par l'auteur, soit aimablement fournis par des locuteurs divers, soit, lorsque cela est indiqué, repris d'ouvrages

traitant de différentes langues. Les références sont toutes données en fin d'ouvrage.

Comme tous les manuels introductifs, celui-ci vise à rassembler des connaissances, en l'occurrence des connaissances dans le domaine de la grammaire générative développées dans les cinquante dernières années. Les questions traitées ici ont donc été élaborées, évaluées, résolues et réexaminées par un grand nombre de chercheurs travaillant dans le cadre générativiste. Afin de ne pas alourdir le texte, les références à tous ces auteurs sont, pour la plupart, insérées en fin de texte. Le lecteur curieux est encouragé à les consulter pour enrichir ses connaissances. Seules certaines références, qui concernent un point précis associé de manière incontournable avec un auteur en particulier, sont données directement dans le texte. A la fin de chaque chapitre, le lecteur trouvera quelques références qui lui permettront d'aller plus loin dans l'étude des points abordés dans le chapitre.

### Pour en savoir plus

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague : Mouton.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass : MIT Press.

Hayes, Bruce. 1989. « Compensatory lengthening in moraic phonology ». *Linguistic Inquiry* 20 : 253–306. [OpenURL Query Data] [Google Scholar].

Hyman, Larry. 1985. « A theory of phonological weight ». (Publications in Language Sciences 19). Foris. [Google Scholar].

McCarthy, John & Alan Prince. 1986. Prosodic morphology 1986. Report No. RuCCS-TR-32. Rutgers University Center for Cognitive Science. Accessible sur : <http://rucss.rutgers.edu/pub/papers/pm86all.pdf>.



## Chapitre 1

# Structure argumentale

Puisque la syntaxe s'occupe de l'organisation de la phrase, il nous faut en tout premier lieu déterminer ce qu'est une phrase. Bien que les énoncés d'un locuteur puissent prendre des formes très variées, nous observons que la notion de phrase est essentielle. La raison principale est qu'une phrase permet d'exprimer une proposition, qui correspond à une situation dans le monde. Prenons un exemple :

(1) a. L'âne d'Anatole arrive.

Cette situation, étant donné un monde défini, peut être évaluée comme vraie ou fausse. Pour exprimer cette situation, il nous faut au minimum un **prédicat** et un **argument**. Le prédicat exprime une propriété, ici *arrive*. Cette propriété s'applique à un individu ou à un objet, ici *l'âne d'Anatole*, qui est l'argument de ce prédicat. La phrase en (1) exprime une situation dans le monde où *arriver* est une propriété de l'âne d'Anatole. Cependant, le terme de *prédicat* peut avoir différentes acceptions. Dans la tradition aristotélicienne, le prédicat correspond à ce que l'on exprime à propos d'un sujet. Dans ce sens, les séquences *arrive* en (1a), mais aussi *a mangé tous les ananas* en (1b) correspondent au prédicat de la phrase :

(1) b. L'âne a mangé tous les ananas.

Une autre acception, issue des travaux du logicien Gotlob Frege, considère comme *prédicat* une entité atomique de la phrase (par exemple le verbe ou l'adjectif prédicatif) qui doit se combiner avec un ou plusieurs autres éléments pour former une proposition. C'est dans ce sens que nous utilisons ici le terme *prédicat*. Il désignera donc le verbe ou, le cas échéant, l'adjectif ou le nom prédicatif de la phrase. En (1b), le prédicat sera donc *mangé*.

Nous considérons donc la phrase comme une séquence linguistique qui est composée au minimum d'un prédicat et d'un argument. Dans ce chapitre, nous examinons les différentes caractéristiques des prédicats et de leurs arguments.

