

LA GRANDE GRAMMAIRE DU FRANÇAIS : PRINCIPES DE CONSTRUCTION

ANNE ABEILLE, ANNIE DELAVEAU, DANIELE GODARD

Abstract. The new reference Grammar of French (La Grande Grammaire du français), which is to be published in 2009, will be approximately 2000 page long, and the result of a collective project (around fifty authors). Like the recent reference grammars of the other Romance Languages, it takes into account the important results of the linguistic research of the past 30 years, while aiming at a non specialist audience (it is not formalized, and uses the traditional syntactic categories). It includes a description of all the syntactic phenomena, as well as semantic and pragmatic insights, specially as they interact with syntax. The analysis concerns the data from standard written French, but also includes data from spoken corpora and regional or non standard French (when accessible). Throughout the grammar, a simple phrase structure grammar is used, in order to maintain a common representation.

1. PRÉSENTATION

Le projet de réaliser une grande grammaire du français a été lancé par A. Delaveau en 2002, avec le soutien du CNRS (Institut de la langue française). Il s'agit d'un projet collectif, réalisé grâce à la collaboration d'une cinquantaine de collègues, français et étrangers, et dirigée par les trois auteurs de cet article¹. Elle est en cours d'écriture, et doit être terminée en 2008, pour être publiée par les Editions Bayard en 2009. La motivation pour ce projet est simple. Alors que les autres langues romanes, ainsi que l'anglais, par exemple, parmi les langues

¹ Voici la liste des auteurs : Anne Abeillé, Inge Bartning, Anne-Marie Berthonneau, Claire Beyssade, Olivier Bonami, Paul Cappeau, Michel Charolles, Bernard Combettes, Elisabeth Delais, Annie Delaveau, Marianne Desmets, Albert di Cristo, Jenny Doetjes, Benjamin Fagard, Szusanna Fagyal, Françoise Gadet, Claire Gardent, Danièle Godard, Jean-Pierre Jaffré, Jacques Jayez, Kerstin Jonasson, Brigitte Kampers-Manhe, Georges Kleiber, Jean-Pierre Koenig, Andres Kristol, Lucien Kupferman, Brenda Laca, Knud Lambrecht, Béatrice Lamiroy, Jean-Marcel Léard, Jean-Marie Marandin, Alda Mari, Ludo Melis, Philip Miller, Christian Molinier, François Mouret, Florence Mourlhon-Dallies, Walter de Mulder, Michèle Noailly, Marleen van Peteghem, Brechtje Post, Laurent Roussarie, Catherine Schnedecker, Henriette de Swart, Liliane Tasmowski, Lucia Tovená, Jesse Tseng, Danièle van de Velde, Co Vet, Carl Veters, Marie-Thérèse Vinet, Robert Vivès, Anne Zribi-Hertz.

germaniques, ont été décrites dans de grandes grammaires durant les vingt dernières années, ce n'est pas le cas pour le français².

Il est vrai que le français ne manque pas de grammaires. D'abord, il a bénéficié au vingtième siècle de la grande grammaire de Damourette et Pichon (1911–1927), qui comporte environ 4 700 pages. Ce travail a eu deux grands mérites. D'une part, les auteurs se sont efforcés de concevoir leur grammaire comme un ensemble cohérent, en mettant au point un ensemble de concepts originaux qu'ils ont définis et utilisés à travers toute la grammaire. D'autre part, ils ne se sont pas limités au « bon français », mais ont voulu intégrer le français tel qu'il se parle, s'appuyant sur des relevés d'énoncés entendus dans la vie quotidienne, produits par des locuteurs n'appartenant pas à une catégorie sociale particulière. Tout le monde, par ailleurs, connaît la grammaire de M. Grevisse (première édition 1936), révisée et complétée par A. Goose (12^{ème} éd. 1986), qui comporte environ 1 750 pages. Cette grammaire est générale, puisqu'elle comporte phonétique, morphologie et syntaxe ; en même temps, elle entre dans de nombreuses analyses de détail, avec des aperçus historiques. Elle ne déconcerte pas le lecteur, car elle s'appuie sur les concepts tels qu'ils sont en général enseignés à l'école (tout en s'efforçant de les définir). Les données sont pour beaucoup empruntées à la littérature, mais ont ceci d'intéressant qu'elles s'ouvrent à des données régionales, et notamment aux « belgicisms ». Il existe enfin des grammaires récentes, visant un public plus universitaire, comme celle de Riegel *et al.* (1994). Cette grammaire couvre également, sous un format compact (670 pages), les différents aspects d'une grammaire. Une des nouveautés est la présence d'un chapitre intitulé *Grammaire et communication*, qui traite de questions comme la référence ou la structuration du texte. Elle tient compte des avancées en linguistique, et s'appuie sur une bibliographie à jour.

Le pari de la grammaire en cours est en quelque sorte de combiner les bonnes propriétés des grammaires précédentes, tout en bénéficiant des avancées théoriques et empiriques de ces dernières décennies. Ainsi, elle doit être complète, mais rester maniable. C'est pourquoi, elle aura environ 2 000 pages. Elle veut intégrer les résultats des recherches fécondes en linguistique, donc, elle doit être cohérente, définir ses concepts, fonder ses analyses sur les propriétés des phénomènes, tout en restant accessible à un public non spécialisé. On se fondera donc sur des concepts établis dans la tradition, comme les catégories syntaxiques, en les définissant soigneusement par des propriétés empiriques observables pour le français, et on introduira des concepts moins connus, là où cela paraît spécialement éclairant. Mais il ne s'agit pas d'une grammaire formalisée, ce type d'ouvrage n'étant pas d'un intérêt général. La conséquence de ces exigences est que la couverture de la

² Voir, pour l'italien, L. Renzi, G. Salvi, et A. Cardinaletti (eds) (1989-1991) pour l'espagnol, I. Bosque et V. Demonte. (eds) (1999), pour le catalan, J. Sola *et al.* (eds) (2002), pour l'anglais, R. Huddleston et G. Pullum (eds) (2002), pour le roumain, le GALR (2005), et un projet en cours, dirigé par C. Dobrovie-Sorin et G. Pană Dindelegan.

grammaire a été réduite du point de vue des types de phénomènes abordés. Il s'agit d'une grammaire qui se concentre essentiellement sur la syntaxe. Ce n'est pas, bien sûr, que les autres aspects de la langue soient totalement absents, mais ils sont présents dans la mesure où ils permettent de répondre aux problèmes posés par les données, c'est-à-dire lorsque la syntaxe est en interaction avec la morphologie, la sémantique ou la pragmatique ; ou bien lorsque le découpage même des phénomènes exige une approche prenant en compte la sémantique. Enfin, la grammaire ne se limite pas au français dit « standard » mais s'étend à d'autres variétés, en s'appuyant sur les corpus qui commencent à exister pour différentes situations et différents genres ; il cherche également à prendre en compte des données de la variation régionale, y compris les français hors de France.

