

La pauvreté des stimuli : quels sont les faits? Trente ans de syntaxe

Denis Bouchard

Volume 30, numéro 1, 2001

1971-2001 : Trente ans de linguistique

URI : <https://id.erudit.org/iderudit/000513ar>

DOI : <https://doi.org/10.7202/000513ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Université du Québec à Montréal

ISSN

0710-0167 (imprimé)

1705-4591 (numérique)

[Découvrir la revue](#)

Citer cet article

Bouchard, D. (2001). La pauvreté des stimuli : quels sont les faits? Trente ans de syntaxe. *Revue québécoise de linguistique*, 30(1), 43–62.
<https://doi.org/10.7202/000513ar>

Résumé de l'article

Il y a un argument qui est demeuré stable au cours des ans en grammaire générative : l'être humain doit être pourvu d'une faculté de langage avec des conditions qui lui sont très spécifiques, puisque certains universaux sont appris sur la base de données très pauvres. Mais cette pauvreté des stimuli est un artefact découlant de la formulation des conditions spécifiques proposées, qui ne permettent pas un lien avec certains stimuli qui sont tout à fait accessibles. Une théorie très parcimonieuse rend compte des relations QU à distance, et tous les stimuli nécessaires pour apprendre cette théorie sont présents. Cette théorie est basée sur des propriétés logiquement antérieures au langage qui touchent à la substance plutôt qu'au formel. La grammaire générative, basée sur le formel, a occasionnellement été amenée à faire appel au substantiel pour arriver à une explication. En prenant systématiquement le parti d'un innéisme adaptatif plutôt que spécifiste, on arrive à des explications de façon plus soutenue.

LA PAUVRETÉ DES STIMULI : QUELS SONT LES FAITS? TRENTE ANS DE SYNTAXE

Denis Bouchard
Université du Québec à Montréal

à Nicolas Ruwet

1. Introduction

La théorie linguistique a beaucoup changé depuis trente ans. Si on compare des articles publiés en 1971 avec ce qu'on lit aujourd'hui, on retrouve très peu de choses qui sont présentées de la même façon. On peut y voir une évolution de la théorie. Une autre façon de mesurer le progrès fait par la théorie, c'est de regarder un des rares éléments qui n'a pas été remis en question. En essayant de comprendre pourquoi cet élément n'a pas changé, et en comparant comment l'ancienne et la nouvelle théorie traitent cet élément, on peut prendre la vraie mesure des changements.

Un élément qui est demeuré stable au cours des ans en grammaire générative, c'est l'argument que l'être humain doit être pourvu d'une faculté de langage avec des conditions qui lui sont très spécifiques, puisque certains universaux sont appris sur la base de données très pauvres, voire même inexistantes dans certains cas : ils nécessitent des données négatives, donc qui ne sont pas accessibles à l'enfant. Comme les principes en question sont très différents de ce qu'on retrouve en biologie et dans les sciences traitant du cerveau (Chomsky 1995, 2000), ils ne pourraient résulter de certains types d'expérience non linguistique ni être des principes généraux, non différenciés de l'esprit (Obenauer 1990). Comme de plus ces principes ne peuvent être inférés de données inaccessibles, ils doivent être disponibles avant toute expérience, donc faire partie du patrimoine génétique.

Je vais montrer que la pauvreté des stimuli n'est pas un fait, mais un artefact découlant de la formulation des conditions spécifiques proposées, qui ne permettent pas un lien avec certains stimuli. Ces stimuli sont tout à fait accessibles, et permettent facilement l'acquisition de conditions qui rendent compte des données. La grammaire générative repose sur une fausse appréhension des données, avec comme conséquence l'élaboration d'une grammaire ne tenant pas suffisamment compte de la substance du langage. C'est la cause principale du fait que la grammaire générative est restée prise dans une ornière descriptiviste, dans des «engineering solutions», de l'aveu même de Chomsky 2000. S'il y a eu peu d'avancement dans l'explication, il y a toutefois eu évolution : le modèle est maintenant poussé à chercher des explications dans des propriétés externes à la grammaire, qui lui sont logiquement antérieures. Cette reconnaissance plus nette que la linguistique n'est pas une *scientia infima*, sans prédécesseur logique, ouvre la voie à une véritable théorie explicative.

2. Pauvreté des stimuli et Grammaire universelle

Je ne peux pas discuter ici de tous les cas où on a posé que des données négatives étaient requises pour révéler un principe. Je prends donc l'exemple typique d'un tel principe pour illustrer mon propos, soit le Principe des catégories vides (PCV) en (1), qui rend compte du paradigme en (2) et (3) tirés de Obenauer 1990, qui présente l'argument de façon claire et concise¹.

- (1) Une catégorie vide doit être gouvernée par une tête lexicale qui lui assigne un rôle thématique ou par un antécédent.
- (2) a. Je sais à qui je pourrai offrir ce livre [t].
b. Je sais à qui ce livre pourrait plaire [t].
- (3) a. Voilà la fille à qui_i je sais lequel_k je pourrai offrir [t_k] [t_i].
b. *Voilà le livre que_k je sais à qui_i [t_k] pourrait plaire [t_i].

Sur la base de telles données, on fait l'argumentation suivante. En (2a) et (2b), les «trous» représentés par les traces [t] satisfont le PCV, de même que les deux traces en (3a). En (3b), toutefois, la trace [t_k] en position sujet n'obéit pas au principe, d'où l'agrammaticalité de la phrase. Or, l'existence du PCV n'est révélée que par des données négatives comme en (3b). De plus, le principe

¹ L'argumentation qui suit vaut *mutatis mutandis* pour d'autres principes de localité comme la Sous-jacence, et si le PCV est dérivé d'autres principes spécifiques au langage, comme la condition d'Impénétrabilité de la phase de Chomsky 2000.

concerne un élément vide de contenu phonologique, donc absent des données linguistiques primaires. Le principe ne peut donc être inféré du matériel linguistique.

On prend aussi pour acquis que les données primaires indiquent à l'enfant de façon «incontestable» que les syntagmes QU peuvent être déplacés en français, et même possiblement très loin comme en (4).

(4) Je sais à qui tu penses que Marie croit que Paul donnera le livre [t].

Toutefois, les données ne renseignent pas l'apprenant sur les limites de ces déplacements. Comme on ne peut conclure qu'une construction est exclue par la grammaire de par sa simple absence des données, on pourrait s'attendre à ce qu'un enfant «surgénéralise» et fasse l'erreur de considérer une phrase telle que (3b) comme étant bien formée. Or, les apprenants ne font pas ce genre d'erreur. On en déduit que «certains éléments de la grammaire développée doivent être déterminés par autre chose que les données rencontrées» (Obenauer 1990 : 79).