## 1.1 Sous-catégorisation

La composition d'une phrase dépend du type de prédicat qu'elle contient. Il existe deux types de contraintes, établies selon des critères sémantiques d'une part et syntaxiques de l'autre. La contrainte sémantique détermine, en fonction du sens du prédicat, le nombre et le type d'arguments qui lui sont associés ; la contrainte syntaxique, elle, définit le type syntaxique des arguments et la relation structurelle qu'ils entretiennent avec le prédicat. Nous commençons par la contrainte syntaxique, illustrée dans les phrases ci-dessous (les phrases précédées d'un astérisque \* sont agrammaticales) :

- (2)
- a. Anatole a vu un âne.
  - b. \*Anatole a vu.
  - c. \*Anatole a vu dans le clocher.
  - d. Anatole a vu le clocher.
  - e. \*Anatole a vu un âne le clocher.

Les phrases en (2) comportent un prédicat qui est de catégorie verbale (voir chapitre 2). Ce prédicat apparaît avec deux éléments de type nominal, *Anatole* et *un âne*. Sur une base empirique mais aussi théorique, il a été proposé que la présence d'un sujet dans une phrase est une propriété universelle des langues naturelles. D'un point de vue syntaxique, il est donc admis que toute phrase a un sujet (qu'il soit exprimé ou non). C'est pourquoi il n'est pas nécessaire d'inclure une mention de sa présence dans une description syntaxique (nous revenons plus loin sur une discussion détaillée des sujets ; pour le moment, nous admettrons que tout locuteur a une idée approximative de ce qu'est un sujet, à savoir l'entité qui possède la propriété exprimée

par le prédicat).<sup>1</sup> Par contre, le verbe *voir* est un verbe **transitif**, qui doit apparaître avec un (deuxième) élément de type nominal (*un âne* en 2a, *le clocher* en 2d). Cet élément est le **complément** du verbe. Le verbe *voir*, comme tout verbe transitif, ne peut, sauf circonstances particulières auxquelles nous revenons ci-dessous, apparaître sans complément (2b). De même, un élément introduit par une préposition (2c) n'est pas un complément possible. Enfin, un verbe transitif ne peut avoir qu'un seul complément (2e).

Le verbe *dormir* est **intransitif** : il ne peut pas prendre de complément nominal.

- (3) a. Berthe dort.  
b. \*Berthe dort des bottes.

Notons que d'autres verbes peuvent demander un autre type de complément :

- (4) a. \*Bastien habite.  
b. Bastien habite à Bordeaux.
- (5) a. Célestin se demande [si Charlotte aime les carottes].  
b. \*Célestin se demande.  
c. \*Célestin se demande [l'amour de Charlotte pour les carottes].

Un verbe comme *habiter* exige un complément de type prépositionnel (4), alors que le verbe *se demander* demande un complément de type phrastique. Certains verbes ont un comportement « mixte », qui leur permet de prendre différents types de complément :

- (6) a. Désiré pense (à haute voix).  
b. Désiré pense à la duchesse.  
c. Désiré pense que la duchesse est jolie.

1 La question du sujet est beaucoup plus complexe qu'il n'y paraît. Nous renonçons délibérément à toute approche qui consisterait – comme on peut le voir dans certains manuels scolaires – à dire que le sujet est « celui qui fait l'action ». Le chapitre 5 donne des pistes plus détaillées quant à la façon d'aborder le problème.

- (7) a. Eugène annonça [qu'il quittait la ville].  
 b. Eugène annonça [son départ].

Enfin, il existe aussi des verbes **di-transitifs**, qui requièrent deux compléments :

- (8) a. Le domestique a donné dix ducats à Désiré.  
 b. \*le domestique a donné à Désiré.  
 c. \*?le domestique a donné dix ducats.
- (9) a. Émile a mis les pieds sur la table.  
 b. \*Émile a mis les pieds.  
 c. \*Émile a mis sur la table.