Dans le reste de cet article, nous nous concentrons sur les deux défis majeurs d'une telle grammaire, que sont la couverture des phénomènes, et la cohérence des analyses.

2. COUVERTURE DES PHÉNOMÈNES

2.1. L'organisation de la grammaire

Cette grammaire comporte une vingtaine de chapitres³. Elle met l'accent sur la syntaxe. Cela se manifeste dans l'organisation des chapitres, puisque la première partie de la grammaire, consacrée à la phrase et à ses constituants, est essentiellement constituée de chapitres organisés autour des catégories syntaxiques majeures, c'est-à-dire celles qui sont la tête d'un syntagme : le verbe, le nom, l'adjectif, la préposition et l'adverbe. En fait, toutes les grandes grammaires des dernières années auxquelles nous nous sommes référées plus haut font également ce choix. Cela peut paraître évident, mais vaut tout de même la peine d'être noté : cela montre la solidité de ces distinctions catégorielles héritées de notre tradition grammaticale, au moins pour ce qui concerne nos langues européennes.

Certains, dans les débuts de la grammaire générative, ont voulu mettre en doute la pertinence des distinctions catégorielles. Il s'agissait de montrer qu'on ne pouvait pas définir ces catégories de manière satisfaisante, qu'il y avait en fait une gradation plutôt que des distinctions nettes (voir, par exemple, Ross 1972). En réalité, il est tout à fait possible de définir les catégories syntaxiques par des propriétés empiriques ; il faut seulement admettre que ce sont les groupements de propriétés et non les propriétés prises individuellement qui sont définitoires. Voici, pour le verbe et le nom, les propriétés sur lesquelles nous nous appuyons. Les noms ont un genre et un nombre autonomes, transmis aux constituants concernés par

³ Chaque chapitre est pris en charge par un responsable, qui est aussi un auteur (voir section 3.1.).

l'accord ; tous les noms peuvent être précédés d'un déterminant (y compris les noms propres), et modifiés par un ajout (par exemple, une relative) ; ils n'ont pas pour complément un syntagme nominal. Les verbes sont fléchis pour les mode, temps, et personne ; ils sont les hôtes des clitiques pronominaux ; ils définissent le nombre, la catégorie et certaines propriétés sémantiques de leurs arguments (réalisés comme sujet et compléments) ; ils admettent un syntagme nominal comme complément.

- (1) a. *Les enfants sont grands.*
 b. *Ce Paul commence à exagérer.*
 c. *le Paul que j'ai connu*
 d. **le combat le terrorisme vs le combat contre le terrorisme*
 e. *Paul vient/viendra/venant ; les enfants viennent/viendront.*
 f. *Paul voit son interlocuteur ; Paul le voit/vous voit.*
 g. *Paul vient/*vient Paris /vient à Paris ; Paul lit (un livre)/*lit Paris.*
 h. *Nous luttons contre le terrorisme ; nous combattons le terrorisme.*

Une autre critique des distinctions traditionnelles a souligné qu'elles occultent ce que des catégories distinctes ou des classes d'items appartenant à des catégories distinctes peuvent avoir en commun. Ainsi, les prépositions et les verbes ont en commun de (pouvoir) prendre un complément SN direct : on dit *combattre le terrorisme*, *contre le terrorisme*, mais non **le combat le terrorisme*, **volontaire le terrorisme*. Ou encore, on peut dire que la distinction entre nom et verbe occulte la parenté étroite entre un verbe comme *répartir* et un nom comme *répartition*, qui partagent non seulement une base morphologique, mais une relation sémantique et une liste d'arguments : répartir SN entre SN. Cependant, pour exprimer les propriétés communes, il existe d'autres moyens que la remise en cause des catégories elles-mêmes. Ainsi, on peut représenter l'organisation des catégories plutôt comme une structure d'objets que comme un simple ensemble⁴. De même, le lexique peut être organisé de telle manière que des relations existent entre des paires, ou plutôt des classes d'items catégoriellement différenciés⁵.

Un chapitre syntaxique mérite un commentaire, car il s'agit d'une innovation : celui qui s'appelle « les prédicats complexes ». Il faut entendre cette expression en un sens syntaxique : on parle de prédicat complexe lorsqu'une combinaison de plusieurs prédicats se comporte, du point de vue syntaxique, comme s'il n'y en avait qu'un. La propriété la plus évidente est celle qui a été appelée la « montée des clitiques ». Les pronoms dits clitiques s'attachent au verbe, en position initiale pour

⁴ Chomsky (1970) propose que les catégories soient définies comme des ensembles de traits. Dans la grammaire syntagmatique HPSG (Pollard, Sag 1994), les catégories forment des hiérarchies, avec un possibilité d'héritage multiple (voir, par exemple, en (5), une hiérarchie partielle).

⁵ Voir les *règles de redondance* des années 1970, et, plus récemment les règles lexicales ; par exemple, dans Wasow (1977), pour la relation entre l'actif et le passif.

ce qui est du français ; ils réalisent des arguments attendus par le verbe. Ainsi, (1f) donne deux types de réalisations différentes de l'argument objet. Avec les prédicats complexes, les clitiques apparaissent non sur le prédicat qui les attend, mais sur la tête verbale qui prend ce prédicat comme complément. Nous parlons ici de « prédicat » plutôt que de verbe, car il peut s'agir d'items appartenant à des catégories autres que les verbes : adjectif (3b) ou nom (4b).

- (1) a. *Paul fait cuire le rôti.*
b. *Paul le fait cuire.*
- (3) a. *Paul est attaché à son idéal.*
b. *Paul y est attaché.*
- (4) a. *Paul a l'impression de comprendre.*
b. *Paul en a l'impression.*

Ce phénomène caractérise un ensemble de structures dont les propriétés communes ne sont pas apparues tout de suite. Il s'agit, pour le français, des auxiliaires de temps, des verbes copules (ce qui inclut la formation du passif), des verbes causatifs, des verbes de perception, et des verbes supports. Le phénomène des prédicats complexes est commun à l'ensemble des langues romanes. Il s'agit d'un système de complémentation différent de la complémentation ordinaire (voir *Paul pense lire ce livre* vs **Paul le pense lire*).

