

ERIC AUDUREAU

THÈSE

1982
AUD

B.U. MLV



D

133 117230 0

1982

AUD

UNIVERSITE DE PROVENCE

THESE DE TROISIEME CYCLE

Mention Logique et Philosophie des Sciences

LES METHODES MATHEMATIQUES

DE LA LINGUISTIQUE

ERIC AUDUREAU

Directeur de Thèse : Monsieur le Professeur GILLES GRANGER



5

AVANT - PROPOS

Nous utilisons constamment dans cette thèse un certain nombre de notions classiques de la Théorie des Grammaires Génératives et de la Théorie des Grammaires Formelles sans en mentionner les définitions. On trouvera ces définitions respectivement dans Chomsky (1965) et Hopcroft et Ullmann (1969). Les notions ne figurant pas dans ces ouvrages sont définies dans le cours du texte. Les notations employées sont celles des conventions dites de Chomsky (Chomsky, 1959). Par ailleurs pour désigner les classes de langages et les classes de grammaires de la hiérarchie de Chomsky, nous employons tantôt l'appellation nominale (régulier, indépendant du contexte, etc...) tantôt l'appellation "indicielle" (de type 3, de type 2, etc...).

PREMIERE PARTIE

PROBLEMES ET LIMITES DE L'APPLICATION DE LA THEORIE

DES GRAMMAIRES FORMELLES

CHAPITRE PREMIER

INTRODUCTION

I. Notre propos, dans les pages qui suivent, est d'étudier les rapports de la Théorie des Grammaires Formelles et de la Théorie des Grammaires Génératives. Usuellement, la théorie des grammaires génératives désigne les modèles linguistiques élaborés par Chomsky et ses disciples, que ceux-ci lui soient ou non fidèles. Cependant, une telle acception est trop restrictive au moins pour deux raisons. Premièrement, la théorie des grammaires génératives est une théorie qui a pour objet les grammaires génératives. De ce fait elle traite d'un domaine du savoir qui dépasse amplement la seule construction de procédés descriptifs des langues naturelles puisqu'elle doit, en outre, par exemple, définir des méthodes rigoureuses pour comparer ces procédés descriptifs. De même, elle doit ordonner ces différents problèmes à une conception clairement définie du langage humain, ce qui implique, en dernier ressort, des choix épistémologiques et philosophiques. Deuxièmement "grammaire générative" ne désigne pas uniquement des systèmes grammaticaux adoptés par Chomsky et ses successeurs, c'est à dire les grammaires génératives transformationnelles, mais tout un ensemble de procédures qui implicitement font appel aux propriétés attribuées de façon explicite par Chomsky au concept de grammaire générative. Postal (1964a) a montré qu'une telle extension du concept de grammaire générative dépassait le stade de la simple hypothèse d'école en réduisant à leur forme générative un nombre important de modèles structuralistes d'analyse du langage.

D'ailleurs, ce problème de la "générativité" des grammai-

res "non-génératives" sera l'objet d'un examen détaillé au cours du troisième chapitre.

Ces précisions terminologiques étant faites, voyons quel est l'obstacle majeur s'opposant au développement de la linguistique mathématique. Cette discussion préliminaire nous permettra de mieux délimiter la ligne directrice à suivre dans notre examen des rapports entre grammaire formelle et grammaire générative. Elle nous permettra également de mieux faire apparaître les motifs de notre étude.

I.I. Mathématisation et inertie symbolique du langage humain

La première tentation à laquelle est soumis qui veut user des mathématiques pour étudier le langage est de prendre comme cadre de référence les sciences où cet usage s'est révélé fécond, c'est à dire essentiellement les sciences physiques. Cependant une telle attitude s'appuie sur un présupposé qui demande à être justifié. En effet, pourquoi croire que les processus de construction des concepts mathématiques de la linguistique doivent être nécessairement analogues à ceux de la physique?

Les situations respectives du physicien et du linguiste présentent pourtant un contraste qui porte immédiatement au scepticisme devant une telle analogie. Le premier, dans sa tentative de description du monde physique, se trouve face à une matière, la Nature, se présentant à lui de façon informe, alors que le linguiste doit analyser un matériau déjà organisé par l'intelligence humaine. Cet état de fait explique que pour le

physicien l'élaboration d'un concept et de sa forme mathématique aillent de pair. Qui plus est, l'une est garante de l'autre dès l'instant où naît, au terme de démarches expérimentales, l'intelligibilité des phénomènes naturels. Les mathématiques participent ainsi de façon synthétique à l'élaboration de la connaissance, puisqu'elles permettent de rassembler tout un ensemble de résultats expérimentaux concordants sous le régime d'une loi unique. Le caractère indissociable des démarches empiriques et mathématiques trouve sa raison d'être dans le fait qu'une loi de physique mathématique est toujours soit vraie soit fausse^(I). Entendons par là que ces lois n'admettent pas d'exceptions dès l'instant où elles sont reconnues comme correctes. Les lois de la linguistique ne possèdent évidemment pas cette propriété puisque généralement elles ne couvrent qu'une partie, plus ou moins importante, de leur domaine d'application. C'est d'ailleurs pour cette raison qu'elles ne peuvent se prêter à une formulation mathématique. Et si elles ne se prêtent pas à une formulation mathématique, c'est en partie parce qu'elles n'ont pas été élaborées dans le cadre d'un processus synthétique analogue à celui des sciences physiques. Ce qui nous intéresse, c'est de comprendre pourquoi la linguistique ne procède pas de la sorte.

Il est clair qu'il ne s'agit pas d'un choix de la part des linguistes, mais plutôt d'une impossibilité pratique. Pour que la linguistique soit à même de formuler des lois synthétiques analogues à celles de la physique, il faudrait qu'elle dispose de méthodes de décryptage des systèmes symboliques que sont

les langues humaines aussi accomplies que le sont les procédures expérimentales de la physique. Ce parallélisme semblera incongru puisque nous rapprochons d'un côté la finalité de la recherche linguistique, et de l'autre un instrument de la recherche physique. Cependant, cette incongruité n'est qu'apparente, car ce que nous souhaitons comparer c'est le degré de délimitation des objets sur lesquels opèrent ces sciences. Les méthodes expérimentales de la physique sont la confirmation même d'une délimitation correcte de ses objets. Inversement, la précarité des théories linguistiques démontrent la fragilité des objets sur lesquels elles raisonnent. L'état incertain dans lequel se trouve la notion fondamentale de la syntaxe, celle de partie du discours, en est un premier exemple. Etrangement, les parties du discours sont demeurées, comme on le sait, pratiquement identiques dans la linguistique contemporaine⁽²⁾ aux partes orationis des grammairiens latins malgré l'évidence de leur inadaptation à une analyse scientifique du langage.

Même un concept moderne comme celui de phonème, bien que défini avec un grand soin théorique, souffre des mêmes insuffisances dès qu'on se propose d'en faire un objet de manipulation expérimentale. La preuve en est donnée avec les résultats désastreux obtenus dans les travaux de synthèse de la parole s'appuyant sur un découpage phonématique de la parole. Selon Guibert (1979, p.127): "Quel que soit le soin avec lequel auront pu être analysés, paramétrisés, et stockés en mémoire les

sons élémentaires⁽³⁾, nous ferons une pénible constatation: la parole sera tout simplement inintelligible. Une première conséquence de cet échec est à rechercher dans les discontinuités qui n'auront pas manqué de se produire aux jonctions interphonémiques.". Pour Guibert, ces considérations "traduisent simplement le fait que les unités phonémiques sont le plus souvent combinées en des unités plus étendues que le phonème. Rien n'est plus faux que de penser que les phonèmes sont comme des perles enfilées les unes à la suite des autres." (Guibert, 1979, p.128). Cette remise en question de la réalité empirique d'un niveau phonologique ne touche pas qu'à ses utilisations techniques. Par exemple De Cornulier (1978), propose de nombreux arguments contre l'idée que le phonème puisse être conçu comme une entité indépendante de la syllabation, ou plus exactement comme une entité à partir de laquelle on peut décrire les processus de syllabation. On notera que De Cornulier, comme Guibert dont les intérêts sont pourtant très différents, conclut qu'on ne peut pas représenter les phonèmes comme les grains d'un chapelet enfilés les uns à côtés des autres si on espère décrire correctement la syllabation (De Cornulier, 1978, pp. 43-44) (4).

Ces problèmes affectent suffisamment la notion de phonème pour que l'on cesse d'y voir un objet significatif de la réalité linguistique. Si la notion de système phonologique est douée d'une efficacité fonctionnelle quelconque, cette dernière semble correspondre à l'activité symbolique des linguistes plutôt qu'à celle des usagers de la langue.

Notre intention n'est pas de faire le procès de la syntaxe ou de la phonologie, mais simplement de rappeler que si phonème et parties du discours peuvent être des catégories commodes pour décrire certains aspects du langage, celles-ci ne peuvent pas pour autant être considérées comme des unités constitutives de ce vaste cryptogramme qu'est le langage. Il y a comme un espèce d'antagonisme entre les symbolismes utilisés par les linguistes pour représenter ce qu'ils croient connaître du langage et le symbolisme objectif du langage qui leur oppose toute son inertie. Il va donc de soi que le linguiste qui veut décrire le fonctionnement du langage à l'aide de lois analogues à celles de la physique ne peut le faire sans avoir découvert au préalable les objets appropriés à cette tâche. En effet l'astronome, par exemple, n'est pas en butte à des difficultés de cet ordre puisqu'il dispose d'objets pertinents, les astres, livrés directement à sa perception. Ces limites imposées à la formulation de lois générales sur le fonctionnement du langage condamnent-elles définitivement la linguistique à être une science où les mathématiques ne joueraient aucun rôle essentiel? C'est l'avis de Hockett (1967, p.50), un des linguistes structuralistes que l'application des mathématiques à la linguistique a le plus préoccupé : "A propos du langage, comme à propos de beaucoup de choses, une grande quantité de nos connaissances n'est pas interprétable sous une forme mathématique. Il est absurde de supposer que la connaissance non mathématique est contrefaite ou de seconde qualité. Nous pouvons, avec de la chance, obtenir une précision supérieure sur le langage par la

un peu
naïf

mathématisation, mais seulement à partir du moment où nous pouvons trouver un modèle mathématique qui corresponde aux faits du langage tels que nous les connaissons déjà de façon non mathématique." Ce jugement de Hockett, qui condamne les mathématiques à une vocation purement ornementale, est acceptable si on envisage l'utilisation des mathématiques dans le cadre de la recherche directe d'objets linguistiques manipulables expérimentalement⁽⁵⁾. Autrement dit le recours aux mathématiques non seulement est superflu, mais il n'est pas envisageable, si sa contribution au développement de la linguistique doit suivre le même chemin que pour la physique.

La question ne nous semble pas close pour autant et notre intention est précisément de montrer que la mathématisation en linguistique ne suit pas les mêmes voies qu'en physique. Ceci du moins pour une branche de la linguistique mathématique: la grammaire formelle.

Comme nous l'avons dit, les langues humaines sont des systèmes symboliques. Or la théorie des grammaires formelles qui traite des propriétés abstraites des grammaires génératives traite indirectement des propriétés de ces systèmes symboliques. Le privilège de la grammaire formelle est que les systèmes symboliques, dès que l'on s'intéresse à leur fonctionnement syntaxique, c'est à dire aux règles d'assemblage des symboles, sont soumis, à un certain niveau de généralité, à un régime commun susceptible d'une étude mathématique. Nous approfondirons cette question au cours du troisième chapitre, mais précisons déjà ce que nous entendons par là.

Notre idée est qu'un système symbolique muni d'une syntaxe est un système formel, que cette syntaxe soit une simple loi de concaténation ou qu'elle soit beaucoup plus complexe.⁽⁶⁾

Par conséquent tout un ensemble de questions traditionnelles concernant les propriétés fondamentales des systèmes formels, comme la décidabilité, par exemple, sont formulables dans le cadre d'une théorie mathématique des grammaires. Le véritable problème devient alors de savoir 1°) si la théorie des grammaires formelles est bien la théorie mathématique des grammaires adaptée à l'étude de ces problèmes et 2°) si ces propriétés fondamentales des systèmes formels peuvent avoir un sens quelconque pour la linguistique. Quelle que soit la réponse réservée à ces deux questions, deux conséquences en découlent pour la théorie des grammaires formelles.

La première est qu'une étude mathématique des grammaires génératives se trouve déjà justifiée en tant que telle par la nécessité de comprendre comment leurs particularités s'ordonnent face à la structure générale des systèmes formels. La deuxième conséquence est que les propriétés mathématiques que l'on parviendra à exhiber dans le cadre d'une telle étude, bien que déterminées partiellement par les langues naturelles, seront avant tout des propriétés des grammaires génératives et l'intérêt linguistique de ces propriétés dépendra de l'intérêt linguistique des grammaires génératives.

1.2. Le champ de la linguistique mathématique

Le problème ainsi posé, on voit que notre projet d'exa-

miner le fonctionnement et les objectifs de la grammaire formelle réduit singulièrement ce que l'on penserait être le champ d'un travail général sur la linguistique mathématique. Ainsi sont éliminés au moins deux domaines importants de cette dernière : la linguistique quantitative avec ses techniques les plus élaborées, comme les modèles inspirés de la théorie de l'information et ce que nous appellerons, faute d'une meilleure dénomination, la sémantique logique, c'est à dire les travaux développés dans le cadre théorique de la grammaire de Montague.

Bien que des raisons pratiques nous conduisent nécessairement à choisir parmi les techniques de formalisation employées en linguistique celles que nous jugeons les plus intéressantes, nous ne saurions expliquer notre silence total sur ces questions par le seul manque de place. Le succès qu'a connu la théorie de l'information chez les linguistes et celui que connaît aujourd'hui l'emploi de la logique mérite en effet au moins que l'on justifie leur élimination.

Influencé par les prédispositions de la grammaire transformationnelle à la formalisation, et motivé par des intentions critiques à son égard, un courant de logicisation de la linguistique s'est développé ces dernières années pour aboutir, finalement, avec la grammaire de Montague, à des modèles linguistiques où tous les instruments descriptifs sont empruntés à la logique mathématique. La grammaire de Montague n'innove pas par son projet. Des travaux comme ceux de Lakoff (1968)

avaient déjà pour but de réduire la grammaire à la logique, ou du moins de défendre la thèse que "les règles de grammaire qui engendrent les phrases grammaticales de l'anglais (...) ne sont pas distinctes des règles reliant les formes de surface des phrases anglaises à leurs formes logiques correspondantes ." (p.12). Cependant, bien que Lakoff estime qu'il faille construire une "logique naturelle" appropriée à la dérivation des phrases grammaticales, la logique des logiciens n'étant pas adaptée à cette tâche, il n'apporte pas d'innovations techniques décisives allant dans ce sens. Contrairement à Lakoff, Montague emploie des outils logiques très élaborés et la sophistication du modèle qui propose une syntaxe, une sémantique et une pragmatique savamment construite et empruntant les techniques des grammaires catégorielles, de la théorie des modèles et de la logique intensionnelle, mériterait un exposé détaillé. Cependant, malgré toute sa sophistication, ce modèle, croyons-nous, ne pourra jamais traiter de phénomènes linguistiques.

L'objectif de Montague est, selon ses propres mots⁽⁷⁾, de suppléer aux lacunes de la grammaire transformationnelle d'obédience chomskyenne qui ne permet d'étudier ni les conditions de validité des énoncés du langage naturel, ni l'inférence logique entre ces énoncés. Or, pour l'auteur ce sont là les problèmes principaux d'une théorie linguistique conséquente.

S'il est indéniable que les moyens par lesquels le langage peut exprimer les raisonnements logiquement valides cons-

tituent un objet légitime de curiosité, il faut admettre que celui-ci relève de la philosophie du langage ou de la philosophie des sciences ou à l'extrême limite, de la logique elle-même, mais certainement pas de la grammaire. On ne verra donc pas une simple coïncidence dans le fait que la majorité des travaux réalisés dans le cadre de la grammaire de Montague portent sur des thèmes classiques de la philosophie des sciences et de la philosophie du langage anglo-saxonne: problèmes liés à la nature de la quantification, de la référence, des pronoms, etc....

Curieusement, Montague est totalement silencieux sur les raisons qui le poussent à croire que le problème de la vérité est crucial pour la grammaire. Sur ce point précis au moins, Lakoff est explicite : "Qu'on le veuille ou non, la plupart des raisonnements qui sont menés dans le monde se font en langage naturel. Et parallèlement, la plupart des usages du langage naturel mettent en jeu un raisonnement quelconque. Il ne devrait pas être trop surprenant de trouver que la structure logique qui est nécessaire pour que le langage soit utilisé comme outil de raisonnement doive correspondre de façon profonde à la structure grammaticale du langage naturel." (Lakoff, 1968, p.I). Cet argument nous semble discutable. Le fait que la "plupart des raisonnements" se fassent en langage naturel n'établit aucune nécessité d'un lien entre logique et langage. Sinon, en transposant l'argument, on pourrait par exemple dé-

qui

bulletins météorologiques sont exprimés en langage naturel.

Notre silence sur la linguistique quantitative est plus délicat à justifier car dans ce cas-ci, nous avons affaire à une véritable investigation sur les propriétés du langage.

Pour ce qui concerne l'application des statistiques proprement dites à la linguistique, nous ne nous y intéresserons pas car celles-ci, au nom de la séparation des statistiques et des mathématiques ne relèvent pas de notre travail qui se borne à l'étude des rapports entre mathématiques et linguistique. Nous ne nions pas pour autant l'intérêt présenté par l'usage des méthodes statistiques en linguistique. Elles ont démontré leur utilité à travers différentes contributions aux études stylistiques, pédagogiques (enseignement du vocabulaire fondamental des langues étrangères) ou bien encore à la sténographie.

La théorie de l'information, elle, on le sait, a donné naissance à plusieurs thèmes d'études linguistiques qu'on peut classer en deux grandes catégories. D'une part, elle a permis de définir certaines grandeurs physiques du langage que l'on peut considérer, à différents degrés, comme linguistiquement significatives. D'autre part, elle a fourni un espèce de modèle théorique auquel a pu se référer, avec plus ou moins de fidélité et une plus ou moins grande identification, tout un courant de la linguistique structurale.

Le premier type d'application a permis de découvrir des propriétés inattendues du langage naturel. L'exemple le plus frappant et le plus fameux en est le taux de redondance, avoi-

sinant 0,50, des lettres de l'anglais écrit tel que l'a calculé Shannon. On pourrait également citer les raffinements apportés par Mandelbrot à la loi de Zipf. La loi de Zipf-Mandelbrot caractérise certaines propriétés stylistiques des textes, comme l'usage des mots "compliqués", par une grandeur physique, la température informationnelle, dont la forme mathématique est analogue à la loi des gaz parfaits. Ces résultats, l'un déconcertant par l'ampleur du phénomène, l'autre inespéré par sa possibilité même, ont permis de supposer que la théorie de l'information fournirait un modèle auquel la linguistique pourrait se référer en toutes circonstances pour trouver les clés explicatives du fonctionnement du langage. Par conséquent, c'est vers une extension de l'application de la théorie de l'information que s'est orienté la linguistique. Par exemple, les relations syntaxiques furent considérées, au niveau de l'organisation des éléments dans les phrases attestées, comme descriptibles par des probabilités de transition associées aux co-occurrences possibles des classes syntaxiques. Il ne suffit pas de s'interroger sur cette hypothèse extrêmement forte seulement au nom du fait qu'elle néglige les problèmes empiriques rencontrés par la tradition grammaticale ou que Chomsky en ait démontré l'inadéquation en tant que modèle mathématique de la syntaxe⁽⁸⁾. On peut toujours estimer en effet que le modèle de la théorie de l'information représente une bonne approximation des dépendances syntaxiques et que, de toute façon, il offre un cadre théorique indéfiniment améliorable sur le plan de son exploitation pratique. Ce qui nous semble plutôt en cause, dans

cette application c'est l'absence totale d'une réflexion sur les raisons d'être de la théorie de l'information et sur l'usage implicite de tout un ensemble de notions dont la pertinence linguistique reste à démontrer. On sait que la théorie de l'information a pour objectif essentiel la recherche de codes permettant la transmission des messages à une vitesse optimale. Les théorèmes de Shannon consistent à affirmer l'existence de tels codes dans les deux cas du canal bruyant et non bruyant. Or l'usage d'un tel modèle pour l'étude du langage s'il ne peut être finalisé dans la même direction que la théorie de l'information elle-même, puisque les théorèmes de Shannon ne peuvent avoir aucune signification linguistique⁽¹⁰⁾, présuppose néanmoins, au minimum, l'identification des langues naturelles à des codes. Hypothèse sans doute intéressante et stimulante, voire même plausible, mais qui mériterait alors une détermination explicite des formes de codage et surtout une délimitation des choses codées. Malheureusement, il ne semble pas que la linguistique ait pu faire des propositions concrètes en ce sens, l'idée que le langage est un instrument de communication dont les expressions permettent de transmettre des expériences ne dépassant pas le niveau de l'analogie. D'ailleurs comment conserver tout son sens à cette analogie quand on note, par exemple, que les ambiguïtés lexicales et syntaxiques s'opposent à la condition d'unicodéchantabilité des codes.

Ces remarques ne condamnent aucunement les travaux des linguistes édifiés à partir de certaines idées clés de la théo-

Bien que notre but ne soit pas de discuter des principes de la grammaire générative dont nous accepterons toutes les thèses, ces pages peuvent être considérées comme relevant de l'épistémologie de cette théorie linguistique. Jusqu'à présent le problème des rapports entre grammaire formelle et grammaire générative semble avoir été réglé des deux façons suivantes: soit on considère qu'il s'agit de deux théories distinctes n'ayant aucune relation entre elles, la théorie des grammaires formelles étant considérée comme une discipline mathématique à part entière traitant de problèmes liés aux "computer sciences". Soit on admet que ces deux domaines entretiennent des rapports, mais alors ceux-ci sont limités dans le temps aux travaux originaux des deux disciplines (Chomsky 1956, 1959, 1961) et on y voit une seule et même science qui pourrait recevoir le nom de syntaxe mathématique. S'il est vrai que la majeure partie des travaux de la grammaire formelle relèvent purement et simplement des sciences informatiques, une quantité non négligeable de ceux-ci concernent assez directement la grammaire générative sans que l'on puisse dire pour autant qu'ils lui soient assimilables ni qu'ils exposent des résultats trop triviaux pour que l'on puisse penser qu'il s'agit là vraiment de mathématiques ⁽¹⁰⁾.

Notre seconde motivation est de discuter la façon dont les linguistes interprètent généralement ces résultats mathématiques. Qu'ils les invoquent pour rejeter des théories concurrentes ou qu'ils les sollicitent pour corroborer les leurs,

on a effectivement souvent l'impression que les linguistes ne s'attachent qu'aux aspects les plus immédiats des propriétés démontrées par les mathématiciens, négligeant toute la signification que peuvent avoir les procédés de démonstration et les contextes dans lesquels sont démontrées et appliquées ces propriétés. Nous proposerons des interprétations personnelles de quelques uns des résultats les plus classiques de la théorie des grammaires formelles, en particulier aux chapitres IV, V et VII, en tenant compte de tous ces aspects.

Une troisième raison qui nous a poussé à étudier ces questions est l'intérêt, pour ne pas dire l'originalité, que présente la théorie des grammaires formelles en elle-même: Il s'agit en effet d'une mathématique purement qualitative qui dans le cadre de ses applications ne vise ni la quantification ni la prédiction. Evidemment, il ne s'agit pas du cas unique de mathématisation de la qualité dans les sciences de l'homme, mais la particularité de la grammaire formelle est d'être consacrée exclusivement à l'étude de la qualité et d'être conçue à cette fin. En effet lorsque celle-ci produit des résultats quantitatifs, par exemple le théorème uvwxy de Bar-Hillel (cf. notre Chapitre VI) ou le lemme de Peters et Ritchie mentionné à la note II du chapitre VII, ceux-ci sont toujours subordonnés à la compréhension de la qualité comme fin. En ce sens la grammaire formelle présente un contraste certain avec les autres mathématiques de la linguistique. Par exemple, s'il est vrai que la loi de Zipf-Mandelbrot permet de comprendre une qualité, celle-ci n'apparaît qu'au niveau de

l'interprétation de données numériques intégrées dans un schéma conceptuel hypothétique: la constance du produit du rang par la fréquence illustre et confirme l'existence d'un principe du moindre effort.

Si comme nous l'avons exposé dans les premières pages de cette introduction la grammaire formelle ne permet pas d'étudier directement les propriétés du langage humain et si elle est une mathématique de la qualité, alors on peut se demander quel est son objet. Nous en avons donné une première idée plus haut en mentionnant le problème de la décidabilité. Par elle-même la décidabilité d'une grammaire ne présente aucun intérêt pour le linguiste. Ce qui devra donc nous préoccuper, avec cet exemple particulier, c'est de comprendre comment le problème de la décidabilité peut devenir pertinent pour la théorie des grammaires génératives. Il s'avère que l'indécidabilité est un critère qui permet de rejeter une grammaire comme empiriquement inadéquate (au sens de Chomsky). C'est donc en tant qu'instrument de contrôle empirique que cette notion devient intéressante. Nous pensons que la théorie des grammaires formelles se prête fréquemment à cette fonction de contrôle sur les méthodes et les concepts de la théorie des grammaires génératives. La quatrième et dernière motivation de notre étude est d'étayer ce point de vue en nous livrant à l'analyse détaillée de quelques exemples précis.

Notre travail est divisé en trois parties. La première, relativement spéculative, est consacrée à l'élucidation de la

nature de la grammaire formelle et de la grammaire générative. En particulier, nous y proposons un critère distinguant ces deux disciplines. Ce critère est mis ensuite à l'épreuve au cours du quatrième chapitre où nous discutons de la conception mathématique de l'ambiguïté structurale.

La seconde partie, plus technique, est destinée à établir notre point de vue sur la théorie des grammaires formelles comme instruments de contrôle de la théorie des grammaires génératives. Nous y examinons principalement des théorèmes de "limitation". Celui de Chomsky sur les grammaires à états finis, ceux de Peters et Ritchie sur les grammaires transformationnelles et enfin celui de Janssen Kok et Meerens sur le programme global de la théorie des grammaires génératives.

Enfin, la troisième partie est dédiée à l'exposition de travaux qui selon nous devraient augurer d'un renouvellement des méthodes mathématiques de la grammaire.

NOTES DU CHAPITRE I

- (1) Si on fait abstraction, bien entendu, du caractère nécessairement approximatif des lois physiques qui dépendent de l'échelle des phénomènes qu'elles décrivent. La vérité dont nous parlons n'a trait qu'aux mesures significatives par rapport à ces échelles.
- (2) Nous nous référons ici toujours à la grammaire générative. La linguistique structurale européenne a proposé des systèmes de classification des parties du discours ne devant rien à Denys de Thrace et aux grammairiens romains. Brøndal (1948), par exemple, a tenté, à partir d'une analyse combinatoire dont les générateurs sont les quatre catégories aristotéliciennes (substance, relation, qualité, quantité) de redéfinir des systèmes classificatoires exhaustifs des parties du discours de plusieurs langues naturelles.
- (3) Ici les "sons élémentaires" sont les signaux acoustiques correspondant de façon bi-univoque aux phonèmes.
- (4) Une des critiques fondamentales que De Cornulier adresse à la phonologie structurale est qu'elle considère comme acquise la segmentation d'un mot en phonèmes. Or on ne peut trouver de procédures effectives de dénombrement des phonèmes sans faire intervenir les règles de syllabation. D'où le paradoxe de la phonologie qui considère la syllabation comme complètement déterminée par la succession d'unités disjointes qui seraient les phonèmes ou les traits pertinents des phonèmes. Voir en particulier p. 55 et sqq. Observons cependant que le reproche, adressé par De Cornulier aux structuralistes, de syllaber uniquement à partir du

phonème n'est pas très fondé. Kurylowicz (1948) par exemple, estime que "dans antrum ou festus la limite syllabique n'est pas située entre deux phonèmes, mais plutôt dans un phonème" (p.197). A propos des géménées et de la distinction, par exemple, entre fato : fatto (italien) : "il ne s'agit pas ici d'une opposition d'éléments (phonèmes), mais de structures (syllabes). Il est en effet impossible de dissocier les deux parties de la géminée (implosive et explosive) de la coupure syllabique (-) pour les opposer à la consonne simple correspondante. L'opposition rigoureuse est t:t-t, non pas t:tt." (P.198). Sur les géménées, Kurylowicz conclut d'ailleurs: "Une orthographe idéale ne se servirait de la gémination (non phonologique) que pour marquer la limite syllabique tombant à l'intérieur de la consonne..." (p.205).

- (5) Peut-être est-il arbitraire de se référer à des manipulations expérimentales dont la nécessité n'est pas liée directement à la recherche linguistique. Cependant, de telles manipulations ne sont pas hypothétiques. Elles sont à l'oeuvre dans toutes les formes de traitement automatique du langage dont la synthèse de la parole n'est qu'un exemple.
- (6) En particulier cette complexité peut être telle qu'il soit impossible d'en donner une description finie.
- (7) Voir les pages introductives des articles "English as a Formal Language" et "Universal Grammar" tous deux réédités dans Montague (1974).
- (8) Notre cinquième chapitre est consacré à cette démonstration et à l'examen de ses conséquences.
- (9) Le premier théorème de Shannon (Shannon et Weaver, 1949, p.59) pour la canal non bruyant ne correspond évidemment pas à la situation du langage humain dont la transmission

est toujours affectée par le bruit. Quant au second théorème (Shannon et Weaver, 1949, p.71) rappelons qu'il ne s'agit que d'un théorème d'existence dont les démonstrations ne donnent pas d'indications sur le mode de construction du code idéal. Pour Shannon, (p;75) une approximation du codage idéal pourrait être telle que le bruit ne perturbe jamais suffisamment le message pour qu'on le confonde avec un autre. Cette observation peut suggérer le rapprochement avec le taux de redondance de l'anglais, mais il suppose alors une intention des usagers de la langue dont la localisation et la forme restent à déterminer.

(10) Pour Schützenberger (1978) la linguistique mathématique actuelle est réduite à ces deux alternatives.

CHAPITRE II

UN CRITERE DISTINGUANT LA GRAMMAIRE FORMELLE

DE LA GRAMMAIRE GENERATIVE

2. La "Théorie Linguistique Moderne", comme Chomsky appelle ce que l'on désigne plus fréquemment par "théorie des grammaires génératives", a été l'objet de nombreuses révisions qui portaient tantôt sur la nature de la grammaire, tantôt sur la fonction de l'activité du grammairien, c'est à dire sur le type de connaissance auquel la linguistique permettrait d'accéder^(I). Cependant, à travers toutes ces réformes subsistent certains traits permanents permettant de reconnaître la théorie linguistique moderne comme une seule et même chose. L'un de ces traits, souvent présenté par Chomsky lui-même comme un objectif primordial, est la volonté de délimiter la classe des grammaires possibles. Concrètement, une telle tâche revient à éliminer des systèmes grammaticaux au nom des critères les plus variés. La vertu unificatrice de cet objectif ne concerne pas seulement les versions successives de théories génératives, mais vaut également sur un plan synchronique : envisagée sous cet angle, grammaire générative et grammaire formelle sont indistinctes.

Selon un sentiment assez confus, partagé par de nombreux commentateurs, cette indistinction sur le plan fonctionnel trouverait son fondement dans une similarité morphologique: grammaire formelle et grammaire générative seraient des mécanismes fondamentalement identiques, fonctionnant comme des systèmes formels et traitant d'une notion de langage ramenée au simple rang d'objet logico-mathématique. Une raison amène spontanément à refuser une telle assimilation, c'est l'évidente différence d'organisation de l'information proposée par les deux disciplines. Dans le

cas de la grammaire formelle, nous avons affaire à des propositions universelles sur les propriétés de systèmes symboliques complexes et élaborés, dans celui de la grammaire générative, le symbolisme est également présent, mais toujours dans une phase d'élaboration, sans jamais atteindre le stade de la constitution d'un système symbolique stable et achevé.

Cependant, un simple constat de visu des différences formelles entre les langages de la grammaire générative et de la grammaire formelle, aussi frappantes soient-elles, n'établit pas pour autant une distinction de nature qui soit essentielle. Aussi la question que nous aborderons au cours de ce chapitre concernera l'existence et l'éventuelle formulation d'un critère distinguant grammaire générative et grammaire formelle.

Dans la recherche de ce critère distinctif, nous considérerons successivement trois possibilités. Mais avant d'exposer pourquoi nous pensons devoir écarter les deux premières, à savoir que la grammaire formelle ne se distingue de la grammaire générative ni parce qu'elle est une mathématique pure, ni parce qu'elle présente un plus haut degré d'abstraction, il nous faut préciser dans quel esprit nous discuterons de ces critères. Nous sommes convaincus, en effet, que la grammaire formelle est une théorie plus abstraite que la grammaire générative et qu'elle est également une théorie mathématique de cette dernière. Mais ce qui nous intéresse ici c'est d'une part de trouver un critère pertinent par rapport à l'activité grammaticale et d'autre part, de vérifier jusqu'à quel point ce critère est fondé. Or, dès

qu'on se place dans une perspective plus pratique, cette dernière exigence peut être négligée. Pour cette raison, dans des chapitres ultérieurs nous procéderons comme si la théorie des grammaires formelles était bel et bien une théorie mathématique des grammaires génératives. La recherche d'un critère distinctif n'est pas pour autant un exercice purement spéculatif. C'est au contraire une étape indispensable pour qui veut comprendre les causes du pouvoir limité de la grammaire formelle sur les faits linguistiques. Nous mettrons d'ailleurs notre critère à l'épreuve au quatrième chapitre et nous verrons qu'il permet d'interpréter des problèmes très précis. Nous voulons simplement dire que nous jugerons tolérable, et parfois même très commode, de nous rallier à l'intuition selon laquelle cette discipline qui énonce des théorèmes sur la morphologie des grammaires ou sur leur capacité générative est bien la théorie mathématique de cette autre discipline qui tente de décrire les langues naturelles.

2.I. La nature mathématique de la Grammaire Générative

Là où les grammaires formelles pourraient se séparer des grammaires génératives, c'est que les premières sont destinées à engendrer des langages formels tandis que les secondes doivent engendrer des langues naturelles. Mais on sait que pour la théorie des grammaires génératives, l'étude des langages naturels se présente comme celle de langages formels particuliers⁽²⁾. S'il y a une différence entre ces deux types de langages, elle apparaîtra au niveau des grammaires puisque les langages humains

sont des systèmes trop riches et trop complexes pour que leur spécificité soit exprimable par une simple définition. Mais si la grammaire générative est conçue comme l'étude de langages formels, alors quelle raison théorique peut nous permettre de la rejeter en dehors des mathématiques? A cette fin, les deux seuls arguments envisageables semblent être les suivants :

1°) La grammaire opère sur des vocabulaires terminaux (les dictionnaires des langues naturelles) qui sont des données empiriques. 2°) La forme de la grammaire est déterminée par des données empiriques (les phrases attestées de la langue) qui se posent continuellement comme des obstacles dont l'assimilation est l'enjeu même du travail grammatical.

2.I.I. Au premier argument on opposera que la notion même de dictionnaire ne diffère pas de celle de vocabulaire terminal d'une grammaire formelle, dès l'instant où on lui impose les conditions suivantes qui semblent être implicitement retenues par la théorie des grammaires génératives:

a) Le dictionnaire est clos. Le dictionnaire ne se présente pas comme une matière empirique brute, mais comme un système -un système lexical - dont l'une des particularités est d'être clos et fini. Cette propriété de clôture, c'est à dire de détermination totale de ce qui est membre du dictionnaire, n'est d'ailleurs pas une réduction propre à l'approche générativiste. Elle correspond plutôt à la nécessité, formulée par Saussure avec les notions de langue et de synchronie, de délimiter provisoirement l'aire d'étude des réseaux de relations qu'entre-

tiennent les éléments d'une structure linguistique. De cette réduction saussurienne on peut dire qu'elle assimile le dictionnaire à un objet formel autonome.

Cependant une telle assimilation ne permet pas d'identifier le dictionnaire au vocabulaire terminal d'une grammaire formelle car une différence essentielle subsiste entre ce dernier et un système lexical. Le vocabulaire terminal d'une grammaire formelle est démunie de ce réseau de relations structurantes qui, dans la perspective saussurienne, constitue l'objet même de l'étude linguistique. Les fonctions significatives qui, dans le cas de l'étude structuraliste, assurent la liaison avec l'empirie sont évacuées d'un tel vocabulaire. Celui-ci est un objet abstrait⁽³⁾.

b) L'étiquetage catégoriel du dictionnaire. Cependant, dire que le vocabulaire d'une grammaire formelle est un objet abstrait ne revient pas à affirmer que celui-ci est réduit à une simple liste de symboles dépourvue de tout caractère structurant ayant un air de famille avec des propriétés du lexique. Le vocabulaire hérite d'une propriété du dictionnaire: comme pour ce dernier ses éléments possèdent une étiquette catégorielle. Cette propriété permet de partitionner en classes d'équivalence le dictionnaire comme le vocabulaire terminal d'une grammaire formelle. La grammaire générative, elle, ne requiert rien de plus de son vocabulaire terminal que ces deux propriétés de clôture et de catégorisation, mais peut-on pour autant, à ce stade, identifier les deux types de vocabulaires? Une dif-

férence apparente semble nous en empêcher. La partition du vocabulaire d'une grammaire formelle s'exprime par et seulement par un certain nombre de productions, appelées règles terminales, de la forme $A \rightarrow \begin{Bmatrix} a_1 \\ : \\ a_n \end{Bmatrix}$ qui en détermine extensionnellement la catégorisation. La catégorie d'un symbole terminal, c'est le symbole de gauche de la règle dans laquelle il apparaît.