Les verbes transitifs ont besoin d'un complément, alors que les verbes intransitifs l'excluent. Le fait que ce soit le type (nominal, prépositionnel, phrastique) de complément, et non pas son sens qui soit déterminant indique que l'on a bien affaire à une contrainte syntaxique. On parle alors de **c-sélection** (sélection de catégorie). Alternativement, on dit que les verbes **sous-catégorisent** pour un certain type de complément. Bien que les grammaires traditionnelles fassent une différence entre les compléments de type nominal illustrés en (2) et les compléments introduits par une préposition (4) ou par une conjonction (5a, 6c, 7a), nous considérerons que ces différents éléments fonctionnent tous comme des compléments, dans la mesure où leur présence est syntaxiquement déterminée par le prédicat.

Notons que la c-sélection a été distinguée d'un autre processus qui utilise un critère **sémantique** (lié au sens) pour sélectionner le type de complément d'un prédicat. Il s'agit alors de **s-sélection** (sélection sémantique). Ainsi, un complément peut être de type *Q* (question), *E* (exclamatif), ou *P* (proposition) tout en appartenant à une même catégorie syntaxique (ici, un complément phrastique) :

- (10) a. Anatole se demande [combien l'âne a d'oreilles]<sub>Q</sub>  
 b. Anatole ne peut pas croire [combien son âne est intelligent]<sub>E</sub>  
 c. Anatole explique [combien d'ânes il a vus]<sub>P</sub>

Inversement, un même type (ici *P*) peut être réalisé par différentes catégories syntaxiques :

- (11) a. Anatole avoue [qu'il aime son âne] *compl. phrastique*  
 b. Anatole avoue [son amour pour son âne] *compl. nominal*

Cependant, nous n'utiliserons que très occasionnellement la s-sélection, et le terme de *sélection* se référera par défaut à la c-sélection.

Les observations données ci-dessus soulèvent un certain nombre de questions, que nous allons examiner ci-dessous.

**Question 1.** La sous-catégorisation ne fait état que des compléments sélectionnés (les éléments obligatoires). Dans les phrases ci-dessous, les éléments en italiques sont-ils sélectionnés ou non ?

- (12) a. Anatole dort à *poings fermés*.  
 b. L'âne a avalé *un panier d'artichauts devant l'échoppe*.

Pour répondre à cette question, on peut appliquer un certain nombre de tests, qui nous permettent de distinguer les compléments sélectionnés des éléments non sélectionnés.

a) omission

Un complément sélectionné ne peut, en principe, pas être omis, comme en témoigne l'exemple (13b), dans lequel il est agrammatical d'omettre *un panier d'artichauts*. Un élément non sélectionné peut toujours être omis (13a, à *poings fermés* ; 13b, *devant l'échoppe*).

- (13) a. Anatole dort (à poings fermés).  
 b. L'âne a avalé \*(un panier d'artichauts) (devant l'échoppe).

Les éléments non-sélectionnés, qui peuvent être omis, sont des **ajouts**. Leur rôle est de modifier le prédicat en ajoutant des informations de lieu, de temps, de manière, etc.

## b) « et cela »

Le complément sélectionné ne peut être séparé du verbe par la locution *et cela*, contrairement aux éléments non sélectionnés (ajouts) :

- (14) a. Anatole dort, **et cela** à poings fermés.  
 b. L'âne a avalé un panier d'artichauts, **et cela** devant l'échoppe.  
 c. \*L'âne a avalé, **et cela** un panier d'artichauts.

Nous voyons que les éléments *à poings fermés* en (14a) et *devant l'échoppe* en (14b) peuvent être séparés du verbe par *et cela*. Ce sont donc des ajouts. En revanche, *un panier d'artichauts* ne le peut pas. Nous en concluons qu'il s'agit d'un complément.

## c) « en faire autant »

La locution *en faire autant* peut remplacer une séquence qui comporte le verbe et les compléments et ajouts qui le suivent :

- (15) a. L'éléphant a [*cassé un vase de Murano dans la boutique hier soir*]  
 et la souris *en a fait autant*.  
 b. L'éléphant a [*cassé un vase de Murano*] (dans la boutique)  
 et la souris *en a fait autant* (dans le musée)  
 c. \*L'éléphant a [*cassé*] un vase de Murano  
 et la souris *en a fait autant* un lustre viennois.