Dire que la grammaire est essentiellement syntaxique ne veut pas dire, bien sûr, que les autres aspects de la grammaire soient absents, simplement que ces types de problèmes ne constituent pas un objet d'étude en eux-mêmes, d'une manière générale. Il n'y a pas d'étude d'ensemble de la structuration du lexique, mais certaines études lexicales sont bien là, comme c'est d'ailleurs le cas généralement dans les grammaires. Il y a des classes de lexèmes (en général, des classes peu larges) qui sont étudiées, et dont parfois les membres donnent lieu à des études de détail : il en va ainsi des déterminants, des quantifieurs nominaux, des verbes têtes de prédicats complexes (dont nous venons de parler), des verbes modaux, des verbes et des adjectifs dits « à montée », dont le sujet n'est pas contraint, mais hérite ses propriétés du sujet attendu par le verbe infinitif complément (comme dans *Il peut pleuvoir* vs *Jean peut venir* etc.), des proformes, de certaines classes d'adverbes et de prépositions, des subordonnants et de certains connecteurs de discours. D'autre part, on étudie aussi certaines alternances verbales. On considère que les verbes actif et passif réalisent le même lexème, c'est-à-dire la même unité abstraite du lexique, caractérisée par un radical phonologique, une relation sémantique, une catégorie syntaxique, et une valence de base. C'est le verbe actif qui fournit la valence de base, qui peut être modifiée par une alternance (pour donner le passif, ou l'impersonnel).

Aucun chapitre n'est consacré à une question morphologique, mais tous les chapitres consacrés à l'étude d'une catégorie comportent une vue d'ensemble sur les classes morphologiques des mots relevant de cette catégorie. Nous nous

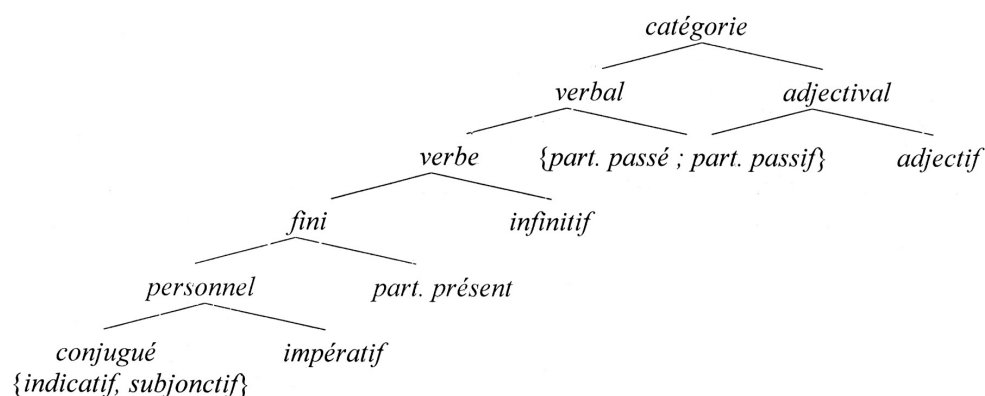
intéressons particulièrement à la morphologie lorsqu'elle entre en interaction avec la syntaxe. Dans notre approche, le mot est l'unité prise en compte par la syntaxe. Mais la distinction n'est pas toujours facile à faire entre une séquence syntaxique et un mot (ou un lexème) complexe, dont la construction fait appel à des procédés qui rejoignent ceux de la syntaxe, mais qui s'est figé pour donner une seule unité syntaxique. La question se pose par exemple pour certaines séquences de la forme 'nom1 de nom2' ou 'adjectif + nom'. On dispose de deux critères :

(i) est-ce que l'objet ainsi dénoté est un sous-type des objets dénotés par nom1 ? par exemple, on peut opposer *pomme de terre* et *pomme de Normandie* : la réponse est positive dans le second cas, mais pas dans le premier. D'une manière plus générale, on se demande donc si la sémantique est compositionnelle ; si elle ne l'est pas, comme dans le cas de *pomme de terre*, c'est une indication que l'on a affaire à un lexème, dans la construction duquel la syntaxe n'a pas à regarder. De même, une séquence comme *chambre noire* est un lexème : il n'est pas clair qu'aujourd'hui une chambre noire dénote un type de chambre, dans la mesure où *chambre* s'est spécialisé pour un type de pièce dans un appartement (celle où on dort), et s'il est vrai qu'une chambre noire n'a pas de fenêtres, elle n'est pas à proprement parler 'noire' (en fait, la lumière est rouge !).

(ii) quel est le degré de figement de la séquence ? Si la séquence est figée, si les éléments ne peuvent être modifiés par un ajout, ni prendre de complément, c'est une bonne indication qu'il s'agit d'un lexème. Ainsi, *pomme de terre* peut être modifié en tant qu'unité (*pomme de terre nouvelle* / *de Bretagne* / *à rissoler*), mais pas les éléments (**pomme ronde de terre*, **pomme de terre calcaire*). En revanche, on peut dire *pomme rouge de Normandie* / *pomme de Basse Normandie*.

Un autre endroit où la morphologie et la syntaxe entrent en interaction est la classification des formes verbales, qui se présente dans cette grammaire comme en (5).

(5)



À la catégorie « verbal » est associée la possibilité de prendre un complément direct (propriété syntaxique), et à la catégorie « verbe » celle d'être l'hôte d'un pronom clitique (propriété morpho-syntaxique), alors qu'à la catégorie « adjectival » est associée la possibilité de se fléchir en genre et nombre dans un système d'accord (propriété morpho-syntaxique). Les participes passé et passif constituent une catégorie mixte, qui hérite des deux types de comportement. Ainsi, d'une part, ils peuvent prendre un syntagme nominal objet (*Paul a lu un livre ; Il a été commandé des tonnes de blé*), et d'autre part, leur morphologie les rapproche des adjectifs. L'infinitif se distingue des formes finies par la position des adverbes négatifs *pas, plus, jamais* (propriété syntaxique) : alors que ces adverbes suivent les formes finies, ils précèdent l'infinitif (*Paul ne vient pas vs ne pas venir*). Le participe présent ne manifeste pas de variation de personnes (propriété morpho-syntaxique). Enfin, les formes indicatives et subjunctives se distinguent de l'impératif en ce qu'elles comportent des variations temporelles en plus des variations de personne (propriété morpho-sémantique).