Dans le cas de la grammaire générative, si une telle représentation demeure par principe réalisable, elle n'est pas pour autant nécessaire. Car ce qui indique la catégorie d'un élément c'est l'information contenue dans le dictionnaire. Mais ce recours au dictionnaire ne signifie pas que la grammaire générative soit élaborée en interférence avec le lexique, puisque le propre du dictionnaire ce n'est pas qu'il classe les mots en catégorie, mais plutôt qu'il les a classés. C'est là la seule incidence du lexique sur le vocabulaire d'une grammaire générative et elle ne peut réintroduire de connexions avec l'empirie car la catégorisation ainsi conçue n'est pas problématisée, elle est simplement donnée a priori.⁽⁴⁾

Si la catégorisation du vocabulaire terminal d'une grammaire générative semble admettre une représentation différente de celui d'une grammaire formelle, cette différence pourrait donc bien reposer sur une équivoque. La grammaire générative opère sur des symboles dont la forme est identique à ceux du dictionnaire. Lorsque la grammaire manipule un de ces symboles, par exemple "chat", elle tient compte dans cette manipulation d'une propriété formelle évoquant un des traits de la définition

du mot "chat" par le dictionnaire, ce trait étant généralement la catégorie du nom (mais il peut aussi être plus complexe, par exemple "nom animé" ou "nom animé masculin", etc...). Mais "chat" n'est que l'homonyme par lequel on désigne tantôt un des éléments pouvant figurer à droite dans la règle " $N \rightarrow \dots$ ", tantôt le "nom masculin; petit mammifère à poil doux, aux yeux oblongs et brillants, à oreilles triangulaires." Ce qui semble être une représentation propre du vocabulaire d'une grammaire générative (le dictionnaire) n'est qu'un moyen commode de se passer d'une représentation essentielle, le sous-ensemble des productions réécrivant les symboles terminaux, dont l'englobement même interdit l'usage. Nous voulons simplement dire par là que soutenir que grammaire formelle et grammaire générative sont distinctes parce que la dernière use d'un dictionnaire dont la détermination est empirique nous semble reposer sur une illusion. Mais nous avons proposé un autre argument pour distinguer grammaire formelle et grammaire générative, ou plus exactement pour ranger la grammaire générative aux côtés des disciplines empiriques.

2.I.2. Les grammaires génératives des langues naturelles sont en perpétuelle révision. Nous ne pensons pas seulement à ces révisions profondes qui amènent les linguistes à modifier les rôles respectifs des composants transformationnels et syntagmatiques, ou des composants syntaxiques et sémantiques, ou des composants syntaxiques et phonologiques, etc... et qui posent

les jalons de la brève histoire de la grammaire transformationnelle. Nous pensons plutôt aux révisions qu'entraîne la recherche quotidienne en grammaire générative et à cet égard celle-ci semble procéder d'une manière qui n'est pas coutumière aux sciences exactes. Dans le cas de ces dernières, la situation semble être la suivante. Lorsqu'un savant propose une nouvelle loi, la communauté scientifique applique ou essaye d'appliquer cette loi aussi souvent que possible et ceci jusqu'au point où la justesse de la loi lui semble douteuse, stade auquel on peut alors songer à lui substituer une nouvelle loi ou un nouveau corps de lois. Pour paraphraser dans les termes les plus simples et les plus concrets qui soient cette description très sommaire de l'activité scientifique, on pourrait présenter les choses ainsi: lorsqu'un physicien expose les résultats de ses travaux, ce n'est pas toujours pour proposer une nouvelle théorie physique, c'est même rarement le cas. La situation est sensiblement différente en grammaire générative où la circonstance qui, par excellence, justifie la présentation de nouveaux travaux est, généralement, la proposition d'une nouvelle hypothèse qui soit se substitue à une hypothèse préexistante, soit met en péril pour des raisons de cohérence interne, un système global de règles.

Une telle situation donne à réfléchir. Si on devait l'interpréter on pourrait, par exemple, en tirer un jugement critique déplorant ce penchant pour la théorie à tout prix ou, moins conventionnellement, en déduire que la grammaire se doit d'avoir

une démarche qui lui soit propre et qu'il est absurde de supposer qu'il existe pour la science des méthodes générales et universelles. Mais ce n'est pas le problème qui nous occupe ici⁽⁵⁾. La seule chose que nous aimerions noter, et qui est vraiment indubitable, c'est la dépendance extrêmement forte existant entre la détermination de la forme de la grammaire et ce que les linguistes appellent "les faits". Car ce qui pousse à modifier la grammaire, c'est bien la volonté de rendre compte de faits inattendus, c'est à dire d'engendrer des phrases imprévues. Nous avons dit "imprévues" et non pas "inobservées". Nous aurions dit "inobservées", notre discussion aurait pu s'arrêter là, car une science qui essaye de rendre compte d'observations est une science empirique et alors nous aurions eu un critère satisfaisant pour distinguer grammaire générative et grammaire formelle: la grammaire générative s'occupe d'observations empiriques et la grammaire formelle s'occupe d'autre chose. La différence principale entre la prévision et l'observation est que la seconde est limitée matériellement tandis que la première ne l'est pas. Ce qui limite matériellement nos observations c'est soit les insuffisances de nos instruments d'observation, soit l'absence de ces instruments. Les hypothèses des sciences empiriques ne visent jamais à rien d'autre qu'à imaginer ce qu'on ne peut pas observer ou ce qu'on ne pourra jamais observer (lorsqu'on s'intéresse, par exemple, aux événements du passé). Or la grammaire générative n'est pas affectée par des limites de cette nature, car aucun des "faits" qui intéressent le grammairien générativiste ne lui est a priori inac-

cessible. Ce n'est pas là une propriété inhérente à la définition de la grammaire générative mais plutôt la conséquence d'un principe épistémologique lié à ce que les générativistes, et Chomsky en particulier, estiment être un des objectifs fondamentaux d'une théorie linguistique: "L'objet premier de la théorie linguistique est un locuteur-auditeur idéal, appartenant à une communauté linguistique complètement homogène, qui connaît parfaitement sa langue et qui, lorsqu'il applique en une performance effective sa connaissance de la langue n'est pas affecté par des conditions grammaticalement pertinentes..."?

Trad.
Milner

(Chomsky, 1965; p.12). Cette conséquence consiste en l'élimina-

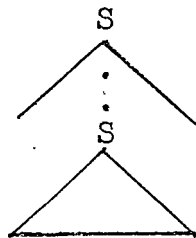
non
séquent

~~tion du recours à des tests comportementaux pour juger de ce~~
qui est un énoncé attesté de la langue. La notion de locuteur-auditeur idéal n'étant qu'un terme transitoire qui permet au grammairien de rester le seul maître dans la décision de ce qui est donné pour grammatical. Il n'y a d'ailleurs rien de contestable dans ce principe puisque le grammairien générativiste se distingue des autres grammairiens par le fait qu'il ne prétend à rien d'autre qu'à décrire son idiolecte.

Mais le "grammatical" dépasse les limites de ce qui est accessible par la seule introspection puisqu'il concerne un ensemble infini d'éléments. Aussi apparaît une contradiction entre cette situation et l'affirmation que nous faisons à l'instant sur l'accessibilité a priori des énoncés intéressant le linguiste. Cependant cette contradiction n'est qu'apparente car le mot clé de notre affirmation est "intéressant". Pour en pré-

ciser le sens nous pourrions dire que la classe des énoncés qui intéresse le générativiste est en gros identique à la classe d'énoncés qu'un grammairien décryptant une langue étrangère serait amené à constituer en recourant à des tests comportementaux. Et s'il fallait donner un sens moins approximatif à "intéressant", on pourrait dire qu'il qualifie la classe des énoncés à partir de laquelle on définit le "grammatical" comme notion exportable aux énoncés inaccessibles à l'observation. Ce qui nous permet de dire que cette seconde caractérisation est plus précise, c'est qu'il est possible de lui donner un contenu technique bien délimité.

Une grammaire générative fait appel à des procédures récursives qui lui permettent de déterminer sur une base non seulement finie, mais encore restreinte (dans un sens que nous préciserons) la nature et la forme de ses mécanismes descriptifs. Théoriquement, elle devrait pouvoir se limiter à l'étude des énoncés dont les indicateurs syntagmatiques ne comprennent pas plus de deux occurrences d'un même noeud récursif dont l'un domine l'autre. Car si toutes les configurations de la forme



sont analysées, les conditions de bonne formation qui en seront déduites vaudront pour les cas où un autre noeud S serait enchassé sous le second, en vertu du fait que le sous-arbre en-

raciné au deuxième S se présentera comme une des configurations déjà étudiées dans le premier cas⁽⁶⁾. Cette analyse reste, bien entendu, schématique puisque la grammaire peut contenir différentes catégories récursives enchassées en alternance.

Ce qui garantit l'homogénéité d'une notion de grammaticalité exportable au-delà de l'accessible intéressant c'est, si l'on veut, une espèce de principe d'induction qui régit l'usage des mécanismes récursifs de la grammaire et dont l'examen exhaustif des structures élémentaires que nous venons de caractériser constitue la base.

Ces considérations nous permettent d'imaginer quelle serait la tâche d'un grammairien tout aussi idéal que le locuteur-auditeur qui l'occupe en dressant l'organigramme de la procédure grâce à laquelle il peut accéder à tous les énoncés intéressants de sa langue :

- I°) Choix d'un corpus restreint d'énoncés grammaticaux au nom du principe que le grammairien peut porter des jugements de grammaticalité.
- 2°) Descriptions grammaticales des énoncés retenus en I°) (attribution d'indicateurs syntagmatiques); élimination des énoncés dont l'indicateur syntagmatique a une profondeur supérieure à 2.⁽⁷⁾
- 3°) Grammaire générative des énoncés retenus en 2°) en conformité avec les descriptions de 2°). (La grammaire générative induite du 2°) étant une grammaire de constituants, elle

peut alors être simplifiée et ramenée à une grammaire transformationnelle faiblement équivalente).

4°) Vérification de la grammaire 3°) par le parcours total du vocabulaire; si un élément du vocabulaire n'entre dans aucune des configurations de 2°), retour à 2°) et enrichissement des descriptions grammaticales soit par l'adjonction de nouvelles catégories, soit par l'adjonction de nouvelles configurations syntaxiques.

Remarquons qu'une telle procédure n'est que l'interprétation des conséquences extrêmes de deux principes pris conjointement: le grammairien doit utiliser des mécanismes récursifs pour décrire à l'aide de moyens finis un objet infini et le grammairien est habilité à porter des jugements de grammaticalité.

La place importante laissée à l'heuristique dans une telle procédure est sans doute le trait qui contribue le plus à l'assimiler à une démarche empirique. En particulier l'étape 4°) de vérification exhaustive des règles sur la totalité du vocabulaire entretient cette confusion de manière tenace, puisqu'elle opère sur un plan constamment tangent à l'empirie. Mais justement cette heuristique n'est pas orientée vers le décriptage de l'empirie mais bien plutôt sur elle-même. Le grammairien, tel que nous le décrivons, est un super-automate engagé dans un processus homéostatique dont le but, pour reprendre le mot de Quine, se définit par ce qu'il fait quand il va vers ce but⁽⁸⁾.

Les sujétions auxquelles est astreint le grammairien dans

ce travail sont d'une nature bien différente de celles auxquelles est soumis le savant qui étudie la Nature. Jamais le grammairien, lorsqu'il suit la procédure de la page précédente, n'est amené à faire d'hypothèses sur la grammaticalité des énoncés que son système engendrera, car à chaque instant il dispose de tous les moyens nécessaires pour vérifier la correction de ses règles. On pourrait en conclure que si les règles proposées par les générativistes sont si fragiles, c'est que ceux-ci font mal leur travail, que contrairement au grammairien idéal de la procédure, ils escamotent la vérification de l'étape 4°). Mais ce serait perdre de vue qu'il y a là un choix pragmatique et que le grammairien réel, à l'opposé du grammairien idéal, non seulement veut arriver à son but, mais veut également y arriver le plus vite possible. Or à ce niveau d'exigences le choix entre l'attitude laborieuse et l'attitude hypothétique ne peut être fait rationnellement et il semble légitime de se laisser guider par le principe du moindre effort.

La principale réserve que l'on pourrait faire devant cette conception de la grammaire porte sur son utilité. En effet nous avons rappelé plus haut que le grammairien générativiste ne s'intéressait qu'à son idiolecte. Or l'adjectif possessif "son" peut prendre dans ce contexte la valeur la plus restrictive qui soit pour ne plus renvoyer à la langue de la communauté linguistique du grammairien, mais véritablement à une connaissance individuelle non objectivée qui se réduirait à la liste des énoncés que le grammairien, en tant que locuteur indifférencié, recon-

*Ca te dést
p. 35
deux*



naîtrait comme grammaticaux. Aussi, quel peut être l'intérêt de "grammaires privées" qui ne visent qu'à décrire non plus une norme à laquelle est soumise la collectivité, mais une norme régissant le comportement individuel alors que c'est en tant qu'il s'écarte de la norme que le comportement individuel devient un sujet d'étude pertinent? C'est que justement, procédant ainsi, la théorie des grammaires génératives fait l'économie d'un pseudo-objet de langue; celle-ci ne prenant corps qu'à partir du moment où s'établit un consensus sur les jugements de grammaticalité délimitant la langue comme l'intersection des idiolectes de chaque locuteur-auditeur. Alors que Saussure et les structuralistes posent l'existence de la langue et la proposent comme objet de connaissance, dans le cas de la linguistique chomskyenne, un tel réquisit n'est pas formulé et ce sont les grammaires génératives qui deviennent constitutives de la langue. Ce n'est pas l'économie ontologique en tant que telle qui est intéressante dans cette dernière perspective, mais plutôt la clarification qu'elle entraîne. Dire qu'il existe un objet langue comme norme à laquelle s'ordonne toute parole, c'est prêter le flanc trop ouvertement à une critique historico-sociologique pour laquelle la dimension sociale du langage reste sa composante principale et qui veut voir en toute norme la mise en place d'une subjectivité masquée. Parler de la langue française c'est, selon ce point de vue, parler d'une langue française celle de la bourgeoisie, des manuels scolaires, de la télévision... Cette critique relève d'ailleurs une difficulté métho-

dologique effective, car la question des niveaux de langue se pose constamment en linguistique descriptive et elle devient même cruciale lorsqu'on a l'ambition de délimiter ce qui est dialectal de ce qui ne l'est pas.

Cette limite ne remet pas en cause la dichotomie saussurienne mais elle fait apparaître que sa valeur est avant tout négative : la langue s'oppose à la parole car elle est dépourvue des caractéristiques qui réglementent l'usage du langage, caractéristiques qui ne doivent d'ailleurs pas être réduites à de seules propriétés accidentelles. Cependant formulée ainsi, la dichotomie ne peut plus être considérée comme l'étape d'une élaboration conceptuelle, elle n'est qu'un précept méthodologique, et s'il est vrai, comme on l'a souvent rappelé, que la distinction performance-compétence évoque la dichotomie saussurienne, ce ne peut être que par cet aspect méthodologique.

En conclusion, les deux arguments que nous avons envisagés pour distinguer la grammaire générative de la grammaire formelle, l'usage du dictionnaire et la sujétion aux énoncés observables, doivent être écartés car leur matérialité, à la manière du relief des trompe-l'oeil, s'évanouit dès que notre regard s'en approche. D'autres types de liens avec l'empirie n'auraient-ils pas pu être envisagés puisque, par exemple, la théorie des grammaires génératives est conçue comme une doctrine psychologique visant à expliquer la nature du comportement verbal? Nous pensons que ce n'est pas le cas, car cette fonction explicative dont veut se charger ici la théorie linguistique échappe à l'ac-

tivité proprement grammaticale. Elle est à la grammaire ce que serait une théorie de l'intelligence mathématique à l'algèbre linéaire. Et si, par une confusion injustifiable, la grammaire se trouvait investie de l'étude de processus psychologiques, celle-ci ne pourrait être conduite, en dernière analyse, qu'avec les méthodes de la grammaire, ce qui nous ramènerait aux cas de figures étudiés plus haut.

2.2. L'abstraction de la grammaire générative

Nous sommes donc toujours à la recherche d'une propriété spécifique distinguant la grammaire formelle de la grammaire générative. Mais avant de poursuivre nous devons nous arrêter sur une conséquence étrange du raisonnement présenté dans les pages précédentes. Au début de celui-ci nous constatons que la grammaire formelle était une théorie mathématique et qu'il suffisait de voir que la grammaire générative était une discipline empirique pour pouvoir distinguer clairement ces deux disciplines. Notre conclusion est donc que la grammaire générative est de la mathématique puisque, étant donnée son indépendance à l'égard des faits observables, il est impossible de l'assimiler à une science empirique. Une telle déduction est paradoxale car à l'évidence la grammaire générative ne présente aucun des signes extérieurs de ce que l'on reconnaît généralement comme étant de la mathématique. En effet, de tels signes n'apparaissent pas dans la forme de la grammaire puisque n'y figurent ni symbolisme élaboré, ni proposition universelle dont la validité reposerait sur la seule cohérence des raisonnements démonstratifs qui y mènent.

Comment nuancer alors notre conclusion si nous ne voulons pas nous opposer à cette évidence manifeste? Peut-être en suggérant que la pensée grammaticale est proto-mathématique. Car aux mathématiques elle emprunte deux attributs fondamentaux: d'une part avec une insistance obsessionnelle, elle recourt à la formalisation, d'autre part elle est avant tout un calcul.⁽⁹⁾ Et ce qui retient d'assimiler la grammaire aux mathématiques, c'est sans doute le côté fragmentaire du formalisme et la nature incongrue et laborieuse du calcul. Car la particularité du calcul grammatical semble être son inconsistance. Chaque mise à l'épreuve des règles dans la dérivation de phrases imprévues révélant leur inadéquation et entraînant leur modification. L'existence de règles préétablies valables aux différentes étapes du calcul, pour chaque variable catégorielle, fait défaut dans la grammaire générative où, contrairement à la grammaire formelle, la délimitation du domaine d'application des règles apparaît soit comme trop précaire, soit insuffisamment abstraite. Cet aspect des choses nous conduit à envisager un autre critère pour distinguer grammaire formelle et grammaire générative : toutes deux sont des calculs, mais la grammaire formelle est un calcul opérant sur un plan d'abstraction plus élevé.

En quoi consiste cette abstraction supérieure de la grammaire formelle? La réponse semble immédiate: les règles d'une grammaire formelle comportent des variables dont les valeurs sont les symboles catégoriels des grammaires génératives. Par exemple, une règle de grammaire formelle de la forme $A \rightarrow BC$

représente le canevas dans lequel peuvent prendre place les triplets suivants de symboles catégoriels (P,SN,SV), (SN, Det,N), etc... Il est vrai que l'on peut métaphoriquement désigner à leur tour ces symboles catégoriels du nom de variable puisqu'ils ont une fonction de représentant de classe syntaxique et que l'on peut leur substituer des symboles de ces classes. Mais il est clair que la réalisation de ces substitutions s'accomplit par l'usage des règles de grammaire (formellement par l'expansion d'une branche de l'indicateur syntagmatique) et qu'elle est par conséquent d'une nature différente des substitutions que l'on opère dans le schéma de construction de la grammaire formelle $A \longrightarrow BC$. Telle est la raison qui permet de déterminer des conditions générales de calculabilité indépendantes des instances particulières liées à chaque grammaire générative. Malheureusement un tel critère, malgré une évidence qui paraît incontestable, ne donne pas entière satisfaction et ceci pour deux motifs:

I°) Généralement, les grammaires génératives comprennent des règles mixtes où figurent, aux côtés des termes catégoriels, des variables de catégories identiques à celles de la grammaire formelle. Cette difficulté pourrait être éliminée en remarquant que justement les règles de la grammaire formelle ne comportent que des variables. Mais une telle réponse ne pourrait être retenue que dans le cas où le rôle des variables dans la grammaire générative serait inexistant ou négligeable comme c'est par exemple le cas dans la formulation classique de la transforma-

tion passive. Or on sait qu'il n'en est rien puisque la présence simultanée de variables et de catégories dans la description structurale d'une règle transformationnelle peut donner lieu à des contraintes très précises dans la dérivation et ceci pour des raisons strictement grammaticales, comme le montre, par exemple, la fameuse étude de Ross (1967).

2°) La distinction entre variable de catégorie (symbole non-terminal d'une grammaire formelle) et catégorie (symbole non-terminal d'une grammaire générative) n'est pas toujours très claire. En effet, les grammairiens ont parfois recours à des symboles non-terminaux dont le statut semble être intermédiaire entre les deux types à peine cités. C'est le cas, notamment, dans les modèles grammaticaux faisant usage de la convention $\bar{X}$ où, dans le but d'uniformiser la notation, sont introduites des catégories dont la nomenclature dépend uniquement de la catégorie "majeure" qui les domine (nom, verbe, adjectif) et de leur place dans l'indicateur syntagmatique au détriment de leur "contenu". Le statut de ces nouveaux symboles est donc déterminé d'un côté par des conditions purement formelles, leur position dans les règles, de l'autre par la nature des catégories grammaticales, au sens traditionnel, les environnant⁽¹⁰⁾.

La mixité des règles et des symboles présents dans les grammaires des langues naturelles s'opposent à une distinction entre grammaire formelle et grammaire générative fondée sur une plus grande abstraction de la première. Cependant la manière

dont nous avons analysé le caractère mixte des catégories de la convention $\bar{X}$ nous permet d'envisager un ultime critère distinctif que nous croyons satisfaisant.

2.3. La grammaire formelle est une syntaxe pure

Ce qui rapproche les catégories utilisées dans la convention $\bar{X}$ des variables de la grammaire formelle, c'est que leur nature est partiellement déterminée par leur position dans les indicateurs syntagmatiques. Si toutes les propriétés de ces catégories dépendaient uniquement de leur position, alors rien ne les différencieraient des variables de la grammaire formelle. Mais nous avons vu que ce n'était pas le cas, car elles héritent de leur voisinage un contenu régissant également leur manipulation. Dès lors nous pourrions dire que la grammaire formelle traite de syntaxe pure et que c'est là ce qui la distingue de la grammaire générative, laquelle traite de pseudo-syntaxe ou plus exactement de sémantico-syntaxe.

Cependant en adoptant un tel critère nous sommes redevables d'une définition de la syntaxe et nous proposerons -très classiquement, croyons nous - la suivante: "La syntaxe est l'étude des règles de combinaison des mots en phrases". Pour obtenir une définition d'origine linguistique contrôlée, il suffit de transposer la précédente dans la terminologie de la grammaire générative en remplaçant "mots" par "éléments du vocabulaire" et phrases par "éléments du langage".

A ce stade, on objectera sans doute que les règles syntaxiques typiques de la grammaire générative satisfont parfai-

tement cette définition comme le fait, par exemple, la règle bien connue $SN \rightarrow Art N$. Mais cette objection n'est pas sérieuse car si on considère cette seconde définition de la syntaxe: "La syntaxe est l'étude des mécanismes grâce auxquels le langage exprime les relations sémantiques unissant les différents éléments de la phrase", on constate que les règles de la grammaire générative la satisfont tout autant. Ce n'est évidemment pas le cas pour la grammaire formelle puisque cette définition n'a aucun sens pour elle. Si la première définition semble s'appliquer au cas de la grammaire générative, c'est simplement parce qu'y figure, si on peut dire, une omission. Cette omission consiste en la non spécification du contenu sémantique des éléments du vocabulaire de la grammaire générative. Notons que notre recours au contenu sémantique n'est pas contradictoire avec l'argumentation développée en 2.I.I., car l'idée que nous défendons dans ce paragraphe est que la présence de propriétés sémantiques dans les éléments du vocabulaire ne permet pas de relier directement la grammaire à l'empirie. La présence du contenu sémantique étant une condition parmi d'autres que le calcul grammatical doit respecter.

En dernier lieu, on pourrait objecter qu'attribuer un contenu sémantique aux catégories relève d'une décision arbitraire et que celles-ci sont plutôt déterminées à partir d'une analyse distributionnelle de type harrissien où la sémantique n'a rien à faire. Toutefois, en arguant de l'origine purement distributionnelle de la classification donnée par le diction-

naire, on néglige deux points: 1°) Au principe de l'étude distributionnelle figure nécessairement le recours au sens. Sinon l'identification des morphèmes serait impossible, 2°) La délimitation des catégories^(II) qui interviennent dans le calcul grammatical doit être achevée puisque l'enjeu de ce calcul est de décrire l'une de ces catégories, la phrase, en fonction d'un choix pré-déterminé de termes descriptifs. Autrement dit, le vocabulaire de la grammaire doit être partitionné en classes closes. Or si ces catégories étaient déterminées distributionnellement, elles ne pourraient satisfaire un tel réquisit car les classifications distributionnelles, déduites à partir de l'examen d'un nombre restreint (ou en tout cas fini) d'occurrences, ne peuvent correspondre parfaitement aux nécessités d'un usage illimité. La classification des vocables dans le dictionnaire doit donc nécessairement dépendre d'une propriété indépendante de leur distribution^(I2).

Notre critère distinctif n'a évidemment aucune fonction exclusive, et il reste vrai que la grammaire formelle est une théorie mathématique des grammaires, comme il reste vrai que l'abstraction y joue un rôle bien plus important que dans la grammaire générative. Toutefois si, en tentant de développer jusqu'à leurs conséquences extrêmes les principes qui, croyons-nous, sont inhérents à la théorie des grammaires génératives, nous avons dégagé l'idée que la grammaire formelle traite de syntaxe pure, ce n'est pas aux fins de faire apparaître des conclusions paradoxales. Au contraire, celles-ci s'accordent

parfaitement avec un certain nombre de réflexions que nous serons amenés à faire sur la nature purement syntaxique des langages formels.

NOTES DU CHAPITRE II

- (1) A cet égard la position de Chomsky a subi d'importantes modifications, puisqu'à l'origine la grammaire était une branche des mathématiques, pour devenir ensuite une branche de la psychologie. Aujourd'hui c'est une branche de la biologie.
- (2) On pourrait objecter au propos général de cette section que puisque les grammaires génératives engendrent des langages formels alors, par définition, la grammaire générative est une activité mathématique. Mais il ne faut pas oublier que c'est la notion de langage formel qui découle de la définition de langage en général proposée par Chomsky (1956, p. 53) et non pas l'inverse. De fait, c'est dans l'unique souci de ne pas préjuger des propriétés du langage humain que Chomsky en donne une définition si générale. Et c'est la généralité de cette définition qui permet de considérer cette notion de langage comme une notion mathématique.
- (3) Par "fonction significative" nous n'entendons pas, bien sûr, les définitions du dictionnaire, mais l'organisation structurale du lexique de façon générale.
- (4) Tel quel ce raisonnement ne vaut pas pour toutes les versions de la grammaire générative. Dans certaines versions de la sémantique générative on a essayé d'exprimer des propriétés lexicales en termes de dérivations transformationnelles, ce qui laisse supposer que la grammaire participe à la constitution de la classification lexicale. Nous ne pensons pas que ces tentatives aient valeur de contre-exemple, et il est sans doute possible de les ramener au cas de figure que nous venons d'exposer. Cependant, pour fixer les idées et dissiper tout malentendu, posons que notre discussion ne vaut que

pour la théorie classique de Chomsky, Fodor, Katz et Postal

- (5) On nous pardonnera cette présentation délibérément ingénue et sommaire du travail scientifique. Notre intention est simplement de souligner le contraste existant à cet égard entre la grammaire générative et les sciences exactes.
- (6) Ces configurations sont celles des structures de surface. Notre raisonnement suppose qu'il est toujours possible de décrire la structure de constituants d'un énoncé retenu comme grammatical. Ce qui est également une hypothèse de la grammaire générative puisqu'une structure de surface produite au terme d'une dérivation transformationnelle est représentable par un graphe planaire.
- (7) Par profondeur d'un indicateur syntagmatique, nous entendons le maximum du nombre de noeuds récursifs figurant sur une branche allant de la racine aux noeuds terminaux.
- (8) Quine (1971, p. 39) Par certains aspects notre raisonnement rappellera celui de Quine (1953, pp. 47-64). Pour une discussion plus détaillée, voir notre chapitre V.
- (9) Un calcul dont les dérivations de la grammaire seraient les formules des phrases grammaticales. L'efficacité de ce calcul pourrait être décrite ainsi: si une formule - une dérivation - peut être appliquée dans un grand nombre de cas (pour un grand nombre de phrases) alors elle sera efficace. Evidemment, comme chaque phrase a au moins une dérivation unique, il faudrait disposer d'une définition formelle de quasi-identité des dérivations pour exprimer correctement une notion d'efficacité du calcul grammatical.
- (10) Ce type de catégorie a été introduit par Chomsky (1970). Jackendoff (1974) donne un exposé détaillé sur les usages de cette notion.
- (11) Harris reconnaît lui-même que les classes paradigmatiques

déterminées par l'analyse distributionnelle ne sont que des approximations dont les limites seraient des classes réduites à un élément unique. (Harris, 1957)

- (I2) Que substituer à l'analyse distributionnelle? Puisqu'à ce jour le problème des parties du discours n'a pas reçu de réponses satisfaisantes, autant en prendre son parti et procéder comme la grammaire générative qui, passant outre aux conclusions du structuralisme, emploie la nomenclature traditionnelle comme étiquetage conventionnel, sans cependant accorder une valeur heuristique à ces catégories. A ce titre, on ne peut manquer de relever le caractère paradoxal de certaines tentatives alternatives, comme celle de la grammaire catégorielle, dont la nomenclature catégorielle est fondée sur une analyse distributionnelle et qui, simultanément, prétendent donner une description finie de la langue infinie.

CHAPITRE III

LES FONCTIONS APPARENTES

DE LA GRAMMAIRE FORMELLE

3. La recherche d'un critère distinctif telle que nous l'avons conduite au cours du chapitre précédent serait superflue si elle n'avait pour but que de rendre compte de l'intuition d'une différence entre grammaire formelle et grammaire générative. Si une telle différence existe elle ne peut porter, en dernière analyse, que sur ce que l'on fait dans le cadre de ces deux théories. Nous proposerons donc dans ce troisième chapitre une première caractérisation du travail accompli par la théorie des grammaires formelles en attendant de mettre plus directement à l'épreuve notre critère distinctif au chapitre suivant. Cette caractérisation, cependant, n'aura rien de définitif puisqu'aussi bien est-ce un des buts globaux de notre travail de comprendre le pourquoi et le comment de la grammaire formelle, mais elle consistera à faire un premier tri parmi les fonctions dont on pourrait croire investie cette science.

Selon un point de vue assez répandu l'apport de la grammaire formelle se réduit à l'accomplissement de deux tâches principales: 1°) Décrire certains instruments d'analyse du langage comme des systèmes formels. 2°) Démontrer que parmi ces instruments - ces systèmes formels donc - certains sont impropres à la description des langues naturelles.

Indéniablement ces deux objectifs sont à l'origine du développement de la grammaire formelle. Mais dans quelle mesure sont-ils atteints et surtout jusqu'à quel point s'agit-il de tâches possibles pour la grammaire formelle? Voilà le type de

questions auxquelles il faut répondre dans un premier temps si on veut évaluer correctement la fonction des méthodes mathématiques de la grammaire.

3.I. Grammaire et système formel

La distinction principale que nous développerons dans notre commentaire sur la première tâche de la grammaire formelle à trait à la localisation de la propriété permettant d'assimiler une grammaire générative à un système formel. Il s'agit en particulier de voir si cette identification a son origine dans des contraintes imposées à la forme de la grammaire, et dont la grammaire formelle procurerait la théorie, ou si l'activité grammaticale ne recèle pas en elle-même certaines exigences qui la conduisent inévitablement à être un système formel.

3.I.I. Les grammaires génératives sont souvent comprises comme des grammaires particulières dont soit la forme (dispositif logico-mathématique), soit la fonction (déterminer les conditions de bonne formation de toutes, et rien que, les phrases grammaticales d'un langage), les distingueraient à la fois des grammaires traditionnelles et des grammaires structuralistes. Cependant, si on prend à la lettre la définition de "grammaire générative" telle que la donne Chomsky dans Syntactic Structures, rien ne permet de faire ces distinctions. Une interprétation stricte de cette définition technique permet de qualifier une grammaire de générative dès qu'elle satisfait les trois conditions suivantes: finitude, exhaustivité, clarté. Il nous semble

inséparable

qu'une telle définition peut par principe être étendue à toute grammaire, qu'il s'agisse d'un manuel d'usage ou d'une étude descriptive de la langue. Mais qu'entend-on ici par clarté? Simplement l'univocité des informations contenues dans la grammaire lorsqu'elle prescrit, décrit ou produit les phrases qu'elle estime appartenir à la langue. Ces informations peuvent être de nature variée : morphologique, syntaxiques, sémantiques, etc..., elles peuvent même porter sur les circonstances de l'énonciation, la connaissance du monde, la chose à laquelle se réfère l'énonciation, sans que soit modifiée l'idée de clarté. La clarté, c'est la possibilité d'isoler des instructions précises à partir de ce que dit la grammaire. L'exhaustivité, quand à elle, est une condition qu'aucune grammaire ne satisfait sans doute vraiment (y compris les "vraies" grammaires génératives), mais que toutes ont vraisemblablement l'ambition de satisfaire. Ajoutons que l'exhaustivité doit être comprise non seulement comme la requête de décrire toutes les phrases du langage, mais encore rien que celles-ci. La générativité, au sens où nous l'avons précisé à la page précédente, au lieu d'être un attribut de grammaires particulières fait partie de la définition générale de la grammaire. (I)

Parmi les conditions de générativité, une seule semble spécifique, c'est le principe d'exhaustivité de la grammaire. Or ce principe suppose de la part du grammairien qui l'applique un dogme qui lui est inhérent; à savoir qu'il est possible de décrire les langues naturelles. On peut transformer ce dogme en

hypothèse lorsqu'on pose le problème, comme l'a fait Chomsky⁽²⁾, en termes mathématiques. Rappelons rapidement l'argument de Chomsky. Les grammaires sont finies (elles comprennent un nombre fini de règles composées d'un nombre fini de symboles), donc il n'y a qu'une quantité dénombrable de grammaires. Les langages sont les parties du langage universel formé sur l'union (finie) des vocabulaires possibles; le langage universel étant infini il y a une infinité non-dénombrable de langages. Il existe donc une quantité non dénombrable de langages pour lesquels il n'existe pas de grammaires. L'hypothèse première de toute théorie de la grammaire est alors que les langues naturelles font partie des objets privilégiés pour lesquels il existe une description possible. On peut interpréter cette hypothèse d'une autre manière. La générativité de la grammaire entraîne sa représentabilité par une machine de Turing, puisque les conditions de clarté et de finitude permettent de décomposer les règles de chaque grammaire en un ensemble fini d'instructions. La classe des langages pour lesquels pourrait exister une grammaire est alors identifiable à la classe des ensembles récursivement énumérables⁽³⁾.

3.1.2. En montrant que la grammaire est nécessairement soumise à cette problématique nous voulons simplement mettre en évidence qu'il y a au coeur de la recherche grammaticale, de quelque obédience qu'elle soit, des motifs (totalement indépendants des choix épistémologiques des linguistes) justifiant une étude mathématique des théories grammaticales. En poussant notre raisonnement jusqu'à ses limites extrêmes, nous pourrions même dire

que la construction du modèle mathématique d'une grammaire est une nécessité. En particulier, la construction d'une grammaire empiriquement satisfaisante, au sens où les descriptions qu'elle fournirait correspondraient à une réalité linguistique quelconque, ne devrait être envisagée qu'une fois établie l'existence d'un système de réécriture illimité produisant les phrases de la langue, même si l'organisation de ce dernier devait être ad hoc et sembler linguistiquement immotivé⁽⁴⁾.

On pourrait objecter que ces considérations n'ont de sens que dans le cas où on attribue à la langue la propriété d'être infinie. Il est vrai qu'en refusant l'idée de l'infinitude de la langue, l'existence d'une description finie devient immédiatement possible, puisqu'il suffit de donner la liste des phrases de la langue pour l'obtenir. Mais si en établissant cette simple liste on satisfait aux conditions de générativité, il est difficile de dire pour autant qu'on a construit une grammaire. Nous voulons dire par là que la définition de la grammaire générative donnée par Chomsky présuppose implicitement que, lorsqu'on l'applique aux langues naturelles, on l'applique à quelque chose d'infini. Et dans ce cas, la caractérisation finie de la langue infinie doit faire intervenir des règles de construction. Il paraît donc légitime d'ajouter à la définition donnée plus haut une quatrième condition que l'on pourrait énoncer ainsi : "... et qui contient des règles". Autrement dit lorsque nous posons la question de la possibilité de la grammaire, c'est à l'existence d'une procédure spécifiant

l'appartenance au langage que nous pensons. Ceci revient à peu près à soutenir que, pour qu'une grammaire soit générative, il faut encore qu'elle ne soit pas réduite à un ensemble vide de propositions. Une fois la grammaire ainsi conçue, la finitude du langage ne permet plus de faire l'économie de l'hypothèse de calculabilité, car comme la théorie des fonctions récursives nous l'enseigne, le problème du "membership" pour les ensembles finis n'est pas nécessairement décidable⁽⁵⁾.