Il n'y a pas de limite sur le nombre maximal d'éléments que l'expression *en faire autant* peut reprendre : en (15a), elle remplace le verbe *cassé*, son complément *un vase de Murano*, et deux ajouts, *dans la boutique* et *hier soir* ; en (15b), elle reprend le verbe et son complément, à l'exclusion de l'ajout. Par contre, l'unité minimale reprise par *en faire autant* est le verbe et son complément sélectionné (15c).

**Question 2.** Certains verbes ont l’air d’avoir deux c-sélections différentes.

- (16) a. La reine boit du lait.  
b. La reine boit.

Y a-t-il deux entrées lexicales (donc deux verbes *boire*) avec des propriétés de sélection différentes ? Ou pouvons-nous considérer que dans l’exemple (16b), il y a un complément implicite ?

Il est évident qu’il y a ici un hiatus entre les propriétés sémantiques du verbe et ses propriétés syntaxiques. Typiquement, (16b), tout comme (16a), implique l’ingestion de liquide. Il faut donc faire la différence entre un **usage transitif** et un **usage intransitif**. Dans l’usage intransitif, nous avons affaire à des verbes transitifs, mais avec un complément qui n’est pas syntaxiquement réalisé. Nous choisirons de considérer ces verbes comme étant sémantiquement des verbes transitifs, mais avec une possibilité de réalisation syntaxique transitive ou intransitive.

## 1.2 Les rôles thématiques

Nous avons examiné les propriétés de sélection syntaxique des verbes, mais il existe d’autres contraintes, qui ne sont pas directement syntaxiques (le # est un signe diacritique qui indique que l’exemple est sémantiquement incorrecte ou bizarre) :

- (17) a. #Bastien surprend le bonnet.  
b. #Le vase de Murano aime la pizza.

Dans les exemples ci-dessus, les contraintes de sous-catégorisation sont respectées : *surprend* et *aime* sont tous deux des verbes transitifs, qui requièrent un complément (en l’occurrence *le bonnet* et *la pizza*). Mais ces phrases ne sont pas acceptables. Comme nous l’avons vu plus haut, un prédicat exprime une propriété, qui peut s’appliquer à un ou plusieurs arguments. Ces arguments peuvent être considérés comme des « participants », qui sont

en relation avec le prédicat. Les relations sont exprimées par des **rôles thématiques**, qui sont déterminés, ou **assignés**, par le prédicat. En (17a), le prédicat *surprend* a besoin d'un participant (argument) qui puisse de façon délibérée ou non surprendre – c'est le cas de *Bastien*, mais aussi d'un participant (argument) qui puisse ressentir de la surprise. Or *le bonnet* est un complément qui ne peut remplir cette condition. De même, le prédicat *aimer* requiert un participant qui puisse ressentir de l'amour, et en (17b), *le vase de Murano* n'est pas un argument qui puisse satisfaire cette exigence.

Il existe un grand nombre de rôles thématiques différents, et les auteurs ne sont unanimes ni sur le nombre, ni sur la dénomination de ces rôles. A titre d'exemple, voici une liste de rôles thématiques possibles :

(18) **Rôles thématiques**

- Agent : animé, fait volontairement l'action
- Cause : animé ou non, est la « cause » de l'action
- Patient : subit un changement d'état, est affecté par l'action.
- Thème : est visé par l'action.
- Instrument : inanimé, utilisé pour l'action.
- But : endroit où se dirige l'action.
- Bénéficiaire : animé, qui bénéficie de l'action.
- Source : provenance de l'action.
- Lieu : endroit où se trouve l'action.
- Expérimenteur (*experier*) : animé, qui fait l'expérience d'une émotion, d'un état psychologique.

Formellement, nous pouvons représenter la structure argumentale d'un verbe de la manière suivante :

- (19) a. Anatole a vu un âne.  
(voir : agent, thème)
- b. Célestin a envoyé une carte postale à sa grand-mère.  
(envoyer : agent, patient, bénéficiaire)