3. Stabilité de l'idée

L'idée d'une faculté de langage avec des conditions très particulières, due à l'absence de stimuli dans certains cas, est un élément central de la grammaire générative depuis le début. Par exemple, *Aspects of the Theory of Syntax* (Chomsky 1965) est l'ouvrage le plus marquant des deux premières décennies de la grammaire générative. Le passage suivant pourrait avoir été écrit aujourd'hui sans en changer un mot :

A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the striking uniformity of the resulting grammars, and their independence of intelligence, motivation, and emotional state, over wide ranges of variation, leave little hope that much of the structure of language can be learned by an organism initially uninformed as to its general character.

Chomsky 1965 : 58

Les données y sont déjà présentées comme étant dégénérées, pauvres, et les grammaires acquises comme ne pouvant s'expliquer ni sur la base de l'intelligence générale, ni sur la base de facteurs psychologiques comme la motivation ou l'émotion (des allusions à peine voilées à Piaget d'une part, et aux behavioristes d'autre part). Cela ne laisserait que des spécifications innées comme explication plausible des faits.

L'idée s'est maintenue tout au long des années florissantes de la grammaire générative. Ainsi, Pinker 1994, un bestseller à vocation de vulgarisation qui reflète très bien l'esprit qui prévalait dans les années quatre-vingt, exprime la stratégie générale à utiliser quand il y a absence de stimuli :

... any no-feedback situation presents a difficult challenge to the design of a learning system [...] How is the child designed to cope with the problem? A good start would be to build in the basic organization of grammar, so the child would try out only the kinds of generalizations that are possible in the world's languages.
Pinker 1994 : 282

Et l'idée continue d'être présente aujourd'hui. Ainsi, Chomsky 2000 propose une condition de localité très spécifique à la faculté du langage pour rendre compte des effets de localité de déplacement qui ne peuvent être déduits des données primaires. Il définit un domaine local, la «phase», qui comprend les éléments en mémoire active et donc accessibles à un point donné de la dérivation : cette phase correspond à la contrepartie syntaxique d'une proposition. Comme l'entièreté de la phase ne saurait être inerte (sinon rien ne pourrait en sortir par déplacement) et que son analyse requiert des déplacements pour expliquer de nombreux faits de relations à distance, mais qu'en même temps il y a des limitations sur les déplacements dont il faut rendre compte, Chomsky propose la condition de localité d'Impénétrabilité de la phase (5) :

- (5) Dans une phase φ avec une tête H, le domaine de H n'est pas accessible aux opérations hors de φ , mais seulement H et sa bordure².

Des opérations peuvent donc relier un élément dans un domaine φ à un élément hors de ce domaine, comme la trace [t] qui est reliée au syntagme à *qui* en (2a), mais ces opérations n'ont accès qu'à certains éléments de la phase. L'Impénétrabilité de la phase ne fait que redire ce qu'on exprimait en termes de Sous-jacence et de Cyclicité il y a trente ans.

On le voit donc, les générativistes ont adhéré sans faillir à l'idée d'un innéisme très spécifique au langage. Cette allégeance a parfois des effets pervers. Parce que l'idée est perçue comme étant incontournable, tout argument concluant à l'innéisme est vu de façon très favorable. Par conséquent, chaque fois qu'un phénomène linguistique semble de prime abord ne pas avoir d'explication simple sur la base de propriétés externes à la langue, comme l'intelligence générale ou des facteurs psychologiques, on élabore une contrainte ou un principe incorporés au patrimoine génétique et spécifiques au langage. L'absence d'explication devient une preuve de plus de la spécificité de la faculté de langage.

² Chomsky définit la bordure («edge») comme étant une structure d'un ou plusieurs Spécificateurs.

Il y a là risque de dérapage, qu'on en vienne à souhaiter ne pas trouver de solution basée sur des propriétés externes, ou du moins qu'on ne cherche pas très fort de ce côté, avec pour conséquence que les solutions de ce type semblent donc improbables.

Non seulement ceci risque d'affaiblir la façon de présenter ses arguments, parce qu'on ne les soumet plus à un certain type de contre-arguments (ne serait-ce que dans l'intimité de sa pensée), mais cela risque aussi de faire rater des explications fortes et élégantes qu'on a pourtant à notre portée immédiate.

4. Fait versus artefact

On nous présente l'absence de stimuli comme un fait incontestable dont il faut rendre compte. Or le problème de l'absence de stimuli n'est qu'apparent et vient de la façon de voir les données dans l'approche formaliste. Les données sont présentées d'une manière qui force la conclusion. Présentées autrement, la conclusion est très différente. Ce n'est pas la pauvreté des stimuli qui nécessite des conditions spécifiques à la faculté de langage, mais les conditions spécifiques proposées qui ne permettent pas un lien avec certains stimuli accessibles. Donc en fait, il n'y a nul besoin de conditions spécifiques à la faculté de langage pour rendre compte du problème puisqu'il n'existe pas.

On nous présente les données de la façon suivante. Premièrement, il serait indisputable que des données linguistiques primaires indiquent à un enfant exposé au français qu'un syntagme *QU* peut être déplacé très loin de sa position de base. Deuxièmement, ces données ne fournissent aucune information quant aux limites de cette possibilité. Troisièmement, l'enfant a pourtant une connaissance inconsciente précise à propos des phrases qui violent les conditions de localité.

Dans cette approche, classique depuis Ross 1967, les déplacements illimités constituent le cas non marqué, et des conditions sont ajoutées pour rendre compte des cas où les dépendances sont restreintes à des domaines fermés. Comme ces domaines fermés ne sont révélés que dans des phrases agrammaticales, ce qui verrouille ces domaines ne peut être inféré des données primaires. On en conclut que certains éléments de la Grammaire sont déterminés par autre chose que les données rencontrées, et donc qu'ils sont probablement innés, et spécifiques au langage. Bref, les principes derrière la «fermeture» de domaines doivent être innés.

Mais il y a une autre façon de voir les données qui permet d'inférer directement à la fois ce qui verrouille certains domaines et ce qui en ouvre d'autres,

et ceci sans faire appel à autre chose que les données rencontrées. En fait, il suffit d'élaborer l'analyse la plus parcimonieuse possible pour arriver à ce résultat.

4.1. Une analyse parcimonieuse

Le langage est une relation entre des éléments de sens et des éléments de forme perceptible. Cette observation, qui remonte aux toutes premières études sur le langage, est universellement reconnue. Elle est à la base de la théorie du signe de Saussure, qui veut qu'un élément de sens soit associé arbitrairement à un élément de forme. La nécessité d'avoir une association avec une forme perceptible découle du simple fait que le langage serait inutilisable si ce n'était pas le cas. L'arbitraire de l'association vient du fait que les formes doivent provenir de notre système articulatoire-perceptuel, et que les sons produits par nos articulateurs phonatoires sont tels qu'ils n'entretiennent pas de relation signifiante ou iconique avec les sens exprimés.