Bien que la grammaire ne soit pas l'unique manière d'aborder l'étude du langage il vaut la peine d'observer que des approches apparemment bien distinctes de celles de la grammaire peuvent lui être réduites dans la mesure où elles entrent dans la problématique de leur possibilité.

Par exemple une des démarches les plus courantes de la linguistique structuraliste consiste non pas à décrire les conditions de bonne formation des énoncés, mais plutôt à étudier le langage sur un plan fonctionnel en tant que système de communication. Dans cette perspective, la priorité de la recherche ne va plus à la structure syntaxique de la phrase mais aux principes organisateurs qui, dans la langue, assurent le codage de l'information et sa transmission sous une forme optimale. Le but de la linguistique est alors de décrire les circonstances permettant à l'acte de communication de se réaliser. Mais cet objectif qui consiste à déterminer les codages possibles d'un signifié dans des formes signifiantes, retombe de nouveau sur le problème de la séparation entre formes possibles et impossibles.

retrouvant ainsi la perspective de la grammaire sous son aspect le plus classique. A moins de décider de prendre comme donnée la communication effectivement réalisée et de se garder d'analyser les conditions explicites sous lesquelles cette dernière n'est pas réalisable. Mais peut-on raisonnablement s'astreindre à cette seule tâche si on a l'ambition de comprendre le fonctionnement du langage?

Harris 3.1.3. Nous avons vu que la façon dont Chomsky avait posé le problème de la grammaire permettait de transformer un a priori sur la possibilité de la grammaire en une hypothèse. Quel traitement la théorie des grammaires génératives réserve-t-elle à cette hypothèse? Le paradoxe de la théorie des grammaires génératives est sans doute qu'elle affirme la vérifiabilité de l'hypothèse de calculabilité tout en affirmant simultanément que cette vérification ne mérite pas d'être faite:

9 { "Evidemment un programme de calculatrice qui parviendrait à engendrer les phrases d'une langue ne serait en lui-même d'aucun intérêt scientifique s'il ne parvenait à jeter quelque lumière sur le type de caractéristiques structurales qui distinguent les langues des ensembles récursivement énumérables. Tant que nous n'avons dit, de la grammaire d'une langue naturelle, qu'elle n'est qu'un système de réécriture non limité, nous n'avons encore rien dit d'intéressant". (Chomsky, 1963, §2)

L'hypothèse de calculabilité n'est donc pas intéressante en soi, seule la forme de la grammaire mérite d'être étudiée. Cette appréciation de Chomsky n'est pas désavouée par l'orienta-

tion actuelle de la grammaire transformationnelle où non seulement ce problème de fondement n'est jamais abordé, mais où de plus la question de la forme précise de la grammaire est reléguée au second plan face à l'étude de propriétés plus générales des systèmes grammaticaux (Principe A sur A, Condition du sujet spécifié, etc...) comme si il était entendu que la construction d'une grammaire ne pose que des difficultés de détails. Du dogme de la descriptibilité nous sommes passés à son équivalent strict, le dogme de la calculabilité, puisqu'on n'essayera jamais de vérifier l'hypothèse de calculabilité. On peut avancer deux explications, non exclusives d'ailleurs, pour comprendre cet acte de foi. La première est à rechercher au niveau de l'éthique du grammairien qui persiste, malgré les échecs successifs des diverses tentatives de description de la langue, à considérer celle-ci comme un objet d'étude raisonnable. La deuxième explication, c'est que l'homme parle une langue infinie, qu'il utilise des moyens finis pour le faire, et que l'on devrait pouvoir, en principe, décrire ce que l'homme peut dire par des moyens finis. C'est plus ou moins ainsi, c'est à dire en supposant que l'esprit humain fonctionne comme une machine de Turing, que les générativistes se déchargent du problème. La question est de savoir si ce transfert sur le plan psychologique élimine la difficulté ou la complique. En effet, dans le cas de l'hypothèse de calculabilité, on voyait très bien quelle était la démarche à suivre pour tester l'hypothèse, tandis que l'on est vraiment indécis sur la procédure qui pourrait mener à la

vérification, ou à la réfutation, d'une telle hypothèse psychologique.

3.1.4. L'identification des grammaires génératives à des systèmes formels repose donc sur des propriétés inhérentes aux grammaires, et on ne saurait voir dans celle-ci la conséquence d'un choix épistémologique impliquant une conception particulière soit de la nature du langage, soit des rapports de la linguistique et des mathématiques. Ce qui ne veut pas dire que la grammaire formelle ne joue aucun rôle dans cette identification, puisqu'elle fournit les instruments conceptuels qui permettent d'une part de la formuler, et d'autre part de la localiser à travers des formes de la grammaire apparemment étrangères à cette problématique. Car il faut garder présent à l'esprit que l'identification d'une grammaire à un système formel n'est pas une simple possibilité théorique, mais qu'il s'agit d'une tâche pratiquement réalisable comme le montre l'étude de Postal (1964a)

D'ailleurs si cette conception algorithmique de la grammaire était la conséquence directe d'un parti-pris, une telle intention serait contradictoire avec les principes épistémologiques explicitement affichés par Chomsky. On voit mal, en effet, comment réconcilier cette vue mécaniciste, qui réduit l'homme parlant à un automate, avec une philosophie du langage résolument anti-behavioriste, centrée sur la créativité linguistique au sens où l'entend Chomsky.

3.2. La théorie des grammaires formelles comme instrument de sélection des grammaires génératives

Venons-en maintenant à la seconde tâche de la grammaire formelle, l'élimination des systèmes formels inadaptés aux besoins linguistiques. Notre propos, dans ce cas-ci, sera différent de celui de la section précédente car nous ne voulons pas nier que la grammaire formelle remplisse cet objectif. Plutôt, l'idée que nous défendrons est que ce point de vue sur la grammaire formelle ne devient acceptable que lorsqu'on se limite à considérer le seul plan des effets de la mathématisation. Autrement dit, nous croyons voir dans la réalisation de cet objectif plus un résultat fortuit que la conséquence d'une pratique élaborée à cette fin précise. Il s'agira donc de montrer ce qu'il y a de réducteur dans une telle description de la grammaire formelle.

3.2.I. Commençons par examiner la première étude axiomatique des grammaires qui soit réellement suggestive, celle proposée par Chomsky (1959). Trois fonctions essentielles peuvent en être dégagées.

I°) Définir une forme générale de la règle de grammaire et constater son identité avec celle des règles des systèmes semi-thueiens. C'est cette fonction qui est exploitée dans l'argumentation montrant que le principe d'exhaustivité présuppose l'existence d'une grammaire pour un langage donné, c'est à dire ce que nous avons appelé plus haut l'hypothèse de calculabilité.

- 2°) Construire une hiérarchie sur la classe des grammaires déterminée par les différentes restrictions que l'on peut imposer à leurs règles. C'est la classification en grammaires à états finis, indépendantes du contexte, etc...
- 3°) Spécifier l'appartenance des grammaires à chacune de ces classes à l'aide de propriétés structurales invariantes par rapport à certaines altérations non significatives des règles. Ceci revient à proposer une méthode pour ramener toute grammaire à sa forme canonique afin de déterminer quelle est sa classe.

Cette axiomatisation délimite la liberté d'usage des grammaires. Elle permet soit de ramener à une propriété formelle simple des paradigmes de variation que l'intuition pourrait considérer comme essentielle - par exemple l'équivalence de la forme normale de Greibach et de la forme normale de Chomsky d'une grammaire indépendante du contexte - soit de réduire à des différences essentielles des variations que l'intuition supposerait minimales - par exemple les grammaires unilatères n'appartiennent pas à la même classe que les grammaires linéaires minimales.

De quelle façon cette axiomatique permet-elle de poser le problème général de l'adéquation des grammaires?

La hiérarchie de Chomsky définit des relations d'inclusions propres sur la classe des grammaires, mais par là même, elle définit également de telles relations sur une sous-classe des langages qui est évidemment la sous-classe des langages

récurivement énumérables. Donc de la même manière que l'on parle des grammaires à états finis, indépendants du contexte, etc... on parle de langages à états finis, indépendants du contexte, etc... Dire alors qu'une grammaire est inadéquate, du point de vue de la grammaire formelle, c'est dire que le langage qu'elle engendre appartient à une sous-classe proprement incluse dans la classe du langage qu'on voudrait lui faire engendrer.

Cette définition de l'inadéquation conduit à éliminer les grammaires en raison de leur incapacité à produire certaines séquences terminales d'un langage donné. Mais le problème inverse pourrait se poser : étant donné la classe du langage produit par une grammaire donnée, ne peut-on pas trouver une autre grammaire, plus restreinte, pour engendrer ce langage? C'est là une situation qui peut apparaître en linguistique. Il suffit pour cela qu'un langage présente des propriétés formelles telles que l'on puisse identifier sa classe indépendamment de toute grammaire. Le problème général de l'adéquation consiste donc à requérir de la théorie des grammaires formelles qu'elle étudie mathématiquement certaines propriétés structurales des langues naturelles pour pouvoir déterminer avec le plus de précision possible la classe de grammaires formelles à laquelle appartiennent les grammaires génératives des langues naturelles.

3.2.2. La grammaire formelle semble s'acquitter de cette tâche en éliminant en haut de la hiérarchie des grammaires insuffi-

Faux !!
samment restreintes et en bas de la hiérarchie des grammaires trop restreintes. Donnons des exemples. Un théorème dû essentiellement à Chomsky, permet de démontrer que les grammaires à états finis ne peuvent décrire les langues naturelles. Ou plus exactement les langues naturelles sur lesquelles on a essayé de vérifier ce résultat, puisque, comme nous l'avons déjà dit, les démonstrations d'inadéquation ne valent jamais que pour un langage donné. Ces grammaires sont donc trop restreintes et il faut choisir, comme grammaires pour les langues naturelles, des systèmes appartenant à une classe plus haute dans la hiérarchie. Inversement, Peters et Ritchie ont démontré que les grammaires transformationnelles construites sur le modèle d'Aspects of the Theory of Syntax pouvaient engendrer n'importe quel ensemble récursivement énumérable. Comme cette propriété contredit l'hypothèse que les langues naturelles sont récursives, il faut éliminer l'emploi de ce type de grammaire au profit de modèles plus restreints.

Mais cette symétrie entre l'élimination par le haut et par le bas repose sur une confusion. Car dans le cas du théorème de Chomsky les grammaires à états finis sont rejetées en fonction d'une observation sur le langage, tandis que dans le second cas les grammaires transformationnelles sont écartées en fonction d'une hypothèse sur le langage qui, elle-même, découle, nous l'avons vu plus haut, d'un a priori sur la grammaire. Alors que dans le cas des grammaires à états finis, il s'agit de l'élimination de ces systèmes comme théories inadéquates du lan-

gage, dans celui des grammaires transformationnelles, on a affaire à l'exclusion d'objets impropres à la théorie des grammaires. Et il est exact que le sens des théorèmes de Peters et Ritchie est tout autre que l'expression d'une condition d'adéquation, telle que nous l'avons définie. Car la différence principale entre ces deux théorèmes n'est pas qu'ils énoncent des propriétés d'objets incomparables, mais qu'ils appartiennent à des niveaux de langage différents lorsqu'on les interprète dans le cadre de la théorie linguistique chomskyenne. En effet, selon Chomsky, il faut distinguer soigneusement entre les divers niveaux de langage présents dans la théorie linguistique : "La théorie linguistique est formulée dans une métalangue par rapport aux langues dans lesquelles les grammaires sont écrites, une métamétalangue par rapport à toute langue pour laquelle une grammaire est construite."⁽⁶⁾ Nous ne pensons pas qu'il faille voir dans cette affirmation la seule expression de la prudence philosophique d'un savant scrupuleux influencé par la fréquentation de la logique et de la philosophie des sciences. Ces distinctions entre niveaux de langage correspondent en fait à un programme de recherche très clairement défini, et la thèse que nous défendrons dans ces pages, et plus particulièrement au cours de la deuxième partie, est que ce programme, parfois considéré comme une aberration épistémologique⁽⁷⁾, a le mérite, du fait de la précision de sa formulation, de se prêter à un contrôle sévère dont la théorie des grammaires formelles est l'instrument principal.

NOTES DU CHAPITRE III

- (1) C'est bien l'intention de toute grammaire de ne pas être confuse et d'être exhaustive, c'est à dire de prévoir toutes les difficultés que peuvent rencontrer ses usagers. La finitude, elle, va de soi
- (2) Chomsky (1956, p.54) Il ne s'agit pas de la part de Chomsky d'un exercice purement spéculatif, celui-ci considère tout à fait plausible que des "langues intéressantes" puissent être grammaticalement indescriptibles.
- (3) Le principe d'exhaustivité impliquant la description de toutes et rien que les phrases grammaticales, il permet donc de se prononcer sur l'agrammaticalité des phrases. Par conséquent, la classe des langages pour lesquels il existe une grammaire est identifiable à la classe des langages rékursifs, puisque eux-mêmes et leur complémentaire sont rékursifs énumérables.
- (4) Cette suggestion pourra sembler doublement absurde:
- 1°) La langue étant infinie, comment vérifier que la grammaire l'engendre effectivement?
- 2°) Au nom de quel principe peut-on s'autoriser à construire une grammaire n'ayant aucune justification linguistique?
- Au premier point on répondra qu'il s'agit exactement du même problème que celui qui est discuté au § 2.1.2. Une telle vérification est possible par induction, sinon l'emploi par la grammaire de procédures rékursives est injustifiable. A la seconde objection on opposera que puisque la question de

l'existence de la grammaire se pose en termes mathématiques, il est nécessaire d'y répondre dans les mêmes termes. De plus pourquoi ne pas imaginer que la construction d'une machine de Turing selon des principes totalement indépendants des considérations grammaticales puisse apporter des idées suggestives à la grammaire. Nous rappelions à la note 9, p. 52 à quel point le traitement des catégories grammaticales était précaire dans les grammaires génératives. Des catégories introduites à la seule fin de résoudre ce problème mathématique n'empirerait pas la situation et aurait au moins l'avantage d'être libérées des déterminations habituelles qu'impose la tradition.

- (5) Harrop (1961) Les procédés employés par Harrop pour définir de tels ensembles peuvent paraître totalement étrangers aux problèmes posés par la construction d'une grammaire. Cependant comment supposer que la grammaire soit réductible à la partie la plus classique de la théorie de la calculabilité sans préjuger de la complexité des opérations mises en oeuvre dans l'activité linguistique humaine.
- (6) Chomsky (1957, note 4, p. 56)
- (7) Nous verrons au chapitre VII que la formulation même de ce programme permet de démontrer mathématiquement qu'il est incohérent sous certaines conditions.

CHAPITRE IV

AMBIGUITE MATHEMATIQUE

ET

AMBIGUITE LINGUISTIQUE

4. La conclusion principale de notre deuxième chapitre était l'idée que la théorie des grammaires formelles étudie des phénomènes de syntaxe pure. Or on pourrait se demander s'il n'y a pas justement un lien entre le fait que la grammaire formelle ne s'occupe que de syntaxe pure et les difficultés qu'elle peut rencontrer lorsqu'elle doit construire une théorie mathématique satisfaisante des phénomènes linguistiques. C'est le problème que nous nous proposerons d'examiner dans ce chapitre à travers un exemple précis, celui du traitement mathématique de l'ambiguïté. Une raison supplémentaire de nous intéresser au problème de l'ambiguïté tient à la place privilégiée qu'il occupe dans la linguistique mathématique (et pas seulement dans la grammaire formelle): c'est une des rares questions, sinon l'unique, qui semble se présenter aux mathématiques comme un objet d'étude original.

Malgré son originalité, sur le plan des résultats obtenus le traitement mathématique de l'ambiguïté n'est que d'un faible intérêt pour l'étude du langage humain. Son apport se résume, en gros, aux deux enseignements suivants : 1°) Le degré d'ambiguïté d'une phrase augmente avec sa longueur; 2°) Il n'existe pas de raison théorique s'opposant à ce qu'une grammaire puisse expliquer toutes les ambiguïtés d'un langage donné. C'est peu de choses, et ce sont surtout des choses que les linguistes savaient déjà avant que les mathématiciens ne se chargent de la question. Nous voudrions chercher à comprendre ici pourquoi ces résultats sont si maigres.

Nous commencerons donc par comparer les définitions et les fonctions de l'ambiguïté dans les théories de la grammaire générative et de la grammaire formelle afin de faire apparaître d'éventuelles incompatibilités entre ces deux notions. Nous signalerons ensuite l'existence d'une solution mathématique idéale au problème de l'ambiguïté et nous conclurons par une discussion des facteurs dont elle dépend.

4.1. Grammaires ambiguës et langages ambigus

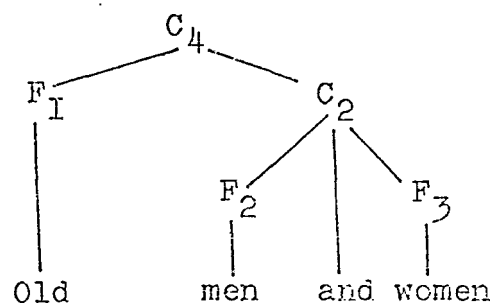
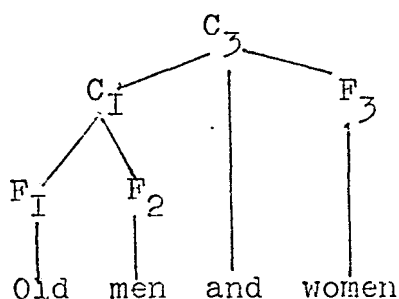
La nécessité de traiter les ambiguïtés syntaxiques est, comme on le sait, à l'origine du choix des grammaires transformationnelles comme grammaires génératives des langues naturelles (Chomsky, 1956). C'est une raison théorique justifiant l'élimination des grammaires de constituants au profit des grammaires transformationnelles. Notons que ce sera, paradoxalement, la prolifération des solutions transformationnelles au problème de l'ambiguïté (entre autres) qui conduira à un nouvel enlèvement de la question : devant la nécessité de contrôler rapidement les conséquences découlant de l'emploi d'un composant transformationnel trop puissant, les théoriciens de la grammaire générative vont s'attacher presque exclusivement à imposer des conditions restrictives à l'usage des transformations. Aujourd'hui donc l'intérêt porté aux moyens - la cohérence du composant transformationnel^(I) - semble l'emporter sur l'intérêt porté aux fins - l'analyse des ambiguïtés syntaxiques jusqu'à l'occulter complètement. Il faut pourtant reconnaître que les motivations originelles de la grammaire transformationnelle

sont extrêmement convaincantes et que même si elles devaient être réduites au statut de principe théorique elles conservent encore aujourd'hui, comme nous le verrons, toute leur valeur.

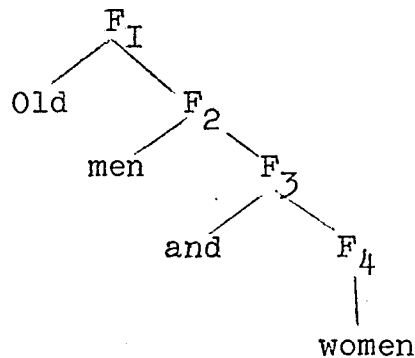
4.I.I. Reprenant à son compte certaines observations de Wells (1947), Hockett (1954) fait remarquer qu'une des difficultés de l'analyse en constituants immédiats tient à ce que différentes procédures d'analyse opérant sur les mêmes constituants terminaux peuvent produire des résultats qui ne sont pas différents. Soient les quatre constructions C_1, C_2, C_3, C_4 et les trois formes F_1, F_2 et F_3 . Supposons que F_1 et F_2 figurent dans C_1 et F_2 et F_3 dans C_2 . Si C_1 apparaît à gauche de C_3 et que celui-ci contient F_3 , et si C_2 apparaît en seconde position dans C_4 et que celui-ci contient F_1 en première position, alors la forme composée $F_1 F_2 F_3$ sera l'unique résultat de deux analyses différentes⁽²⁾.

Ce qui pour Hockett était une difficulté propre à certains modèles d'analyse en constituants (les modèles par "items et arrangement"), devient pour Chomsky un des traits essentiels qui les distingue des grammaires à états finis et les rend supérieurs à celles-ci. L'énoncé "Old men and women", pris comme exemple par Hockett peut recevoir avec l'analyse en constituant:

*non
man:
c'est plus
"naturel"*



tandis qu'avec une grammaire à états finis seule lui est attribuable la description



Il ne s'agit pas là d'une simple opposition de points de vue sur les possibilités des grammaires de constituants⁽³⁾. Ce qui sépare avant tout Chomsky de Hockett, c'est l'introduction de la notion formelle d'indicateur syntagmatique comme représentant d'une classe de dérivations équivalentes. La notion de dérivation étant implicitement présente mais inexprimée chez Hockett. Ainsi l'ambiguïté de "Old men and women" ou de "They are flying planes" est expliquée par l'homonymie de construction et l'homonymie de construction est expliquée par les deux dérivations non équivalentes. Les grammaires à états finis n'offrent aucune possibilité d'analyser les ambiguïtés en termes d'homonymie de constructions puisqu'elles ne disposent que de règles unilatères⁽⁴⁾.

De la même manière les grammaires transformationnelles peuvent rendre compte de l'ambiguïté d'énoncés du type "The shooting of the hunters" par homonymie de constructions transformationnelles. "The shooting of the hunters" pouvant être obtenue soit par l'application de la transformation

SN - V $\implies$ The - V + ing - of + SN

à la phrase noyau "The hunters shoot", soit par l'application des transformations

$SN_I - V - SN_2 \implies \text{The} - V + \text{ing} - \text{of} + SN_2 - \text{by} + SN_I$
et $\text{The} - V + \text{ing} - \text{of} + SN_2 - \text{by} + SN_I \implies \text{The} - V + \text{ing} - \text{of} + SN_2 - \emptyset$
à la phrase noyau "They shoot the hunters".

Ce type d'ambiguïté ne pouvant être interprété en termes d'homonymie de constructions par une grammaire de constituants, ces grammaires ne peuvent être des grammaires génératives de l'anglais.

De nombreuses observations ont confirmé que l'ambiguïté n'était pas un phénomène marginal dans les langues naturelles et que l'argument de Chomsky avait une portée dépassant largement celle d'une démonstration "théorique" de l'inadéquation des grammaires de constituants. L'élimination de l'ambiguïté par les constructions syntaxiques semble donc être un critère empirique fondamental de sélection des grammaires. Si on s'appuie sur les faits que nous venons d'exposer, deux conditions générales doivent être remplies par une théorie linguistique voulant rendre compte des ambiguïtés : 1°) Les grammaires doivent produire des indicateurs syntagmatiques et non pas de simples séquences de symboles; 2°) Les grammaires doivent avoir des transformations. Cependant ces conditions sont déduites d'observations trop limitées proportionnellement à leur généralité. A priori, rien ne permet de supposer qu'elles sont suffisantes pour garantir l'élimination de toutes les ambiguïtés syntaxiques et qu'il ne faille pas recourir à d'autres moyens formels pour traiter celles-ci.

La théorie mathématique de l'ambiguïté donne des éléments de réponse à ce problème. Cependant, avant d'exposer ces réponses, nous pensons qu'il faut éclaircir la signification mathématique de l'ambiguïté car le transport de cette notion du cadre de la grammaire générative à celui de la grammaire formelle s'accompagne d'abus de langage, désormais consacrés par l'usage, qui rendent leur interprétation délicate.

4.I.2. Dans le cas des langues naturelles, l'ambiguïté est une propriété des phrases du langage . Une grammaire n'est pas ambiguë lorsqu'elle attribue plusieurs constructions homonymes à une phrase ambiguë. Ainsi les grammaires de constituants ne sont pas ambiguës pour la description de "They are flying planes", mais elles le sont pour "The shooting of the hunters". Comme on peut, de façon naturelle, associer à chaque phrase un degré d'ambiguïté (un nombre d'interprétations possibles), par commodité on appelle degré d'ambiguïté d'une grammaire le nombre de constructions homonymes qu'elle peut attribuer à une phrase. C'est le sens qu'a ce terme dans la théorie des grammaires formelles. Mais il est clair que son interprétation exacte est "degré de non-ambiguïté", puisqu'il correspond au nombre de descriptions différentes qu'attribue la grammaire à une séquence terminale. Les expressions d'un langage formel ne peuvent, en effet, présenter en tant que telles aucune ambiguïté puisque ce sont des objets purement syntaxiques. Une séquence d'un langage formel ne devient ambiguë qu'à partir du moment où une grammaire en donne plusieurs constructions homonymes. Aussi ne peut-on parler d'ambiguïté d'un langage formel que par rapport à une grammaire

donnée, ce que les auteurs ne manquent jamais de faire remarquer. Cependant, il semble clair que les précautions employées par les mathématiciens dans la définition de l'ambiguïté sont avant tout d'ordre pédagogique. Nous croyons quant à nous que cette prudence didactique n'épuise pas toute l'équivocité de la notion de langage formel ambigu. Car dire que cette propriété est relative, signifie que l'on peut s'appuyer sur un repère fixe, un langage donné L , pour pouvoir la définir. Or tout ce que l'on sait du langage on ne peut l'apprendre que par sa grammaire, et ceci vaut en particulier lorsqu'on affirme que G est une grammaire de L , c'est à dire lorsqu'on identifie L et $L(G)$. D'une façon plus générale, il est même impossible de savoir de quel type est la grammaire de L , sauf dans le cas très restreint des langages réguliers, puisqu'on ne connaît pas de caractérisation structurale des langages dépendants du contexte, indépendants du contexte, etc..., qui soient totalement indépendantes des grammaires⁽⁵⁾. La seule chose que l'on puisse alors dire c'est que, tout en engendrant L , G l'engendre de façon ambiguë. Mais dans ce cas, l'ambiguïté n'est plus une propriété du langage relativement à une grammaire donnée, mais une propriété intrinsèque de la grammaire.

De même, le concept d'ambiguïté inhérente est trompeur, car il laisse supposer qu'il y a là une propriété absolue des langages, alors qu'il ne s'agit que d'une propriété relative des grammaires. Lorsque nous disons que le langage indépendant du contexte $L = \{ a^n b^m a^p / n=m \text{ ou } m=p \}$ est inhéremment ambigu,

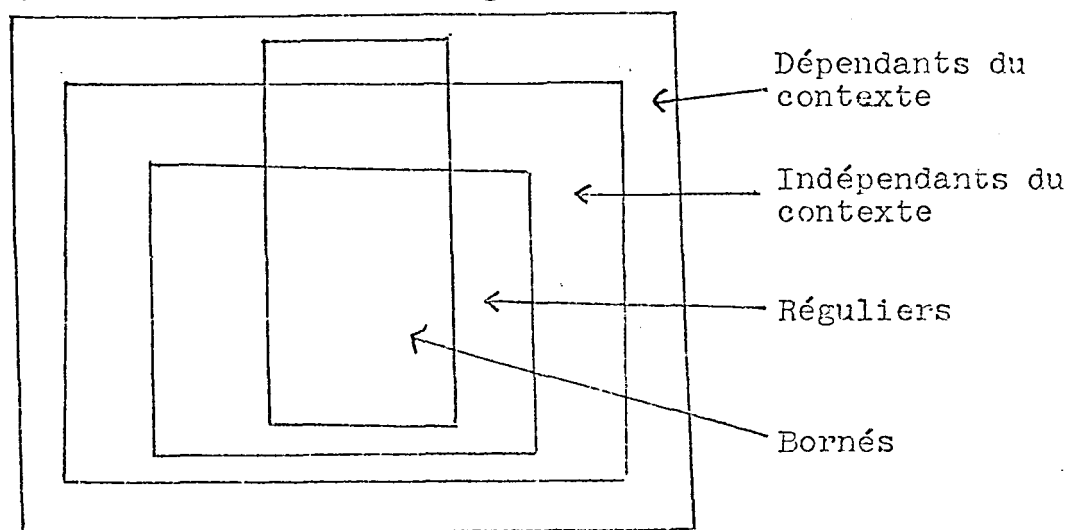
qu'affirmons-nous? Simplement que c'est l'un des objets qui vérifient le théorème de Parikh : "Il existe un langage indépendant du contexte L tel que toute grammaire indépendante du contexte de L est ambiguë." Comme on le voit, pour asserter que L est inhéremment ambigu il faut d'abord poser qu'il est indépendant du contexte. Mais être indépendant du contexte, pour un langage, tout comme être ambigu, est une propriété relative: celle-ci n'étant connue qu'une fois établi l'existence d'une grammaire indépendante du contexte pour L. La mention faite, dans le théorème, de l'existence de L est superflue puisqu'elle ne joue aucun rôle logique. Si on voulait débarrasser de toute équivocité l'énoncé du théorème de Parikh, il faudrait le reformuler ainsi : "Il existe une grammaire indépendante du contexte ambiguë G telle que toutes les grammaires indépendantes du contexte faiblement équivalentes à G sont également ambiguës." Le type d'information apporté par ce théorème pourrait alors être décrit dans les termes suivants. Lorsqu'on compare des grammaires, on dispose de deux critères principaux, l'équivalence faible et l'équivalence forte. L'équivalence forte n'est guère intéressante puisqu'en pratique elle revient à l'identité des grammaires. Inversement, l'équivalence faible est une propriété trop générale pour être grammaticalement significative. Il est donc souhaitable de trouver des critères originaux permettant d'établir une typologie des grammaires sur la base de caractéristiques structurales plus riches que l'équivalence faible et moins contraignantes que l'équivalence forte.

Or on peut estimer, si toutefois on accepte notre formulation du théorème de Parikh, que l'ambiguïté inhérente, en tant qu'elle détermine des classes de grammaires équivalentes, est un de ces critères. On notera d'ailleurs que cette interprétation s'accorde assez bien avec une remarque de Chomsky et de Schützenberger (1965, pp.123-4). Introduisant les grammaires comme générateurs de séries formelles de puissance, ceux-ci estiment que leur point de vue consistant à représenter un langage comme un ensemble de couples (s, n) (où s est une séquence et n son degré d'ambiguïté) est intermédiaire entre les deux points de vue précédemment cités. Cette remarque signifie simplement que lorsqu'on parle de l'ambiguïté inhérente, on exprime nécessairement des propriétés de la grammaire plutôt que des propriétés du langage.

Concluons ces réflexions en insistant sur le fait qu'elles ne demeurent valides que tant que l'on ne dispose pas de caractérisation des langages indépendants du contexte qui ne dépende pas des grammaires. Ce qui nous permet de répondre à deux objections que l'on pourrait éventuellement opposer à notre propos. I°) La condition nécessaire et suffisante proposée par Ginsburg (1966, th.6.I.I., p.195) pour qu'un langage indépendant du contexte soit inhéremment ambigu semble être un démenti flagrant à notre point de vue puisque celle-ci ne fait intervenir que des opérations algébriques et ensemblistes sur les expressions du langage à l'exclusion de toutes considérations grammaticales. Cependant ne perdons pas de vue que la terminologie adoptée

par l'auteur est extrêmement restrictive puisque par définition celui-ci identifie les langages formels aux langages indépendants du contexte, économisant ainsi une spécification qui fait intervenir la grammaire de façon essentielle et qui demeure implicitement présente. D'autre part cette condition nécessaire et suffisante ne vaut que pour une classe restreinte de langages (donc de langages indépendants du contexte), les langages bornés.

2°) En second lieu on pourrait objecter que les langages bornés, précisément, sont susceptibles d'une caractérisation effectivement indépendante de tout recours à la grammaire. Dans ce cas particulier, nous reconnaissons que nos remarques tombent immédiatement. Cependant, nous ferons observer que les langages bornés occupent une position singulière dans la zoologie des langages formels puisque cette classe qui contient des langages dépendants du contexte (non indépendant du contexte) et des langages indépendants du contexte n'inclue pas la totalité des langages réguliers. On pourrait dire de façon imagée qu'elle coupe transversalement la classification de Chomsky, comme le montre la figure suivante:



La possibilité d'une caractérisation d'une classe de langages, dès que l'on déborde le cas extrêmement simple des langages à états finis où les grammaires sont si pauvres qu'on peut virtuellement les identifier aux langages, n'offre donc pas nécessairement la garantie d'une relation significative à la théorie des grammaires puisque dans le cas particulier des langages bornés on ne voit pas quelle pourrait être la forme canonique de la grammaire les engendrant, ou tout au moins sous quel aspect cette forme canonique serait comparable à celle des grammaires formelles "classiques". Or, c'est seulement quand elle peut être rattachée en un mode quelconque à un type grammatical qu'une caractérisation des langages devient intéressante, car interprétable et linguistiquement utilisable.

Ainsi pouvons-nous faire une distinction entre l'ambiguïté au sens de la grammaire formelle et au sens de la grammaire générative. Ceci non pas en nous plaçant sur le seul plan des effets d'une terminologie conventionnelle qui dénomme a contrario l'homonymie de construction comme ambiguïté, mais en relevant une différence essentielle entre les deux définitions de cette notion. Dans le cas de la grammaire générative l'ambiguïté est une propriété des expressions du langage, dans le cas de la grammaire formelle il s'agit d'une propriété des grammaires. L'origine de cette divergence réside dans le fait que les langages formels sont des objets purement syntaxiques qui voient, par là même, leurs propriétés structurales intrinsèques insuffisamment déterminées pour qu'il soit possible

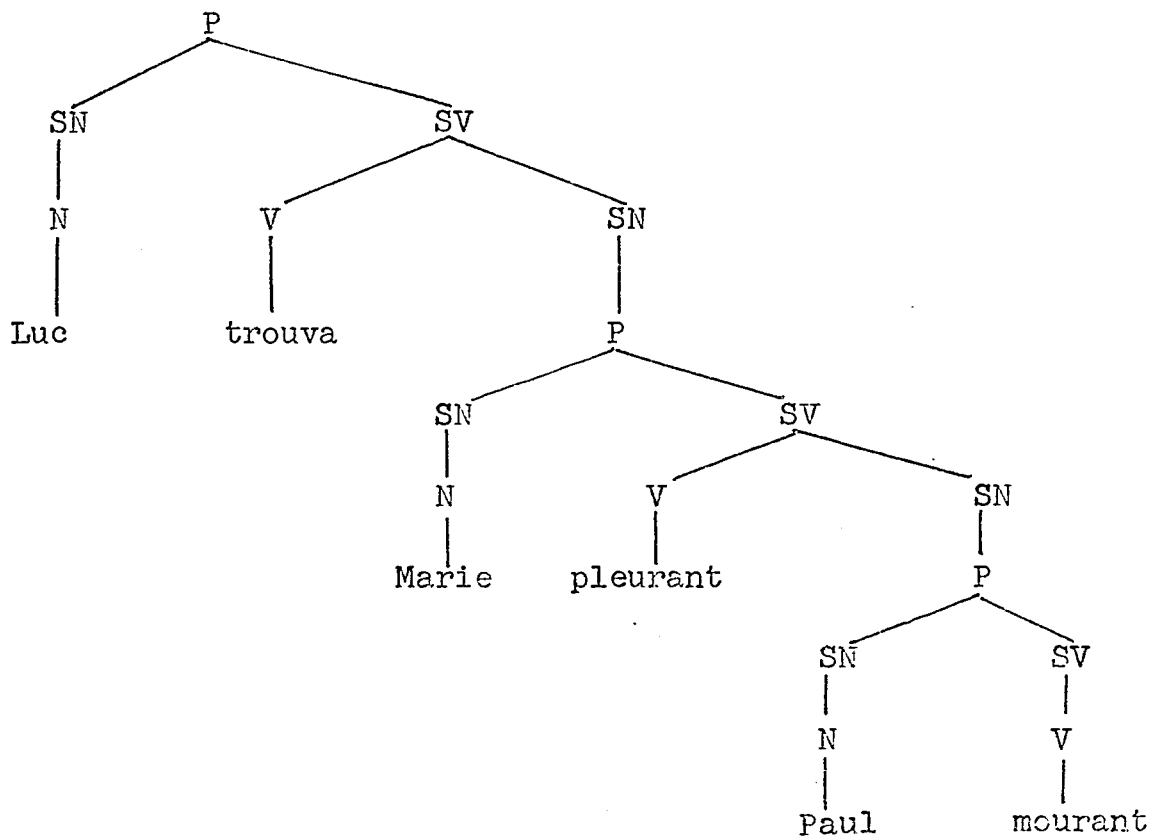
d'attribuer des interprétations (de quelque type que ce soit) à leurs expressions. A l'opposé, les expressions des langages humains pouvant présenter naturellement plusieurs interprétations possibles et parmi celles-ci certaines assimilables aux conséquences des homonymies de constructions syntaxiques car la flexibilité du syntaxique dans la grammaire générative lui permet d'absorber du sémantique, c'est à dire de ramener une fraction du sens à des régularités distributionnelles. C'est précisément cette intégration du sémantique par le syntaxique qui conduit à un enrichissement structural autorisant l'identification d'un niveau syntaxique d'ambiguïté.