Il est clair également que la sémantique et la pragmatique ont un rôle important dans la grammaire. De ce point de vue, l'ouvrage n'a pas la prétention d'être complet. Mais, en l'état actuel des connaissances, il est probable que cela s'avérerait une impossibilité⁶. Même dans la première partie, certains chapitres sont sémantiquement définis : il en va ainsi des chapitres sur la détermination et la quantification nominale, sur les proformes, sur la négation, et sur les questions de temps, aspects et modes. Pour le premier, il nous a paru plus judicieux de séparer les questions concernant les classes de nom et la construction du syntagme nominal, de la question sémantique posée par le syntagme nominal. Sous le premier aspect, on peut rapprocher les questions posées par le nom de celles qui sont posées par le verbe ou l'adjectif, même si, bien entendu, il y a des différences : on regarde comment l'expression tête (le nom, le verbe, l'adjectif, la préposition) est complété par des arguments ou modifié par des informations supplémentaires. Mais l'autre chapitre pose des questions radicalement différentes, puisqu'il s'agit de voir par quels moyens syntaxiques et lexicaux une expression linguistique dénote une ou plusieurs entités dans le domaine de référence, plutôt que des situations ou des propriétés comme le font les verbes et les adjectifs. Le chapitre des proformes est défini lui aussi sémantiquement : il s'agit d'un ensemble d'expressions qui ne sont pas autonomes mais renvoient soit à une expression dans le dialogue, soit à des entités présentes dans la situation de discours. Il y a plusieurs raisons pour lesquelles nous les avons réunies dans un même chapitre. D'abord, cela permet de mettre l'accent sur ce phénomène qui caractérise les langues naturelles. D'autre part, le phénomène est transcatégoriel : bien qu'il concerne essentiellement les expressions nominales, il ne s'y restreint pas. On trouve des

⁶ Pour des résultats concernant les grandes questions de sémantique pour le français, voir F. Corblin et H. de Swart (eds) 2004.

proformes pour des syntagmes prépositionnels (*où, auquel, en, y, lui*), pour des expressions prédicatives (le clitique *le*, pour les attributs prédicatifs), pour des verbes (*le faire, en faire autant, faire de même*), pour des phrases (*oui, non, si*). Enfin, même si on avait voulu grouper les proformes avec les catégories dont elles relèvent (ce que l'on a fait pour les prophrases), cela aurait posé un problème de maniabilité et de lisibilité. En effet, le français, comme les langues romanes dans leur ensemble, ont des proformes que nous appelons ici 'clitiques pronominaux', et dont le statut prête à discussion : est-ce que ce sont des mots existant en syntaxe et s'appuyant sur d'autres mots parce qu'ils sont déficients phonologiquement, ou bien est-ce que ce sont des affixes, c'est-à-dire des éléments du verbe ? quelle est la répartition entre ceux-ci et les formes pronominales fortes ? Cette discussion (morpho-syntaxique) n'a rien à voir avec les problèmes posés par le nom.

La seconde partie de la grammaire (de la phrase au discours) contient d'abord quatre chapitres consacrés aux types de phrase. Cette classification fait appel à la fois à la syntaxe, la sémantique et la pragmatique. Le premier étudie les phrases déclaratives, interrogatives, exclamatives et impératives. Or, cette typologie repose d'abord sur l'association entre des ensembles de formes et des contenus. Ainsi, le contenu d'une déclarative est une proposition, celui d'une interrogative est une question, celui d'une impérative est une visée, et celui d'une exclamative est une proposition contenant une expression de haut degré (voir, par exemple, Ginzburg, Sag 2000). En même temps, ces types de phrases ont également à voir avec de la pragmatique, et plus précisément, des actes de langage. Les phrases, sauf les impératives, peuvent être racines ou subordonnées. Lorsqu'elles sont racines, elles sont associées à un acte de langage. Or, il existe une association par défaut entre un type de phrase et un type d'acte de langage. Ainsi, les déclaratives sont assertées, les interrogatives sont questionnantes, les impératives donnent des ordres, et les exclamatives donnent des exclamations. Les deux chapitres suivants concernent des phrases ajouts : les relatives sont ou peuvent être des ajouts au syntagme nominal, alors que les circonstancielles sont ajouts à la phrase. La différence est accompagnée d'une différence sémantique : alors que les relatives dénotent une proposition à propos de l'antécédent, les circonstancielles établissent une relation entre des propositions ou des situations ; les différents types de circonstancielles sont eux-mêmes classés sémantiquement comme le veut la tradition (hypothétiques, causales, finales, concessives, temporelles), chacun, en fait, étant susceptibles de réalisations syntaxiques variées. Enfin, les phrases dont l'étude constitue le dernier de ces chapitres, les comparatives et les consécutives, sont différentes syntaxiquement et sémantiquement, puisqu'elles reposent sur des systèmes de corrélation.

Les chapitres qui suivent concernent également des questions d'interface. La coordination, qui est un phénomène transcatégoriel, est sans doute envisagée essentiellement d'un point de vue syntaxique, mais elle touche aussi aux questions d'interprétation, notamment au travers des problèmes des structures elliptiques.

L'ordre des mots est un endroit où se manifestent des contraintes de différents types, faisant intervenir non seulement des différences catégorielles, mais aussi des structururations discursives, et prosodiques. Un chapitre montre comment la phrase est traversée par des phénomènes qui s'expliquent par la structuration du discours ou du dialogue. Enfin, les deux derniers chapitres, sans doute plus originaux, articulent la syntaxe avec la prosodie d'un côté, et les codes de l'écrit, de l'autre.

2.2. Les données

Cette grammaire a l'ambition de ne pas être normative, et donc de décrire, autant que possible, des variétés et des usages qui peuvent différer de ceux du français standard. Nous nous efforçons donc d'intégrer des données caractéristiques de certaines régions, ou de français non standard, c'est-à-dire rejetées par une norme définie socialement et non linguistiquement. Nous nous efforçons aussi de traiter de certaines données de français parlé, qui peut parfois exploiter largement des structures peu présentes à l'écrit.

Des collègues spécialistes de la variation et des données orales, Françoise Gadet et Paul Cappeau, nous apportent leur contribution sur ces points. Françoise Gadet s'est chargée de rassembler les observations empiriques qui ont été faites, pour l'ensemble de la grammaire, et coordonne le travail d'experts pour les variations régionales (particulièrement en Belgique, au Québec et en Suisse).