Toutefois, il n'y a pas que les unités de sens qui doivent être associées à une forme perceptible; les relations de sens doivent aussi être transmises de façon perceptible : la langue doit indiquer dans une forme perceptible qu'un élément porteur de sens entre en relation avec un autre. Dans une phrase comme (4), la relation sémantique qui nous concerne, c'est celle entre *à qui* et *donnera*. Il faut déterminer exactement quelle forme perceptible correspond à cette relation de sens et comment la grammaire génère le tout.

En français, quand le syntagme n'est pas QU, la relation sémantique qu'un objet indirect entretient avec un verbe s'exprime en prenant les formes perceptibles correspondant à ces deux éléments et en les juxtaposant dans un ordre donné comme en (6).

(6) Paul le donne à Pauline.

La Juxtaposition est un moyen parmi les quatre que permet la physiologie articulatoire des humains pour exprimer oralement qu'une relation sémantique existe entre deux éléments³. La Juxtaposition est l'assignation d'un rôle fonctionnel

3 Les trois autres moyens sont : 1° la Surimposition, comme lorsqu'un élément B est une modulation surimposée sur A pour exprimer affirmation, question, exclamation en français, ou des sens lexicaux et des fonctions grammaticales dans une langue à tons; 2° le Marquage du dépendant, comme le marquage casuel d'un argument; 3° le Marquage de la tête, comme le marquage du prédicat dans les langues polysynthétiques.

Notre constitution physiologique impose ces choix paramétriques à la faculté de langage. L'arbitraire du signe et les grandes classes typologiques qui en découlent ont donc une origine essentiellement physiologique. Comme ces propriétés physiologiques sont très saillantes, il suffit de très peu d'expérience pour établir les valeurs paramétriques, pour déterminer laquelle des quatre possibilités est en jeu dans un cas particulier. Voir Bouchard (à paraître) pour plus de discussion.

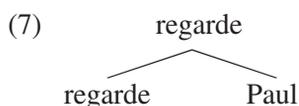
à la linéarisation temporelle des mots et morphèmes. On exprime habituellement ce rôle fonctionnel de la Juxtaposition par une projection syntagmatique. La structure syntagmatique n'est donc pas une primitive de la théorie, mais est dérivée du fait que deux mots ne peuvent être simultanés, qu'ils apparaissent dans un ordre temporel irréversible, asymétrique. Cette idée est un fondement de l'approche de Kayne 1994 basée sur son «Linear Correspondence Axiom». L'idée joue déjà un rôle important chez Tesnière 1959, inspiré par Saussure, comme on le voit dans le passage suivant :

*Le principe fondamental de la transformation de l'ordre structural en ordre linéaire est de transporter les **connexions** de l'ordre structural en **séquences** de l'ordre linéaire, de façon que les éléments qui sont en connexion dans l'ordre structural se trouvent en voisinage immédiat sur la chaîne parlée.*

Tesnière 1959 : 20

La linéarisation temporelle découle de propriétés physiologiques, du fait qu'on ne peut articuler deux mots en même temps. L'explication de ce fait relève donc de sciences qui sont logiquement antérieures à la théorie linguistique, comme la physiologie et la physique de l'articulation phonatoire, l'acoustique, etc. Il est donc des plus parcimonieux de fonder la théorie sur des notions d'ordre et de structure syntagmatique, puisque ces notions sont des axiomes de coût nul, étant logiquement antérieurs.

Pour rester à ce niveau de parcimonie, il faut veiller à ce que non seulement la mise en relation soit minimale, mais que les éléments mis en relation soient aussi restreints au strict nécessaire. En plus de la relation de Juxtaposition, le minimum requis dans la structure syntagmatique, ce sont des éléments lexicaux à mettre en relation. L'approche idéale s'en tiendrait donc à ces éléments et n'inclurait pas dans le vocabulaire de la théorie des étiquettes de niveaux comme X' ou X'' , puisque ces éléments sont redondants : il est suffisant de dire que lorsque deux éléments sont mis en relation par Juxtaposition, l'un ou l'autre étiquette la structure résultante⁴. Ainsi, un syntagme verbal comme *regarde Paul* reçoit la structure suivante :



4 Voir Chomsky 1994 : 11. Bouchard 1982 : 43-44, note 6, fait déjà allusion à cette idée : «labelling of syntactic nodes is done by lexical insertion.» Quant aux différences de niveau, elles encodent le fait que des constituants établissent des relations différentes avec la tête. Mais ces distinctions sont complètement inférables de l'ordre et du contenu sémantique des syntagmes. Il en va de même pour la distinction entre complément et modifieur. Voir Bouchard (à paraître).

Le plus simple quant à la projection, c'est de ne rien dire, de n'ajouter aucune condition à la théorie restreignant la projection à des traits particuliers de l'élément lexical. Tout projette donc : en (7), le *regarde* supérieur n'attribue pas seulement sa nature catégorielle au syntagme, mais tout son contenu sémantico-syntaxique.

Cette approche minimale de la structure syntagmatique a deux caractéristiques qui expliquent directement le phénomène des relations à distance et leurs contraintes illustrées en (2)-(4). Premièrement, la Juxtaposition définit une localité très stricte comme cas de base, essentiellement la relation d'adjacence⁵. Deuxièmement, l'étiquetage par les éléments lexicaux rend toutes les propriétés de la tête accessibles au niveau du syntagme. Ceci a pour effet d'étendre le domaine de la tête de façon très précise. Par exemple, dans le schéma en (8), la tête C prend le constituant [A B] comme dépendant.

(8) C [_A A B]

Comme A étiquette le constituant [A B] et que A projette tous ses traits au nœud adjacent de C, A est un nœud adjacent «étendu» de C, en ce sens que les spécifications de A sont accessibles à C. Le domaine de A est donc étendu, mais de façon très précise : ce sont les spécifications lexicales de la tête qui sont accessibles au niveau du syntagme, donc les propriétés permanentes comme les spécifications de sélection. Ainsi, reprenons notre exemple de relation «à distance».

(9) Je sais à qui tu penses que Marie croit que Paul donnera le livre.