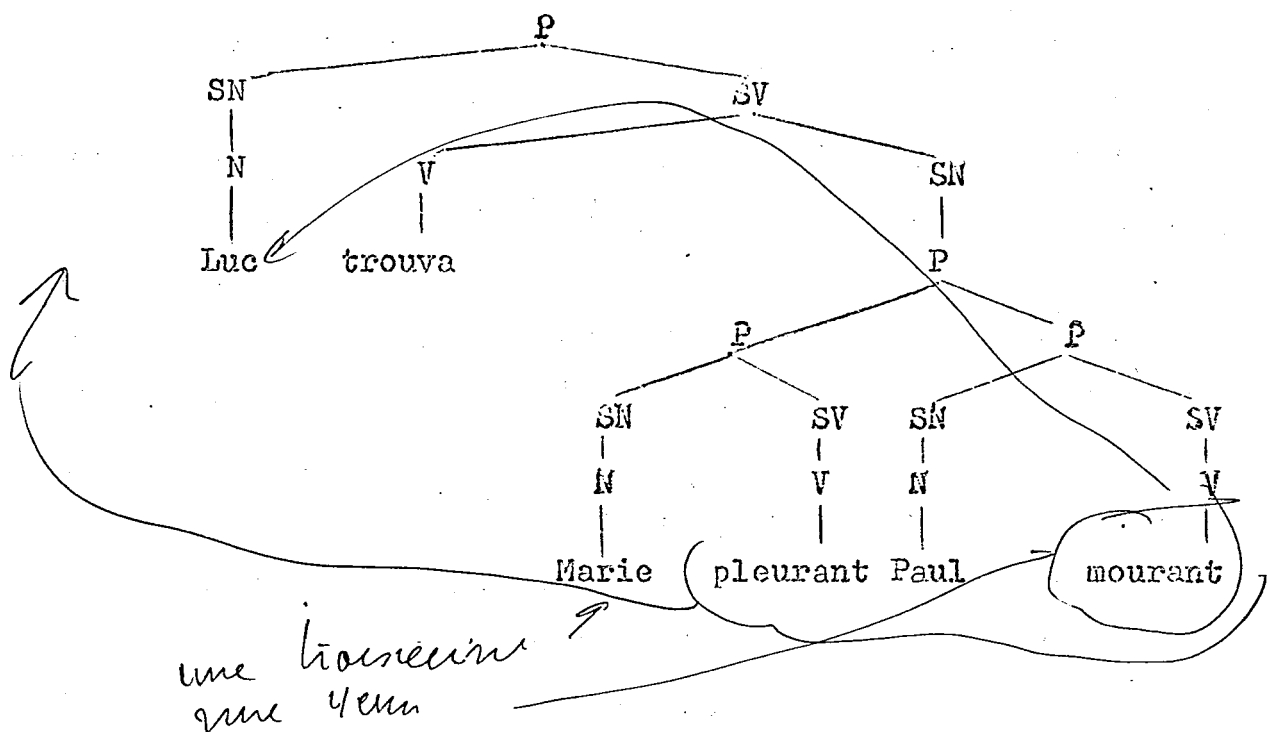
4.2. Le traitement mathématique de l'ambiguïté

Pour faire apparaître la divergence entre ces deux conceptions de l'ambiguïté nous avons dû nous interroger sur les fondements de leur construction conceptuelle. Aussi la question demeure de savoir si cette différence fondamentale n'est pas neutralisable lorsqu'on se place au niveau d'une technique du traitement de l'ambiguïté. A cette condition plutôt que de considérer la discordance entre ces deux notions d'ambiguïté comme une incompatibilité épistémologique essentielle, il faudrait y voir les conséquences des principes d'une stratégie pour laquelle l'élimination de certains "contenus" prime toujours sur un certain type d'harmonie conceptuelle. Pour réfléchir à cette question nous allons maintenant tenter d'évaluer

à travers un exemple précis, la portée du traitement mathématique de l'ambiguïté.

4.2.I. Dans le cas de l'étude des langages de programmation le problème est parfaitement déterminé⁽⁶⁾: il faut construire des grammaires qui ne soient pas ambiguës. Sinon dans la compilation, la machine mise en présence d'une expression pouvant recevoir des analyses différentes ne "saurait" pas comment choisir celle qui convient. Mais pour les langues naturelles les choses se présentent différemment. Ce que l'on recherche, ce sont des grammaires qui tout en fournissant des descriptions structurales en nombre suffisant ne produisent pas pour autant des dérivations en quantité excédentaire. Prenons un exemple. La phrase "Luc trouva Marie pleurant Paul mourant", peut recevoir les deux descriptions suivantes:





A ces descriptions syntaxiques correspond la
grammaire composée des règles

$P \rightarrow SN \text{ } SV + PP$

$SN \rightarrow N + P$

$SV \rightarrow V+V \text{ } SN$

En posant $N=a$ et $V=b$, on peut associer à ce système de règles le système d'équations suivant (voir Gross et Lentin (1970) pour les détails techniques relatifs à la construction de tels systèmes et Chomsky et Schützenberger (1963) pour un exposé théorique sur ces questions):

$$\begin{cases} P = NV \cup PP \\ N = a \cup P \\ V = b \cup bN \end{cases}$$

$abbaab + abbaaba + abbaabb + abbaaba + abbaaba + abbaaba + \dots +$
 $abbaab + abbaaba + abbaabb + abbaaba + abbaaba + \dots + abbaab +$
 $abbaaba + abbaabb + abbaaba + abbaaba + \dots + abbaab + abbaaba +$
 $abbaabb + \dots + abbaab + abbaaba + abbaabb + \dots + abbaaba + \dots$
 $= ab + aba + abb + 3 abab + abba + abba + 2 ababa + ababb + 2 abba +$
 $ababa + ababb + 3 ababa + 2 ababa + abba + abba +$
 $ababb + 2 ababa + abba + 2 ababa + 2 ababa +$
 $ababb + ababb + abba + abba + abba +$
 $abba + abba + 2 ababa + abba + 2 ababa +$
 $abbaabb + 5 ababa + abba + abba + abba +$
 $abba$

On constate que les degrés d'ambiguïté augmentent rapidement avec la longueur des phrases puisque la grammaire peut produire jusqu'à cinq dérivations distinctes pour une séquence terminale de huit termes. Le fait que le calcul soit opéré à partir de symboles pré-terminaux peut laisser croire que les degrés calculés sont supérieurs aux degrés réels puisqu'on néglige ainsi les réalisations distinctes des verbes (passé simple et participe présent). En fait ce phénomène est compensé par les ambiguïtés supplémentaires qu'introduisent ces réalisations distinctes. Par exemple "Luc trouvant Marie pleurant Paul mourant" a trois interprétations différentes. D'ailleurs on remarque que le degré d'ambiguïté de "ababab" qui correspond à la séquence pré-terminale de la phase précédente et de notre premier exemple est égal à trois, ce qui est inférieur à la réalité. Ce "retard" est cependant comblé par le

calcul du sixième ordre d'approximation du langage engendré par le système dont on voit qu'il attribuera un degré d'ambiguïté égal à six pour "ababab" (produit de "ab" par "3 abab" et inversement). Le calcul du sixième ordre d'approximation étant le dernier qui produit des "ababab".

4.2.2 La méthode algébrique permet donc de déterminer facilement les différents degrés d'ambiguïté attribués par une grammaire à chacune des séquences qu'elle engendre. Remarquons que l'utilité de la méthode n'est pas totalement redondante, car si dans l'exemple que nous avons choisi l'ambiguïté de la grammaire est visible "à l'oeil nu" sur une simple inspection des règles, ce n'est pas nécessairement le cas pour une grammaire qui comprend plusieurs centaines de règles comme le serait sans doute la grammaire d'une langue naturelle (7). L'intérêt de la méthode algébrique est dans le fait qu'elle donne un renseignement précis sur une propriété particulière. Elle n'a évidemment pas le privilège exclusif d'indiquer combien de dérivations différentes une grammaire donnée attribue à une séquence donnée, puisqu'une simple énumération des graphes le ferait tout aussi bien, mais elle est plus efficace. Elle répond en effet à la seule question du degré d'ambiguïté sans avoir besoin de passer par la construction d'indicateurs syntagmatiques complets, évitant ainsi de donner une information superflue par rapport au problème traité. La théorie algébrique des langages indépendants du contexte convient donc parfaitement à la construction méthodique des langages de programmation. Convient-elle à l'étude des grammaires génératives? Les grammaires génératives sont déterminées à partir d'analyses de phrases

particulières. Une fois établies les règles de la grammaire, celles-ci sont transformées en systèmes d'équation dont les solutions donnent le nombre de constructions homonymes. Mais si on est certain que cette méthode permet de retrouver ce qu'on savait déjà avant de l'appliquer, à savoir, que les degrés d'ambiguïté calculés correspondent à ceux observés sur les phrases qui ont permis de construire la grammaire, dès que le calcul devient prédictif, le sens de ses prédictions nous échappe car nous ne savons plus si le degré d'ambiguïté correspond à une propriété de la grammaire ou à une propriété du langage décrit. Pour que la méthode algébrique du calcul des ambiguïtés puisse avoir un intérêt linguistique il faudrait l'assortir d'études statistiques sur le langage. Celles-ci pourraient, par exemple, porter sur le rapport entre la longueur des phrases et leur degré d'ambiguïté. La comparaison des résultats de l'étude statistique et de l'étude algébrique fournirait alors une mesure d'évaluation de la grammaire. C'est, pensons-nous, le seul moyen de conserver un sens linguistique à la théorie mathématique de l'ambiguïté, car l'appropriation de la notion d'ambiguïté par la théorie des grammaires formelles s'effectue d'une façon telle, que sa signification linguistique originale est totalement dissoute dans l'abstraction. Il ne faut cependant pas regretter le caractère nécessairement approximatif qu'introduirait l'usage de descriptions statistiques pour traiter un problème dont la formulation initiale laissait supposer qu'il pouvait recevoir une solution exacte comme dans le cas des langages de programmation. Cet aspect est commun à pratiquement toutes les théories mathématiques des phénomènes empiriques puisqu'il concerne la mesure d'une grandeur. Simplement faut-il souligner le caractère

paradoxal d'une situation où il semble que l'on dispose du modèle mathématique d'un phénomène empirique encore insuffisamment étudié. En effet, si dans le cas des langages de programmation la question de l'ambiguïté est sujette à une solution exacte, c'est que le problème posé concerne uniquement une qualité de la forme. Les apparences peuvent laisser croire que c'est également le cas pour les langages humains, mais nous avons déjà exposé qu'il s'agissait ici d'une forme enrichie par l'intégration de contenus, ce qui explique l'imperfection d'un modèle mathématique élaboré dans la perspective d'un idéal formel. Cependant nous avons constaté que si les contenus ne pouvaient être réduits totalement à une représentation syntaxique qui tend à leur être équivalente ceci ne signifie pas pour autant que la mathématisation se heurte à un obstacle irrémédiablement insurmontable. Les déterminations qu'imposent les contenus au problème de l'ambiguïté sont susceptibles elles aussi d'une représentation mathématique pertinente. Evidemment, comparé à l'intégration progressive et continue du sémantique par le syntaxique opérée par la grammaire générative, la rupture brutale entre la méthode "absolutiste" de la grammaire formelle et les techniques statistiques peut gêner. Mais il faut bien voir que cette discordance interne au traitement mathématique de l'ambiguïté est le prix à payer pour neutraliser l'incompatibilité épistémologique entre l'ambiguïté mathématique et l'ambiguïté linguistique.

4.3. Les difficultés méthodologiques d'une solution théorique

Jusqu'à présent en exposant en quoi les modèles mathématiques étaient inadaptés aux besoins de la grammaire, nous avons tenté de mettre en relief que, au niveau de connaissance où se trouve actuellement la grammaire formelle, une théorie de la syntaxe pure était trop pauvre pour analyser la pseudo-syntaxe sémantiquement enrichie de la grammaire générative. Mais encore faut-il comprendre la nature de cette pauvreté. Il ne s'agit pas tant d'une carence des moyens formels face aux besoins grammaticaux que de l'évanouissement de certaines dimensions significatives des phénomènes grammaticaux. Dans cette dernière section nous constaterons une fois encore cette discordance entre le point de vue de la grammaire formelle et celui de la grammaire générative. Cependant, ici, elle se présentera sous une forme paradoxale dont l'intelligibilité, contrairement aux problèmes exposés dans les sections précédentes, ne pourra être atteinte par le seul recours à l'incompatibilité entre les deux notions d'ambiguïté.

4.3.I. A la fin du paragraphe 4.I.I., nous présentions les conclusions de l'argumentation de Chomsky dans les termes suivants: pour pouvoir rendre compte des ambiguïtés, une grammaire doit assigner des descriptions structurales aux éléments du langage et incorporer des transformations. Unissant ces deux réquisits, il y a une caractéristique formelle qui, face au problème de l'ambiguïté, amène à les considérer comme deux variantes d'une unique condition : l'univocité de l'interpré-

tation des dérivations. Exprimée autrement, cette condition revient à imposer, au niveau méta-grammatical, l'usage d'une méthode non-ambigüe d'analyse syntaxique pour traiter les ambiguïtés syntaxiques. Or cette exigence d'univocité restreint drastiquement les choix de la forme des règles grammaticales. A un premier stade d'approximation, cette condition conduit à écarter les grammaires dépendantes du contexte au profit des grammaires indépendantes du contextes, puisque seules ces dernières se prêtent à un décodage univoque des informations structurales. Pour cette raison nous avons pu, dans notre commentaire sur le traitement mathématique de l'ambiguïté, nous limiter au cas de ces dernières, malgré le fait qu'elles ne soient jamais utilisées en tant que telles par les générativistes. C'est seulement dans cette mesure que l'on peut dire que le recours à des grammaires indépendantes du contexte est nécessaire pour le traitement de l'ambiguïté. Les exemples du type "The shooting of the hunters" qui exige un traitement transformationnel montre que ces grammaires n'ont pas les moyens suffisants pour traiter intégralement les ambiguïtés syntaxiques. Chomsky (1957, p.97 n.6), bien qu'argumentant en faveur d'un traitement transformationnel fait remarquer qu'on pourrait considérer que l'ambiguïté de "The shooting of the hunters" provient de celle de "to shoot" qui est à la fois transitif et intransitif. En adoptant le principe d'une distinction lexicale entre deux verbes "to shoot", les deux interprétations seraient représentées par deux indicateurs syntagmatiques distincts qui se différencieraient par

la présence des symboles catégoriels Vtr et Vint dans la même configuration. Cependant, cette solution n'est pas satisfaisante puisque ce dont on veut rendre compte, en définitive, c'est que "hunters" est tantôt sujet, tantôt objet.

Mais il y a un argument mathématique qui semble encore plus décisif pour adopter un traitement transformationnel de l'ambiguïté. Il découle d'une observation faite par Peters et Ritchie (1971, p.498) : pour tout langage récursivement énumérable, il existe une grammaire transformationnelle non ambiguë, quel que soit le degré d'ambiguïté des éléments du langage. Malheureusement, ce résultat ne signifie pas nécessairement que la solution transformationnelle soit parfaite. La remarque de Peters et Ritchie est le corollaire d'un théorème qu'on peut sommairement interpréter comme prescrivant l'usage de transformations effaçant des morphèmes sémantiquement pleins. Bien qu'on ne puisse affirmer avec certitude que la non-ambiguïté des grammaires transformationnelles soit liée directement à la présence de transformations d'effacement, c'est cependant probablement le cas. Car les effacements permettent d'engendrer des dérivations de longueur non bornée, donc d'associer (8) un nombre illimité de dérivations à toute séquence terminale. On peut voir ce procédé à l'oeuvre dans la dérivation transformationnelle suivante (9).

The	cow-boys		shoot		the hunters	⇒
	SN _I		V		SN ₂	

The		shooting		of the hunters		by the cow-boys	⇒
Det		V+ing		of + SN ₂		by + SN _I	

The	shooting	of the hunters	
Det	V+ing	of + SN ₂	ø

proposée par Chomsky (1957) où l'homonymie de construction de la phrase dérivée introduit un effacement illicite, car source d'excès de capacité générative.

Si la remarque de Peters et Ritchie était encore vraie pour la classe des grammaires transformationnelles décidables, mathématiquement le problème de l'ambiguïté recevrait une solution optimale. En effet, on disposerait ainsi d'une classe de grammaires engendrant des langages rékursifs produisant des indicateurs syntagmatiques et capables de traiter toutes les ambiguïtés observées et observables dans les langues naturelles.

Mais, d'une part, rien ne permet de supposer que de tels objets mathématiques existent, d'autre part, même s'ils existaient, la construction d'une telle grammaire pour une langue naturelle resterait à faire. A ce stade apparaîtraient alors des difficultés que les mathématiques ne peuvent plus traiter. Dans le modèle transformationnel idéal que nous venons d'évoquer, aucune raison de principe ne peut s'opposer au traitement transformationnel de toutes les ambiguïtés. En particulier, les ambiguïtés purement lexicales comme celles d'"amateur", dans "Max joue au football en amateur"⁽¹⁰⁾ peuvent recevoir un traitement transformationnel. Il est évident qu'une telle solution serait arbitraire et, ce qui est plus grave, qu'elle compliquerait démesurément la grammaire. Le problème ici est de trouver des critères corrects de triage

entre les ambiguïtés syntaxiques et les ambiguïtés lexicales. Gross (1975) montre que de telles décisions ne sont pas toujours faciles à prendre. L'ambiguïté de certains verbes comme "exclure" dans "Paul exclut qu'on favorise Pierre" opposée à la non-ambiguïté de "La situation exclut qu'on favorise Pierre", peut laisser croire que celle-ci est d'origine syntaxique (présence d'un sujet humain dans la première phrase). Cependant l'observation des différentes propriétés distributionnelles d'"exclure" conduisent à distinguer deux verbes distincts malgré des sens extrêmement voisins. Gross admet qu'il n'est pas très satisfaisant de procéder ainsi, mais il fait remarquer qu'une solution transformationnelle ne serait pas meilleure car il faudrait introduire une transformation qui ne pourrait opérer sur les verbes ayant des propriétés syntaxiques identiques à celles d'"exclure" (P.139-40). C'est un cas parmi beaucoup d'autres qui montre que le choix entre une solution syntaxique ou lexicale comprend souvent une part d'arbitraire.

Ainsi l'existence théorique d'une solution mathématique au problème de l'ambiguïté est sans incidence sur la pratique linguistique. On pourrait même penser qu'elle est source de difficultés nouvelles car elle instaure une doctrine libérale autorisant tous les excès, dont le traitement transformationnel des ambiguïtés lexicales n'est même pas un cas limite^(II).

4.3.2. Ces remarques nous permettent de faire une première évaluation des rapports entre la grammaire formelle et la grammaire générative qui dépasse le seul problème de la grammaire for-

melle comme théorie de la syntaxe pure. On pourrait dire que la linguistique mathématique apporte deux types de connaissances, qualitatives et quantitatives. Dans le cas du qualitatif c'est, bien souvent, sous la forme de théorèmes d'existence que les mathématiques nous livrent des informations sur la théorie des grammaires. Elles peuvent alors répondre à des questions totales du type : est-il possible d'adopter telle règle, tel modèle, telle procédure, etc...? Comme généralement ces théorèmes d'existence ne sont pas assortis de procédures constructivistes permettant de déterminer l'objet dont l'existence est assertée, il s'ensuit que les réponses aux questions précédentes ne sont intéressantes que lorsqu'elles sont négatives, c'est à dire lorsqu'on a affaire à des théorèmes de "non-existence". Ces réponses négatives sont significatives et utiles car elles permettent d'éliminer des choix conceptuels contradictoires soit avec les nécessités empiriques, soit avec les dispositions internes de la théorie, contraignant ainsi l'ensemble des choix possibles. Inversement, les réponses positives ont généralement peu de significations car elles indiquent seulement que certaines options théoriques ne sont pas logiquement interdites, mais sans donner pour autant de renseignements sur la façon dont on doit procéder pour exploiter la liberté d'usage de ces options. Cependant dans le cas particulier du traitement transformationnel de l'ambiguïté, il est clair que le manque de signification du théorème a son origine dans l'indétermination du problème posé. Il serait en effet déraisonnable d'attendre de la grammaire formelle une réponse ins-

tructive alors qu'on ne dispose pas d'une définition linguistique précise de l'ambiguïté syntaxique en toute circonstance. Nous aurons l'occasion de constater dans les chapitres ultérieurs que lorsque les notions linguistiques sont nettement circonscrites, il devient possible soit de construire des modèles grammaticaux satisfaisant des conditions mathématiques données, soit de vérifier que des grammaires particulières respectent ces conditions.

Les rapports de la grammaire générative à une mathématique quantitative posent un problème sensiblement différent. D'abord, notons qu'une telle forme de connaissance ne prend son sens que si elle est orientée vers l'empirie puisqu'elle concerne la mesure d'une grandeur - le degré d'ambiguïté des expressions d'un langage dans notre cas particulier. Or s'il est vrai, comme nous l'avons soutenu dans un chapitre précédent que la grammaire générative repose sur des principes épistémologiques ayant pour conséquence son indépendance à l'égard des faits observables, on comprendra immédiatement quelle est l'origine du hiatus entre un modèle mathématique exprimant des mesures et la grammaire générative. L'objection que tout phénomène linguistique accessible à l'aide de moyens statistiques ou probabilistes relève nécessairement de la Performance, n'est pas ici recevable. En effet, nous avons bien constaté que l'incertitude sur le degré d'ambiguïté, dont on ne sait s'il exprime une propriété de la grammaire ou une propriété du langage, pouvait être éliminée par une enquête statistique ne privilégiant d'aucune manière des usages particuliers. L'étude de l'ambiguïté

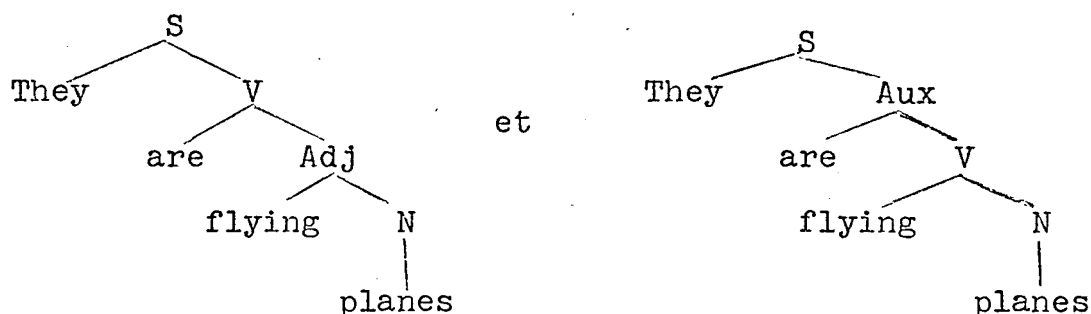
té n'est d'ailleurs pas l'unique voie ouverte aux méthodes "statistico-algébriques". Il serait aisé d'en suggérer d'autres. La théorie probabiliste des grammaires génératives pourrait par exemple permettre d'élaborer une explication de l'inacceptabilité des phrases grammaticales en termes de probabilité de production.

Toutefois il serait inopportun de conclure à la seule responsabilité de la théorie des grammaires génératives dans le demi échec de cette tentative de mathématisation. C'est avant tout d'une définition mathématique trompeuse de la notion de langage ambigu que proviennent les imperfections de la théorie mathématique de l'ambiguïté. D'une façon plus générale, on peut estimer que le concept même de langage formel est trop pauvre pour être considéré comme un modèle satisfaisant du langage humain. Ceci ne veut pas dire que les rapports de la grammaire formelle et de la grammaire générative sont condamnés à la vacuité, mais seulement que la théorie des grammaires formelles n'est pas un instrument efficace de connaissance directe du langage humain. On pourrait croire qu'une fois privé de cette possibilité une linguistique mathématique n'a plus de raison d'être. La seconde partie de notre travail s'attachera à montrer qu'il n'en est rien et qu'une théorie mathématique des grammaires peut être fort utile, parfois même indispensable, pour résoudre les problèmes méthodologiques, théoriques, voire épistémologiques que pose la construction d'une discipline linguistique visant à la rigueur.

NOTES DU CHAPITRE IV

- (1) Pour ce qui concerne l'actualité récente ce problème ne se pose même plus puisque le composant transformationnel des grammaires est pratiquement réduit au néant Cf. Chomsky (1979b)
- (2) Hockett (1954, p.217). Les formes sont simples ou somposées. Les morphèmes sont des formes simples, les formes composées consistent en deux ou plusieurs constituants immédiats figurant dans une construction et formant un constitué . (p.214).
- (3) Le terme de grammaire ne trahit pas la pensée de Hockett. Il est toujours étonnant de constater que l'auteur de "The State of Art" ait pu écrire "A grammatical description (...) sets forth principles by wich one can generate any number of utterances in the language, in this sense it is operationnally comparable to the structure of that portion of human being wich enables him to produce utterances in a language : i.e. to speak" (p.257). Ou encore plus loin (p.232): "A model must be productive when applied to a given language the results must make possible the creation of an indefinite number of valid utterances".
- (4) L'analyse de la notion d'homonymie de construction pose cependant un problème de fondement qui semble escamoté par Chomsky. On ne comprend pas pourquoi la possibilité d'être distincts qui est attribuée à deux indicateurs syntagmatiques, si ceux-ci ne diffèrent que par la nature de leurs symboles catégoriels, est refusée aux graphes des grammaires à états finis. Ainsi, d'après la note 6 de la page 97 de Chomky (1957) "The shooting of the hunters" pourrait éventuellement recevoir deux indicateurs syntagmatiques différents sur la base d'une distinction entre "to shoot" tran-

sitif et "to shoot" intransitif, mais



ne sont pas considérés comme des dérivations distinctes, alors qu'elles répondent au même critère

- (5) A l'exception de types très particuliers de langages, les langages bornés sur lesquels nous reviendrons plus loin.

Par ailleurs il est vrai qu'il existe un théorème dit de "caractérisation des langages indépendants du contexte" du principalement à Chomsky (1962) et Stanley (1965). Mais ce théorème qui exprime que tout langage indépendant du contexte est l'image homomorphique de l'intersection d'un langage régulier et d'un langage de Dyck, suppose déjà établie l'indépendance du contexte du langage en question. L'examen de la démonstration fait d'ailleurs apparaître le rôle crucial qu'y jouent les productions de la grammaire engendrant ce langage.

- (6) Dans la mesure où on peut considérer un langage de programmation comme un langage indépendant du contexte. On sait que, par exemple, la forme normale de Backus telle qu'elle est présentée dans le rapport de Naur (1960), n'est pas un modèle exact de l'Algol.

- (7) Dans notre exemple, nous avons vu que les ambiguïtés apparaissaient au quatrième ordre de dérivation malgré le nombre restreint de règles. Il est probable qu'avec un nombre élevé de règles (et de symboles catégoriels) il faille attendre un ordre d'approximation beaucoup plus grand.

- (8) Voir le chapitre VI pour des commentaires détaillés sur les

théorèmes de Peters et Ritchie et sur la décidabilité des grammaires transformationnelles.

- (9) Le remplacement du sujet réel "cow-boy" par un sujet possible dans le modèle dit "classique", afin de respecter le principe de conservation du sens sous transformation n'affecte pas nos remarques comme le montre le théorème de Peters et Ritchie. Le fait capital étant ici l'émondage dans l'arbre de la branche dominant le sujet, que celui-ci soit "cow-boy" ou postiche.
- (10) Certains sens de "amateur" peuvent être analysés comme des dérivés transformationnels de constructions syntaxiques complexes (Benvéniste, 1969), mais certainement pas tous.
- (II) On sait que c'est sur la base de principes analogues que s'est développée la sémantique générative.

DEUXIEME PARTIE

LA THEORIE DES GRAMMAIRES FORMELLES COMME INSTRUMENT

DE CONTROLE DE LA THEORIE DES GRAMMAIRES GENERATIVES

CHAPITRE V

L'INADEQUATION DES GRAMMAIRES A ETATS FINIS

5. A la fin du troisième chapitre, nous rappelions les distinctions proposées par Chomsky différenciant les niveaux de langage présents au sein d'une théorie linguistique. On pourrait dire qu'à chacun de ces niveaux de langage correspond, en adoptant, librement cette fois, une autre distinction terminologique chomskyenne, l'une des formes d'adéquation que doit satisfaire toute théorie linguistique intéressante : adéquation explicative, descriptive, observationnelle. Dans notre interprétation l'adéquation observationnelle relève du niveau métalinguistique, c'est à dire de ce que la grammaire affirme directement du langage naturel. L'adéquation descriptive, elle, concerne le niveau méta-grammatical, sa satisfaction équivaut à celle des rapports entre la forme et les objectifs de la grammaire. Les contraintes déterminées par la satisfaction de l'adéquation descriptive sont exprimées dans un méta-langage par rapport à la grammaire, c'est à dire dans le langage de la théorie, et elles ont trait principalement à la construction de la grammaire considérée comme un système global, à la différence des contraintes déterminées par l'adéquation observationnelle qui concernent l'ajustement des règles de la grammaire aux données. Enfin l'adéquation explicative détermine l'organisation entière de la théorie dans un langage méta-théorique. A l'intérieur de celui-ci sont formulés les choix épistémologiques fondamentaux guidant la démarche théorique. A ce titre ce méta-langage n'est pas susceptible d'un contrôle réglé par des principes d'une rigueur comparable à celle mise en oeuvre dans les

autres méta-langages. Cependant les options qui y sont prises peuvent être la conséquence, de façon indirecte, de déductions purement logiques opérées au niveau des autres méta-langages. Simplement le langage métathéorique ne peut prétendre a priori à une justification ultime de ses assertions qui soit autre que de "bon sens". En définitive, la caractéristique du langage méta-théorique serait qu'il ne peut être le langage objet d'aucun méta-langage et que par suite aucun principe de déduction ne peut y être formulé.

Relier, comme nous le faisons, niveaux de langage et formes d'adéquation suppose des définitions de l'adéquation explicative et descriptive nettement distinctes de celles proposées par Chomsky. On sait que pour ce dernier une telle distinction repose essentiellement sur l'opposition entre données externes et principes internes. L'adéquation explicative n'est intéressée que par la détermination de principes réglés internes permettant de définir "une procédure pour évaluer des grammaires en compétition" (Chomsky, 1965, p.31). D'une façon plus générale c'est sur la base de l'adéquation explicative que s'effectuent les choix parmi les termes satisfaisant l'adéquation descriptive. Sur un plan pratique, cette dernière - qui englobe l'adéquation observationnelle comme aspect marginal - concerne la majeure partie du travail accompli à l'intérieur de la théorie linguistique puisqu'en relève la définition de la plupart des termes empiriques, par exemple celle de la notion de "phrases possibles", et des termes théoriques,

par exemple celui de "grammaire générative". Comme nous le disions à l'instant, l'adéquation descriptive, au sens de Chomsky, est fonction des données externes. C'est la raison pour laquelle nous ne pouvons accepter en tant que telle cette distinction entre adéquation explicative et adéquation descriptive. En effet, rappelons que dans notre interprétation de la théorie des grammaires génératives, les données externes sont assimilées à des données internes auxquelles on ne peut accéder qu'au moyen de la théorie. Par exemple, la définition extensionnelle de la notion de "phrase possible" revient à l'énumération d'énoncés retenus par le linguiste au nom de son seul jugement de grammaticalité, énumération qui peut toujours être envisagée sous le mode formel de l'ordre lexicographique^(I).

En proposant une interprétation personnelle des différents types d'adéquation que doit satisfaire une théorie linguistique, nous ne sommes pas guidés par des intentions polémiques, ni même n'espérons-nous introduire de la sorte des clarifications épistémologiques. Notre propos est simplement de délimiter une classification des problèmes traités par la grammaire formelle, les rendant interprétables dans le cadre de la "théorie linguistique moderne". Or il nous semble que les distinctions entre adéquation observationnelle, descriptive et explicative, telles que nous les avons formulées, fournissent une telle classification. Ce sera la thèse principale de cette seconde partie que de présenter la théorie des grammaires formelles comme un instrument de contrôle des conditions minimales d'adéquation. Sous quelle forme et jusqu'à quel point

la grammaire formelle accomplit et peut accomplir cette tâche? Telles sont les questions auxquelles nous tenterons de répondre dans ces pages où nous consacrerons un chapitre à chacun des niveaux de langage en examinant un problème particulier propre à chaque forme d'adéquation.

Cependant chacune des questions abordées au cours de ces trois chapitres présente en elle-même un intérêt spécifique susceptible d'illustrer des aspects particuliers des rapports de la théorie des grammaires génératives et de la théorie des grammaires formelles échappant à la perspective générale dans laquelle nous venons de nous placer. Aussi ne manquerons-nous pas de nous intéresser à ces questions incidentes lorsqu'elles se présenteront.

C'est par l'examen d'un exemple célèbre que nous exposerons en quoi la grammaire formelle peut jouer un rôle crucial dans la construction de tests éliminant des grammaires ne satisfaisant pas l'adéquation observationnelle. Il s'agit de la démonstration, donnée par Chomsky (1956) puis par Bar-Hillel et Shamir (1960), du fait que l'anglais ne peut être décrit par une grammaire à états finis.

Les "Three Models for the Description of Languages" retiennent l'attention à plus d'un titre. Historiquement, on peut y voir l'acte de naissance de la grammaire formelle. L'approche mathématique des phénomènes linguistiques y est sans précédent et diffère radicalement, par exemple, de celle de la théorie mathématique de la communication. Dans ce dernier cas, l'idée de

Shannon était d'appliquer à un système de communication non privilégié, le langage naturel, des méthodes éprouvées par ailleurs sur la seule base de l'analogie de certains problèmes selon toute vraisemblance communs aux langues naturelles et aux systèmes de télé-communication, tels que la recherche d'un codage optimum, la mesure de la redondance, etc... Inversement, l'approche mathématique adoptée par Chomsky est immanente à l'investigation des propriétés syntaxiques des langues naturelles, ceci même si ce caractère immanent est dissimulé par la superposition d'abstractions nécessairement réductrices qu'implique une telle approche.

Evoquer la question de l'abstraction, c'est sans doute toucher le point sensible, doublement sensible même, de la démarche de Chomsky, car une approche aussi novatrice que celle-ci suppose une élaboration conceptuelle surmontant deux ordres de difficultés. D'une part, faire accepter le bien fondé de notions linguistiques inhabituelles, d'autre part mettre sous une forme mathématique satisfaisante une intuition lumineuse pour l'expression de laquelle n'existe pas de cadre de référence établi. Deux démarches parallèles que nous allons maintenant examiner dans le détail et dont, vingt-ans après, on peut à peine dire qu'elles sont arrivées à leur terme puisqu'encore aujourd'hui la notion de grammaticalité, une des idées clés sur laquelle s'appuie la démonstration de Chomsky, n'est pas toujours comprise et que par ailleurs le théorème de Chomsky n'a reçu une démonstration satisfaisante que tout récemment.

5.I. Sens et portée de la notion de "grammaticalité"

On sait à quelles controverses la notion de grammaticalité, avec son pendant l'acceptabilité, a pu donner lieu. Notre intention n'est pas de donner une contribution complémentaire à ce débat, auquel nous n'avons rien à ajouter qui n'ait déjà été dit, mais simplement d'essayer de situer les divers arguments avancés par rapport à la question qui nous intéresse; c'est à dire leur rôle dans la démonstration de Chomsky.

Mais auparavant nous exposerons brièvement, une fois encore, les arguments linguistiques avancés par l'auteur pour justifier l'élimination des grammaires à états finis comme procédé inadapté à la description de la syntaxe de l'anglais.

5.I.I. Chomsky fait observer que les trois constructions syntaxiques suivantes sont employées naturellement dans la langue anglaise et qu'il n'existe aucune raison d'en limiter l'usage dans une grammaire de l'anglais:

- (1) If S_1 then S_2
- (2) Either S_3 or S_4
- (3) The man who said S_5 arrives tomorrow

où S_1 , S_2 , etc... désignent des phrases grammaticales déclaratives quelconques.

L'emploi de ces schémas syntaxiques n'admet pas de restrictions. C'est à dire que les cas particuliers suivants:

- (I') If (1) then S_2 = If if S_1 then S_2 S_2
- (I'') If (I') then S_2 = If if if S_1 then S_2 then S_2 then S_2
- .
- .
- .

doivent être considérés comme des phrases grammaticales. Evidemment il s'agit là de faits marginaux, et ce que Chomsky veut relever c'est le caractère général de ces processus itératifs - rékursifs selon la terminologie de la grammaire générative - dans les constructions de la forme (1), (2) et (3). Le fait pertinent ici, celui que traite la grammaire par un schéma formel simple, c'est la dépendance syntaxique régissant l'usage des couples de termes (If, then), (Either, or), (The man who said, arrives tomorrow). Si le premier terme d'un de ces couples figure dans une phrase, le second terme doit nécessairement y figurer, et ceci dans une position que les règles syntaxiques déterminent complètement. Ainsi ces schémas de règles interdisent la production des énoncés suivants:

If the man who said S arrives tomorrow $\left\{ \begin{array}{l} \text{when} \\ \text{or} \\ \text{John} \end{array} \right\}$ it will rain

le paradigme des morphèmes, ou séquences de morphèmes pouvant figurer devant "It will rain" étant réduit à "then"⁽²⁾. C'est cette propriété, combinée avec le fait que le nombre d'éléments pouvant figurer entre les deux termes du couple (ici "If" et "then") est illimité, qui exige d'une grammaire générative de l'anglais qu'elle soit plus puissante qu'un automate à états finis. En effet, un tel automate devrait disposer d'un nombre d'états croissant exponentiellement avec le nombre d'éléments interposés entre les termes du couple, ce qui serait contraire à sa définition. Pour qu'un automate à états finis puisse théoriquement engendrer l'anglais, il faudrait donc renoncer à l'hypothèse de la non-limitation des phrases enchassées.