Nous nous servons de deux sources de données : nous nous servons d'une part des jugements d'acceptabilité basés sur les intuitions de locuteurs, et d'autre part, de corpus, écrits ou oraux. La première manière d'obtenir des données est courante en linguistique, et reste indispensable pour tester les possibilités d'un système. En effet, pour asseoir une régularité en linguistique, il faut pouvoir faire la différence entre ce qui est grammatical (ce qui fait partie du système de la langue) et ce qui est agrammatical. Or, par définition, les corpus ne fournissent que des données attestées, et donc en principe possibles à l'intérieur de la langue. On peut imaginer qu'il sera possible, un jour, de déduire les agrammaticalités de l'inexistence des formes dans les corpus. Pour cela, il faut que se développent de très grands corpus, représentant des textes de genres variés, ainsi que des enregistrements dans des situations variées. Pour les genres de l'écrit, on a une idée assez claire de ce qu'ils peuvent être, mais il faut disposer d'une ontologie des situations pertinentes pour l'oral, qui ne semble pas encore être disponible. Il y a un effort dans ce sens, mais il faut encore du travail d'analyse et de recueil des données. En l'état actuel, les jugements d'acceptabilité restent indispensables. Cependant, il faut être conscient de ce qu'ils reflètent plutôt les données standard, et sont beaucoup moins fiables pour ce qui concerne les autres usages. On sait, en effet, depuis les études de Labov (1972) qu'il y a une différence sensible entre ce que les locuteurs pensent qu'ils disent, et ce qu'ils disent effectivement. D'autre

part, les corpus sont indispensables pour préciser des points sur lesquels les jugements sont incertains, et pour avoir accès à la variété des usages particulièrement les usages régionaux.

Pour assurer au mieux la prise en compte de la variation, nous avons donc établi une liste de corpus recommandés. Pour l'écrit, il s'agit d'abord de Frantext, qui est le corpus le plus important et que l'on peut interroger par genre, et par date. Pour que les données concernent le français contemporain, nous avons limité les textes à ceux qui sont après 1950. De plus, pour éviter de prendre en compte des jeux stylistiques qui peuvent brouiller les contraintes de langue, nous avons recommandé aux auteurs d'éviter les textes littéraires, sauf pour les phénomènes où ce type de texte est approprié (par exemple, les narrations sont le mieux susceptibles de donner des informations concernant les emplois des temps). A ce corpus s'ajoutent les corpus dirigés vers les variétés régionales, en particulier le corpus de l'Equipe du Trésor de la Langue française de l'université Laval au Québec, et plusieurs corpus de journaux. Pour les corpus oraux, nous avons recommandé le livre de Blanche-Benveniste *et al.* (2002), qui contient un choix de textes parlés transcrits, et celui de Cresti, Moneglia (2005), qui rassemble des corpus relevant de situations variées, dans plusieurs langues romanes. Pour la Belgique, il existe une banque de données orales (centre de recherches Valibel). Des corpus radiophoniques existent également en France et en Belgique.

Il faut cependant rester modestes. Retravailler l'ensemble des données pour l'ensemble de la grammaire à la lumière des corpus est une tâche en soi, et ce n'est pas l'objectif de ce travail. Les auteurs ne peuvent effectuer que des sondages, là où ils ont une idée de ce que les données ou les problèmes peuvent être. Nous n'entrerons dans le détail des données que sur les sujets qui sont suffisamment connus, comme l'apparition ou l'absence du préfixe verbal *ne* dans la négation (*Paul (ne) viendra pas*), ou les relatives dites non-standard : les relatives en *que*-pronom (*!une fille que tu la vois qu'une ou deux fois par semaine*, Deulofeu (1981), les relatives dans lesquelles la forme du subordonnant n'est pas en correspondance avec la forme attendue du constituant manquant (*%un renseignement que j'ai besoin, !Ces messieurs dont contrairement à ce que vous pouvez penser j'ai une grande admiration*, corpus d'Aix).

On voit la conception de la grammaire d'une langue qui sous-tend nos choix. La grammaire constitue bien un système d'ensemble, mais un système où il y a du jeu, des sous-systèmes en concurrence sur tel ou tel point. Cette conception est soutenue depuis longtemps par les historiens des langues, et par les socio-linguistes (particulièrement affirmée par W. Labov dans les années 70), mais largement admise aujourd'hui. On sait qu'il serait vain d'essayer de découper des petites grammaires homogènes qui représenteraient des variétés du français. En fait, ces

variétés n'existent pas en tant que telles : les locuteurs ont leur propre palette de variations, qui ne constitue pas un système, ni même un ensemble cohérent de sous-systèmes.

3. COHÉRENCE DES ANALYSES

Il n'y a pas de grammaire sans cohérence des analyses. Or, ce problème qui se pose à toute grammaire d'ensemble, se pose de manière bien plus aiguë pour ce projet. En effet, il réunit une cinquantaine d'auteurs, qui ont accepté de faire bénéficier la grammaire de leur compétence sur un sujet particulier. Une telle situation peut facilement aboutir à des incohérences. D'abord, à cause de la situation de la discipline : si les concepts de base comme les catégories et les fonctions sont assez bien partagés, leur définition n'est pas toujours suffisamment ferme pour être admise par tous, et la terminologie même n'a pas (encore) donné lieu à un consensus. Ce problème est aggravé par le fait que les auteurs appartiennent à des approches parfois notablement différentes, avec des habitudes terminologiques, et même des analyses différentes pour le même phénomène. Nous expliquons ici la façon dont nous cherchons à tenir l'objectif de la cohérence.

3.1. Organisation du travail

Deux choix ont été faits concernant l'organisation du travail, qui visent à minimiser les risques d'incohérence. D'une part, nous bénéficions de la collaboration active d'une quinzaine d'auteurs, qui ont bien voulu, outre leur travail comme auteurs, assurer la responsabilité d'un chapitre. Ces collaborateurs sont : Olivier Bonami, Paul Cappeau, Michel Charolles, Bernard Combettes, Véronique Dahlet, Elisabeth Delais, Marianne Desmets, Jacques Jayez, Brenda Laca, Ludo Mélis, Philip Miller, Michèle Noailly, Liliane Tasmowski, Jesse Tseng, Carl Vetters, Anne Zribi-Hertz. Cela devrait permettre d'assurer une première cohérence à l'intérieur de chaque chapitre. D'autre part, les auteurs et les responsables de chapitres ont rédigé des synopsis détaillés, qui ont été rassemblés et mis à la disposition de l'ensemble des auteurs.

Malgré ces précautions, nous ne pouvons éliminer la possibilité que certains textes au moins soient à reprendre sur tel ou tel point. De plus, étant donné les habitudes terminologiques parfois imprécises dans notre communauté, il semble difficile que les auteurs (y compris les directeurs de l'ouvrage!) soient absolument constants dans leur usage. Une ou plusieurs lectures attentives seront nécessaires, même lorsque la cohérence des analyses est assurée. Bien entendu, cette reprise se fera en concertation avec les auteurs. De plus, quelques linguistes extérieurs à l'entreprise ont accepté de relire tout ou partie de l'ouvrage, et de nous faire bénéficier de leurs critiques. Il s'agit de Claire Blanche-Benvéniste, de Jean-Claude Chevalier et de Bernard Fradin.