Ici, le syntagme à *qui* est en relation avec son nœud adjacent, le syntagme *tu penses que Marie croit que Paul donnera le livre*. Supposons que la tête de ce syntagme est *penses*⁶. Puisque *penses* projette et étiquette le syntagme adjacent de à *qui*, à *qui* a accès aux spécifications de *penses*; en particulier, les spécifications de sélection de *penses* sont accessibles pour à *qui*. Le verbe *penses* a à son tour un complément et les spécifications de la tête du complément lui sont accessibles. Et ainsi de suite jusqu'au constituant qui a *donnera* à sa tête : par transitivité de toutes ces relations tête-complément, les spécifications de sélection de *donnera* sont accessibles pour à *qui*. Le syntagme à *qui*, de sa position en tête de phrase, peut donc établir directement un lien syntaxique avec le verbe avec lequel il entretient une relation sémantique.

5 J'utiliserai donc le terme «nœud adjacent» pour traduire l'expression anglaise «sister-node», par analogie avec «angles adjacents : angles ayant même sommet, un côté commun, et situés de part et d'autre de ce côté» (Dictionnaire Larousse).

6 Si on croit qu'il y a une projection avec un COMP et une autre avec un INFL, ceci ne fait qu'ajouter des étapes au raisonnement, mais n'en change pas la teneur essentielle. Il faut simplement alors prendre en compte la relation de chacune de ces têtes avec son complément.

Remarquez que *à qui* établit son lien syntaxique avec *donnera* sans qu'il y ait de trace en jeu, ni comme nœud adjacent de *donnera* ni dans aucune position intermédiaire entre *à qui* et *donnera*⁷. Le strict minimum est donc suffisant : ce n'est pas avec une position que *à qui* doit établir une relation, mais avec le verbe *donnera*. Les traces ne jouent donc aucun rôle dans cette mise en relation. L'emploi standard des traces cherchait à exprimer que la relation entre le syntagme QU et son verbe s'établit par la fusion du syntagme QU avec ce verbe à un certain niveau de représentation autre que celui de la surface. Mais en prenant en compte l'effet de projection des têtes, le domaine de ce avec quoi le syntagme *à qui* se fusionne en (9) est étendu de telle sorte qu'il comprend le verbe *donnera*.

Bien que ce domaine soit étendu, il ne l'est que d'une façon bien spécifique découlant des éléments en jeu. Il faut que la succession de dépendances ait une certaine permanence afin d'être saillante et récupérable dans la mémoire de travail pendant que le syntagme QU est en attente d'être déchargé. C'est ce qui se produit quand l'extension du domaine ne se fait que lorsque les relations tête-dépendant en jeu impliquent de la sélection obligatoire, de sorte que chaque relation de sélection est encodée lexicalement dans une tête : ceci rend la relation saillante et récupérable, et l'extension est donc transparente.

Cette extension de domaine correspond essentiellement à la notion de Chaîne V («V-chain») de Bouchard 1984, qui éliminait toutes les traces intermédiaires entre le syntagme QU et sa position de base; Bouchard 1987 se départit même des traces de base. La relation de sélection entre tête et complément joue un rôle central dans l'analyse des relations à distance de plusieurs autres approches; ainsi, elle est exprimée par le suscrit («superscript») dans la Projection par percolation («Percolation Projection») de Kayne 1983 et par le marquage θ dans le Marquage L de Chomsky 1985.

(10) Projection par percolation (Kayne 1983) :

A est une projection par percolation de B si et seulement si A est une projection de B, ou A est une projection de C, où C porte le même suscrit que B et gouverne une projection de B, ou une projection par percolation de B.

(11) Marquage L (Chomsky 1985) :

α L-marque β si et seulement si α est une catégorie lexicale qui θ -gouverne β .

(α θ -gouverne β si et seulement si α est une catégorie de niveau zéro qui θ -marque β , et que α , β sont adjacents.)

⁷ Il n'y a pas non plus de traits spéciaux comme le SLASH en grammaire de structure syntagmatique généralisée (Gazdar et coll. 1985).

Mais ces deux conditions sont spécifiques aux déplacements et supposent une série de relations par déplacement qui reproduisent les relations déjà établies indépendamment par fusion et projection des têtes. Ce ne sont donc pas des hypothèses minimales.

On retrouve aussi l'idée d'extension de domaine sous une autre forme chez Chomsky 2000. Dans sa quête de minimalisme, il suggère de simplifier la composante transformationnelle, de n'utiliser que Fusion («Merge») plutôt que Fusion+Effacement (c'est-à-dire le déplacement). Par exemple, une phrase comme (12a) a la représentation en (12b) avant l'Épellation («Spell-out»).

(12) a. A proof is likely to be discovered.

b. A proof is likely [a proof to be [discovered a proof]]

Il n'y a qu'un seul élément *a proof* d'introduit dans la dérivation, mais il est fusionné à trois endroits différents. Chacune des «occurrences» de *a proof* est définie comme étant son contexte de fusion, soit son nœud adjacent. Ainsi, en (12b), les trois «occurrences» de *a proof* seraient *is likely to be discovered*, *to be discovered* et *discovered*. Chomsky définit une chaîne non plus comme une séquence, mais comme un ensemble, s'appuyant de façon explicite sur le fait qu'une «occurrence» plus haute d'un élément comme *a proof* comprend celles qui sont plus basses : *is likely to be discovered* comprend *to be discovered*, qui à son tour comprend *discovered*. Cette utilisation des notions de nœud adjacent et de chaîne ressemble à l'hypothèse d'extension que j'ai présentée ci-dessus. Toutefois, il y a des différences importantes. D'abord, un élément «à distance» a un traitement spécial : il est fusionné à plusieurs reprises dans la dérivation (ce qui crée le problème de déterminer où il sera prononcé, exigeant donc encore une autre stipulation). Ensuite, la notion de chaîne conserve la redondance entre la série de relations nécessaires entre têtes et nœuds adjacents d'une part, et une série de relations exactement parallèles et superflues établies par les fusions multiples du même élément. De plus, les occurrences multiples qui constituent les chaînes requièrent des stipulations afin de définir des domaines pour obtenir des effets de cyclicité : les notions de «phase» et «d'impénétrabilité» jouent ce rôle. En somme, même si l'opération de déplacement ne fait techniquement plus partie des mécanismes de la théorie, ce qui le remplace réintroduit des mécanismes de même «coût» sous une autre forme.