5.1.2. Les arguments de Chomsky n'ont jamais fait l'unanimité

parmi les linguistes. Reich (1969), par exemple, estime que les formes $I^n S_1$ (then S_2)ⁿ autorisées par l'analyse de Chomsky ne peuvent être assimilées à des phrases grammaticales. La distinction entre grammaticalité et acceptabilité qui permettrait d'expliquer pourquoi ces phrases ne sont pas reconnues comme bien formées par les locuteurs anglais lui semblent empiriquement injustifiable. La notion de "phrase grammaticale" étant équivalente à celle de "phrase produite par la grammaire", il n'y aurait comme issue à ce cercle vicieux que le jugement arbitraire du linguiste se substituant à tout critère objectif permettant de distinguer l'agrammatical de l'inacceptable. Nous verrons plus loin comment doit être appréciée l'accusation de circularité. Quant au reproche d'arbitrarité, il semble au premier abord confirmé par les tests psychologiques de grammaticalité. Des locuteurs placés devant des énoncés de la forme (3) avec un grand nombre d'enchassements ou a fortiori devant les énoncés de la forme (I^n) liraient ceux-ci avec une intonation dont l'absence de contour est identique à celle qui résulte de la lecture d'une suite quelconque de mots dépourvus de toute structure grammaticale (Miller, 1962). Cependant, Miller et Chomsky (1963) estiment que de telles expériences ne sont pas probantes puisqu'elles font intervenir de nombreux facteurs totalement étrangers à la connaissance implicite des règles grammaticales. Les contingences de la performance pouvant être progressivement éliminées si on autorise les sujets de l'expérience à utiliser certains moyens auxiliaires leur permettant d'éliminer les restrictions de mémoire, de temps, de

disponibilité, etc... Miller et Chomsky ne se prononcent pas sur la nature de ces moyens auxiliaires, mais Langendoen (1977) interprète leurs suggestions de la façon suivante : "... si on donne à une personne le temps de réfléchir sur le statut de ces phrases (...), et si on lui donne également de l'espace auxiliaire non structuré de calcul (par exemple un crayon et du papier) pour analyser leurs propriétés, elle trouvera acceptable un plus grand nombre de phrases qu'elle ne le ferait dans les conditions ordinaires d'utilisation du langage"⁽³⁾. L'acceptation de phrases de longueur arbitraire dépendrait donc seulement de la grandeur du temps et de l'espace (dimension de la feuille de papier, longueur du crayon?) mise à la disposition des sujets. Malheureusement, comme le fait remarquer Langendoen, cette hypothèse n'est généralement pas vérifiée dans des limites raisonnables de temps et d'espace. Celui-ci estime cependant qu'il n'en faut rien conclure puisqu'il est toujours possible que l'échec de ces locuteurs soit imputable au fait qu'ils ne savent pas utiliser leur connaissance implicite de la grammaire dans l'analyse des phrases testées⁽⁴⁾. Autrement dit, la grammaticalité de la plus inacceptable des phrases peut être reconnue par n'importe quel locuteur si, comme au singe dactylographe de Borel on lui laisse assez de temps et de papier pour qu'il parvienne à y gribouiller les principes de la grammaire générative!

Une seconde objection due entre autres à Levelt (1974) concerne non pas l'excès de capacité générative des schémas syntaxiques proposés par Chomsky, mais au contraire leur carac-

rière insuffisamment restrictif. Si ces règles étaient correctes, alors les phrases du type "John drank his coffee then he left", "John drank his coffee then he left then he met his sister", etc... seraient agrammaticales, de même que la phrase "If John sleeps he snores". Ces faits devraient amener à penser que l'ensemble des phrases de la forme $If^n S_1 (then S_2)^m$ est constitué de phrases grammaticales, or un tel ensemble peut être produit par une grammaire régulière, ce qui invalide les démonstrations de Chomsky et de Bar-Hillel et Shamir.

5.I.3. Ces objections semblent reposer sur un certain nombre de malentendus. Mentionnons en premier lieu la confusion entretenue par le choix d'exemples, dans la discussion que nous venons d'évoquer, où le statut de grammaticalité est refusé à certaines phrases en fonction du fait qu'elles débutent par une suite de "if" plutôt qu'en fonction du fait qu'elles comprennent des phrases auto-enchassées, ce qui constitue quand même le point essentiel de l'argumentation de Chomsky. Par ailleurs, ce que Chomsky avance, ce n'est pas tant que la co-occurrence de "if" et "then" (ou des termes de n'importe lequel des autres couples) est un phénomène grammatical, mais plutôt que la co-occurrence de couples interdits est impossible. Le respect de ces données est une condition nécessaire d'adéquation observationnelle du phénomène, ce n'est évidemment pas une condition suffisante. Il suffit de remarquer qu'une grammaire indépendante du contexte permet d'engendrer les phrases des formes (1) (2) et (3). Si la production de ces phrases était une condition suffisante d'adéquation descriptive, c'est à dire

si elle présentait le maximum de complexité syntaxique requise par une grammaire de l'anglais, on pourrait estimer que celui-ci est un langage indépendant du contexte, proposition que Bar-Hillel et Shamir démontrent comme fausse avec des moyens identiques à ceux utilisés pour démontrer l'inadéquation des grammaires régulières. Les critiques de Reich et de Levelt semblent formulées comme si Chomsky, réfutant l'adéquation des grammaires régulières, voulait établir celle des grammaires indépendantes du contexte. Ce qui n'est évidemment pas le cas.

D'autre part, pour Chomsky, ce qui est intéressant dans la formulation des schémas (1), (2) et (3), c'est leur simplicité. Le choix d'incorporer de telles règles dans la grammaire, même s'il est coûteux du fait de la production de phrases inacceptables est économique sur le plan de l'organisation de la grammaire puisqu'il évite la spécification de contraintes ad hoc sur la structure interne des phrases enchassées.

La présence potentielle d'"exceptions" à ces règles pourrait faire penser que la grammaire générative se retrouve dans une situation identique à celle de la grammaire traditionnelle. Ce serait là un rapprochement douteux car dans la grammaire scolaire l'exception est la marque d'une irrégularité. On peut même interpréter comme une explication satisfaisante de l'inacceptabilité d'une phrase parfaitement construite selon les règles de la syntaxe, son analyse grammaticale en termes du nombre d'auto-enchassements.

La question du choix d'un critère de grammaticalité reste cependant ouverte puisque, pour Chomsky, les tests psychologiques d'acceptabilité ne peuvent remplir une telle fonction. Notons que cette position n'est pas une simple conséquence de la distinction entre compétence et performance (d'ailleurs formulée postérieurement à la distinction grammatical - acceptable), mais qu'elle découle plutôt de l'existence potentielle, notée par Quine, d'énoncés inaccessibles à l'observation. Quine, on le sait, fait une distinction entre la classe des énoncés "qui se présentent, se sont présentés ou se présenteront aussi bien dans les limites de l'observation professionnelle qu'en dehors de celle-ci" et la classe infinie des énoncés qui pourraient être prononcés, laquelle inclue évidemment la précédente⁽⁵⁾.

Pour pouvoir rendre compte des énoncés inobservables de cette classe, la seule solution possible, pour Quine comme pour Chomsky, est d'expliquer ce qui pourrait être sur la base de ce qui est " et en outre sur la simplicité des lois avec lesquelles nous décrivons et extrapolons ce qui est".⁽⁶⁾ L'inacceptabilité des phrases présentant un nombre trop élevé d'auto-enchassements pourrait peut-être nous faire croire que les règles de la grammaire sont mal choisies, mais cela ne semble pas être le cas, car il faudrait alors considérer ces énoncés comme phrases observées. Ce qu'ils ne sont pas, puisque, par principe, ils se situent au-delà de la classe des énoncés matériellement observables. Ils sont rejetés par les locuteurs com-

me serait rejetée toute phrase de plus de mille mots. Donc, même si une phrase est lue avec une intonation semblable à celle qui résulte de la lecture d'une suite quelconque de mots, la différence entre ces deux "énoncés" est que dans le cas du premier la grammaire permet de lui attribuer une structure sur la base d'observations.

Un autre argument peut être invoqué pour justifier la production par la grammaire de phrases ayant un nombre illimité d'auto-enchassements et plus particulièrement de phrases de la forme (I^n) et (2^n) , il s'agit de la traductibilité des énoncés du calcul propositionnel dans le langage naturel. Refuser le statut de grammaticalité aux phrases précédentes reviendrait à voir dans le langage symbolique des mathématiques le support autonome d'un mode d'énonciation échappant totalement à tout décodage verbal⁽⁷⁾. La traduction des expressions mathématiques dans le langage commun donne d'ailleurs lieu

à des cas typiques de ce que les linguistes considèrent comme des faits de performance. Il est par exemple peu probable que la façon dont les algébristes italiens du XVIème siècle énonçaient l'équation du second degré puisse être considéré comme un énoncé acceptable.

Enfin, l'objection de Levelt selon laquelle l'ensemble des constructions avec "if" et "then" est de la forme $If^n S_I (then S_2)^n$, et que comme tel il peut être engendré par une grammaire régulière n'est d'aucune portée puisque la validité des schémas (2) et (3) n'est pas concernée. De plus, les constructions faisant intervenir "if" et "then" présentent malgré

tout des dépendances illimitées, même si les deux conjonctions ne sont pas toujours strictement co-occurentes, si on interprète cette contrainte syntaxique de façon "négative" comme nous l'avons fait plus haut (voir la note 2).

Ceci dit, il est indéniable que la notion de grammaticalité fait problème. Dans le second chapitre nous l'interprétons strictement comme un justificatif pour le grammairien qui ne s'intéresserait qu'à son idiolecte. Cette interprétation n'exprimait qu'un sentiment personnel à l'égard de la pratique de la grammaire générative telle que nous pouvons l'observer. Cependant, si on la relie à la volonté d'éliminer la langue comme objet possible de connaissance nous trouvons, aujourd'hui plus que jamais, une confirmation de ce point de vue chez Chomsky (1979a) pour qui seule la grammaire est digne d'intérêt⁽⁸⁾.

La question demeure toutefois de savoir s'il est possible de définir une notion de grammaticalité valable pour un langage qui ne serait pas identifiable au seul idiolecte du grammairien. En particulier la partition des phrases acceptables en grammaticales et agrammaticales est-elle possible sans préjuger de la nature de la grammaticalité? Pour Chomsky (1957), comme pour Quine, la notion de "phrase grammaticale" devant être définie en termes de phrases observées, et les phrases observées ne pouvant, par la force des choses, être autres que les phrases acceptables (par un informateur), l'identification du grammatical et de l'acceptable semble aller de soi tant qu'elle concerne "ce qui est". Mais alors, qu'est-ce qui nous

permettra, lorsque l'informateur nous refusera son assentiment sur une phrase "qui pourrait être" (c'est à dire produite par la grammaire), de distinguer si celui-ci la refuse en fonction de son inacceptabilité ou de son agrammaticalité? La grammaire ne peut plus se prononcer sur ce "qui pourrait être" et doit se réduire à la description de "ce qui est".

Si nous nous sommes arrêtés un peu longuement sur ces controverses, ce n'est pas parce que nous estimons que le contenu des schémas (1) (2) et (3) est d'un intérêt particulier. Ce qui nous semble plutôt significatif dans cette polémique, ce sont les résistances à une tentative de réduction du problème, que toute la pratique linguistique rend pourtant légitime. Si la substance même des observations de Chomsky était inexacte, si la dépendance des phénomènes de co-occurrence n'était pas illimitée, la limite supérieure observée de telles dépendances serait suffisamment grande pour accepter, malgré tout, les conclusions de Chomsky. Entre la proposition théorique "Il n'existe pas d'automates à états finis pouvant engendrer l'anglais" et le constat pragmatique "On ne pourra jamais construire un automate à états finis pouvant engendrer l'anglais", la seule différence c'est que l'une appartient à une théorie et l'autre à une pragmatique...

5.2. De la mise en forme d'une intuition à sa démonstration

Sur le plan mathématique, la démarche de Chomsky se décompose en quatre étapes.

I°) définir une propriété combinatoire, la n-dépendance il-

limitée, qui n'est satisfaite par un langage que si celui-ci n'est pas régulier.⁽⁹⁾

2°) définir des classes de langages non-réguliers, comme la classe des langages en miroir, la classe des langages n-dépendants ou la classe des langages de copie.

3°) identifier l'ensemble des phrases produites par les schémas (1), (2) et (3) de la section précédente à un langage en miroir.

4°) affirmer qu'un langage contenant un sous-langage qui n'est pas régulier, n'est pas lui-même régulier.

FAUX Les étapes (2) et (4) de cette démonstration reposent sur des raisonnements incorrects comme l'ont remarqué récemment Langendoen (1977) pour la première et Levelt (1974) pour la seconde. Les étapes (1) et (3) elles, laissent informulés certains détails non négligeables.

5.2.1. Après avoir défini la n-dépendance, Chomsky fait observer que le langage en miroir $L_2 = \{aa, bb, abba, baab, abbbba, \dots\}$ et le langage de copie $L_3 = \{aa, bb, abab, baba, aabbaab, \dots\}$ présentent de telles dépendances pour n arbitraire, mais il ne dit rien de semblable pour le langage $L_I = \{aa, bb, aabb, \dots, a^n n^n \dots\}$ qu'il donne pourtant comme exemple d'un langage non-régulier.

Il apparaît en effet que la n-dépendance est une condition trop forte pour être satisfaite par tous les langages non-réguliers (elle est suffisante mais pas nécessaire). C'est peut-être ce qui conduit Chomsky à identifier l'anglais à un langage en mi-

roir du type L_2 . Le choix d'un codage dans L_I (if = or = the man who said = a, then = either = b, the man... = arrives ... = C) présentant nécessairement un certain degré d'arbitraire on comprend mal les raisons qui peuvent pousser à retenir L_2 plutôt que L_I , sinon que ce dernier ne présente pas de n-dépendance pour n arbitraire⁽¹⁰⁾. Le fait que dans la littérature l'argumentation de Chomsky soit toujours présentée avec une identification de l'anglais à $L_I^{(II)}$ semble confirmer le caractère plus naturel, car plus simple, de cette identification, les auteurs donnant généralement aux propos de Chomsky une interprétation qui n'est pas la leur. A cet égard, il est peut-être significatif que Bar-Hillel et Shamir (1960) aient choisi de redonner une nouvelle démonstration du théorème de Chomsky en identifiant l'anglais à un langage n-dépendant. Pour ce fait ils ne caractérisent plus les langages par les propriétés structurales de leurs éléments, mais par les classes de congruence de ces éléments. Ils utilisent ainsi un outil mathématique plus précis que la n-dépendance puisque, pour un langage, la propriété d'être représentable comme une union finie de classes de congruence est une condition nécessaire et suffisante pour appartenir à la classe des langages réguliers⁽¹²⁾. Bar-Hillel et Shamir qui raisonnent directement sur les phrases de l'anglais font remarquer que leur démonstration ^{ne} serait plus valable si devait être adjointe à la grammaire la règle : "si x et y sont des phrases et x ne débute pas par "if", alors "if x then y" est aussi une phrase". Il faut s'arrêter un ins-

tant sur cette remarque. Elle semble indiquer qu'un certain principe de compensation règle les rapports des mathématiques et de la linguistique. Dans le cas de la démonstration de Chomsky, le théorème s'applique sur un domaine d'objets dont l'extension n'est pas affectée par des contre-exemples particuliers comme celui mentionné par Bar-Hillel et Shamir. Mais par contre certaines classes de langages échappent au domaine d'application de la démonstration de Chomsky, puisque leurs membres ne peuvent être caractérisés par la n -dépendance illimitée. Inversement, la méthode de Bar-Hillel et Shamir vaut pour tous les ensembles de phrases de la forme (1), (2), (3) et pour ceux obtenus par combinaison de ces formes, mais une seule règle linguistique "exceptionnelle" peut bloquer la démonstration.

Sur un autre plan, la comparaison des démonstrations de Chomsky et de Bar-Hillel illustre peut-être les remarques faites par Granger (1968)⁽¹³⁾ sur le rapprochement entre les notions saussuriennes d'axe paradigmatique et d'axe syntagmatique, et les deux manières dont on peut caractériser l'état d'un automate fini. Chomsky faisant usage des relations syntagmatiques de co-occurrences dans les fonctions de transition, Bar-Hillel et Shamir raisonnant sur les classes paradigmatiques naturellement associables à chaque état d'un automate fini. Toutefois nous ne pensons pas que cette caractéristique "témoigne de la profondeur de la conception saussurienne" comme l'écrit Granger. Il nous semble au contraire que c'est parce que les vues de Saussure sont simplificatrices qu'elles sont repré-

sentables par une machine aussi élémentaire (et inadéquate) qu'un automate à états finis. Si on désire conserver l'analogie entre les vues de Saussure et la structure des automates - et nous pensons que celle-ci est justifiée - alors il faut se rallier au point de vue de Chomsky qui voit dans la négligence des phénomènes syntaxiques le point faible de la conception saussurienne de la langue.

Une autre difficulté mineure apparaît dans l'étape (3) de la démonstration de Chomsky, c'est l'identification elle-même de phrases de la forme $x_1 a_1 \dots x_n a_n z b_1 y_1 \dots b_n y_n$ avec des séquences soit de la forme $c_{i_1} c_{i_2} \dots c_{i_n} c_{i_n} \dots c_{i_2} c_{i_1}$ soit de la forme $c^n d^n$ sans que soit mentionnée l'application de recodage dans le langage formel auquel est identifié le langage naturel. Nous savons aujourd'hui que de telles substitutions laissent invariantes la nature des langages si ceux-ci sont des langages réguliers ou des langages indépendants du contexte (Bar-Hillel, Perles et Shamir, 1961, Th.3.3. et son corollaire p.126-8), mais a priori, qui nous dit que les morphismes implicitement utilisés dans ces codages sont conservateurs de structures? Cette observation prend tout son sens si on se souvient 1°) que la classe des langages dépendants du contexte n'est pas close pour les substitutions faisant intervenir le mot nul (c'est l'opération $\Psi(x_i) = e$ utilisée par Chomsky), 2°) que dans le cadre de "Three models", la hiérarchie de Chomsky (1959) n'existe pas encore et qu'en particulier la distinction entre langages dépendants et indépendants du contexte n'est pas formulée. Poten-

tiellement, donc, une identification comme celle à laquelle procède Chomsky peut transformer un langage de constituants en un langage récursivement énumérable non dépendant du contexte. Bien que ce phénomène soit sans incidence dans le cas particulier de la démonstration de Chomsky, on ne pourrait utiliser de procédés analogues si on devait démontrer qu'une grammaire dépendante du contexte a une capacité générative trop restreinte pour engendrer un langage donné. Ces remarques ne remettent donc pas en cause la validité de la démonstration de Chomsky, tout au plus soulignent-elles l'aspect inachevé de ce qui est plus près de la mise en forme d'une intuition que d'une démonstration parfaite. Cependant, ce caractère inachevé est encore plus manifeste lorsqu'on examine les étapes 2 et 4 de la démonstration.

5.2.2. Nous avons vu plus haut que Levelt par exemple, estime que l'ensemble $\underline{\text{If}^n S_1} \underline{(\text{then } S_2)^m}$ correspond de façon plus précise à l'anglais que l'ensemble $\underline{\text{If}^n S_1} \underline{(\text{then } S_2)^n}$. Il est clair que le second est inclus dans le premier. On a là un exemple de langage indépendant du contexte inclus dans un langage régulier. En général ce n'est pas parce qu'un langage contient un sous-langage indépendant du contexte ou dépendant du contexte qu'il est lui-même indépendant ou dépendant du contexte. L'exemple le plus trivial de ce fait est celui de l'inclusion de tout langage dans le langage universel construit sur le même vocabulaire.

Observons que cette généralisation hâtive ne peut être

attribuée uniquement à Chomsky puisque, jusqu'à ce que Levelt la signale, elle était reprise par pratiquement tous les linguistes s'exprimant sur la question⁽¹⁴⁾. Il est permis de supposer que pour les linguistes mathématiciens, cette précision allait sans dire et que nous avons encore affaire ici à un abus de langage plutôt qu'à un manquement à la rigueur. Cependant, le ton convaincu de Ruwet exposant la démonstration de Chomsky semble laisser peu de place à cette hypothèse : "On peut donc rencontrer dans les langues naturelles des dépendances emboîtées les unes dans les autres, sans que l'on puisse fixer une limite théorique au nombre d'emboitements successifs. Il en résulte que les langues naturelles comportent un ensemble infini de phrases qui ne peuvent pas être engendrées par une grammaire à états finis; les langues naturelles ne sont donc pas des langages à états finis."⁽¹⁵⁾

Les difficultés que nous avons mentionnées jusqu'à présent demeurent ambiguës. Il est toujours délicat de juger si elles proviennent d'une absence de formulation ou d'une absence de perception des propriétés des objets mathématiques. Le cas de l'étape 2 de la démonstration de Chomsky est cependant beaucoup plus clair : il y a bien ici une propriété qui n'a pas été perçue. Que celle-ci soit satisfaite dans le cas particulier de cette démonstration ne doit pas nous faire sous-estimer l'importance de cette lacune. Si on veut donner le nom de théorème à la proposition de Chomsky, et par là même instituer une méthode générale de démonstration de l'inadéquation des gram-

maires rudimentaires, il devient indispensable de formuler soigneusement cette étape du raisonnement pour éviter de démontrer des choses fausses.

Cette étape concerne la caractérisation des langages n -dépendants et des langages en miroir. Les langages en miroir sont de la forme générale $L = \{ x\tilde{x}/x \in F \}$ où F est un langage régulier et x l'image inversée de x . Du moins est-ce sous cette forme que Langendoen suppose qu'ils doivent être caractérisés. Le langage constitué des séquences suivantes: $abba, abbaabbaabba, abbaabbaabbaabbaabba, \dots$ répond à la définition générale des langages en miroir bien que la forme de ces expressions, $(abba)^{2n+1}$, montre que celui-ci est n -dépendant puisqu'il est de la forme générale $\{ xu^n w v^n y / u \neq e \text{ et } v \neq e \}$ mais c'est évidemment aussi un langage régulier. La démonstration de la non-régularité de l'anglais reposant sur l'identification d'une partie de celui-ci soit avec un langage miroir soit avec un langage n -dépendant, il est indispensable de disposer de critères garantissant que ces langages sont bien en miroir ou n -dépendants. Cette restriction reste valable lorsqu'on identifie la classe des langages de copie à une sous-classe des langages strictement dépendants du contexte comme le montre Langendoen avec l'exemple du langage n -dépendant $\{ a^n b a^n b \}$ qui a l'aspect d'un langage de copie. Ce dernier exemple est important puisque Postal affirme dans sa démonstration sur le Mohawk, que tout langage de copie n'est pas indépendant du contexte.

Ces incorrections ne pouvaient pas être relevées à l'époque de la publication de l'article de Chomsky. C'est l'étude intensive des grammaires formelles durant la décennie qui suivit la parution de "Three Models" qui nous permet aujourd'hui de voir ce qu'il y a d'approximatif dans cette première tentative d'étude mathématique de la syntaxe. Examinons brièvement quelles modifications doivent être apportées à la démonstration de Chomsky pour la rendre parfaitement rigoureuse. L'étude des propriétés booléennes des différentes classes de langages définies par la hiérarchie de Chomsky (1959) conduit au théorème suivant;

"Pour chaque classe C_i de la hiérarchie de Chomsky, si l'intersection d'un langage quelconque L , avec un langage de type 3, R , appartient à C_i , alors L ne peut appartenir à une classe C_j où $j > i$."

Pour être précis, il faudrait dire qu'il s'agit ici non pas d'un théorème, mais de quatre théorèmes différents dont les énoncés sont regroupés en un seul⁽¹⁶⁾. Nous introduisons cette précision avec l'expression "de la hiérarchie de Chomsky", une formulation trop générale comme celle que donne Langendoen ; pouvant laisser penser qu'ici i et j sont des variables dont le domaine pourrait s'élargir en fonction d'un raffinement de la hiérarchie. L'application directe du théorème montre que, l'intersection de l'anglais et de l'ensemble $\text{if}^n S_1$ ($\text{then } S_2$)^m étant le langage de type 2, $\text{if}^n S_1$ ($\text{then } S_2$)ⁿ, l'anglais ne peut pas être régulier. Remarquons que la forme de

l'expression mathématique dans son aspect négatif, "ne peut être", respecte très bien l'intention de la démonstration de Chomsky telle que nous la reprécisions plus haut : la nécessité d'utiliser une grammaire de type 2 pour décrire l'anglais n'implique pas qu'une telle grammaire suffise à cette tâche. On peut donc ainsi réduire une partie des objections de Levelt et de Reich à une argumentation en partie illogique.

Ce résultat ne présente cependant pas un très grand intérêt mathématique puisqu'il est le corollaire immédiat des propriétés booléennes des langages. On ne peut même pas envisager une extension qui donnerait un critère similaire d'identification pour déterminer si un langage est de type 2 ou non, puisque cette classe n'est pas close sous l'intersection. Toutefois il faut reconnaître, croyons-nous l'intérêt d'une méthode générale qui parvient à réduire certains problèmes d'interprétation empirique, et les polémiques pouvant y être associées, à la simple vérification d'une propriété mathématique élémentaire.

Les critères de caractérisation des langages n-dépendants et des langages en miroir, en revanche, exige une analyse beaucoup plus abstraite des propriétés de ces langages. Contrairement au cas précédent où la solution mathématique était disponible et ne demandait qu'à être exprimée de façon appropriée, c'est un nouveau théorème qui est ici requis. Les conditions nécessaires et suffisantes proposées par Langendoen sont les suivantes:

1°) un langage n -dépendant $L = \{uv^nwx^ny/u \neq e, x \neq e\}$ est de type 2 sauf si $w = e$ et $v = x$.

2°) un langage en miroir $L = \{x\tilde{x} / x \in R \text{ où } R \text{ est de type 3}\}$ est de type 2 sauf s'il existe une séquence r et deux séquences q et s telles que $R = \{qr^ns / r=r, s \neq e \text{ ou } s\tilde{s}=r\}$

On constate que la présence d'un symbole central distinctif dans les séquences du langage (que celui-ci soit en miroir ou n -dépendant) suffit pour que le langage ne soit pas de type 3. Cette condition peut être satisfaite par identification, par exemple, aux langages formels $\{a^ncb^n\}$ ou $\{xc\tilde{x}\}$.

Si l'interprétation naïve des théorèmes de Langendoen se traduit par une propriété extrêmement simple (la présence d'un marqueur central dans les phrases du langage) il ne faut pas déduire pour autant qu'il s'agit là d'une contrainte ad hoc. Ces résultats sont les conséquences d'une analyse détaillée des propriétés les plus fondamentales des langages indépendants du contexte. Les démonstrations de Langendoen utilisent principalement un théorème (le "pumping lemma")^(I7) dont l'origine est double. Formellement, celui-ci est apparenté à un lemme de Bar-Hillel, Perles et Shamir connu sous le nom de "théorème uvwxy" dont la fonction principale est de permettre la construction du théorème de solvabilité de la finitude des langages indépendants du contexte. On peut considérer le "pumping lemma" comme un cas particulier du théorème uvwxy puisque tous deux caractérisent à l'aide de méthodes combinatoires identiques les langages formels par des constantes numériques. La seconde

origine de ce théorème est dans la recherche d'un critère pratique de délimitation des opérations arithmétiques réalisables par un automate à états finis. Deux problèmes dont on ne pouvait prévoir que les solutions permettraient un jour de régler une question de méthodologie de la grammaire générative

5.2.3. Nous pouvons tirer trois conclusions de cette analyse de la démonstration du théorème de Chomsky.

1°) D'abord elle justifie a posteriori l'étude des propriétés formelles des grammaires, quand bien même celle-ci serait menée sans finalité particulière. Nous avons vu qu'une démonstration correcte du théorème n'était possible qu'au prix de l'existence de plusieurs résultats (propriétés booléennes des langages formels, théorème uvwxy, clôture sous la substitution, etc...) établis pour des raisons indépendantes non seulement du théorème de Chomsky, mais encore indépendantes de toute application immédiate à la linguistique.

2°) Ensuite s'il est vrai, comme nous l'avons dit au début de ce chapitre, qu'en formulant son théorème Chomsky devait du même coup inventer un cadre de référence, la théorie des grammaires formelles, pour exprimer celui-ci, on constate que plus l'exploration de cette théorie va de l'avant et plus elle s'enrichie, plus ses problèmes se présentent comme des cas particuliers de questions générales dont la théorie mathématique existe déjà. C'est l'exemple du "pumping lemma".

3°) Enfin, si nous nous intéressons à la signification générale du théorème de Chomsky, c'est à dire à ce qu'il représente en

tant qu'essai de mathématisation que pouvons-nous en penser? On aime souvent dire que les mathématiques permettent d'expliquer et de prévoir des faits. Ce théorème n'explique aucun fait, il semblerait même que ce soit plutôt l'inverse. Quand aux prédictions qu'il autorise, la question est vite réglée : il ne prévoit rien car il n'y a rien à prévoir. En fait ce théorème légalise, si on peut dire, l'attitude des linguistes qui, lorsqu'ils pratiquent la grammaire, ont la tendance naturelle à croire que les relations syntaxiques entre les éléments d'une phrase ne peuvent être déterminées à partir de l'observation d'une simple partie de cette phrase.

NOTES DU CHAPITRE V

- (1) Enuméré ainsi, l'ensemble des phrases d'un langage, quel qu'il soit, est toujours récursif sauf si le grammairien ne sait pas se prononcer sur la grammaticalité
- (2) Peut-être est-il plus large. En particulier le morphème nul peut sans doute y figurer. Cependant quel que soit l'extension de ce paradigme, l'important est de remarquer qu'il est complètement déterminé par le "if" en tête de phrase.
- (3) Langendoen (1977, p.165), nous traduisons littéralement l'expression "auxiliary unstructured computation space"
- (4) "... it could be maintained that people do not know how to integrate external unstructured computation space with their internal computation space, and that if they did, they would be able to determine grammaticality of sentences" (p.166). Rappelons toutefois que Chomsky (1957) (note 6, Ch.8) est très clair sur ce point : "... les tests opérationnels pour les notions linguistiques peuvent demander à l'informateur une réponse, mais non qu'il exprime son opinion sur son comportement, son jugement sur la synonymie, la distinction phonologique, etc..."
- (5) Quine (1953, Ch.3, § 3)
- (6) Selon Quine la grammaire trouve ici même une de ses justifications: "Il n'est pas en effet satisfaisant de dire qu'une séquence signifiante soit simplement une séquence de phonèmes quelconques prononcée par un Naturkinder quelconque de la vallée choisie par notre grammairien" (p.52)
- (7) Nous ne voulons pas dire que le sens des expressions symboliques soit garantie par la possibilité de la traduction dans la langue vernaculaire. Notre remarque exprime simplement qu'il est toujours possible de faire correspondre une expression non symbolique à une expression symbolique.

- (8) "... various questions would remain to be settled as to what we would choose to call a "language", but they do not seem to be particularly interesting, because the notion "language" is derivative and relatively unimportant".
(Chomsky 1979a,p.33)
- (9) En simplifiant à l'extrême, on pourrait caractériser la n-dépendance comme un critère définissant les substitutions, sur les parties propres des expressions d'un langage, qui conservent la bonne formation de ces expressions.
- (10) En effet la permutation des symboles d'une partie initiale d'une séquence de L_I peut laisser cette séquence inchangée.
- (11) Cf. par exemple Gross (1972,pp.126-27), Langendoen (1977, p.161), Brainerd (1971, p.177), Levelt (1974, Vol.2,p.25)
- (12) L'idée de caractériser les langages réguliers par l'index de leur classe d'équivalence est due à Nérøde (1958).
- (13) Granger (1968,p.150). Ruwet (1967,p.93) fait une remarque analogue mais en se référant plutôt au fonctionnement de l'automate : "... chaque transition correspond à une transition dans la phrase, (...) les diverses transitions admises entre deux états successifs correspondent à une classe paradigmaticque".
- (14) Gross (1972,pp.126-27), fait cependant remarquer que l'argument traditionnel est incomplet tant que l'on a pas montré que le sous-ensemble des phrases avec enchassement n'est pas lui-même sous-ensemble d'un langage régulier. Il illustre sa remarque avec l'exemple de l'homonymie de description de phrases coordonnées et de phrases enchassées :
"John, Paul, work, slept, ate," est de la forme SN^nV^m et
"The rat the cat killed disappeared" de la forme SN^nV^n . Le deuxième ensemble est donc "plongé" dans le premier.

(I5) Ruwet (1967, p.96)

(I6) Chomsky et Miller (1958) ont montré que les langages de type 3 formaient une algèbre de Boole, Bar-Hillel, Perles et Shamir (1961) que la classe des langages de type 2 était close sous l'intersection avec les langages réguliers Landweber (1963) que la classe des langages de type I était close sous l'intersection

(I7) Arbib (1969, p.88)

CHAPITRE VI

LA DECIDABILITE DES GRAMMAIRES TRANSFORMATIONNELLES

ET LA CONJECTURE DE PUTNAM

6. Dans la recherche des propriétés formelles des grammaires, la nécessité de recourir à des démonstrations mathématiques ne s'impose pas toujours avec évidence. Il est, par exemple, certain que ce ne sont pas les résultats discutés au chapitre précédent qui ont amené les linguistes à rejeter des modèles d'analyse en constituants aussi évidemment élémentaires que les grammaires à états finis. Les considérations empiriques prouvaient amplement leur limitation. Dans ce cas, la contribution des mathématiques n'est pas négligeable puisque, comme nous l'avons exposé, elle instaure une connaissance d'une autre nature; mais elle n'est pas déterminante si on se restreint au seul problème de la recherche d'un critère d'adéquation pour ces grammaires.

Toutefois il est des circonstances où la démonstration des propriétés mathématiques d'une grammaire s'avère déterminer aussi les conditions sine qua non de son emploi. La tripartition des niveaux de langages présentés au début de cette seconde partie peut nous permettre de comprendre quelles sont ces circonstances. Dans le cas de l'élimination des grammaires à états finis, nous avons affaire à une propriété du langage naturel qui devait être exprimée au niveau méta-linguistique, dans la grammaire. Il était donc normal que de simples observations sur le langage naturel puisse permettre de prendre une décision à cet égard qui soit fondée. Mais il est aisé d'imaginer une situation plus complexe où le grammairien aurait à se prononcer sur une propriété globale du système grammatical inexprimable à l'intérieur de ce système.

Pour au moins deux raisons une telle éventualité semblera bien hypothétique. D'abord, pourquoi envisager que la description grammaticale des énoncés d'une langue naturelle puisse exiger que l'on se prononce sur des propriétés autres que celles impliquées directement par la description elle-même? Ensuite n'est-ce pas se laisser gagner par un pan-logicisme un peu gratuit que de penser que

les problèmes linguistiques soient solubles dans des termes analogues à ceux de la métamathématique?

Le meilleur moyen de répondre à ces objections est de passer directement à l'analyse d'un problème précis dont la solution, vitale pour la théorie des grammaires transformationnelles, ne peut être appréhendée à travers la pratique linguistique mais uniquement par une approche où le niveau métalinguistique est pris comme langage objet et donc, où la nécessité d'un traitement mathématique s'avère indispensable puisque la solution du problème suppose une définition mathématique de la grammaire.