3.2. Choix du modèle grammatical

Comme nous le savons, il existe dans notre discipline des traditions variées, et les auteurs de cette grammaire n'échappent pas à cette dispersion. D'un autre côté, notre objectif est de faire en sorte que l'ouvrage soit largement lu : comme nous l'avons souligné dès le début, il n'est pas destiné aux seuls professionnels de la langue, ni même aux étudiants en langue française ; il devrait pouvoir être consulté par un large public de gens intéressés par les questions de langue. Il est donc exclu que la grammaire relève d'un formalisme particulier. Nous avons donc opté pour une grammaire syntagmatique simple, qui est une grammaire dite de surface : il n'y a pas de structure profonde, qui serait éventuellement différente de la structure apparente, il n'y a pas de catégorie syntaxique vide ; on ne parle pas non plus de cas, dans la mesure où les formes nominales ne portent pas de marque morphologique de cas en français moderne. En revanche, les structures proposées doivent pouvoir être justifiées par des ensembles de propriétés : autrement dit il s'agit bien d'une grammaire raisonnée, comme les travaux de linguistique de ces dernières années nous en ont donné l'exigence.

Cette grammaire utilise, pour la syntaxe, une liste de catégories et de fonctions. Toute expression doit pouvoir être analysée en ces termes. Les concepts ont été soigneusement définis par des propriétés, comme on l'a vu à la section précédente pour le verbe et le nom.

(6) Liste des catégories (lexicales) et des fonctions

CATÉGORIE		FONCTION	
Adjectif	<i>jaune, grand, illustre, national, supposé,</i>	Ajout	<i>Content de son sort, Paul souriait.</i> <i>Si tu viens, tu verras.</i> <i>Très bien !</i>
Adverbe	<i>facilement, à peine, tout à coup,</i>	Attribut	<i>Il est <u>médecin/ intelligent/ contre le projet/ très bien.</u></i> <i>Ça l'a rendu <u>heureux.</u></i> <i>Je te retrouve <u>rue Lepic.</u></i>
Complémenteur	<i>que, si, comme (ajout temps-cause), quand (ajout temps), parce que, bien que, lorsque,</i>	Complément oblique	
Conjonction de coordination	<i>et, ou, ni, mais, car, soit, %ainsi que, %comme</i>	Coordonné	<i>Paul <u>et</u> Marie, livres <u>et</u> dossiers.</i>
Déterminant	<i>définis (le, son, ce), indéfinis et quantifieurs (un, un de ces Nsg, un certain, certains, chaque, de ce/son, du, plusieurs, quel, quelque, zéro, je ne sais / on ne sait quel), terme de choix libre (n'importe quel)</i>	Extrait	<i><u>A Paul,</u> je fais entière confiance.</i> <i>Un problème <u>sur lequel</u> il faut encore travailler</i> <i><u>Combien</u> ça vaut ?</i> <i><u>Comme</u> il exagère !</i> <i><u>Plus</u> tu lis, <u>plus</u> tu comprends</i>

Interjection	<i>hélas, bravo, bof, bon,</i>	Marqueur	<i>commencer <u>à</u> lire, je crois <u>que</u> Paul vient. Paul regarde <u>le ciel</u>. Il est venu <u>des visiteurs</u>.</i>
Nom	<i>table, ouvre-boîtes, chambre noire, Paul, président, destruction, gentillesse,</i>	Objet direct	
Préposition	<i>à, de, avec, pour, sans, avant, après, dessus, sur,</i>	Périphérique	<i><u>Paul</u>, tu le connais ? La mienne, <u>de</u> voiture, marche bien. <u>Ce genre de feux</u>, vaut mieux appeler les pompiers. <u>Pour le théâtre</u>, on peut y aller. <u>cet/ un/ quel</u> étudiant, plusieurs étudiants ; <u>trois questions, semblables</u> aventures ; <u>quantité</u> de livres ; <u>une centaine</u> de livres.</i>
Pronom	<i>personnel (lui, moi, je, tu, nous, vous, eux, on), démonstratif (cela, ceci, ce), indéfini (quelqu'un, quelque chose), quantifieur (tout, rien, chacun, personne), relatif (qui, quoi, lequel), interrogatif (que, qui, quoi, lequel)</i>	Spécifieur	
verbe	<i>dormir, entrer, voir, manger, donner, être, penser, dire, sembler,</i>	Sujet	<i><u>Les enfants</u> sont là. <u>Que tu sois venu</u> me fait plaisir. Le <u>livre</u> de Paul, Paul <u>viendra</u>, <u>content</u> de son sort, très <u>bien</u>,</i>
		Tête	

Les catégories listées en (6) sont lexicales (elles sont associées à un mot) ; les catégories syntaxiques non-lexicales sont le syntagme et la phrase. Les expressions relevant des catégories lexicales peuvent être la tête (c'est une fonction) d'un syntagme. Elles sont alors soit complétées par les valents qu'elles attendent (sujet, objet direct, complément oblique, attribut, spécifieur), ou modifiées par un ajout. Nous avons illustré chaque concept de quelques exemples. Les expressions qui illustrent la fonction sont soulignées. Certains exemples illustrent la lexicalisation de séquences, comme les adverbes *à peine*, *tout à coup*, les compléments *parce que*, *bien que*, le nom *chambre noire*, le déterminant *n'importe quel*, le pronom *quelque chose*. Comme on voit, nous n'avons pas de catégorie « conjonction de subordination ». C'est que les termes ainsi désignés sont répartis entre les compléments et les séquences analysables en préposition + phrase avec un marqueur *que*. *Parce que* et *bien que* sont des séquences lexicalisées, formant un mot ; en effet, rien ne peut séparer les éléments (**parce, justement, il a été prévenu trop tard* vs *parce que, justement, il a été prévenu trop tard*). cela n'est pas vrai de la séquence *avant/après que*, qui se laisse analyser en une structure préposition tête du syntagme et phrase complément (*avant, justement, que tu ne t'éloignes*).

Nous sommes très attentifs à la distinction entre la catégorie et la fonction. Cela mérite d'être souligné particulièrement en ce qui concerne le phénomène de la détermination, où la terminologie donne lieu à confusion. Nous distinguons entre le phénomène sémantique de la détermination (qui permet à une expression nominale de dénoter des entités), la catégorie syntaxique de déterminant et la fonction de spécifieur. Les éléments en fonction de spécifieur assurent la détermination sémantique. Ainsi, un syntagme nominal comportant un spécifieur est capable de fonctionner de façon autonome, pour dénoter des entités, ce qui n'est pas le cas en général des noms sans déterminant (**J'ai lu livre vs J'ai lu trois livres, J'ai lu ce livre*). Mais ce n'est pas toujours le cas qu'un syntagme nominal demande un spécifieur, même pour être déterminé. Ainsi, un nom propre peut être intrinsèquement déterminé (*Paul viendra*), et une construction coordonnée peut assurer ce résultat : *Rangez ici livres et cahiers*. Par ailleurs, les déterminants se distinguent des autres expressions en fonction de spécifieur par deux propriétés corrélées : ils ne se combinent pas entre eux, et ils n'ont pas d'autre fonction que d'être spécifieurs. Un élément comme un adjectif (d'identité comme *semblable*, ou cardinal comme *trois* en (6)) peut se combiner avec un spécifieur : *les trois questions, une semblable erreur*, etc. Ce ne sont donc pas des déterminants, mais des adjectifs qui peuvent être en fonction de spécifieur.