En comparaison, dans l'approche que je présente, le minimum de parcimonie nous amène à l'établissement direct d'un lien entre le syntagme QU et son verbe dans un exemple comme (9), conditionné par une succession de relations locales indépendamment nécessaires. Cette analyse ne fait pas que

décrire les faits : elle est fondée sur des propriétés de notre système articulatoire-perceptuel qui sont logiquement antérieures à la théorie linguistique, de sorte qu'elle répond à des considérations de théorisation idéale. Cette analyse la plus parcimonieuse de la structure syntagmatique explique la nature de l'extension de domaine qu'on retrouve dans les langues, c'est-à-dire la localité découlant de l'interaction entre les relations d'adjacence induites par la Juxtaposition et la sélection par les têtes. L'analyse explique parce qu'elle établit une relation causale entre les faits et des propriétés fondées indépendamment, logiquement antérieures : le type de localité qui se dégage est nécessaire et fait du sens, étant donné ces propriétés. Il n'y a nul besoin de conditions spéciales pour exprimer cette localité puisqu'elle découle de propriétés logiquement antérieures de notre physiologie et d'une simple condition de récupérabilité. Ainsi, la cyclicité découle directement de la Juxtaposition : il ne peut y avoir de «sauts» par une Juxtaposition itérative. De même, le fait que le PCV, la Projection par percolation, le Marquage L définissant le domaine de la Sous-jacence font tous appel à la relation de sélection entre tête et complément n'est pas une série de découvertes fortuites : dans mon analyse, le rôle de cette relation s'explique par la nature même de la Juxtaposition et la nécessité de la sélection pour assurer que chaque dépendance est récupérable dans la mémoire de travail pendant que le syntagme QU est en attente. Le fait que les sujets soient moins accessibles que les objets pour les relations à distance n'est donc pas un fait du hasard qu'il faut platement décrire en ajoutant à la liste des conditions : cela découle directement de différences de sélection, l'objet étant sélectionné lexicalement, mais non le sujet.

De plus, l'analyse s'étend directement à des cas qui n'ont jamais été bien couverts par le PCV ou la Sous-jacence. Par exemple, dans la série de relations tête-complément servant à relier un syntagme QU à son verbe, si un des verbes impliqués a une relation de modification avec son SN adjacent plutôt qu'une relation de sélection, comme c'est le cas pour les verbes *murmurer* et *crier* exprimant la manière de parler (voir Stowell 1980 et Bouchard 1984), alors l'extension de domaine est brisée.

(13) *Qu'est-ce que Paul a dit qu'elle a **murmuré** que Paul a vu?

Ces verbes ne sont pas un «pont» entre le syntagme QU et son verbe parce que la relation avec le SN n'a pas la saillance requise pour permettre la récupérabilité de la relation. Dans la même veine, le complément d'un N ou une phrase relative modifiant un N n'établissent pas une relation de sélection

avec le N, et ceci brise aussi l'extension de domaine (la contrainte du SN complexe de Ross 1967)⁸ :

- (14) a. *Qui connais-tu l'homme qui a vu?
 b. *Qui connais-tu la rumeur que Paul a vu?

Les constructions impliquant de la coordination sont également difficiles pour une analyse structurale, mais s'expliquent directement dans une analyse où la sélection joue un rôle central. Ainsi, la violation en (15) de la contrainte sur les structures coordonnées (Ross 1967) vient du fait que c'est toute la structure coordonnée qui est sélectionnée, pas seulement une sous-partie, tandis que les constructions «Across-the-board» comme (16) (Williams 1978) impliquent des compléments avec des relations de sélection identiques.

- (15) a. *Qui est-ce que Jean a vu [_{NP} Marie et ___]?
 b. * Qui est-ce que Jean a vu [_{NP} ___ et Marie]?

- (16) Qui est-ce que Jean a vu ___ et Paul a frappé ___?

Bref, l'analyse nous permet d'aller au-delà de la description cas par cas et de comprendre pourquoi les choses se passent comme elles le font.

8 Une observation de Erteschik-Shir 1981, s'appuyant sur Cattel 1979, montre bien que la contrainte porte sur la saillance et la récupérabilité plutôt que sur la structure. L'opacité d'un SN quant à une relation à distance avec un syntagme QU varie selon de subtiles différences de contexte. Par exemple, alors que (ib) est acceptable, (iib) est perçu comme déviant :

- (i) a. John wrote a book about Nixon.
 b. Who did John write a book about?
 (ii) a. John destroyed a book about Nixon.
 b. #Who did John destroy a book about?

Erteschik-Shir attribue ces contrastes à une condition de Dominance, qui dit qu'un constituant ne peut être «extrait» que s'il est dominant :

- (v) Dominance (Erteschik-Shir 1981) :

Un constituant c d'une phrase S est dominant dans S si et seulement si le locuteur entend diriger l'attention de l'interlocuteur à l'intention de c, en émettant S.

L'objet de *about* est dominant en (i) parce que le contenu du livre est très pertinent pour le verbe *wrote*. Mais en (ii), il n'est pas dominant parce qu'avec *destroy*, le locuteur veut normalement que l'interlocuteur porte son attention sur l'acte de destruction plutôt que sur le contenu du livre. Toutefois, comme le fait remarquer Bouchard 1995 : 363-382, dans un contexte où on sait que John est farouchement anti-Nixon, et que sur le contenu du livre est la raison pour laquelle il veut le détruire, alors l'objet de *about* est dominant et (iib) devient acceptable. Dans une analyse structurale, c'est un résultat surprenant. Dans une analyse basée sur la saillance et la récupérabilité de la relation, un tel effet de Dominance va de soi.

4.2. Une autre façon de voir les données

La parcimonie théorique nous amène à voir les données sous un autre jour. Premièrement, des données linguistiques primaires indiquent à un enfant exposé au français que, parmi les quatre moyens permis par la physiologie articulatoire des humains pour exprimer les dépendances du type objet direct/indirect, c'est la Juxtaposition qui est utilisée dans cette langue⁹. Deuxièmement, des données primaires indiquent aussi qu'il y a extension de domaine en français pour les syntagmes QU, comme nous l'avons vu ci-dessus¹⁰. Troisièmement, cette extension est fortement contrainte : la Faculté de langage étant minimale et n'admettant donc que les éléments lexicaux comme unités primitives, l'apprenant

9 C'est un choix paramétrique qui est facile à établir pour l'apprenant puisqu'il implique des propriétés très saillantes et facilement détectables de notre physiologie : très peu de données suffisent pour déterminer le choix fait par la langue. Par exemple, si un apprenant se trompe et attribue l'expression d'une relation entre A et B à l'ordre linéaire dans une langue à Marquage de cas, les propriétés en jeu sont très saillantes, et l'apprenant aura des indices concrets pour s'apercevoir de son erreur et identifier le bon mode d'encodage. Le Problème épistémologique et le Problème du crédit discutés par Dresher 1999 ne se posent donc pas.