Alors qu'une des hypothèses fondamentales de la grammaire générative exige des langues naturelles qu'elles soient des ensembles récursifs ⁽¹⁾, la classe des systèmes grammaticaux utilisés par les linguistes, les grammaires transformationnelles de la théorie dite "classique", est insuffisamment restreinte à cet égard puisqu'elle englobe des grammaires dont la capacité générative est excessive; tous les ensembles récursivement énumérables pouvant être engendrés par ces grammaires. On a parfois interprété hâtivement ce fait, démontré par Peters et Ritchie (1973a), comme censurant de façon nette et définitive l'usage des grammaires transformationnelles. Nous verrons au cours de ce chapitre à quel point une telle interprétation est sommaire, puisque déjà, sur un plan strictement technique, les théorèmes de Peters et Ritchie signifient que parmi les grammaires transformationnelles, certaines engendrent bien des ensembles récursifs. Une des origines de ces erreurs d'interprétation réside vraisemblablement dans la difficulté du problème mathématique lui-même. Construire un théorème d'adéquation pour les grammaires transformationnelles est en effet, pour le mathématicien, une tâche bien différente de la démonstration de l'inadéquation des grammaires à états finis ou des grammaires indépendantes du contexte. Dans ces deux derniers cas, on dispose d'une délimitation

pré-mathématique de l'objet à sanctionner au nom d'un critère d'adéquation, tandis que pour les grammaires transformationnelles ce n'est pas le cas. Effectivement, donner une définition mathématique des grammaires transformationnelles et énoncer le théorème "tout langage récursivement énumérable est engendré par une grammaire transformationnelle", n'est guère instructif si on ne sait pas spécifier quelles propriétés sont impliquées dans cette définition. Car le fait qu'on soit ici à la recherche d'un théorème d'adéquation, et non plus d'un théorème d'inadéquation comme pour les grammaires à états finis, exclut qu'on considère la classe des grammaires transformationnelles dans toute sa généralité. Et toute la difficulté est de trouver parmi les propriétés formelles de ces grammaires celles qui sont linguistiquement pertinentes, et une fois ceci fait, de dégager parmi ces propriétés pertinentes celles qui pourront garantir la récursivité des langages engendrés. Une telle tâche revient finalement à construire une typologie des grammaires transformationnelles qui soit de plus linguistiquement significative, deux aspects du travail mathématique dont on faisait¹ l'économie dans l'étude des grammaires à états finis.

6.1. La conjecture de Putnam

Avant de passer à l'analyse des contraintes mathématiques garantissant la récursivité des grammaires transformationnelles, nous discuterons brièvement des problèmes épistémologiques que pose l'hypothèse de récursivité des langues naturelles. Nous avons déjà abordé d'un point de vue théorique cette question au cours du premier chapitre, mais ici, c'est sous un autre aspect que nous aimerions l'envisager, c'est à dire comme un fait dont doit rendre compte la grammaire et non plus comme hypothèse inhérente à l'activité grammaticale.

Certains auteurs ont en effet avancé l'idée que le problème était empirique. Soit que les langues naturelles étaient de fait des ensembles récursifs et que par conséquent les grammaires transformationnelles devaient se limiter à engendrer de tels ensembles. Soit que les langues naturelles possédaient une certaine propriété que la grammaire devrait nous permettre de découvrir. et si, malgré tous nos efforts pour restreindre la capacité générative des grammaires transformationnelles, celles-ci continuaient à engendrer des ensembles récursivement énumérables non récursifs, alors il faudrait y voir la preuve qu'il s'agit là d'une caractéristique des langues humaines.

6.1.1. Pour Putnam, qui semble être le premier à avoir fait remarquer que les grammaires transformationnelles pouvaient engendrer tous les ensembles récursivement énumérables, trois arguments peuvent appuyer l'idée que les langues naturelles sont récursives:

(1) L'autonomie du langage ("The self-containedness of language"). C'est à dire "le fait que les locuteurs peuvent probablement classer les phrases comme acceptables ou inacceptables, déviantes ou non déviantes, etc..., sans référence aux contextes extra-linguistiques "... "Il semble qu'en accomplissant cette tâche on utilise implicitement quelque chose ressemblant à une procédure effective". (Putnam, 1961; p. 98)

(2) La possibilité d'opérer sur des phrases dépourvues de sens ("The usability of non-sense sentences") dans la procédure de classification qui permettrait de considérer cette classification comme "quelque chose d'effectif ou de mécanique (et donc que l'ensemble des phrases grammaticales est récursif, du moins si nous acceptons la thèse de Church)".

(Ibid. p.100)

(3) "The teachability of grammar" et l'indépendance relative de cette

aptitude par rapport au niveau d'intelligence: "... Une aptitude de ce type (qui) peut être acquise par pratiquement n'importe qui ou peut-être utilisée par pratiquement n'importe qui indépendamment du niveau d'intelligence, a certainement un caractère quasi-mécanique" (Ibid., p.101). En acceptant le principe q'une machine de Turing est un modèle du classificateur, il devient légitime de rechercher des modèles grammaticaux dont les mécanismes sont représentables par des fonctions récursives.

Putnam observe par ailleurs que "la caractérisation générale donnée par Chomsky d'une grammaire transformationnelle est beaucoup trop large. Il serait aisé de montrer que tout ensemble récursivement énumérable peut être engendré par une grammaire transformationnelle au sens de Chomsky" (Ibid., p.102). Il faut donc restreindre la classe des grammaires transformationnelles possibles de sorte qu'elle ne puisse produire que des langages récursifs.

Toutefois les restrictions proposées par Putnam ne sont pas très satisfaisantes, comme il semble d'ailleurs le reconnaître lui-même: "On pourrait imposer deux restrictions à toute grammaire d'une langue naturelle: (a) que pas plus de n_1 mots puissent être effacés par une transformation d'effacement, et (b) que pas plus de n_2 transformations d'effacement puissent apparaître dans la dérivation d'une séquence terminale, où n_1 et n_2 sont des constantes dépendant du langage. Cependant la seconde restriction semble ad hoc et peu attirante." Remarquons que ces deux conditions n'en font qu'une puisque toute transformation se laisse décomposer en transformations élémentaires (cf. Chomsky (1956)), et qu'une transformation qui efface n_1 éléments est, selon la définition même des transformations, équivalente à une suite de n_1 transformations effaçant un élément. La condition (b) de Putnam avec $n = n_1 \times n_2$ serait suffisante.

Cependant, comme nous le verrons plus loin, une telle restriction n'est pas seulement ad hoc, elle est tout simplement inacceptable.

En conclusion, voici, dans les termes mêmes de l'auteur, ce que l'on pourrait appeler la conjecture de Putnam: "Voilà donc une question ouverte intéressante (et sans doute très difficile): définir une notion de "grammaire transformationnelle":

- (1) suffisamment large pour tous les objectifs linguistiques
- (2) dégagée de restrictions "artificielles" comme celle qui limite le nombre d'effacements dans une dérivation
- (3) telle qu'un théorème d'"élimination des coupures" puisse être énoncé pour tous les langages transformationnels" (Putnam, 1961;p.103)

La problématique adoptée par Putnam appelle un certain nombre de commentaires, en particulier en ce qui concerne les arguments invoqués pour établir le caractère récursif des langues naturelles.

6.1.2. Commentant le premier argument, Putnam affirme qu'en posant une relation entre le travail de classification sur la langue et une procédure effective, il s'appuie sur des hypothèses très générales concernant la nature du cerveau humain ("In this connection I am, of course, relying on certain very general hypotheses as to the character of human brain") (Ibid., p.98). Cette option psychologiste est confirmée par l'énoncé même du troisième argument: "The teachability of grammar and the relative independance of intelligence level of this skill". Il est difficile de savoir exactement ce que Putnam entend par "teachability of grammar". Nous pouvons supposer qu'il ne s'agit pas de la possibilité d'enseigner les règles de la grammaire, car dans ce cas l'avis de Putnam ne serait certainement pas partagé par la grande majorité des maîtres d'école. Il est plus probable que c'est de "learnability of language"

Q dont il veut parler ici, sans que toutefois on puisse assimiler ce terme à la "faculté de langage" ^{chomskyenne} chomskyenne, avec qui elle semble partager la caractéristique d'être la propriété de tous les hommes, même les "plus hébétés et stupides, sans excepter les insensés". En effet, la parenté de ces deux notions semble s'arrêter là, puisqu'il est improbable que pour Chomsky la faculté de langage ait un caractère "quasi-mécanique". En second lieu, bien que cette remarque soit anachronique d'un point de vue terminologique, les arguments de Putnam semblent plutôt établir que l'ensemble des phrases apprises par un locuteur sous des conditions dites "de performance" est récursif, question qui n'est pas liée nécessairement à la nature intrinsèque de la grammaire comme théorie de la compétence ⁽²⁾. Cette confusion provient, à notre avis, de la tentative de légitimer par des arguments psychologiques, ou comportementaux, le fait que l'ensemble des phrases d'une langue doit être récursif. Si nous excluons de cette problématique l'étude d'un lien éventuel entre "apprentissabilité" et récursivité, ce n'est pas parce que nous pensons qu'elle est inintéressante ou qu'elle pose un faux problème. Ce type de question est d'ailleurs étudié d'un point de vue formel par plusieurs auteurs qui proposent des modèles faisant intervenir des paramètres spécifiques à la situation d'apprentissage, la nature des données apprises, etc... ⁽³⁾.

Ce que nous voudrions plutôt souligner, c'est la confusion ou l'identification de cette question avec celle, clairement posée dans "Syntactic Structures" ⁽⁴⁾, de la définition de la grammaire comme mécanisme produisant des ensembles récursifs. Cette confusion est d'ailleurs maintenue, à des degrés divers, par la majorité des auteurs qui se sont exprimés sur ce point. Soit qu'ils estiment, comme Putnam, qu'une grammaire non décidable ne rend pas compte de l'intuition, partagée par la plupart des linguistes, de la récursivité

des langues naturelles. Soit qu'ils considèrent que des langages récurrents énumérables ne peuvent pas être appris, étant donnée une certaine définition de l'"^aapprentissage" (Leveit, 1974). Soit encore qu'ils déduisent de l'étude "empirique" des langues naturelles (c'est à dire de la construction de modèles grammaticaux très abstraits) le caractère récursif ou non de celles-ci: les langues sont des ensembles récurrents (ou rec-énumérables) parceque les grammaires produisent des ensembles récurrents (ou rec-énumérables). Ce qui laisse supposer que la récursivité d'un langage est un fait que devrait nous faire découvrir sa grammaire (Lapointe, 1977; Peters, 1973; Peters et Ritchie, 1973a).

Ces différents malentendus nous permettront de répéter une fois de plus que la véritable nécessité de la récursivité de l'ensemble des phrases d'une langue est d'ordre méthodologique. Une grammaire est un mécanisme qui engendre un ensemble de phrases. Si cet ensemble n'est pas récursif, il n'existera pas de procédure de décision pour y déterminer la grammaticalité d'une phrase quelconque. Et a fortiori, il n'existera pas d'algorithme pour déterminer quelle description structurale une grammaire donnée peut attribuer à une phrase donnée. Evidemment, ce qui, dans ce cas, est récursif, c'est le langage formel produit par la grammaire et la décision de déduire la récursivité de la langue décrite par cette grammaire ne vaut que si l'identification entre le langage engendré par la grammaire et la langue naturelle qu'elle ambitionne de décrire n'est pas mis en cause.

Peut-être faudrait-il mentionner une dernière interprétation équivoque de ce que nous entendons par "récursivité des langues naturelles" pour mieux circonscrire notre problème. Des logiciens et philosophes du langage ont observés que certains énoncés font apparaître



des "suites de constituants quantifiés, reliés par une conjonction, telles que, bien que dans chacune des suites il y ait une dépendance linéaire entre les quantificateurs, les deux suites elles-mêmes ne sont pas reliées par une telle dépendance"⁽⁵⁾. De telles constructions feraient donc intervenir des branchements de quantificateurs et comme telles constitueraient une classe de phrases qui ne serait même pas récursive énumérable, puisqu'on démontre que la classe de toutes les vérités logiques employant des branchements de quantificateurs est équivalente à un fragment du calcul de second ordre qui n'est pas lui-même récursivement énumérable.

Ces considérations sont intéressantes et elles illustrent bien le degré de complexité des phénomènes dont doit rendre compte une sémantique qui aurait pour objet l'analyse des conditions de vérité dans la langue. Elles demeurent cependant étrangères au problème qui nous occupe, la forme algorithmique de la grammaire ne concernant ici que les modes de formation des phrases. L'étude de Gabbay et Moravcsik ne tend d'ailleurs qu'à montrer, comme ils le signalent eux-mêmes, qu'il y a "une intéressante asymétrie entre les composants syntaxiques et sémantiques, puisque l'ensemble des phrases valides n'est pas récursif énumérable tandis que l'ensemble des phrases bien formées est récursif"⁽⁶⁾. Remarquons qu'il suffit cependant d'observer trivialement que les mathématiques toutes entières sont exprimables dans les langues naturelles pour constater cette asymétrie.

L'ensemble des suggestions faites par Putnam dans sa conjecture conduit enfin à une situation paradoxale. On ne pourrait, en les acceptant, se contenter de modifier les définitions ou l'organisation interne d'une grammaire transformationnelle, elles entraîneraient fatalement une transformation complète des objectifs du système. Dans la première version de la grammaire transformationnelle, celle à laquelle s'adressent les remarques de Putnam

la fonction récursive (au sens d'itérative) de la grammaire est assurée par le composant transformationnel. Le nombre d'indicateurs syntagmatiques produits par le composant de base est donc fini. Les transformations généralisées qui transforment plusieurs indicateurs syntagmatiques en un seul sont donc responsables de l'infinitude des langages engendrés. Si nous observons que toutes les transformations généralisées mentionnées dans Chomsky (1957) font intervenir des effacements, nous voyons immédiatement qu'une limitation de l'usage des effacements équivaut à une limitation de l'usage des transformations généralisées, donc à la production d'un ensemble fini de phrases. Indépendamment du rôle que joue le principe de l'infinitude des langues dans l'organisation de la théorie linguistique chomskyenne, cette idée est difficilement soutenable, puisqu'elle reviendrait par exemple, à assigner une limite supérieure aux nombres de coordinations possibles.

Mais en dehors de ces raisons théoriques, il y a un argument majeur d'un tout autre ordre, pour refuser les solutions de Putnam, c'est que les transformations, dans le modèle de "Syntactic Structures" relient entre elles non pas des descriptions d'une même phrase prise à des stades différents, mais des phrases différentes. Ainsi, restreindre l'usage des transformations ne revient pas à sélectionner une forme de description ou d'engendrement parmi d'autres, mais à se priver délibérément de la possibilité d'engendrer certaines phrases.

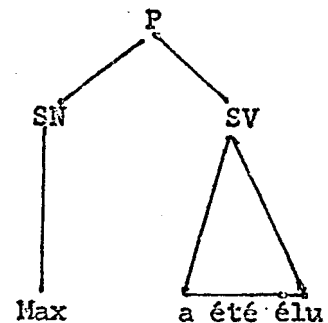
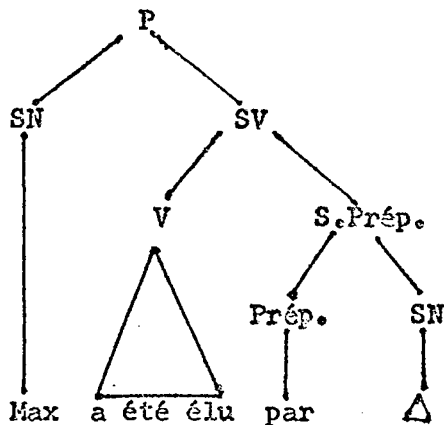
6.2. Le principe de Récupération des Effacements

Cette impossibilité de principe disparaît avec la seconde formulation de la grammaire transformationnelle où les transformations ne sont plus que des opérations permettant de relier des structures profondes interprétées sémantiquement à des structures de surface. Les transformations

doivent dans ce cas respecter le principe, dit de Katz et Postal (Katz et Postal, 1964), qui leur impose d'être conservatrice de sens. Si ce principe n'était pas respecté, l'idée d'interpréter sémantiquement les structures profondes n'aurait en effet aucun sens. Les transformations d'effacement doivent donc satisfaire le principe de Katz et Postal ou plus précisément, elles doivent satisfaire une condition spécifique qui n'est qu'un cas particulier de ce principe, la condition de "Récupération des Effacements" ("Recoverability of Deletion"). Cette condition, également formulée par Katz et Postal, répond à des motivations d'ordre syntaxique et sémantique. C'est grâce à ce principe, qu'étant donnée une transformation qui efface ou substitue des éléments d'un indicateur syntagmatique, on peut remonter de l'indicateur dérivé à l'indicateur de structure profonde de façon unique, par un recouvrement unique. "On peut avoir une idée de la force de l'hypothèse que l'on fait lorsqu'on impose que toute composante syntaxique puisse assurer la reconstruction univoque des indicateurs syntagmatiques sous-jacents, si l'on remarque que c'est ce principe qui rend récursif l'ensemble des phrases engendrées par une composante syntaxiques" (Katz et Postal, p.80).⁽⁷⁾

4.2.1. Le principe de récupération des effacements s'énonce à peu près ainsi: une transformation comprenant une transformation élémentaire d'effacement ne peut s'appliquer que si : a) les termes effacés figurent ailleurs dans l'indicateurs syntagmatique sur lequel s'applique la transformation; ou si b) les termes effacés appartiennent à une liste spécifiée d'éléments (par exemple une classe restreinte de terminaux).

Par exemple la transformation d'effacement du syntagme verbal qui permet de dériver "Max aime le cinéma et Luc aussi" à partir de "Max aime le cinéma et Luc aime le cinéma" s'applique en respectant la condition (a) tandis que dans le cas suivant Effacement de l'Agent respecte (b)



en effaçant l'élément spécifié Δ qui joue le rôle de sujet profond "impersonnel", puisque Max dans "Max a été élu" n'est que le sujet apparent de la phrase. L'utilisation de l'élément Δ est ici justifié par le fait que si on avait un agent "réel" du type "les électeurs", alors il y aurait une partie du sens qu'on ne pourrait pas récupérer, violant ainsi le principe de Katz et Postal. Comme il n'y a que le premier pas qui coûte, on pourrait supposer que l'agent est un constituant plus gros du type "les électeurs qui sont toujours aussi bêtes", "les électeurs qui sont toujours aussi bêtes et disciplinés", "que par le passé", etc... et il y aurait une infinité de constructions à partir desquelles "Max a été élu" pourrait être dérivé.

Selon Chomsky, un tel principe devrait restreindre la capacité générative des grammaires transformationnelles de façon qu'elles n'engendrent que des ensembles rékursifs, bien que toutefois il demeure plus prudent sur ce point que Katz et Postal (Chomsky 1965, note 37, ch.I). Ceci même si, comme il le remarque, le principe de Récupération des Effacements ne va pas sans poser un certain nombre de problèmes linguistiques. Par exemple s'il bloque la dérivation de "Je vois le ~~xx~~ les hommes font la cuisine ~~/homme~~", il entraîne nécessairement la réflexivisation de "Max ennue Max", interdisant

la production de cette dernière phrase sinon au prix de l'introduction d'indices référentiels et de toutes les difficultés liées à l'utilisation de ce genre de notions. D'autre part, la dérivation de "Anne est plus intelligente que Luc" est également impossible puisque, en structure profonde, "Luc est intelligent" (cf. Chomsky 1965, ch.4, §2.2). Néanmoins, si on néglige ces problèmes, Récupération des Effacements (R.E.) doit être une condition suffisante pour rendre les grammaires décidables. Il est en effet admis par les linguistes, même si à l'époque d'"Aspects" on ne dispose pas encore d'une démonstration mathématique de ce fait, que tout ensemble récursivement énumérable est ^{engendré} ~~engendré~~ par une grammaire transformationnelle privée de R.E.. D'ailleurs il suffit de disposer d'une "définition convenable de la grammaire transformationnelle" (pour reprendre le terme de Putnam) pour démontrer cette propriété. Ce résultat est établi par Ginsburg et Partee (1969) pour les grammaires transformationnelles à base indépendante du contexte et par Salomaa (1971) pour les grammaires transformationnelles à base régulière.

Remarquons que tant que R.E. reste une condition que doivent satisfaire les grammaires transformationnelles, le fait de savoir si une grammaire respecte ce principe est récursivement insoluble. Donc l'existence d'une telle condition externe ne répond pas à son objectif qui est de restreindre les grammaires de telle sorte qu'elles soient des procédures de décision. Peters et Ritchie (1973 a) ont donc incorporé R.E. dans la définition même des transformations suivant, semble-t-il, en cela l'observation de Chomsky: "... cette condition d'identité n'a jamais à être énoncée dans la grammaire, puisqu'elle est une condition générale sur le fonctionnement des grammaires. (...) En plaçant la condition d'identité en dehors des grammaires, nous sommes par conséquent en mesure de formuler strictement les analyses structurales définissant les transformations comme des conditions booléennes sur l'analysabilité, et ainsi de restreindre fortement la puissance de la

théorie de la grammaire transformationnelle." (Chomsky, 1965, n.13, ch.3).

6.2.2. Cette dernière possibilité s'avère malheureusement irréalisable puisque Peters et Ritchie (1973a) ont démontré le théorème suivant:

Tout langage récursivement énumérable est engendré par une grammaire transformationnelle à base dépendante du contexte et réciproquement.

Nous donnons une esquisse de la démonstration de ce théorème car elle montre bien, intuitivement, où se situe la faiblesse des grammaires transformationnelles.

Notons d'abord que la réciproque est immédiate puisque, par définition, les grammaires transformationnelles sont des systèmes semi-thueiens. On sait que pour tout langage récursivement énumérable L , il existe un système de réécriture illimité G , tel que $L = L(G)$. A partir de V_t et V_n , les vocabulaires terminaux et non terminaux de G , construisons V'_t et V'_n tels que $V'_t = V_t \cup b$ et $V'_n = V_n \cup B$. Soit G' possédant le même symbole initial que G et les règles suivantes:

I. Si $\varphi \rightarrow \psi$ est une règle de G , alors:

- a) $\varphi \rightarrow \psi$ est une règle de G' , si $l(\psi) \geq l(\varphi)$ (où l est la fonction "longueur des séquences".)
- b) $\varphi \rightarrow \psi B^n$ est une règle de G' avec $n = l(\varphi) - l(\psi)$ si $l(\varphi) > l(\psi)$.

II. c) $BA \rightarrow AB$ est une règle de G' pour tout $A \in V_n$.

d) $B \rightarrow b$

On constate que pour toute séquence terminale x dérivée dans G , on peut dériver xb^m dans G' . Comme les règles de G' sont non raccourcissantes, c'est une grammaire dépendante du contexte. En choisissant un composant

transformationnel T réduit à l'unique transformation $t: X - b \Rightarrow X - \emptyset$, on pourra par application cyclique de t effacer tous les b en bout de phrase (nous omettons de préciser que G' est équivalente à une grammaire G'' qui introduit récursivement autant de S initiaux que de b terminaux, permettant ainsi au cycle transformationnel de s'appliquer.) Donc $\langle G \rangle = (G', T)$ est la grammaire recherchée.

Remarquons que cette grammaire ne viole pas R.E. puisque b peut appartenir à la liste des éléments spécifiés mentionnés dans la condition (b) de ce principe.

La question qui vient naturellement à l'esprit face à ce résultat peu encourageant est de se demander si celui-ci ne découle pas directement des propriétés combinées des grammaires dépendantes du contexte et des transformations d'effacement, étant donné la "puissance" des règles dépendantes du contexte. En d'autres termes, si on choisit pour composant de base un mécanisme moins commode, mais moins puissant qu'une grammaire dépendante du contexte, cette différence de capacité générative aura-t-elle une incidence sur la classe des langages transformationnels? Il n'en est rien puisque Peters et Ritchie (1971) ont démontré un théorème analogue au précédent où "indépendante du contexte" remplace "dépendante du contexte". Ils ont même établi le résultat suivant, encore plus étonnant (Peters et Ritchie, 1971, th1) que nous pourrions appeler le "théorème de la base minimum".

Pour l'alphabet $a_1, \dots, a_n$ les cinq conditions suivantes sont équivalentes:

- 1) L est un langage récursif énumérable
- 2) L est engendré par une grammaire transformationnelle dont la base est une grammaire indépendante du contexte.

- 3) L est engendré par une grammaire transformationnelle dont la base est une grammaire linéaire
- 4) L est engendré par une grammaire transformationnelle dont la base est une grammaire linéaire unilatère
- 5) L est engendré par une grammaire transformationnelle dont la base est constituée des seules règles: $S \rightarrow S\#$, et $S \rightarrow a_1, \dots, a_n b\#$, où b et # sont des terminaux n'appartenant pas à $a_1, \dots, a_n$, # étant le symbole de frontière.

Bien que ce théorème aille à l'encontre de ce que pouvaient espérer les linguistes, son contenu demeure très intéressant. Il confirme, comme l'affirmait Chomsky (1970), que le choix présenté aux linguistes entre l'enrichissement du composant de base et l'enrichissement du composant transformationnel est sans effet sur la capacité générative de la grammaire. Question qui opposait, rappelons-le, les partisans de la Sémantique Générative et ceux de la théorie classique d'"Aspects". Peters et Ritchie font d'ailleurs observer que même lorsque le composant de base est fixé, les variations qu'il faut imposer au composant transformationnel sont véritablement minimales, une seule transformation semblant dépendre réellement de la nature du langage particulier engendré.

Le théorème rend également caduc le problème de la "base universelle" qui consiste à rechercher une classe restreinte d'indicateurs syntagmatiques de structures profondes qui serait valable pour toutes les langues naturelles. En effet, trivialement, une seule règle produisant la liste des éléments du vocabulaire terminal est suffisante pour engendrer toute phrase formée sur ce vocabulaire. Evidemment, la "base universelle" est construite de façon à satisfaire des conditions empiriques de tout ordre: psychologiques, observation des structures de surface, économie de la grammaire, etc... et c'est dénaturer le problème que de le poser aussi crûment. Il faut simplement garder

présent à l'esprit que le fait qu'une grammaire transformationnelle soit capable d'engendrer une langue naturelle ne peut être en aucun cas un argument pour justifier le choix d'un composant de base donné. En particulier, cette capacité n'est preuve ni de la pertinence psychologique des structures profondes, ni de la valeur universelle des catégories utilisées dans ce composant de base.

Ces ultimes considérations démontrent tout l'intérêt d'une étude mathématique des systèmes grammaticaux. Le théorème précédemment mentionné était établi avec l'intention précise d'analyser la capacité générative des grammaires transformationnelles, or nous avons vu qu'il se prête de façon inattendue à des interprétations totalement étrangères à cet objectif. Aussi deux points méritent d'être observés devant cette situation. D'abord se trouve partiellement confirmée l'interprétation que nous proposons plus haut de la grammaire formelle comme instrument de contrôle des concepts formulés dans la théorie linguistique puisque le théorème de Peters et Ritchie permet de se prononcer sur le choix de tels concepts. Ensuite se trouve démentie l'opinion selon laquelle la grammaire transformationnelle serait une théorie trop compliquée et trop dépendante des aléas de l'étude empirique pour être susceptible d'une étude mathématique allant au-delà de la formalisation de propriétés triviales comme le pense, par exemple, Schützenberger (1978).

6.2.3. Le principe de récupération des effacements ne garantissant pas la décidabilité d'un langage transformationnel, quelle condition ajouter à la grammaire pour arriver à cette fin? L'analyse de la démonstration du théorème sur les grammaires à base dépendante du contexte peut laisser penser que R.E. n'est pas assez fort en ce sens que les effacements autorisés par la condition (b) ne sont pas vraiment récupérables. La possibilité d'effacer cycli-

quement les séquences b^n , quelque soit n , fait perdre toute trace de l'histoire dérivationnelle de la séquence terminale. C'est le défaut formel des grammaires transformationnelles. Voyons comment il se traduit, sur le plan computationnel, au niveau de la procédure elle-même.

Pour déterminer si une séquence quelconque x appartient au langage engendré par $G = (B, T)$, il faut:

- (1) Enumérer successivement les arbres bien formés sur le vocabulaire de B .
- (2) Etant donné un arbre bien formé, décider s'il est le produit par T (ce qui est toujours possible puisque (B) est récursif, sinon retourner en (1), si oui aller en (3)).
- (3) Appliquer les transformations de T à l'arbre et déterminer si oui ou non la dernière ligne de la dérivation obtenue est une structure de surface bien formée ayant x pour séquence terminale.

On constate alors que si x n'appartient pas à $L(B, T)$, la procédure ne s'arrêtera jamais, cherchant toujours parmi des structures profondes plus longues un arbre correspondant à x . Ce qu'il faudrait, pour rendre la procédure effective, c'est qu'à chaque séquence x soit associée une limite supérieure à la longueur de la structure profonde pouvant lui être attribuée avec cette limite égale à zéro lorsque x n'appartient pas à $L(B, T)$. Dans la formulation de Peters et Ritchie (1973a), la "longueur" d'une structure profonde peut s'interpréter en termes du nombre de cycles transformationnels pouvant lui être appliqué, c'est à dire du nombre de sous-phrases enchassées. Donc l'existence d'une méthode effective permettant d'obtenir la limite supérieure du nombre de cycles pour toute séquence x est une condition nécessaire pour que $L(B, T)$ soit décidable. C'est égale-

ment une condition suffisante car si $L(B,T)$ est décidable étant donné un x quelconque, l'application de la procédure attribuera une structure profonde à x , donc une limite supérieure au nombre de cycles.

Peters et Ritchie définissent alors la fonction cyclique f_g d'une grammaire transformationnelle G comme:

$$f_g(x) = \begin{cases} 0 & \text{si } x \notin L(G) \\ \text{le plus petit } s \text{ tel que } x \text{ ait une structure profonde avec} \\ & s \text{ sous-phrases.} \end{cases}$$

Moyennant cette définition ils établissent le théorème suivant (Corollaire 6.6. Peters et Ritchie, 1973a):

Les trois conditions suivantes sont équivalentes pour toute grammaire transformationnelle:

- (a) $L(G)$ est décidable
- (b) La fonction cyclique f_g est récursive
- (c) f_g est majorée par une fonction récursive

Ce théorème est obtenu au prix de l'établissement d'une série de lemmes démontrant, de façon constructiviste, l'existence de machines de Turing qui réalisent les étapes de la procédure décrite à l'instant. Ces lemmes, combinés à la propriété fondamentale des règles récursives de la grammaire qui produisent des terminaux $\#$ à chacune de leur application, conduisent à un théorème affirmant l'existence d'une machine de Turing pour toute grammaire transformationnelle qui accepte un mot $x\#^S$ si, sous-jacent à x il existe un arbre contenant au plus s sous phrases.

Ces résultats, bien que ne répondant pas de façon totalement satisfaisante au problème général posé par Putnam, montrent malgré tout que l'utilisation de transformations d'effacement n'entraîne pas nécessai-

rement l'indécidabilité du langage produit. Ce qui n'est guère étonnant, la majorité des grammaires transformationnelles utilisées par les linguistes étant, selon toute vraisemblance, des procédures de décision. En effet, le théorème de la "base minimum" mentionné plus haut n'est qu'un théorème de la fausse base minimum dans la mesure où cette base peut produire pour des séquences dont la structure de surface est très simple, par exemple la séquence " $a_1 a_2$ ", des structures profondes d'une complexité démesurée comme la suivante:

$$[S [S \dots [S a_1 a_2 \neq] \dots \neq] \neq]$$

De plus, même la simplicité de la règle $S \longrightarrow a_1 \dots a_n b \neq$ est illusoire dans la mesure où elle laisse penser qu'une grammaire ne contenant qu'elle (plus $S \longrightarrow S \neq$) est moins complexe qu'une grammaire contenant une liste de n règles $S \longrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.

La dérivation de " a_1 " implique avec la première base $n-1$ transformations d'effacement tandis qu'elle n'en nécessite qu'une avec la seconde. Evidemment, mathématiquement, la distinction entre ces deux systèmes de règles reste sans effet (voir la note 8). Mais cette identification, justifiée si l'on considère simplement l'équivalence faible, masque de profondes différences, qui même si elles ne sont pas susceptibles de recevoir une expression mathématique immédiate sont partiellement formalisables. La notion d'efficacité d'un algorithme, telle qu'elle est étudiée dans les recherches contemporaines sur la théorie de la calculabilité pourrait fournir une thématisation appro-

priée à ces propriétés potentielles. Peut-être que par là même il serait possible de donner un contenu formel à la notion de mesure de simplicité d'une théorie. La mesure de simplicité, dans un usage intuitif, étant bien un critère qui guiderait le choix entre les deux grammaires que nous comparons.

La complexité injustifiée des structures profondes introduites par le modèle mathématique repose également sur deux autres facteurs : l'interprétation réductionniste de la fonction du symbole de frontière de phrase $\neq$, et un usage artificiel, démesuré de ce symbole comme "compteur" des cycles transformationnels.

La fonction linguistique du $\neq$ est triple: 1°) Délimiter le domaine d'application des transformations, 2°) Déterminer un ordre partiel d'application des transformations 3°) Faire jouer au composant transformationnel un rôle de filtre. Le modèle mathématique ne traduit fidèlement aucune de ces fonctions. On peut même estimer qu'il les interprète de façon incorrecte. Dans l'usage linguistique, les $\neq$ imposent une limite maximum au nombre de transformations applicables pour déterminer la grammaticalité d'une séquence terminale. C'est donc également un moyen de sélectionner les structures profondes. Au contraire dans le modèle de Peters et Ritchie, les $\neq$ sont introduits aux seules fins du calcul d'une séquence terminale quelconque. Cette inversion des fonctions du symbole de frontière est clairement visible dans la démonstration du théorème de la "base minimum" donnée par Peters et Ritchie, elle l'est également dans

l'esquisse de démonstration du théorème donné page 147. Remarquons par ailleurs que les auteurs définissent une "convention de réduction" (analogue à la règle de "Tree Pruning" de Ross) pour les noeuds non récursifs. Celle-ci permet d'éliminer les branches de la forme $[_A [_B [_A \omega]_A]_B]_A$ où la double analyse de ω comme un constituant A est redondante. Cette convention linguistique est naturelle, aussi devrait-elle être généralisée aux symboles récursifs, interdisant à une suite de S de se dominer exhaustivement les uns les autres comme dans la structure $[_S [_S \dots [_S \omega \#_S] \dots \#_S] \#_S]$, de tels cas ne se rencontrant jamais dans les analyses linguistiques⁽⁹⁾. En ce sens une interprétation négative du théorème de la "base minimum" n'est peut-être pas légitime. Nous aurons d'ailleurs l'occasion de constater à quel point les symboles de frontière jouent un rôle déterminant lorsque nous examinerons, au sixième chapitre, un autre modèle mathématique des grammaires transformationnelles étudiées par Peters et Ritchie.

6.3. Propriétés globales et propriétés locales

De ces remarques il ressort que le problème était mal posé puisqu'à la question "totale" : les grammaires transformationnelles satisfaisant le principe de récupération des effacements engendrent-elles des ensembles récursifs? La réponse n'est pas "oui" ou "non" mais "parfois oui, parfois non". C'est uniquement dans la mesure où il permet de rejeter la question totale que le théorème sur la fonction cyclique devient intéressant, puisqu'il réoriente le problème vers une analyse loca-

le des propriétés mathématiques des grammaires.

Ce qu'il faut maintenant rechercher c'est une détermination réaliste de la fonction cyclique d'une grammaire. Cette détermination passe nécessairement par un retour à l'étude du fonctionnement des systèmes linguistiques. L'objectivation formelle ne peut s'auto-alimenter avec les informations qu'elle produit elle-même, le retour à l'objet empirique, ici le mécanisme grammatical, est nécessaire lorsque le modèle mathématique a épuisé par une gestion optimale (et malheureuse, nous l'avons vu) les données initialement fournies. L'un des aspects les plus originaux des travaux de Peters et Ritchie est sans doute l'étude du degré de complexité d'une grammaire caractérisé par la "place" de sa fonction cyclique dans les différentes hiérarchies de fonctions récursives. Cette étude suggère naturellement une stratégie d'analyse locale des grammaires. Etant donnée une grammaire particulière on peut tenter d'estimer, à partir de l'analyse de son fonctionnement, la localisation de sa fonction cyclique parmi les fonctions récursives.

6.3.I. Pour Peters la question peut être formulée ainsi : quelles propriétés doit avoir G si $f_g(x)$ croît rapidement avec x ? Examinons un moment le sens de cette question. Si nous disposions d'un principe R.E'. qui, contrairement à R.E., laisse toujours une trace, dans la structure de surface, d'un noeud récursif présent en structure profonde, alors le problème serait réglé. En effet, d'une part toute sous-phrasé (c'est à dire toute portion d'arbre dominée par un S et ne contenant

aucun S) est nécessairement de longueur inférieure à une constante k qui ne dépend que du composant de base. D'autre part $R.E'$ garantit que le nombre de S dans la structure profonde d'une séquence terminale quelconque x , est inférieur ou égal à la longueur de x , $l(x)$. Donc la longueur de toute structure profonde serait majorée par la fonction linéaire $kl(x)$ et la fonction cyclique majorée par $l(x)$. Ce qui remplit les conditions du théorème puisque les fonctions linéaires sont récursives.

Malheureusement les effacements requis par les descriptions grammaticales ne permettent pas que l'on adopte un principe du genre de $R.E'$. Considérons les propriétés syntaxiques des verbes anglais "promise", "threaten", "purport", etc... telles que les analyse Peters (1973). Soient les phrases suivantes:

- 1) Their sitting down promises to steady the canoe
- 2) Their sitting down's promising to steady the canoe threatens to spoil the joke
- 3) Their sitting down's promising to steady the canoe's threatening to spoil the joke purports to follow the script.

Dans (1) le sujet et l'objet de "promises" sont tous deux des phrases compléments, de plus ces phrases sont également dominées par un noeud S . Il y a donc plus de deux S dans (1). A son tour (1) peut être elle-même complément du sujet dans (2), et ainsi de suite. Les phrases $x_1, x_2, \dots, x_i, \dots$ ainsi construites possèdent au moins $2^i S$ dans leurs structures profondes.