Nous avons également défini des concepts relevant des autres aspects de la grammaire que la syntaxe, sans chercher à couvrir complètement le champ des possibilités. Le nombre des concepts morpho-syntaxiques et prosodiques est assez petit, car il s'agit de cerner ce qui va être utilisé dans l'ensemble de la grammaire. Les concepts sémantiques et pragmatiques sont en plus grand nombre, ne visent pas à l'exhaustivité ni à guider l'analyse, mais plutôt à assurer la cohérence terminologique.

3.3. Choix d'analyses

Certains choix d'analyse ont été imposés pour l'ensemble de la grammaire, lorsque les travaux linguistiques laissaient plusieurs possibilités ouvertes, et que le choix en un endroit avait des incidences sur le reste du travail. L'origine de ces choix est diverse : il a pu venir d'un auteur, d'un groupe d'auteurs après discussion, ou des directeurs de l'ouvrage. Nous commentons ici la définition de la phrase, l'analyse des formes dites clitiques pronominaux, et le traitement des ellipses.

La phrase a reçu une définition fonctionnelle. Est analysée comme phrase une expression dont la tête est un prédicat qui est saturé pour son sujet. Par prédicat, nous entendons le contenu d'un item lexical, et, par extension, l'item lexical lui-même. Cette définition s'oppose donc à la définition de la phrase comme syntagme verbal : il est vrai que le verbe est en général la tête de la phrase, mais seulement parce que la plupart des phrases sont verbales. Nous admettons donc l'existence de

phrases averbales, c'est-à-dire dont la tête est une catégorie autre qu'un verbe : un adjectif (*admirable*) en (7a), une préposition (*à*) en (7b), et un nom (*professeur*) en (7c).

- (7) a. *Admirable, ce comportement.*
 b. *Tous à la Bastille !*
 c. *Paul une fois professeur, toute la famille se détendit.*

Est aussi une phrase (verbale) une expression dont la tête est un impératif. On pourrait penser que les impératifs ne sont pas saturés pour le sujet, mais, en fait, ils le sont car ils ne se combinent jamais avec un sujet en français (le syntagme nominal en (8a) est un vocatif, car il n'y a pas accord entre les personnes du verbe et de l'expression nominale), et le sujet est toujours interprété comme une personne de discours, dépendant étroitement de la forme verbale : les impératifs en français sont donc des formes « pro-drop », dont l'argument « sujet » lui est fourni par la forme même. En revanche, une expression comme celles de (8b) ne sont pas considérées comme des phrases. En effet, l'adjectif *debout* en (8b) attend un sujet qui ne lui est fourni que par la situation non linguistique ; le syntagme prépositionnel *au diable* est ambigu. Ou bien, il attend un sujet (il est à peu près synonyme de *Allez au diable !*), et donc est un prédicat non saturé. Ou bien il s'agit d'une interjection, qui n'est pas saturée pour son sujet, car la question du sujet ne fait pas sens pour ce type d'expression.

- (8) a. *Allez jouer dehors, les enfants !*
 b. *Debout ! Au diable !*

La phrase est donc distincte de l'énoncé : les énoncés sont des expressions qui sont associées à un acte de langage. Ces expressions peuvent être des phrases, si ce sont des phrases racines ; mais elles peuvent également être des expressions non-phrastiques. Ainsi, les expressions de (8b) constituent des énoncés, associés à un ordre ou une exclamation, bien qu'elles ne soient pas syntaxiquement des phrases. Les phrases averbales de (7a, b) sont associées à un acte de langage car ce sont des phrases racines, même si l'acte associé à ce type de phrase doit souvent être spécifié par la prosodie et le contexte.

Un autre choix qui a des conséquences à plusieurs endroits de la grammaire, et singulièrement dans le chapitre concernant la phrase, est le statut des pronoms dits « clitiques » (Kayne 1975). Traditionnellement, on oppose la série des pronoms personnels forts aux pronoms faibles. Alors que les premiers (*lui, elle, eux, nous1, vous1*) se comportent comme des syntagmes nominaux ordinaires du point de vue de leur distribution, les seconds (*il, elle, on, le, la, les, lui, leur, en, y, nous2, vous2*) ont des propriétés spécifiques, phonologiques, syntaxiques et morpho-syntaxiques, en français comme dans l'ensemble des langues romanes. La

question qui se pose est la suivante : est-ce que ces formes sont à analyser comme des mots syntaxiques, ou bien est-ce que ce sont des réalisations affixales de certaines propriétés de valence des verbes ? Dans le second cas, par exemple, une forme comme *le voit* en (1f) est un mot-verbe, saturé pour son complément d'objet (il n'attend pas d'objet, alors que le lexème LIRE dit que ce verbe attend un sujet et un objet). Miller (1992) a donné des arguments pour dire que l'analyse comme affixe est la plus à même de rendre compte de cet ensemble de propriétés spécifiques. Cette analyse n'est pas admise par l'ensemble de la communauté des linguistes. L'auteur de la section (P. Miller) a tranché de la manière suivante : les clitiques en fonction sujet sont considérés comme des mots (et non comme des affixes), alors que les autres sont des affixes. La raison est la portée sur la coordination : les clitiques compléments n'ont pas portée sur une coordination de verbes, alors que les clitiques sujets peuvent avoir portée sur une coordination (de verbes et de syntagmes verbaux, voir respectivement (9c) et (9d)).

- (9) a. *Paul achètera et lira les romans de Balzac.*
 b. **Paul les achètera et lira.*
 c. *Il [[_vachètera] [_{vet} lira]] les romans de Balzac.*
 d. *Il [[_{sv}achètera ces textes] [_{svet} les lira]].*

En conséquence, nous n'avons pas besoin de la notion de « clitiques » pour ces pronoms (autrement qu'en un sens phonologique), puisque l'on a d'un côté des mots, et de l'autre, des affixes. Ce choix a des répercussions pour la structure de la phrase. En effet, si l'analyse de Miller (1992) avait été suivie jusqu'au bout, cela voulait dire que le français est une langue « pro-drop », autrement dit, que les phrases dont le sujet est un pronom personnel sont dépourvues de sujet syntaxique, et qu'on ne doit pas dire, comme on l'admet généralement, que les phrases avec un verbe conjugué demandent un sujet en français. L'analyse proposée est donc, à cet égard, conservatrice.