10 «... contrairement, par exemple, au japonais», dit Obenauer 1990 : 79. Or, ce n'est pas tout à fait juste. Ce qui distingue le japonais du français, ce n'est pas que les syntagmes QU ne peuvent apparaître «hors position», mais plutôt la façon d'obtenir les relations à distance : ce n'est pas par extension de domaine que le japonais obtient ces relations, mais à cause du fait qu'il n'encode pas les fonctions grammaticales par la Juxtaposition, mais par le Marquage de cas. Il en résulte un ordre libre des arguments entre eux (une approche à base universelle unique est obligée de postuler une opération de Brouillage («Scrambling») pour en rendre compte). Un argument d'une phrase enchâssée peut même précéder ceux du verbe principal, y compris si cet argument est un QU comme en (ib), où *nani-o* est un SN QU relié au verbe enchâssé, mais apparaît devant *John-ga*, un argument du verbe principal.

- (i) a. John-ga Mary-ga nani-o katta ka sitteiru.
 John-NOM Mary-NOM quoi-ACC acheta Q sait
 John sait ce que Mary a acheté.
- b. Nani-o John-ga Mary-ga katta ka sitteiru.
 quoi-ACC John-NOM Mary-NOM acheta Q sait
 John sait ce que Mary a acheté.

L'ordre libre du SN QU est comme celui de tous les SN en japonais : il dépend du Marquage de cas, qui permet d'identifier la fonction grammaticale sans recours à des propriétés structurales dérivées de la Juxtaposition. Parce qu'il est QU, il faut que sa portée soit déterminée en plus de sa fonction grammaticale. C'est le marqueur *ka* qui le fait en se plaçant à droite du verbe visé. Ainsi, en (i), *ka* indique que *nani-o* a portée sur le complexe fonctionnel du verbe enchâssé *katta*. A noter que contrairement au japonais, un QU à l'initiale d'une phrase principale en français ne peut avoir une portée uniquement sur la subordonnée et doit porter sur la principale :

- (ii) *Quel livre Jean a demandé que Marie a acheté?

L'ordre des SN QU est donc en fait plus libre en japonais qu'en français. Voir Bouchard 2001 pour plus de détails et des références.

n'a pas d'autre choix que d'attribuer cette extension de domaine à l'accessibilité aux spécifications de sélection que permet la projection des têtes lexicales.

La parcimonie théorique suppose donc un apprentissage très différent de celui suggéré par une théorie incorporant des conditions spécifiques comme le PCV. Confronté à des données comme (2) à (4), l'apprenant a comme point de départ (et d'arrivée) que les relations sont strictement locales. L'apprenant n'étend un domaine que s'il a des données (positives) qui l'amènent à le faire, c'est-à-dire des données lui indiquant que la langue permet l'extension de domaine. Les restrictions sur cette extension sont entièrement dérivées de notions des plus élémentaires. Il n'y a donc aucun besoin de dispositifs comme le PCV pour écarter les phrases agrammaticales comme (3b) : de telles constructions ne se retrouvent jamais dans les données primaires, donc il n'y a pas d'indication dans les données que les domaines devraient être étendus de cette façon, et la théorie minimale n'offre aucun fondement qui permettrait ce genre de relation à distance. Il n'y a donc aucune raison pour que l'apprenant fasse l'erreur d'étendre un domaine de cette façon. Sa connaissance inconsciente précise à propos des phrases qui violent le PCV provient simplement du fait que ces phrases ne tombent ni sous la localité stricte, ni sous les conditions qui étendent les domaines de localité stricte.

Si on adopte la théorie la plus parcimonieuse, il n'y a pas de problème d'induction parce qu'il n'y a pas de problème de pauvreté des stimuli. Il n'est pas nécessaire de poser un principe particulier au langage comme le PCV pour rendre compte du fait que les apprenants ne font pas l'erreur de considérer une phrase telle que (3b) comme étant bien formée. Pour commettre cette erreur, il faudrait qu'un enfant «surgénéralise» une règle de déplacement libre. Or, une théorie parcimonieuse, de concert avec les propriétés logiquement antérieures de la Juxtaposition en jeu, n'admet pas une telle règle et impose une localité sévère qui rend compte des faits¹¹.

Les deux approches offrent des réponses très différentes à la question de savoir pourquoi la Grammaire a des effets comme ceux de (2) à (4). En grammaire générative, cela dépend de mécanismes comme les transformations et les conditions particulières au langage qui les contraignent. Mais on s'est très peu posé la question quant à la raison pour laquelle le langage comporterait de

11 Contrairement au Marquage, qui permet des relations plus libres que celles admises par l'accessibilité aux spécifications de sélection par la projection des têtes lexicales, comme nous l'avons vu pour le japonais à la note 10. Le Marquage est parfois le seul mode d'encodage utilisable parce que le type d'information à transmettre ne permet pas l'emploi des autres modes. Par exemple, la coréférence entre éléments d'une phrase à une autre est toujours encodée morphologiquement et non structurellement, parce que la Juxtaposition et la Surimposition ne peuvent opérer à l'interphrastique.

tels mécanismes. Miller et Chomsky 1963 offrent une réponse fonctionnelle : ces mécanismes seraient dus à des considérations d'emploi du langage, comme faciliter le passage, marquer la distinction entre la sémantique de base (thématique) et la sémantique plus orientée vers le discours (rhématique). Chomsky 1995 : 317 reprend cette idée en une phrase. Chomsky 2000 en discute brièvement comme étant possiblement une motivation externe au déplacement, justifiant ainsi cette «imperfection» dans la théorie. Toutefois, l'argumentation est un peu circulaire puisqu'on ne nous donne nulle part de raisons indépendantes pour séparer la sémantique en deux sortes, ni pour croire que l'une d'entre elles serait externe à la faculté de langage. De plus, la nécessité de poser des conditions particulières est attribuée à la pauvreté des stimuli; or, en fait, c'est la formulation des conditions spécifiques proposées qui ne permet pas un lien avec les stimuli présents.

Une grammaire adaptative comme celle que je présente offre une toute autre réponse. Les effets comme ceux de (2) à (4) doivent être exactement tels qu'ils sont dans une théorie minimale qui prend appui sur des propriétés externes non pas vagues comme la distinction souhaitée entre deux sémantiques, mais des propriétés logiquement antérieures précises de notre système articulatoire et de notre système cognitif¹².

5. Tout est en place

Tous les stimuli nécessaires pour l'apprentissage des contraintes en jeu sont donc présents. S'il est peu vraisemblable qu'on puisse expliquer les faits sur la base de l'intelligence générale ou de facteurs psychologiques comme la motivation ou l'émotion, cela ne veut pas dire qu'il ne reste comme solution que des principes particuliers très différents de ce qu'on trouve dans les autres sciences. Il y a d'autres propriétés, logiquement antérieures à la grammaire, qui permettent d'élaborer une théorie beaucoup plus parcimonieuse si on tient compte de tous leurs effets.