Si $s_i = 2^i$ alors on a $f_g(x_i) > s_i$ quelque soit i . D'autre part la longueur de chacune de ces phrases (exprimée en mots) est $l(x_1) = 8$, $l(x_2) = 13$, $l(x_3) = 18$, $l(x_i) = l(x_{i-1}) + 5$ et $l(x_i) = 5i + 3$. On a donc $s_i = 2^i \gg 2^{\frac{5i+3}{8}} = 2^{1i/8}$ et $f_g(x_i) > K^1 \propto K^{\frac{8}{\sqrt{2}}}$. Donc la fonction cyclique n'est pas une fonction linéaire de la longueur des séquences terminales, elle est au mieux (si on parvient à la majorer convenablement) une fonction exponentielle de cette longueur. Bien qu'on ne puisse pas déduire de cette démonstration que l'anglais n'est pas un langage dépendant du contexte, elle établit qu'un automate linéairement borné ne peut pas accéder aux structures sous-jacentes des phrases anglaises.

Une grammaire transformationnelle (de l'anglais) bien construite serait donc représentable par une machine de Turing d'une classe assez bien délimitée puisqu'elle devrait appartenir au complément de la classe des automates linéairement bornés dans la sous-classe des machines de Turing effectives.

Les conséquences de ce résultat sont importantes pour la théorie transformationnelle. On peut rapprocher la démonstration de Peters de celles qui établissent l'inadéquation des grammaires régulières et des grammaires indépendantes du contexte. Elle complète ainsi le projet critique de la grammaire transformationnelle à l'égard du structuralisme, en rejetant le plus puissant de ses modèles. Deux aspects distinguent cependant la démonstration de Peters de celles de Chomsky, Postal, etc... Premièrement, elle est plus satisfaisante car elle

s'appuie sur la critique d'un type de concept, la description structurale, qui est fondamental pour les générativistes. Les analyses de la capacité générative forte étant, Chomsky l'a répété maintes fois, les seules vraiment intéressantes. Deuxièmement elle est plus fragile. Les démonstrations de la linguistique mathématique s'articulent toujours sur des concepts linguistiques, leur validité dépend de la pertinence de ces concepts. Dans le cas de la démonstration de l'inadéquation des langages à états finis il faut admettre un certain nombre de définitions générales comme celles de "grammaire", "langage", ... Ici, l'argumentation de Peters ne suppose pas simplement une adhésion à ces définitions générales, puisqu'elle implique aussi l'accord avec des analyses linguistiques propres à une grammaire particulière. A chacune de ces médiations est associée une certaine incertitude d'adéquation conceptuelle, les incertitudes se multipliant comme les médiations se superposent .

6.3.2. C'est toujours au moyen de l'étude locale des propriétés des grammaires que Peters définit une condition pour que leur fonction cyclique soit majorée par une fonction récursive. La "Propriété du Survivant" ("Survivor Property") conçue pour éviter les effacements arbitraires discutés dans les pages précédentes pourrait s'énoncer ainsi : si un arbre est l'entrée d'un cycle transformationnel quelconque, alors l'arbre de sortie de ce cycle contient plus de noeuds terminaux que n'importe lequel de ses sous-arbres déjà transformé.

En définissant la profondeur maximale s , d'un arbre comme

le plus grand nombre de S rencontrés sur tout chemin simple allant du sommet de l'arbre aux terminaux, la propriété du survivant garantit que le nombre m de terminaux est toujours plus grand ou égal à la profondeur maximale de la structure profonde. Si $t_1, \dots, t_n$ est une dérivation transformationnelle et que la sortie de cette dérivation t_n , est telle que $l(t_n) = m$, alors, en remarquant que chaque cycle transformationnel ne peut raccourcir, au plus, son entrée que d'un facteur $k^{(I0)}$, nous avons $l(t_1) \leq mk^{s(t_1)} \leq mk^m$. Ce qui, combiné avec l'observation que le nombre maximum de noeuds non terminaux d'un arbre dominant une séquence terminale de longueur m est au plus égal à $q(2m-1)^{(II)}$ (où q est le nombre d'éléments du vocabulaire non terminal), donne $s(t_1) \leq 2mk^{m-1}$. On déduit donc qu'une séquence terminale x de longueur m produite par une grammaire ayant la propriété du survivant a une structure profonde contenant moins de $(2k^2)^m S$ et que pour tout x , $f_g(x) \leq (2k^2)^{m(x)}$, ce qui est la majoration recherchée.

La propriété du survivant est donc une condition suffisante pour garantir la récursivité des langages transformationnels. Cependant, le fait "qu'une revue des grammaires transformationnels que les linguistes ont construit pour différents langages naturels indique qu'elles possèdent toutes la propriété du survivant" ^(I2) autorise-t-il à interpréter cette propriété comme une condition plus forte de recouvrement des effacements ("stronger condition of recoverability of deletion")? Evidemment non, puisque pour une grammaire, posséder

la propriété du survivant n'est qu'une propriété accidentelle⁽¹³⁾. Pour qu'elle devienne une propriété essentielle, il faut lui donner le statut de condition sur les grammaires dans la théorie linguistique. Mais procéder ainsi ne résoud pas le problème car la question de savoir si une grammaire satisfait Condition du Survivant est récursivement insoluble puisqu'elle suppose l'inspection des dérivations de chaque séquence du langage engendré.

La solution idéale consisterait à trouver une "astuce" pour transformer cette condition en propriété inhérente d'une manière analogue à celle employée par Peters et Ritchie pour R.E.. Ici la difficulté tient au fait que les contraintes recherchées portent sur les dérivations - et non plus sur les seules transformations -, donc sur l'application de cycles transformationnels consécutifs, et qu'ils paraît impossible de trouver une contrainte plus forte que la propriété exprimant que chaque cycle raccourcit son entrée d'un facteur déterminé. De plus, toute propriété dérivationnelle exprimable dans le modèle mathématique de Peters et Ritchie ne peut être que globale, c'est à dire concerner des rapports entre indicateurs syntagmatiques pris dans leur totalité. Or les restrictions souhaitées concernent, d'une certaine façon, les "contenus" des indicateurs syntagmatiques d'une dérivation, puisqu'elles expriment une dépendance entre la branche de profondeur maximale des arbres et les arbres eux-mêmes. C'est donc vers une délimitation des configurations possibles d'indicateurs syntag-

matiques qu'il faut également s'orienter. Comme ces configurations dépendent, dans une certaine mesure, des règles de base, la question pourrait naturellement se poser de savoir s'il n'y a pas des conditions initiales de configurations dont la base seule serait responsable. Cependant ceci paraît exclu a priori étant donné le théorème de la base minimum et les propriétés générales de la hiérarchie de Chomsky. En effet, puisque l'on recherche ici des conditions contraignant la capacité générative faible des grammaires transformationnelles pourquoi espérer que les modifications du composant de base puissent affecter celle-ci, la démonstration de Peters et Ritchie valant pour la base la plus restreinte qui soit.

De plus on peut supposer, par expérience, que toute variation envisageable, même importante sur la forme des règles, pourvu qu'elle entre sous la forme canonique d'une classe définie de grammaire de constituants, serait sans effet sur la capacité générative. L'étude purement mathématique des grammaires de constituants fournit de nombreux exemples d'équivalence de systèmes apparemment fort différents. Qu'on songe, pour ne citer que les exemples les plus connus, à l'équivalence des grammaires catégorielles et des grammaires indépendantes du contexte ou à l'équivalence des grammaires catégorielles unidirectionnelles et bi-directionnelles. Il est vrai qu'il existe certaines sous-classes de langages indépendants du contexte incomparables entre elles⁽¹⁴⁾ - c'est à dire pour lesquelles la relation d'inclusion ne définit pas un ordre total -, mais

ces cas marginaux, du fait de leur inclusion stricte dans la classe des langages indépendants du contexte, ne peuvent entrer en ligne de compte pour le problème nous occupant ici, puisqu'il semble indubitable que les composants de base utilisés par les linguistes dans les grammaires transformationnelles aient une capacité générative au moins égale à celle des grammaires indépendantes du contexte non restreintes.

Mais on doit observer que toutes ces considérations sur la typologie des grammaires formelles sont articulées uniquement sur une propriété particulière qui, nous l'avons signalé à plusieurs reprises et nous y reviendrons dans des chapitres ultérieurs, n'est pas essentielle. La raison majeure pour laquelle l'étude de la capacité générative faible occupe une place prépondérante dans la théorie des grammaires formelles tient vraisemblablement beaucoup moins à l'intérêt qu'elle présente par elle-même qu'au fait qu'elle soit la plus facilement accessible à une analyse mathématique. Sans doute est-il paradoxal d'émettre des réserves à l'égard du rôle de la capacité générative faible précisément dans un cas où elle est concernée au premier lieu. La raison pour laquelle il nous semble légitime de formuler cette critique, même dans le cas présent, est que la sujétion à l'étude de la capacité générative faible est telle qu'elle en vient à faire perdre de vue ce qu'on pourrait appeler un bon sens "gestaltiste" selon lequel le tout n'est pas nécessairement égal à la somme des parties. En clair pourquoi supposer que la capacité générative d'une grammaire générative soit égale à celle de son composant de base plus

celle de son composant transformationnel. Il semble en effet lorsqu'on analyse les travaux de Peters et Ritchie que ne soit pas envisagée la possibilité d'une dialectique des contraintes entre composant de base et composant transformationnel. Ceux-ci étudient l'effet des variations de la base tout en "fixant" le composant transformationnel, ou bien l'inverse, rejetant ainsi l'éventualité que la solution du problème puisse résider ailleurs que dans l'un de ces deux cas de figure. Pourtant le fait que les transformations s'appliquent à des indicateurs syntagmatiques, et non pas à de simples séquences de symboles laisse imaginer que la forme de ces objets puisse être une caractéristique aussi importante que leur mode d'énumération.

La mathématisation semble victime d'une détermination habituellement typique de l'étude empirique. Elle ne parvient pas à surpasser une catégorisation limitative de son objet, certes légitimé par un usage solidement éprouvé, mais conçue pour des motifs finalement indépendants du problème ici posé. Aussi ne s'étonnera-t-on pas de savoir que c'est en retournant à l'examen détaillé des grammaires utilisées par les linguistes et en se dégageant de la problématique adoptée par Peters et Ritchie, que Lapointe (1977) a pu proposer une réponse positive à la conjecture de Putnam.

6.3.3. Les conditions découvertes par Lapointe sont à la fois extrêmement simples et empiriquement tout à fait plausibles. Elles reviennent à distinguer dans le vocabulaire auxiliaire de la grammaire un type particulier de symboles dont on pourrait dire qu'ils sont purement lexicaux, en ce sens qu'ils apparais-

sent uniquement en position préterminale et qu'ils partitionnent le vocabulaire terminal en classes disjointes. Evidemment, la "plausibilité empirique" de cette dernière condition peut laisser perplexe puisqu'elle suppose résolu le problème de la classification du dictionnaire en parties du discours, mais n'oublions pas que cet a priori n'implique n'aucun engagement allant au-delà des principes épistémologiques fondamentaux de la grammaire générative⁽¹⁵⁾.

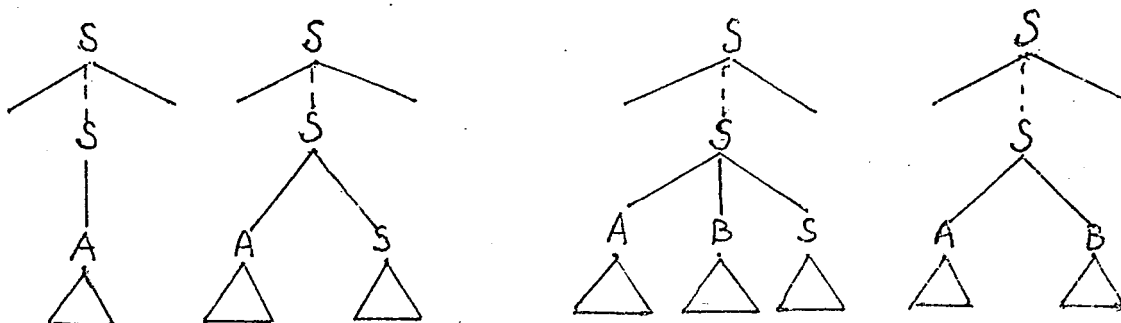
Les conditions sur le composant de base sont les suivantes:

- 1°) Toute sous-phrasé minimale doit contenir au moins une catégorie lexicale majeure (N, V, Adj, ...)
- 2°) Tout composant de base, $\mathcal{B}$, doit être de la forme normale suivante:
 - a) Quel que soit $a \in V_t$, il existe un A unique appartenant à V_n tel que $\mathcal{B}$ contient $A \rightarrow a$ et
 - b) Quel que soit $A \in V_n$, si $\mathcal{B}$ contient la règle $A \rightarrow a$, alors $\mathcal{B}$ ne contient aucune règle de la forme $A \rightarrow \sigma D \omega$ où $D \in V_n$ et $\sigma, \omega \in (V_n \cup V_t)^*$

La première de ces contraintes assure que les noeuds S ne seront pas produits en nombre arbitraire comme l'autoriserait la stricte interprétation mathématique de la règle $S \rightarrow S$. Cette possibilité était largement exploitée, nous l'avons vu, dans les démonstrations des théorèmes de Peters et Ritchie. La seconde contrainte garantit que les éléments terminaux d'un indicateur syntagmatique ne seront jamais dominés par des noeuds récursifs. Donc lorsque ces terminaux

sont effacés (entraînant ainsi l'effacement de leur mère) aucun noeud récursif ne disparaît avec eux.

Cependant l'effet essentiel de ces restrictions n'apparaît réellement que lorsqu'elles sont combinées à la troisième contrainte proposée par Lapointe, portant sur les transformations et qui est un principe de récupération des effacements renforcé. Celui-ci énonce qu'un terme ne peut être effacé que si la description structurale de la transformation d'effacement comporte un terme analysant une soeur non récursive du terme effacé. Les figures suivantes montrent les effets de cette condition:



Effacement de A impossible Effacement de A ou B possible

La conjonction de ces trois restrictions entraîne la propriété du survivant. En effet, observons d'abord, que pour toute dérivation $t_1, \dots, t_n$, chaque indicateur syntagmatique t_i , avec $1 \leq i \leq n$, a au moins une catégorie lexicale sinon nous aurions un t_i identique à celui de la figure 1 qui ne pourrait provenir que d'un t_n par application de l'effacement de a. Ce qui est impossible puisque l'unique soeur de A est récursive (Fig.2).

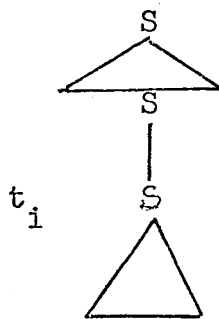


Fig. I

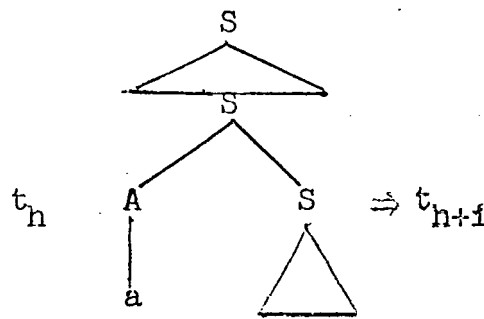


Fig. 2

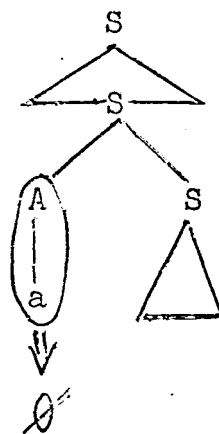
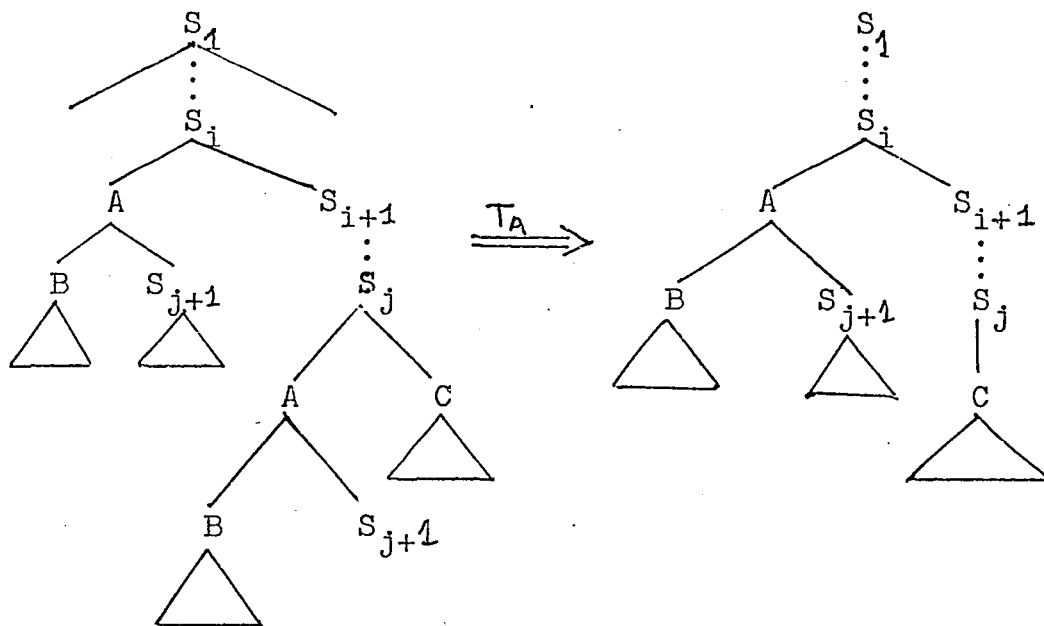


Fig. 3

Ensuite, nous pouvons remarquer que si s est la profondeur maximale de t_1 et $\mathcal{J}$ est la fonction qui associe à t_1 le nombre total de noeuds récurrents qu'il contient alors la relation $s \leq \mathcal{J}(t_1)$ est toujours vraie. C'est par définition évident pour $\mathcal{J}(t_1)$. Supposons vérifié que $s \leq \mathcal{J}(t_1)$. Pour obtenir $s > \mathcal{J}(t_{i+1})$ des noeuds S doivent nécessairement avoir été effacés sur la branche de profondeur maximale. Mais si un S est effacé il doit laisser ailleurs dans l'arbre une copie de lui-même sous peine de violer R.E. Donc $s \leq \mathcal{J}(t_{i+1})$. La différence entre une grammaire transformationnelle respectant R.E. renforcé est illustrée par la figure suivante:

T_A : Descr. Struct. : $A - X - [S A - C]$
 Chang. Struct. : $I - 2 - \emptyset - 4$



Si on n'avait pas disposé de R.E. renforcé qui exige que A ait une soeur non cyclique, ici C, figurant dans la description structurale de l'effacement, alors S_j aurait été effacé en même temps que A et la relation $s \leq \mathcal{J}(t_i)$ n'aurait plus été vérifiée. Donc, toute grammaire respectant ces trois contraintes produira toujours des indicateurs syntagmatiques terminaux possédant un nombre de noeuds terminaux supérieur ou égal à la profondeur maximale de leur structure profonde.

La question de savoir si une grammaire respecte ces contraintes est naturellement décidable puisqu'elle repose sur une inspection des règles de base et des règles transformationnelles qui sont en nombre fini. Lapointe fait d'ailleurs un inventaire des règles d'effacement usuellement utilisées dans le modèle d'"Aspects" et montre que celles-ci respectent les conditions requises.

6.3.4. Il nous semble que deux points essentiels peuvent être relevés dans cette analyse du problème de Putnam. Le premier est l'étroite co-détermination des concepts linguistiques et des concepts mathématiques. Les premières formalisations de la grammaire transformationnelle laissaient supposer que tout ensemble récursif énumérable pouvait être engendré par une grammaire transformationnelle car elles ne tenaient compte que des propriétés les plus générales des grammaires transformationnelles. C'est à dire celles qui dans la typologie des grammaires formelles permettait de les distinguer des autres types de grammaires génératives. Or, ces propriétés, pour être spécifiques, n'épuisent pas toutes les caractéristiques des grammaires transformationnelles telles qu'elles apparaissent dans l'usage proprement linguistique. C'est la sollicitation, d'abord brutale, de la mathématique ("Comment restreindre les effacements?") qui impose une série supplémentaire de structurations progressive de l'objet grammaire ("Comment rendre les fonctions cycliques récursives?", "Comment respecter la propriété du survivant?") faisant apparaître des propriétés mathématiques complexes là où il n'y avait que des commodités de routine dictées par les nécessités empiriques. Le second point est la correspondance entre les propriétés mathématiques des grammaires transformationnelles et d'autres domaines plus classiques de la mathématique. On pourrait croire que la très forte détermination de la théorie mathématique des grammaires transformationnelles par les exigences particulières des descriptions linguistiques conduise nécessairement à une théorie close sur elle-même, sans

lien avec les autres disciplines mathématiques, n'ayant à la limite que des rapports très généraux avec la théorie des grammaires formelles. Ce n'est pas le cas. L'idée fondamentale de Peters et Ritchie de caractériser les grammaires transformationnelles par leur fonction cyclique permet au contraire de rapprocher la linguistique mathématique de domaines d'études plus généraux comme celui des fonctions récursives. La caractérisation d'une fonction récursive par sa complexité de calcul, que ce soit dans les termes de la théorie de la calculabilité ou dans les termes de la théorie des automates est un des thèmes essentiels de la recherche en théorie de la récursivité depuis les travaux de Grzegorzcyk (1953) et de Ritchie (1963). L'approche de Peters et Ritchie ne fait qu'introduire comme cas particulier de ce domaine de la mathématique l'étude des grammaires transformationnelles.

NOTES DU CHAPITRE VI

(1) Rappelons que nous avons discuté du statut de cette hypothèse au cours du troisième chapitre.

(2) La question de savoir quel type d'automate permet de simuler le comportement linguistique est différente selon qu'on s'intéresse à la performance ou à la compétence. Rappelons que pour le premier de ces cas Miller et Chomsky (1963) estiment que les mécanismes mis en jeu ne devrait pas requérir une capacité généralisatrice faible excédent celle d'un automate à états finis. Plus récemment Chomsky (1979a) exprimait son scepticisme quant à l'analogie entre le comportement linguistique et une procédure effective, allant jusqu'à mettre en doute le fait qu'un sujet puisse se prononcer effectivement sur la grammaticalité des énoncés qu'il perçoit. Position sans doute contradictoire avec celle de Miller et Chomsky (1963) mais certainement pas avec l'idée que l'automate simulant la compétence engendre des ensembles récurrents.

(3) L'étude de Gold (1967) est à l'origine de ce type de travaux, repris et développés par d'autres auteurs notamment Wexler et Hamburger (1973) et Hamburger et Wexler (1973).

(4) "L'objectif fondamental de l'analyse linguistique d'un langage L est de séparer les séquences grammaticales qui sont des phrases de L, des séquences agrammaticales qui ne sont pas des phrases de L (...). La grammaire de L sera ainsi un mécanisme engendrant toutes les séquences grammaticales de L et aucune des séquences agrammaticales." (Chomsky, 1957; § 2.1)

(5) Gabbay et Moravcsik (1974); voir également Hintikka (1974) pour une argumentation similaire.

(6) Il n'est pas certain que l'argumentation développée par Gabbay et Moravcsik soit correcte. Voir à ce sujet Guenther et Hoepelman (1975).

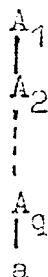
(7) Ce sont toutefois des considérations purement grammaticales qui sont à l'origine de cette notion. Chomsky (1964, p.994) l'introduit pour dériver les propositions relatives. Pour cette raison on pourrait dire que l'adoption du principe de Katz et Postal par Chomsky (1965), est plutôt une conséquence de l'adoption du principe de récupération des effacements que l'inverse.

(8) Bien que ce composant de base ne soit pas régulier il est trivialement équivalent à une grammaire régulière. L'équivalence des conditions (1) et (5) n'est donc qu'un nouvel énoncé (avec une nouvelle démonstration) du théorème déjà démontré par Kimball (1967).

(9) Faire engendrer de telles structures par la base n'est qu'un procédé ad hoc destiné à empêcher la règle de Tree Pruning de s'appliquer sur des suites de S se dominant exhaustivement les uns les autres.

(10) K est complètement déterminé par les différents paramètres suivants: nombre de symboles auxiliaires, nombre de transformations, nombre maximum de termes figurant dans les transformations et longueur maximum des séquences terminales figurant dans les conditions structurales des transformations.

(11) Comme on le montre facilement par induction. Si $m = 1$, le plus grand arbre pouvant dominer la séquence terminale "a" a la structure:



Supposons que la relation soit vérifiée pour tout $1 \leq m$. L'arbre dominant la séquence terminale de longueur m peut être décomposé en deux arbres disjoints dominant chacun une séquence terminale de longueur inférieure à m , soit l' et l'' les longueurs de ces séquences, on a $l' + l'' = m$. Donc $q(2l' - 1) + q(2l'' - 1) = q(2m - 2) < q(2m - 1)$ est le nombre maximum de noeuds contenus dans l'arbre dominant la séquence terminale de longueur m .

- (I2) Peters (1973, p.382)
- (I3) Peters relève lui-même des contre-exemples dans son article. Voir également Lapointe (1977) pour une formulation légèrement différente évitant ce genre de difficultés.
- (I4) Par exemple les grammaires séquentielles et les grammaires métalinéaires. Voir Chomsky (1963, §4.I.)
- (I5) Voir à ce sujet la discussion du chapitre I.

CHAPITRE VII

UN THEOREME DE LIMITATION

POUR LE PROGRAMME DE CHOMSKY

7. Dans les deux chapitres précédents, nous avons tenté de montrer comment la théorie des grammaires formelles permettait d'intervenir dans la clarification et le contrôle d'assertions portant sur le langage et la grammaire. Nous allons maintenant compléter ces chapitres en exposant de quelle façon elle permet d'évaluer des propositions concernant l'organisation générale de la "théorie linguistique moderne".

Dans le premier cas, la grammaire formelle permettait de répondre à une question du type suivant: "Etant donnée une certaine propriété p du langage, la grammaire G possède-t-elle la propriété P nécessaire à la description de p ?". Dans le second cas, c'est à la question: "Etant donnée la définition D de la notion de grammaire, le système descriptif possédant les propriétés $p_1, \dots, p_n$, répond-il à la définition D ?". Pour le problème examiné dans ce chapitre, la question à laquelle permettra de répondre la grammaire formelle sera sensiblement différente. Elle ne portera plus, comme dans les cas précédents, sur la correspondance entre la définition d'un objet et les fonctions qui lui sont attribuées. Il s'agira plutôt de montrer l'incompatibilité des conditions régissant le fonctionnement supposé de la théorie des grammaires génératives.

Nous disions au début du chapitre V, que le langage dans lequel est formulée la théorie étant au dernier étage des niveaux métalinguistiques, il ne pouvait être réglé par des principes d'une rigueur comparable à celle à l'oeuvre dans les autres métalangages. Nous allons voir que c'est précisément

parce que les propositions de Chomsky sur l'organisation de la théorie linguistique impliquent une telle rigueur qu'elles peuvent être démontrées comme irréalisables.

7.1. Le programme de Chomsky

Dans les préliminaires méthodologiques d'"Aspects" , Chomsky définit dans les termes suivants le programme de recherche qui, selon lui, devra être suivi par une "théorie de la structure linguistique qui vise à l'adéquation explicative":

- (I) une théorie phonétique universelle qui définisse la notion de "phrase possible".
- (II) une définition de la notion de "description structurale"
- (III) une définition de la notion de "grammaire générative"
- (IV) une méthode propre à déterminer la description structurale d'une phrase, étant donné une grammaire
- (V) un procédé pour évaluer des grammaires en compétition.

En apparence, les objectifs de ce programme semblent modestes, à l'exception, peut-être, du point (I) (mais ceci tient à une formulation qui, nous allons le voir dans un instant, demeure vague et générale par rapport à l'idée que Chomsky entend exprimer). Les points (II) et (III) ne requièrent qu'une définition des objets dont la théorie traite. Le point (IV) exige l'effectivité des procédures descriptives construites dans la théorie. Seul le cinquième point, enfin, semble plus délicat à respecter car, contrairement aux autres, sa formulation ne comporte aucune suggestion sur la méthode à

suivre pour le satisfaire. Il correspond en fait, si on se réfère à la discussion qu'en donne Chomsky (1957, Ch.6), à la plus faible des exigences qu'on puisse imposer à une théorie quelconque sur le choix de ses méthodes descriptives^(I).

A cette caractérisation exprimée en termes linguistiques, Chomsky associe une formulation équivalente en termes mathématiques :

- (I) une énumération de la classe $s_1, s_2, \dots$ des phrases possibles,
- (II) une énumération de la classe $SD_1, SD_2, \dots$ des descriptions structurales possibles,
- (III) une énumération de la classe $G_1, G_2, \dots$ des grammaires génératives possibles,
- (IV) la caractérisation d'une fonction f telle que $SD_f(i, j)$ soit la description structurale assignée à la phrase s_i par la grammaire G_j , i et j étant arbitraires,
- (V) la caractérisation d'une fonction m telle que $m(i)$ soit un entier, associé à la grammaire G_i , et en définisse la valeur (la valeur la plus basse étant, disons, indiquée par le nombre le plus haut).

Si on compare un à un chacun des points de ces deux formulations on peut observer que les précisions introduites par la description mathématique soit affaiblissent les exigences du programme, soit les laissent inchangées.

On voit que les conditions exprimées dans les trois premiers points réduisent le problème des définitions des notions

de phrase , de description structurale et de grammaire à celui de leur définition extensionnelle. Les deux derniers points, par contre, sont identiques dans les deux formulations. La seconde définition du point (V) précisant simplement que le procédé pour évaluer des grammaires en compétition doit se traduire par un codage dans les entiers, requiert qui n'introduit aucune contrainte supplémentaire tant que l'on n'a pas d'indications plus précises sur la nature de la fonction m .

Si maintenant on considère individuellement chacun de ces objectifs dans leur formulation mathématique, on constate qu'ils sont en principe réalisables, même si leur modestie, malgré les remarques du paragraphe précédent, semble avoir disparu. L'énumération de la classe des phrases possibles est réalisable par un informateur (ou un grammairien) auquel on présente dans l'ordre lexicographique la suite des séquences formées sur le vocabulaire de sa langue. La description des phrases possibles, la classe de ces phrases étant donnée, ne présente en elle-même aucune difficulté théorique car il ne s'agit que d'analyser en constituants immédiats chacune de ces phrases. Le mode d'énumération des grammaires génératives possibles, lui, dépend de l'interprétation donnée à "possible". A cet égard on peut envisager deux éventualités. Soit "possible" a trait à la nature intrinsèque de l'objet "grammaire générative", comme le suggère le rapprochement avec la première formulation. Il ne s'agit alors que d'énumérer la classe des systèmes de réécriture en fonction de leur définition mathématique. Mais dans ce cas d'une part, on perd le "bénéfice" de la

définition extensionnelle et, d'autre part, la théorie est contrainte à énumérer des grammaires génératives n'engendrant pas les langues naturelles. Soit "possible" doit être compris comme "possible pour les langues naturelles". Dans ce dernier cas, qui est le plus vraisemblable, l'énumération sera encore différente de celle des phrases et des descriptions structurales car elle dépendra d'hypothèses sur la nature de la grammaire⁽²⁾ qui présélectionneront la classe des grammaires possibles. Cependant ces hypothèses ne devraient pas s'opposer à une énumération constructiviste analogue au cas précédent. La quatrième exigence qui porte sur la décidabilité des grammaires, dépend nécessairement de ce qu'est une grammaire générative possible. Elle hérite donc de l'ambiguïté de ce dernier terme. Notons simplement que si le deuxième sens de "possible" autorisait la présence, dans l'énumération du point (III), de grammaires indécidables l'exigence (IV) n'aurait pas de sens. Si par ailleurs cette dernière éventualité ne se présente pas, le point (IV) est également satisfait. En fait, cette condition semble contenue dans le point (III).

Pour Chomsky une théorie remplissant les quatre premières conditions satisfait l'adéquation descriptive, la satisfaction de l'adéquation explicative n'étant obtenue que si de surcroît est remplie la cinquième condition. Bien qu'aucune indication ne soit donnée sur la nature de la fonction m , on voit que celle-ci ne peut être considérée en elle-même et qu'elle dépend des quatre exigences précédentes; puisque l'évaluation de la gram-

maire dépend de la nature des DS qu'elle attribue aux s et qu'elle pourrait également dépendre de la rapidité avec laquelle f trouve $SD_{\hat{f}}(i,j)$. On ne verrait d'ailleurs pas la nécessité d'une définition formelle pour l'évaluation des grammaires si celle-ci ne renvoyait pas à une procédure effective. Comme de plus la fonction m induit un bon ordre total sur la classe des grammaires énumérées en (III) tout en dépendant de la fonction définie en (IV), la condition (V) revient à exiger que les conditions (III) et (IV) soient accomplies par des procédures décidables. Donc ce qui rend ce programme de recherche contraignant ce ne sont pas les exigences prises individuellement, mais l'ensemble de ces exigences avec les dépendances qu'il y a entre elles. En particulier les dépendances entre les conditions (II) et (III) et entre les conditions (III) et (IV) telles qu'elles découlent respectivement des conditions (IV) et (V).

Ce sont les dépendances entre ces exigences qui sont à l'origine d'un théorème de limitation remettant théoriquement en cause la possibilité de réaliser ce programme. Toutefois, avant d'exposer le contenu de ce théorème nous examinerons les motivations qui sont à l'origine du programme de Chomsky.

7.1.2. Ce qui étonne en premier lieu dans ce programme c'est le contraste entre son caractère mécanisciste et le cadre épistémologique résolument anti-béavioriste dans lequel il s'intègre. La description de la théorie linguistique que nous venons de commenter n'étant rien d'autre que "ce qui est impliqué par la construction d'un "modèle d'acquisition" pour le

langage (Chomsky, 1965, p.50). La conception de l'acquisition du langage défendue par Chomsky étant celle que l'on sait. Peut-être n'y a-t-il pas là un véritable paradoxe mais ce qui semble assuré c'est que la formulation mécaniciste du programme ne devrait pas découler directement des thèses innéistes de Chomsky.

L'origine de cette conception des théories scientifiques semble en fait remonter à des préoccupations indépendantes et antérieures, exprimées dans "Syntactic Structures" et surtout dans "The Logical Structure of Linguistic Theory". Dans la préface autobiographique de ce dernier ouvrage, Chomsky (1975, p.33) s'exprime sur ce point dans les termes suivants: "J'étais particulièrement impressionné par le travail de Nelson Goodman sur les systèmes constitutifs. Par son aspect général ce travail était d'une certaine manière semblable à celui de Harris, il me semblait fournir le cadre intellectuel approprié pour la recherche de procédures taxinomiques que je considérais alors comme centrales pour une théorie linguistique. Mais en développant sa critique de l'induction, Goodman semblait indiquer une direction assez différente suggérant l'inadéquation de principe des approches inductives."

La référence explicite au projet constitutionniste de Goodman n'équivaut évidemment pas à une conformation à cette philosophie de la connaissance. En particulier, si le rejet de l'empirisme affiché par Chomsky est en accord avec la pensée de Goodman lorsqu'il concerne la critique des démar-

ches inductives il devient par contre plus difficile de défendre l'idée d'une correspondance étroite entre la pensée du philosophe et celle du linguiste sur le rôle joué par le "monde sensible" dans une théorie scientifique. Car ce qui demeure énigmatique dans le programme de Chomsky, c'est le mode de détermination des termes empiriques fondamentaux, les phrases attestées, auquel on ne trouve aucune référence claire et précise. On peut toutefois interpréter ce silence comme un assentiment tacite à la position exposée antérieurement par Chomsky (1957) où la notion de phrase possible est définie en termes de phrases observées. Dans ce cas il devient possible de reconnaître une influence du projet constitutionniste sur l'épistémologie chomskyenne en relevant dans cette dernière les trois caractéristiques suivantes:

- 1°) Le recours nécessaire au "monde sensible" dans l'énumération des phrases possibles. (3)
- 2°) Le rejet d'une approche inductive pour la définition des termes théoriques comme, par exemple, celui de "grammaire générative"
- 3°) L'étroite et nécessaire dépendance logique entre les différents objectifs de la théorie linguistique.

En dehors du problème général que pose le principe de l'application d'une épistémologie à la démarche scientifique, comme si la philosophie des sciences avait une valeur prescriptive que la science pourrait utiliser avec bénéfice, la na-

ture algorithmique du programme de Chomsky correspond-elle à une pratique scientifique effective? Malgré le rejet des approches inductivistes la conception chomskyenne du travail scientifique partage avec celles-ci un principe discutable, comme l'a fait remarquer Putnam (1974). Putnam, qui voit là l'occasion de rapprocher Popper, Carnap et Chomsky, dénie l'existence d'un lien entre théorie et prédiction jouant un rôle central dans la déduction scientifique. En ce sens inductivistes et déductivistes adopteraient le même schéma directeur qui serait: "regardez les prédictions impliquées par la théorie; vérifiez si ces prédictions sont vraies". Dans une telle optique on comprend que les déductivistes, comme Chomsky, puissent nier que l'accumulation de l'expérience perceptuelle puisse conduire aux théories puisque dans le cadre du schéma "théorie implique prédiction" on peut faire varier librement les hypothèses de la théorie jusqu'à ce qu'elles impliquent des prédictions correctes. Mais le fondement de cette attitude dépend de la validité pratique du schéma cité à l'instant. Nous sommes convaincus, comme Putnam, que les théories en elles-mêmes n'impliquent pas de prédictions⁽⁴⁾.