Nous donnons un dernier exemple d'analyse adoptée à travers la grammaire. Il s'agit du traitement des ellipses. Par ellipse, nous entendons informellement les expressions dans lesquelles la combinatoire des catégories ne suffit pas à nous donner une information sémantique complète (nous laissons de côté, bien sûr, le recours au contexte pour interpréter les expressions déictiques). Nous distinguons, comme c'est l'habitude, entre l'ellipse extra-grammaticale, lorsque l'information est récupérée du contexte non-linguistique, et l'ellipse grammaticale. Dans le premier cas, la syntaxe n'a rien à dire, et l'information à suppléer concerne l'interface avec la sémantique ou la pragmatique. Reprenons l'énoncé de (8b) *Debout !* Nous avons simplement un syntagme adjectival qui attend un sujet. La dénotation de ce sujet est à trouver dans la situation de discours, puisqu'il s'agit du ou des interlocuteurs. De manière analogue, la combinatoire bien connue de *commencer* avec un complément nominal ne dénotant pas un événement, comme

dans *commencer un livre*, est syntaxiquement complète : *commencer* prend un syntagme nominal objet, comme il le fait dans *commencer la moisson*. Nous traitons cette ellipse dans le lexique : le lexème ‘commencer’ prend comme objet un syntagme nominal qui dénote soit un événement, soit un individu. Dans le premier cas, l’interprétation peut se faire directement, dans le second, la description même du lexème dit qu’il faut suppléer un prédicat verbal, à spécifier selon les propriétés du contexte d’occurrence (*lire, écrire, copier, relier* etc.). Ce qui intéresse la syntaxe, c’est l’ellipse grammaticale, illustrée en (10a, b), dans laquelle l’information doit être récupérée du contexte linguistique :

- (10) a. *Achète les moins chers.*
 b. *Qui vient ?*
 – *Paul.*

En (10a), nous avons un syntagme nominal sans nom ; dans (10b), la réponse est représentée par un syntagme nominal, alors que l’interprétation est celle d’une proposition, comme cela aurait été le cas si l’on avait eu une phrase complète. Ce qui est commun aux deux cas est l’absence de la tête syntaxique : le nom ou le verbe. Cette propriété caractérise les ellipses grammaticales. Dans la mesure où nous l’avons, dit, nous choisissons une grammaire de surface, sans catégorie vide, nous adoptons une représentation qui utilise le concept de ‘fragment’ (voir, par exemple, Ginzburg et Sag 2000) : nous acceptons des catégories syntaxiques sans tête. Les deux exemples de (10) sont représentés comme en (11) :

(11)

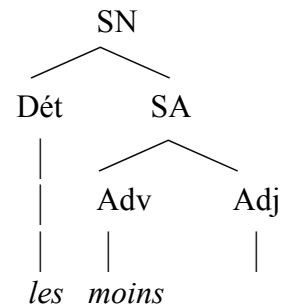
a.
Phrase

|
N

|
Paul

chers

b.



REFERENCES

Blanche-Benveniste, C., C. Rouget, F. Sabio, 2002, *Choix de textes de français parlé, 36 extraits*, Paris, Champion.

- Bosque I., V. Demonte (eds), 1999, *Gramatica Descriptiva de la Lengua Española*, 3 vols, Madrid, Real Academia Española.
- Chomsky, N., 1970, « Remarks on Nominlaization », in : R. Jacobs, P. Rosenbaum (éds), *Readings in English Transformational Grammar*, Waltham, Mass., Ginn, 184–221. (trad. française dans N. Chomsky, *Questions de sémantique*, Paris, le Seuil, 73–131).
- Corblin, F., H. de Swart (éds), 2004, *Handbook of French Semantics*, Stanford, CSLI.
- Cresti, E., M. Moneglia (éds), 2005, C-ORAL-ROM, *Integrated Reference Corpora for Spoken Romance Languages*, Amsterdam, J. Benjamins.
- Damourette, J., E. Pichon, 1911–1946, *Des Mots à la Pensée ; essai de grammaire de la langue française*, Paris, Editions Vrin.
- Deulofeu, J., 1981, *Perspective linguistique et sociolinguistique dans l'étude des relatives en français. Recherches sur le français parlé* 3, 135–193.
- GALR 2005, *Gramatica Limbii Române*, 2 vols, Bucarest, Editura Academiei Române.
- Ginzburg, J., I. A. Sag., 2000, *Interrogative Investigations*, Stanford, CSLI Publications.
- Grevisse, M., A. Goosse, 1993–1997, *Le Bon Usage*, Paris et Louvain-la-Neuve, DeBoeck-Duculot (13e édition).
- Huddleston, R., G. Pullum (eds), 2002, *The Cambridge Grammar of the English Language*, 1 vol., Cambridge, Cambridge University Press.
- Kayne, R. S., 1975, *French Syntax, the transformational cycle*, Cambridge, Mass., MIT Press (traduction française, *Syntaxe du français*, Paris, le Seuil, 1977).
- Labov, W., 1972, « The Study of Language in its social context », in *Sociolinguistic Patterns*, Philadelphie, the University of Pennsylvania Press, 183–259.
- Miller, P., 1992, *Clitics and Constituents in Phrase Structure Grammar*, New York, Garland.
- Pollard, C., I. A. Sag, 1994, *Head-driven Phrase Structure Grammar*, Chicago et Stanford, University of Chicago Press et CSLI.
- Renzi, L., G.-P. Salvi, A. Cardinaletti (eds), 1989–1991, *Grande Grammatica italiana di consultazione*, 3 vols., Bologne, il Mulino.
- Riegel, M., J.-C. Pellat, R. Rioul, 1994, *Grammaire Méthodique du français*, Paris, PUF.
- Ross, J. R., 1972, « The category Squish : Endstation Hauptwort », in : P.M. Peranteau, J. N. Levi, G. C. Phares (éds), *Proceedings of the Eighth Meeting of the CLS*, 316–328.
- Solà, J. et al. 2002, *Gramàtica del Català contemporani*, 2 vols, Barcelone, Editorial Empúries.
- Wasow, T., 1977, « Transformations and the Lexicon », in : P. Culicover, T. Wasow, A. Akmajian (éds), *Formal Syntax*, New York, Academic Press, 327–360.