On peut se demander pourquoi cette théorie n'a pas été proposée plus tôt, d'autant plus que tous ses principaux éléments sont déjà en place depuis un certain temps. Ainsi, on trouve déjà chez Postal 1972 un argument en faveur d'une théorie très parcimonieuse (très semblable au programme minimaliste, d'ailleurs). Une approche très minimale envers la structure syntagmatique, semblable à celle dérivée de la Juxtaposition, apparaît indépendamment à plusieurs

12 Bouchard (à paraître) présente plusieurs autres effets des propriétés logiquement antérieures du système articulatoire et du système cognitif sur la Grammaire.

endroits (la grammaire catégorielle, Bouchard 1979, Chomsky 1994). On a vu que Chomsky 1994 propose de n'avoir que le minimum comme étiquettes dans les structures, soit les éléments lexicaux, idée qu'on peut aussi retracer en grammaire catégorielle (voir aussi Bouchard 1982). L'idée que les relations strictement locales sont non marquées et que les relations à distance sont marquées est déjà présente de façon explicite chez Koster 1978. Enfin, l'idée de chercher une motivation externe aux mécanismes théoriques qui ne sont pas minimaux est en filigrane dans la théorie depuis Miller et Chomsky 1963, avec une tentative d'effets précis chez Kayne 1994 avec son «Linear Correspondence Axiom», idée retraçable chez Tesnière, voire chez Saussure.

Malgré la présence de tous ces éléments pendant tout ce temps, on ne les trouve pas présentés de façon intégrée, systématiquement reliés entre eux : on les retrouve plutôt par bribes ici et là, alors qu'ils auraient pu donner une théorie cohésive et parcimonieuse dès le début. C'est trop étonnant pour être un hasard. Si on a trouvé ces propriétés, c'est qu'avec une exposition constante aux données, et souvent avec un sens de la langue scintillant comme chez J.R. Ross, des éléments finissent par se révéler par un système implicite plutôt que par une théorie explicite. Par contre, si ces éléments sont restés disparates, c'est que leur systématisation dépend de propriétés logiquement antérieures au langage qui touchent à la substance plutôt qu'au formel. Or, la grammaire générative a très peu pris ce genre de propriétés en compte à cause de la façon dont elle s'est développée. Dès le début, le principal objectif a été d'expliquer comment un locuteur en vient à connaître les faits de langue. Comme Chomsky 1995 : 4 le rappelle, pour ce faire, il fallait mieux comprendre le fait que le langage comporte l'emploi infini de moyens finis, reprenant l'aphorisme de Humboldt. Chomsky dit que ce sont des progrès dans les sciences formelles qui lui ont permis d'aborder ces problèmes d'une façon nouvelle. Ceci a eu pour conséquence que les faits de langue sont étudiés presque exclusivement en termes de leurs propriétés formelles. Ainsi, la quantité prodigieuse de travail voué à l'identification des restrictions universelles dans le langage est centrée sur le fait que l'éventail de langues naturelles possibles est restreint en regard de toutes les possibilités logiques qu'offre un système computationnel.

Cette façon d'aborder le langage a pour résultat qu'on exprime d'abord les propriétés qui entrent facilement dans un tel formalisme, comme les propriétés touchant à l'ordre. Une fois qu'on a exprimé ces principales propriétés, il ne reste guère qu'à tenter d'en raffiner la description pour plus de précision et d'élégance formelle. Ainsi, depuis Chomsky 1975 [1955], combien n'avons-nous pas connu de reformulations du passif, du déplacement *QU*, du placement de l'accord et du temps sur les formes verbales; et depuis Ross 1967, que de

reformulations de conditions de localité et de chaînes. Ces reformulations, quoiqu'utiles au niveau métathéorique, ne nous informent guère plus sur les faits.

La prise en compte presque exclusive du formel n'est pas une propriété inhérente à la grammaire générative. Chomsky 1990 : 147 le dit bien : «The serious problem is to learn more, not to formalize what is known to make unmotivated moves into the unknown.» Même que la formalisation peut nuire en faisant perdre de vue les effets de la substance dépendant de propriétés logiquement antérieures, comme le constatent Chomsky et Halle 1968 : 400 :

The problem is that our approach to features, to rules, and to evaluation has been overly formal. Suppose, for example, that we were systematically to interchange features or to replace [αF] by [$-\alpha F$] (where α is +, and F is a feature) throughout our description of English structure. There is nothing in our account of linguistic theory to indicate that the result would be the description of a system that violates certain principles governing human languages. To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made use of the fact that the features have intrinsic content.

Pourtant, dans la pratique, c'est le jeu de la formalisation et de la reformulation qui l'a emporté, au point que Hale et Reiss 2000 peuvent affirmer dans la revue phare du générativisme qu'il faut préférer une théorie de la phonologie qui est strictement formelle et qui ne fait pas référence au contenu de substance. Pour Hale et Reiss, une théorie phonologique peut admettre certaines combinaisons de traits qui sont phonétiquement impossibles. Selon eux, ce serait même coûteux de bloquer ces combinaisons en phonologie parce que cela redoublerait des principes phonétiques. Je suis d'accord en principe. Toutefois, s'il s'avère que tous les effets d'un outil théorique sont toujours éliminés par un principe de substance indépendant, la méthode scientifique pose que cet outil théorique ne devrait pas faire partie de la théorie, sinon n'importe quoi peut être admis même si c'est sans effets tangibles. En fait, même s'il y a des effets, mais très limités, alors que l'outil est d'application très large, c'est habituellement signe qu'il existe une meilleure alternative.

L'étude de cas types présentée dans cet article montre que des propriétés de substance, qui sont logiquement antérieures à la théorie linguistique, procurent une théorie générale plus explicative. Chomsky 1990 : 145 reconnaît qu'il est douteux que l'étude du langage puisse se faire en isolation, sans tenir compte de l'acquisition, de l'emploi, et des mécanismes physiques. On a vu que dans Chomsky 2000, l'imperfection formelle qu'est le déplacement l'a amené à chercher des facteurs externes pour l'expliquer. Le bref regard que nous venons de jeter sur les conditions de localité nous apprend que, si nous avons adopté

l'approche basée sur les propriétés externes dès le départ (ce qui aurait dû être le cas puisqu'elle est plus parcimonieuse), nous aurions obtenu une explication systématique de la localité, au lieu d'une liste de conditions descriptives enchâssée dans une Grammaire universelle qui est une sorte de mystère génétique.

L'approche basée sur le formel a de toute façon été amenée à faire appel au substantiel pour arriver à une explication. Nous l'avons fait dans la pratique : nous devrions le faire par principe. Le cerveau, dans lequel la faculté de langage est établie, contient aussi un système conceptuel avec ses caractéristiques propres, et il est situé dans un corps humain qui a un système sensorimoteur particulier. Ces deux propriétés ont pour conséquences immédiates que le sens est restreint à des notions qui peuvent être reliées à notre système conceptuel, et que la sorte de forme qui peut servir au langage est contrainte par les propriétés physiologiques de notre système articulatoire-perceptuel. Ces propriétés sont logiquement antérieures à la théorie linguistique, en ce sens qu'elles doivent préexister puisque l'objet d'étude de la linguistique les présuppose. Cette reprise de l'observation traditionnelle que le langage est une relation entre forme et sens soulève une question de fond : la faculté de langage est-elle entièrement déterminée par ces conditions externes, ou contient-elle des principes qui lui sont particuliers au delà de ces conditions ? La grammaire générative a répondu qu'une quantité importante de conditions particulières sont nécessaires. Mais en même temps, elle a été progressivement amenée à s'appuyer sur des propriétés externes. Je propose une démarche systématique en ce sens, pour des raisons de principe voulant qu'on prenne des propriétés logiquement antérieures à celles à l'étude comme base axiomatique. La réponse à la question est alors très différente : la faculté de langage a des propriétés bien à elle, mais elles sont en grande partie, et peut-être même complètement, déterminées par un nombre limité de propriétés logiquement antérieures. Les propriétés langagières ne sont donc pas mystérieusement particulières et différentes de tout ce qu'on trouve en biologie et dans les autres sciences traitant du cerveau, mais sont préexistantes et adaptées. La nature des objets mis en relation – les éléments de sens et de forme – détermine le type de relations qui peuvent être établies entre eux. Plutôt qu'une Grammaire universelle spécifiste, c'est une Faculté de langage innéiste adaptative qui émerge alors.

On peut croire qu'en prenant résolument le parti d'un innéisme adaptatif plutôt que spécifiste, les années à venir nous promettent d'arriver à des explications de façon plus directe et sans méandres inutiles.

Références

- BOUCHARD, Denis 1979 *Conjectures sur une grammaire indépendante du contexte pour les langues naturelles*, mémoire de maîtrise, Université de Montréal.
- BOUCHARD, Denis 1982 *On the Content of Empty Categories*, thèse de doctorat, MIT.
- BOUCHARD, Denis 1984 *On the Content of Empty Categories*, Dordrecht, Foris.
- BOUCHARD, Denis 1987 «A Few Remarks on Past Participle Agreement», *Linguistics and Philosophy* 10 : 449-474.
- BOUCHARD, Denis 1995 *The Semantics of Syntax*, University of Chicago Press.
- BOUCHARD, Denis 2001 «The concept of ‘universal’ and the Case of Japanese», *Lingua* 111 : 247-277.
- BOUCHARD, Denis (à paraître) *French Adjectives and Universal Grammar*, Oxford, Elsevier.
- CATEL, Ray 1979 «On Extractability from Quasi-NPs», *Linguistic Inquiry* 10 : 168-172.
- CHOMSKY, Noam 1965 *Aspects of the Theory of Syntax*, Cambridge (Mass.), MIT Press.
- CHOMSKY, Noam 1975 [1955] *The logical structure of linguistic theory*, New-York, Plenum.
- CHOMSKY, Noam 1985 *Barriers*, collection Linguistic Inquiry Monographs, Cambridge (Mass.), MIT Press.
- CHOMSKY, Noam 1990 «Topic...Comment : On Formalization and Formal Linguistics», *Natural Language and Linguistic Theory* 8 : 143-147.
- CHOMSKY, Noam 1994 «Bare Phrase Structure», *MIT Occasional Papers in Linguistics* 5, Cambridge (Mass.), MIT, Département de linguistique.
- CHOMSKY, Noam 1995 *The Minimalist Program*, Cambridge (Mass.), MIT Press.
- CHOMSKY, Noam 1999 «Derivation by Phase», *MIT Occasional Papers in Linguistics* 18, Cambridge (Mass.), MIT, Département de linguistique.
- CHOMSKY, Noam 2000 «Minimalist inquiries : The framework», dans R. Martin, D. Michaels, J. Uriagereka et coll., *Step by step : Essays on minimalist syntax in honor of Howard Lasnik*, Cambridge (Mass.), MIT Press, p. 89-155.
- CHOMSKY, Noam et Morris HALLE 1968 *The Sound Pattern of English*, Harper & Row, New-York.
- DRESHER, Bezalel Elan 1999 «Charting the Learning Path : Cues to Parameter Setting», *Linguistic Inquiry* 30 : 27-67.
- ERTESCHIK-SHIR, Nomi 1981 «More on Extractability from Quasi-NPs», *Linguistic Inquiry* 12 : 665-670.
- GAZDAR Gerald, Ewan KLEIN, Geoffrey PULLUM et Ivan A. SAG 1985 *Generalized Phrase Structure Grammar*, Cambridge University Press.
- HALE, Mark et Charles REISS 2000 «Substance Abuse and Dysfunctionalism : Current trends in phonology», *Linguistic Inquiry* 31 : 157-169.
- KAYNE, Richard 1983 «Connectedness», *Linguistic Inquiry* 14 : 223-249.
- KAYNE, Richard 1994 *The Antisymmetry of Syntax*, Cambridge (Mass.), MIT Press.
- KOSTER, Jan 1978 *Locality Principles in Syntax*, Dordrecht, Foris.

- MILLER, George et Noam CHOMSKY 1963 «Finitary models of language users», dans R.D. Luce, R. Bush, E. Galanter et coll., *Handbook of mathematical psychology II*, New-York, Wiley.
- OBENAUER, Hans-Georg 1990 «De quelques malentendus récurrents concernant la grammaire générative», *Recherches linguistiques de Vincennes* 19 : 73-106.
- PINKER, Steven 1994 *The Language Instinct. How the Mind Creates Language*, New-York, William Morrow & Co.
- POSTAL, Paul 1972 «The Best Theory», dans Stanley Peters et coll., *Goals of Linguistic Theory*, Englewood Cliffs, Prentice-Hall, p. 131-170.
- ROSS, John Robert 1967 *Constraints on Variables in Syntax*, thèse de doctorat, MIT. Publiée avec révisions en 1986 sous le titre *Infinite Syntax!*, New-York, Ablex.
- STOWELL, Tim 1980 «Complementizers and the empty category principle», *Proceedings of NELS XI*, Université d'Ottawa, p. 345-363.
- TESNIÈRE, Lucien 1959 *Éléments de syntaxe structurale*, Paris, Klincksieck.
- WILLIAMS, Edwin 1978 «Across-the-Board Rule Application», *Linguistic Inquiry* 9 : 31-43.