Les prédictions trouvent en fait leur origine dans la conjonction d'une théorie et d'assertions auxiliaires reposant sur des suppositions inexprimables dans le cadre de la théorie. Ces suppositions correspondent, si l'on veut, aux intuitions des scientifiques et aux aléas qui déterminent l'orientation de leur pensée. L'aléatoire de ces suppositions étant inverse-

ment proportionnel à la familiarité expérimentale avec les phénomènes analysés. En définitive que l'on estime que les hypothèses d'une théorie ne soient que des conjectures provisoires ou que l'on conçoive que ces hypothèses soient justifiées par leur degré de probabilité élevé on n'élimine pas pour autant les difficultés posées par la logicisation de la démarche scientifique qui réduit le travail scientifique à des procédés purement logiques descriptibles par des algorithmes.

A cet égard on ne peut manquer d'évoquer le rôle que jouèrent les théorèmes de limitation dans les recherches sur les fondements des mathématiques. Une des interprétations que l'on peut donner aux théorèmes de Gödel est que le fonctionnement des théories mathématiques ne peut être décrit d'une façon purement logique, la notion de démonstration étant trop riche pour être représentée par des algorithmes. A fortiori, la même chose devrait se vérifier pour les sciences autres que les mathématiques puisque la déduction scientifique est d'une nature encore plus complexe et plus riche que la déduction mathématique qui n'est que le cas limite de la précédente. Il n'est donc guère étonnant que la théorie linguistique soit affectée par un théorème limitant les opérations qu'on peut y effectuer, dans la mesure où ces opérations sont définies avec précision.

7.2. Un théorème de limitation

Comme nous le disions quelques pages plus haut, ce qui

peut choquer dans la description des objectifs d'une théorie linguistique proposée par Chomsky c'est le parti-pris qui consiste à préjuger des voies méthodologiques que devrait suivre la recherche. Parti-pris qui trouve sa légitimation dans l'idée qu'une théorie scientifique convenable se doit de suivre les préceptes édictés par une philosophie des sciences. Cette attitude trouve cependant, dans les cas particuliers de la formalisation proposée par Chomsky, sa contrepartie positive en offrant le moyen de juger de la possibilité théorique de sa réalisation. Toutefois ici comme dans d'autres cas la démonstration de la possibilité mathématique du programme de Chomsky demeurerait sans signification pratique car elle laisserait ouverte la question de sa mise en oeuvre. C'est seulement dans la mesure où le programme peut être sanctionné par un théorème de limitation qu'un tel théorème peut être instructif.

En réinterprétant un résultat de Dekker (1953), Janssen Kok et Meertens (1976) nous rappelle que par un raisonnement diagonal classique on peut démontrer qu'il n'existe pas de sous-classe G de la classe des grammaires génératives satisfaisant les trois conditions suivantes :

- (i) G est récursivement énumérable
- (ii) il existe une méthode pour décider si une phrase donnée peut être engendrée par une grammaire donnée de G
- (iii) pour tout langage récursif il existe une grammaire dans G .

La première condition ne fait que traduire l'énumérabilité

des grammaires possibles. La seconde est la traduction stricte du point (IV) du programme et équivaut selon les auteurs⁽⁵⁾ à l'assertion selon laquelle les langues naturelles sont récursives. Dès lors, on ne comprend plus pourquoi ils estiment nécessaire, "afin d'échapper au risque d'exclure des langues naturelles possibles", d'adjoindre la troisième condition. D'une part celle-ci est superflue par rapport à la seconde condition, d'autre part, elle impose absurdement à la théorie d'énumérer des grammaires engendrant des langages composés de phrases agrammaticales (les complémentaires des langages composés de phrases grammaticales étant eux-mêmes récursifs). C'est d'ailleurs cette propriété qui est utilisée dans l'argument diagonal où est défini constructivement un ensemble récursif composé de phrases refusées par chaque grammaire de l'énumération. L'intérêt de ce théorème réside donc uniquement dans la délimitation du problème qu'il entend traiter.

Réexaminons les conditions du théorème. La première condition, d'abord, exprimant l'existence d'une sous-classe récursivement énumérable, G , de la classe des grammaires génératives est superflue puisqu'à partir de l'énumération de la classe complète la condition (ii) sélectionnant les grammaires décidables suffit à déterminer la sous-classe restreinte. Par ailleurs les restrictions sur la classe des grammaires possibles sont plus fortes que celles proposées par Janssen, la classe n'étant pas récursivement énumérable, mais récursive (ceci est dû à la finitude des grammaires et à leur mode de construction)

qui plus est, si la classe des grammaires possibles n'était pas récursive alors l'exigence (V) du programme de Chomsky n'aurait pas lieu d'être formulée comme nous le disions plus haut. La seconde condition du théorème de Janssen étant une traduction exacte du point (IV) du programme nous pouvons l'adopter telle quelle. Cependant pour pouvoir se prononcer sur la réalisabilité du programme, il convient d'intégrer les contraintes exprimées par le point (V). Il semble raisonnable de dire que l'on aura $m(i) = 0$ lorsque f ne sera pas calculable pour G_i , c'est à dire lorsque G_i ne sera pas une procédure de décision. De plus, étant donné notre interprétation de la première condition du théorème de Janssen, Kok et Meertens, m est une fonction définie et calculable pour tout i . L'ensemble de ces remarques peut alors être résumé par la proposition suivante:

Il existe une procédure de décision permettant de déterminer pour tout i si G_i est une procédure de décision.

Dans le cas de l'interprétation de "grammaire générative possible" comme "mathématiquement possible" la proposition précédente n'est rien d'autre, par la thèse de Church, qu'une formulation particulière du Problème de l'Arrêt pour les machines de Turing et elle est évidemment fausse.

Si par contre on comprend "grammaire possible" comme "possible pour les langues naturelles", alors l'énumération des G_i dépend des hypothèses possibles sur la structure des grammaires génératives pour les langues naturelles. Mais de deux choses l'une; ou ces hypothèses sont énumérables ⁽⁶⁾ et,

hormis l'absurdité empirique d'une telle supposition, nous retrouvons le cas précédent⁽⁷⁾, ou ces hypothèses ne sont pas énumérables (au sens mathématique du terme) et alors on doit nous expliquer ce que signifie, pour une théorie linguistique, d'énumérer la classe des grammaires génératives possibles.

Evidemment ce théorème de limitation ne signifie pas qu'il soit impossible de déterminer si une grammaire est une procédure de décision. Nous avons d'ailleurs vu dans le chapitre précédent que, bien que difficile, ce problème avait été résolu pour un modèle particulier de la grammaire transformationnelle. Le théorème remet simplement en cause l'idée que la recherche de conditions métaagrammaticales de décidabilité puisse se faire selon des méthodes générales et que cette généralité permette d'évaluer, en un sens mathématique, des grammaires concurrentes.

NOTES DU CHAPITRE VII

- (1) Ce choix s'oppose à celui des conceptions de la théorie comme procédure de découverte des grammaires à partir de corpus de données et comme procédure de décision sur la correspondance entre grammaire et corpus
- (2) Ces hypothèses pourraient par exemple être les suivantes:
la grammaire n'a pas de règles essentiellement dépendantes du contexte, la grammaire comprend plusieurs symboles ré-cursifs, etc...
- (3) Nous parlons ici des intentions explicites du programme.
Que ces intentions soient respectées est un autre problème (Cf. notre chapitre II)
- (4) Si on donne à "implique" le sens qu'il a en logique mathématique
- (5) A ce niveau de raisonnement une telle équivalence est acceptable. On se réfèrera à la section 6.I.2. pour un commentaire sur ce problème
- (6) Etant entendu qu'une hypothèse équivaut à une condition formelle sur la nature de la grammaire
- (7) Puisqu'il s'agit de voir si l'adjonction d'une hypothèse h_i à la grammaire G_j entraîne ou non la décidabilité de cette dernière.

TROISIEME PARTIE

PERSPECTIVES POUR LA THEORIE MATHEMATIQUE

DES GRAMMAIRES

CHAPITRE VIII

VERS UNE THEORIE MATHEMATIQUE

DES

LANGAGES TRANSFORMATIONNELS

8. Le lecteur familiarisé avec la théorie des grammaires formelles pourra penser que, jusqu'à présent, notre plaidoyer en faveur de la théorie mathématique des grammaires n'a guère été convaincant. En effet, si on voulait bien nous concéder que nous sommes parvenus à dégager des interprétations de cette théorie lui conférant une signification linguistique intéressante, on pourrait toujours nous objecter que les résultats discutés dans les chapitres précédents ont été choisis à dessein et que loin d'être exemplaires de ce qui intéresse la linguistique, ils ne correspondent qu'à quelques cas marginaux parmi la masse des travaux consacrés à cette discipline. Celle-ci étant généralement orientée soit vers l'étude de problèmes d'informatique, soit vers celle de questions de mathématiques pures. Si nous admettons que telle est bien la situation, nous tenterons toutefois de nuancer cette présentation des choses en faisant remarquer que l'éventuelle impression de panorama exhaustif des relations entre mathématiques et grammaire générative que pourraient laisser les pages précédentes tient aussi au fait que nous avons discuté des résultats les plus connus en la matière.

Or, si ces résultats sont particulièrement connus cela tient plus à ce qu'ils concernent des problèmes critiques internes à la linguistique, qu'à ce qu'ils proposent une approche mathématique intéressante en elle-même, soit par l'originalité de ses méthodes, soit par la découverte de propriétés inatten-

dues. Aussi nous voudrions montrer, dans cette troisième partie, que la théorie des grammaires formelles échappe à l'espèce de dualisme qui voudrait que lorsqu'elle se livre à une exploration mathématique gratuite des systèmes grammaticaux, ne répondant à aucune sollicitation linguistique particulière, elle ne parvienne qu'à contribuer à une théorie mathématique ne présentant aucune perspective d'application à la grammaire générative.

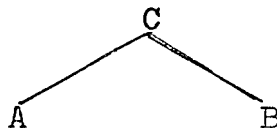
Au terme des deux chapitres qui vont suivre s'imposera une constatation : les études mathématiques fructueuses de la grammaire, au sens où elles concernent les grammaires génératives effectivement utilisées par les linguistes, ne sont possibles qu'une fois que la grammaire formelle s'est affranchie de certaines des déterminations imposées par sa problématique traditionnelle.

8.I. Sur l'existence des langages transformationnels

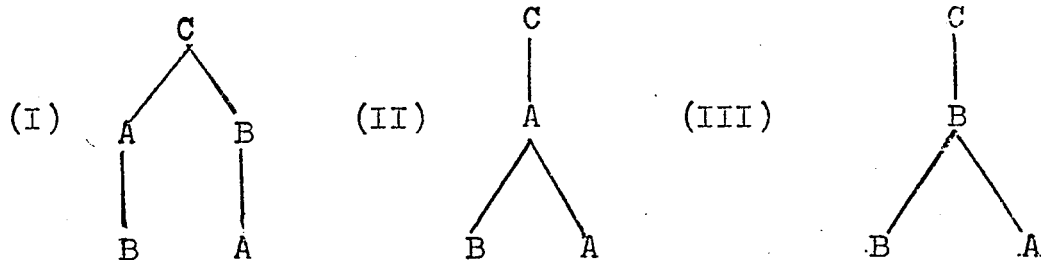
Les quelques travaux consacrés à l'étude mathématique des grammaires transformationnelles, ceux de Ginsburg, Kimball, Peters et Ritchie, Salomaa, auxquels on peut ajouter ceux de Joshi partagent en commun un même objectif : démontrer que tout ensemble récursivement énumérable est engendré par une grammaire transformationnelle. Une telle propriété étant commune à ces différents modèles, on pourrait en conclure que les grammaires transformationnelles ne sont qu'un moyen parmi d'autres pour exprimer la notion de calculabilité et qu'elles sont dépourvues de propriétés les distinguant par exemple des machines de

Turing. C'est de cette équivalence que nous allons maintenant discuter.

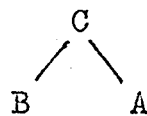
8.I.I. Selon Chomsky (1959,p.128) les transformations sont des opérations trop complexes pour que les grammaires transformationnelles puissent être étudiées d'un point de vue analogue à celui selon lequel on étudie les différentes grammaires de constituants. Les grammaires de type 0, c'est à dire la classe entière des systèmes semi-thueiens, étant comprise également comme grammaire de constituants. Cette limitation tient au fait que les transformations sont des opérations "dépendantes de la structure" contrairement aux règles de différentes grammaires de type 0, 1, 2 et 3. Ces dernières règles ne portent effectivement que sur des symboles ou des suites de symboles, étant aveugles aux étapes précédentes de la dérivation, alors que les transformations sont des opérations stochastiques dont l'application dépend des étapes antérieures de la dérivation. Ainsi il y a une différence fondamentale entre la dérivation de la séquence BA par la grammaire de constituants $\{C \longrightarrow AB, AB \longrightarrow BA\}$ et la grammaire transformationnelle $\{C \longrightarrow AB, [A] [B] \Rightarrow [B] [A]\}$. La première permet de construire à partir de l'indicateur syntagmatique:



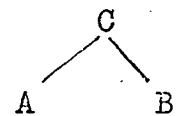
les différents indicateurs syntagmatiques dérivés:



sans donner d'indication sur le choix du développement particulier qui doit être retenu^{et} surtout, dans le cas du choix de (1) qui est la dérivation linguistiquement la plus naturelle, en attribuant d'incohérents étiquetages de constituants. Le constituant B étant analysé comme un A et réciproquement. Inversement la règle transformationnelle qui permute les constituants donnant



à partir de



évite

ces deux difficultés. Cependant, la distinction entre règle de réécriture et règle transformationnelle tient-elle encore si on mène l'analyse formelle jusqu'à son terme? Cette distinction entre les règles est en effet justifiée par les différents usages empiriques qu'en font les grammairiens. C'est parce qu'elles ont des fonctions diverses dans la grammaire que leur sont attribués des statuts formels distincts, mais a priori ce ne sont pas des singularités formelles qui impliquent des comportements formels différents. En particulier rien ne semble s'opposer à ce que l'on voit dans les grammaires transformationnelles un cas particulier de grammaire de type 0 dont le vocabulaire comprendrait des parenthèses étiquetées permettant ainsi de conserver l'information contenue dans les gra-

phes en arbre. Il est en effet connu que les arbres ne sont qu'un des moyens possibles de représentation des indicateurs syntagmatiques et qu'une représentation par parenthèses étiquetées leur est équivalente. Prenons un exemple. Soit la grammaire de constituants G:

$P \longrightarrow SN \ SV$
 $SN \longrightarrow Det \ Nh + Det \ N$
 $SV \longrightarrow V \ SN$
 $V \longrightarrow mange$
 $Nh \longrightarrow garçon$
 $N \longrightarrow gâteau$
 $Det \longrightarrow le$

Cette grammaire permet la dérivation suivante: $P \xRightarrow{(1)} SN \ SV$
 $Det \ Nh \ SV \xRightarrow{*} Det \ Nh \ V \ SN \xRightarrow{*} Det \ Nh \ V \ Det \ N \Rightarrow le \ garçon \ mange \ le \ gâteau.$

Nous pouvons alors définir une grammaire transformationnelle G', en adjoignant à G la règle de transformation passive: (2)

$SN - V - SN \Rightarrow SN - est - V + pp - par \ SN$
 $1 \quad 2 \quad 3 \quad 3 \quad est \quad 2 + pp - par \ 1$

La transformation est bien différente d'une simple règle de réécriture puisqu'elle ne porte pas sur les sous-séquences terminales "le garçon", etc... mais sur ces sous-séquences et sur les étiquettes de constituants les plus basses les dominant exhaustivement. Il est cependant possible de construire une grammaire de type 0, G'' qui donne autant d'in-

formations structurales que la grammaire G' . Ainsi G'' :

$$P \rightarrow [{}_P \text{SN SV}]_P$$

$$\text{SN} \rightarrow [{}_{\text{SN}} \text{Det Nh}]_{\text{SN}} + [{}_{\text{SN}} \text{Det N}]_{\text{SN}}$$

$$\text{SV} \rightarrow [{}_{\text{SV}} \text{V SN}]_{\text{SV}}$$

$$\text{V} \rightarrow [{}_{\text{V}} \text{mange}]_{\text{V}}$$

$$\text{Det} \quad [{}_{\text{Det}} \text{le}]_{\text{Det}}$$

$$\text{Nh} \quad [{}_{\text{Nh}} \text{garçon}]_{\text{Nh}}$$

$$\text{N} \quad [{}_{\text{N}} \text{gâteau}]_{\text{N}}$$

$$P [{}_{\text{SN}} \text{Det Nh}]_{\text{SN}} [{}_{\text{SV}} \text{V}]_{\text{SV}} [{}_{\text{SN}} \text{Det N}]_{\text{SN}} [{}_{\text{SV}} \text{V}]_{\text{SV}} \Rightarrow [{}_{\text{PSN}} \text{Det N}]_{\text{PSN}} [{}_{\text{SV}} \text{est} [{}_{\text{V}} \text{V}]]_{\text{SV}} \text{pp}$$

$$\text{par } [{}_{\text{SN}} \text{Det Nh}]_{\text{SN}} [{}_{\text{SV}} \text{V}]_{\text{SV}} \text{p}$$

est bien de type 0 (elle est même de type 1 puisque ses règles "conservent" la longueur) et elle donne autant de renseignements sur la structure que la grammaire transformationnelle G' . Il semble donc qu'à une grammaire contenant des règles transformationnelles on puisse associer une grammaire de type 0 fortement équivalente. La spécificité formelle des grammaires transformationnelles, s'il y en a une, ne peut alors pas résider dans le fait que celles-ci comprennent des règles dépendantes de la structure. Mais si la classe des grammaires transformationnelles est un objet mathématique distinct de la classe des systèmes de réécriture illimités, elle ne peut l'être que de deux façons: soit elle est formellement incomparable à la classe des grammaires de type 0, soit elle est une sous-classe propre de celle-ci.

Nous venons de voir que si une de ces possibilités est réalisable, ce ne peut être que la seconde.

Mais cependant, si on veut établir une véritable spécificité des grammaires transformationnelles, il faut faire intervenir d'autres propriétés que la seule forme de leurs règles. Car si la classe des grammaires transformationnelles se trouvait être une sous-classe propre de la classe des grammaires de type 0 (T_0) cela permettrait de les distinguer de cette dernière mais ce serait insuffisant pour dire que ce n'est pas une classe de grammaires de constituants parmi les autres. Les critères de distinction doivent donc être recherchés ailleurs que dans les règles, c'est à dire dans l'organisation de ces règles à l'intérieur de la grammaire. En dehors de la présence de règles de constituants et de règles transformationnelles, quelles sont les propriétés des grammaires transformationnelles? Il y en a deux essentielles: 1) les grammaires transformationnelles comprennent des conditions d'admissibilité sur les séquences produites, des filtres, imposant à celles-ci de ne pas contenir le symbole #, marqueur de frontière de phrases; 2) les grammaires transformationnelles imposent un ordre partiel sur les règles transformationnelles, leur permettant de s'appliquer cycliquement.

8.1.2 On peut alors poser le problème de la spécificité des grammaires transformationnelles en d'autres termes. Les grammaires transformationnelles seront distinctes des grammaires de constituants si elles définissent une sous-classe propre de T_0 et si cette inclusion propre est la conséquence soit de la présence des filtres, soit de la présence du cycle trans-

formationnel, soit de la présence des filtres et du cycle transformationnel.

La première possibilité ne mérite pas d'être étudiée, puisque, la présence des filtres, sans cycle transformationnel, dans une grammaire, vu la possibilité de représentation des règles transformationnelles par des règles de type 0, permet d'identifier trivialement les grammaires transformationnelles à des grammaires de type 0. De plus, les filtres augmentant la capacité générative des grammaires transformationnelles, ils ne peuvent être à l'origine de la spécification d'une sous-classe propre de la classe des langages de type 0.

Nous avons déjà analysé la troisième possibilité au chapitre VI, lorsque nous avons exposé les théorèmes de Peters et Ritchie, et nous avons alors vu que la double présence des filtres et du cycle transformationnel avait des conséquences importantes. Mais ces conséquences étaient précisément opposées à celles que nous recherchons actuellement puisque c'est par l'usage intensif des filtres que les auteurs démontraient que tout ensemble récursivement énumérable était engendré par une grammaire transformationnelle filtrante quelle que soit la nature de sa base. La conclusion est donc que la troisième possibilité ne permet pas d'établir la spécificité des grammaires transformationnelles puisqu'au contraire elle permet de démontrer que grammaire transformationnelle et grammaire de type 0 sont les mêmes

objets. Le fait que l'on puisse déterminer des conditions de récursivité décidables pour les langages transformationnels engendrés par des grammaires filtrantes ne change rien à cette situation. Les conditions proposées par Peters et Ritchie et Lapointe reviennent en effet à formaliser certaines propriétés empiriques des grammaires transformationnelles. Leur problème étant de déterminer un modèle mathématique qui soit le plus représentatif possible des grammaires utilisées par les linguistes afin de vérifier si celles-ci engendrent des langages récursifs. Mais ces conditions de récursivité ne découlent pas de propriétés mathématiques inhérentes aux grammaires transformationnelles et elles permettraient sans doute de définir une classe de grammaires de constituants (c'est à dire ne contenant que des règles de réécriture) intermédiaire entre la classe des grammaires de type 1 et celle des grammaires de type 0, comme c'est le cas ⁽³⁾ pour les grammaires transformationnelles récursives. Ce qui nous intéresse ici est différent. Nous voudrions savoir si les grammaires transformationnelles possèdent une propriété particulière, une "transformationnalité" pour ainsi dire, partagée par la classe entière de ces systèmes ou plus exactement par une des sous-classes définies par la combinatoire élémentaire de propriétés présentée deux pages plus haut.

La seule possibilité restante est que les grammaires transformationnelles définissent une sous-classe propre de T_0 et que cette inclusion soit la conséquence de la présence

du cycle transformationnel. Nous avons vu, toujours au chapitre VI, que cette possibilité n'était pas réalisable pour les grammaires transformationnelles à base dépendante du contexte. Ce qu'il faut donc étudier, ce sont les propriétés mathématiques des grammaires transformationnelles non-filtrantes à base indépendante du contexte.

8.1.3. Telle que nous venons de l'introduire, cette étude se présente comme un problème de mathématiques pures. Question intéressante en soi pour qui voudrait tester la validité de l'affirmation de Chomsky présentée plus haut. Dans la perspective des rapports entre grammaire formelle et grammaire générative qui nous occupe, l'intérêt du problème dépend cependant de ses liens éventuels avec la réalité grammaticale. Or il apparaît que la majorité des grammaires transformationnelles employées par les linguistes, d'une part possèdent des filtres, d'autre part ont une base dépendante du contexte. L'usage fréquent que font les linguistes de ces propriétés permet-il toutefois d'affirmer qu'elles sont liées à l'analyse transformationnelle de façon essentielle?

Nous avons déjà signalé que dans la démonstration de Peters et Ritchie la fonction des filtres était opposée à celle qu'elle a dans les analyses linguistiques. Les grammairiens voient dans les filtres les moyens d'éliminer des séquences terminales mal formées, les mathématiciens, inversement, les utilisent, par un calcul artificiel, pour dériver une séquence quelle que soit sa forme. Mais en procédant ainsi

ils ne font qu'exploiter une possibilité que rien, dans l'usage linguistique, n'interdit. Donc le problème d'une utilisation linguistique raisonnable des filtres peut se poser, et il se pose. Kuroda (1969, §3) dans son analyse de la distribution des adverbes "even" et "also" fait remarquer que l'adhésion stricte au principe de Katz et Postal conduit à un emploi des transformations filtrantes réduisant celles-ci à un dispositif ad hoc. Si l'usage des filtres est parfaitement légitime dans le cas de la relativisation où ils permettent de dériver (1) à partir de (2)

(1) A man whom John knows is standing there

(2) A man # John knows that man # is standing there
alors qu'ils bloquent la dérivation de (3)

(3) A man # John knows that boy # is standing there
il n'en est pas de même dans le cas du blocage des phrases du type de (4)

(4)* The storm even destroyed even his farm.

Le seul moyen de ne pas produire cette phrase, si on obéit au principe de Katz et Postal, étant de lui attribuer une structure profonde de la forme (5):

(5) The storm even # the storm destroyed even his farm # destroyed his farm.

puisque la différence de sens entre (6) et (7)

(6) The storm even destroyed his farm

(7) The storm destroyed even his farm

empêche de relier transformationnellement ces deux phrases

et oblige à poser dans le composant de base deux règles indépendantes du contexte: $VP \rightarrow (\text{even}) V NP$ et

$NP \rightarrow (\text{even}) \text{Det } N.$ (4)

Pour Kuroda le filtrage des phrases simples n'est pas conceptuellement satisfaisant puisque celles-ci, en principe, devraient pouvoir être dérivées au terme d'un seul cycle transformationnel.⁽⁵⁾ Il n'est pas non plus normal, d'un point de vue formel, de faire intervenir des filtres "qui ne jouent absolument aucun rôle positif dans la dérivation de la forme superficielle des phrases" comme c'est le cas pour le blocage de (4). Formellement, l'emploi des filtres est légitime d'une part lorsqu'on traite des phrases complexes où plusieurs noeuds S sont enchassés en structure profonde, d'autre part, lorsqu'ils permettent de sélectionner dans le cadre d'un processus général et fréquent de dérivation, comme la relativisation, des phrases bien formées.

A ces remarques appelant à un usage modéré de la notion de filtres, nous pouvons ajouter certaines observations de Perlmutter (1970) où celui-ci montre que la dérivation de certaines structures de surface mal formées ne peut être bloquée par les filtres. Il faut alors avoir recours à des contraintes sur les structures profondes qui sont par nature différentes des conditions d'admissibilité liées à la présence des # et qui jouent le rôle de filtres locaux.

Relativement au problème qui nous intéresse, le sens

des analyses de Kuroda et de Perlmutter est qu'il est illogique de faire fonctionner tout le dispositif grammatical pour obtenir des séquences dérivées qui seront bloquées et qui doivent être dérivées uniquement pour pouvoir être bloquées. Implicitement, c'est un critère de meilleur rendement du mécanisme transformationnel que recherchent Kuroda et Perlmutter, critère qui devrait être substitué à un filtrage ad hoc et peu efficace.

L'étude des grammaires non filtrantes ou peu filtrantes n'est donc pas réalisée pour satisfaire uniquement la curiosité des mathématiciens, mais également parcequ'elle se présente comme une alternative empiriquement intéressante. La seconde partie de la question que nous posions plus haut sur l'intérêt de l'étude de grammaires à base indépendante du contexte tombe immédiatement devant un théorème démontré par Peters et Ritchie (1969): "Toutes les grammaires dépendantes du contexte fonctionnant en reconnaissance sont équivalentes à des grammaires indépendantes du contexte". Pour dire les choses autrement: il n'y a pas d'analyse en constituants immédiats qui soit essentiellement dépendante du contexte⁽⁶⁾. Résultat apparemment sans grande signification si on veut voir dans les grammaires génératives des mécanismes de production, plutôt que de reconnaissance, comme le suggère leur présentation usuelle dans la littérature. Mais il ne faut pas oublier que ces systèmes ne sont pas - ou ne devraient pas être - complètement arbitraire, et que leur forme résulte, dans une certaine mesure, d'une pratique.

Cette pratique, c'est ce que les anglo-saxons appellent le "parsing", c'est à dire l'analyse des énoncés se présentant à l'observation. Analyse dont la propriété fondamentale est de pouvoir être représentée par un graphe planaire, comme l'exprime le théorème de Peters et Ritchie juste cité. Or il serait vain de supposer que les systèmes de règles de réécriture puissent être conçus en toute indépendance à l'égard des techniques et des catégories de l'analyse des énoncés. Celle-ci détermine le moule dans lequel est nécessairement formé tout système de réécriture qui par-là même conserve une "analycité inhérente" dans son fonctionnement en production. Telle est la raison pour laquelle il est toujours possible d'associer un indicateur syntagmatique aux dérivations des grammaires dépendantes du contexte habituellement employées dans les grammaires transformationnelles.

Avant d'en venir à l'exposé des propriétés mathématiques des grammaires non-filtrantes nous voudrions conclure cette section par une dernière remarque. Un rapprochement entre ce qui va suivre et le chapitre VI pourrait laisser penser que notre intention est de comparer deux types d'études mathématiques d'un objet unique qui serait la grammaire transformationnelle. Il n'en est rien; les grammaires discutées au chapitre VI et celles qui nous intéressent ici sont des systèmes aussi différents les uns des autres que peuvent l'être les grammaires indépendantes et dépendantes du contexte. Ceci ne veut pas dire que l'objet "grammaire transformationnelle" ne puisse être appréhendé selon des techniques mathématiques.

ques différentes conduisant à des structurations diverses. La théorie des grammaires formelles nous en donne d'ailleurs un exemple avec les théorèmes de Kimball (1967) et de Peters et Ritchie (1971) sur les grammaires transformationnelles à base régulière où le premier auteur démontre par des méthodes algébriques ce que les seconds démontrent par des méthodes combinatoires⁽⁷⁾.

8.2. Propriétés remarquables des grammaires localement filtrantes.

Comme nous le disions plus haut, la motivation empirique principale de l'étude des grammaires non-filtrantes est la recherche de conditions de fonctionnement optimal. Avec une grammaire contenant des filtres on peut être amené à construire la dérivation complète de phrases, jusqu'à l'application de l'ultime cycle transformationnel, pour voir celles-ci bloquées par les filtres, situation qui ne peut jamais se produire avec une grammaire non-filtrante.

A la différence d'efficacité, intuitivement ressentie, avec laquelle ces grammaires énumèrent leurs langages la théorie de la calculabilité permet d'associer un concept mathématique, celui d'énumération prédictive. Un langage L est dit prédictiblement énumérable s'il est fini ou s'il existe une machine de Turing Z et une fonction élémentaire récursive⁽⁸⁾ f telles que

(1) Z énumère L

(2) Z énumère au moins $(n+1)$ membres de L en utilisant au plus $f(n)$ cases du ruban de Z pour tout n .

Cette définition formalise bien l'idée, sinon d'un rendement optimal, du moins de conditions minimales d'efficacité de ~~xxxxxxxx~~ rendement d'une machine de Turing. La classe des langages prédictiblement énumérables étant une sous-classe propre de celle des langages récursivement énumérables, il devient intéressant de formaliser, dans le cadre de la grammaire cette fois, l'idée d'efficacité d'énumération car celle-ci devrait sans doute se traduire par une restriction sur la classe des langages engendrables donnant alors un contenu objectif à la distinction entre grammaires filtrantes et grammaires non-filtrantes.

Plutôt que de s'intéresser aux grammaires non-filtrantes Peters et Ritchie (1973b, p.183) se proposent d'étudier les grammaires localement filtrantes, et cela pour deux raisons.

(1°) Les grammaires localement filtrantes incluent comme cas particulier les grammaires non-filtrantes.

2°) Certaines transformations peuvent agir de façon post-cyclique. Ceci pour éviter de bloquer par filtrage la dérivation de $[_S \omega [_S \psi [\psi \#]_S]_S \#]_S$ à cause de la présence de # à droite de ψ dans la séquence déjà transformée $[_S \psi [\psi \#]_S]_S$ sur laquelle peut éventuellement s'appliquer une transformation post-cyclique⁽⁹⁾ (par exemple la transformation qui convertit un constituant étiqueté S en une proposition subordonnée).

Une grammaire G est alors dite localement filtrante si toute dérivation D de G satisfait la condition suivante:
Si la sortie d'un cycle quelconque de D contient un # , alors

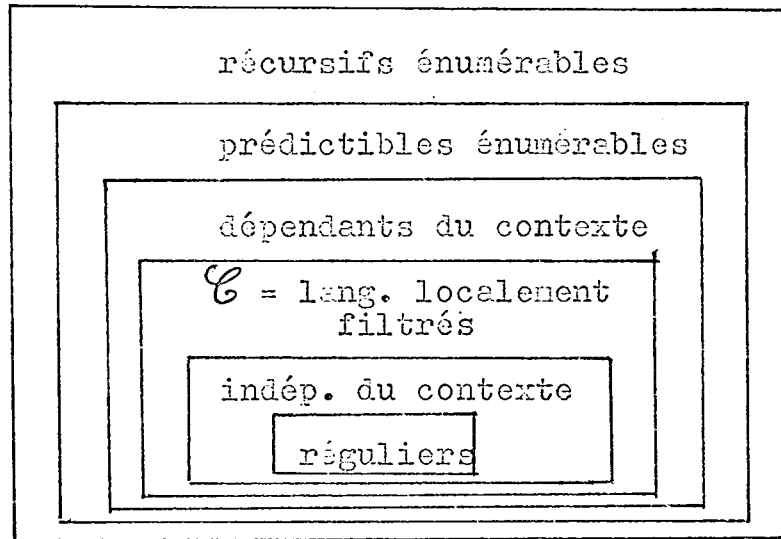
la dernière ligne de D contient un #.

La classe $\mathcal{C}$ des langages engendrés par les grammaires localement filtrantes à base indépendante du contexte possède un certain nombre de propriétés remarquables dont nous inventerions les plus significatives. En premier lieu examinons le contenu de $\mathcal{C}$. Elle contient, trivialement, tous les langages indépendants du contexte et des langages dépendants du contexte, mais pas la totalité de ceux-ci. Considérons les langages suivants: $L_1 = \{a^n / n=2m, m \geq 0\}$ et $L_2 = \{a^n / n = 2^{2^m}, m \geq 0\}$. Ce sont des langages dépendants du contexte puisqu'ils sont acceptés par un automate linéairement borné. La grammaire G_1 contenant les règles de base $\{S \rightarrow aS, S \rightarrow a\}$ et la transformation $\{a - S \Rightarrow S - S\}$, produit des structures profondes $[_S a [_S a \dots [_S a]_S \dots]_S]_S$ qu'elle convertit en séquences terminales de la forme $\{a^n / n = 2m, m \geq 0\}$ et elle est bien localement filtrante puisqu'elle est non-filtrante (en effet elle ne contient pas de #). Mais le théorème de caractérisation de la classe $\mathcal{C}$ nous montre que L_2 ne peut être membre de celle-ci:

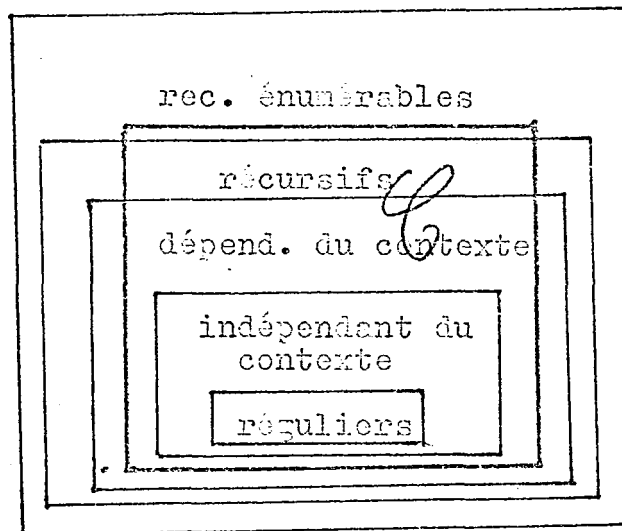
Si L est dans $\mathcal{C}$, alors L est prédictiblement énumérable et si L est infini, alors il existe une constante K telle que pour tout n il existe un x appartenant à L vérifiant $K^n \leq l(x) < K^{n+1}$

Le langage L_2 n'appartient pas à $\mathcal{C}$ car la longueur de ses éléments croît trop vite. Cependant L_2 est un langage prédictiblement énumérable (10). Avec ces renseignements sur $\mathcal{C}$ nous pouvons essayer de situer cette classe de langages formels

dans la hiérarchie de Chomsky à l'aide de la figure ci-dessous en nous fiant au principe selon lequel les restrictions sur la grammaire sont représentées par des conditions d'inclusions sur la classe des langages correspondants:



Ce diagramme correspond au type de classification classique où "selon les critères plus ou moins restrictifs que l'on introduit on obtient toute une hiérarchie (de classes de grammaires) qui s'ordonne par inclusion" (Gross et Lentin, 1970, p.77). Cependant ce diagramme est faux car parmi les propriétés de $\mathcal{C}$ que nous n'avons pas encore mentionnées figure la suivante: $\mathcal{C}$ contient un langage non-récursif (11). On ne peut donc représenter la classe $\mathcal{C}$ par un diagramme où ne figurent que des inclusions, et d'ailleurs comme Peters et Ritchie le soulignent $\mathcal{C}$ ne s'insère dans aucune hiérarchie connue de langages formels. Le diagramme convenable pour représenter $\mathcal{C}$ en fonction de la hiérarchie de Chomsky est le suivant:



Pour en terminer avec la présentation des propriétés mathématiques de la classe des langages transformationnels localement filtrés signalons que celle-ci n'est pas close sous l'intersection avec les langages réguliers. En cela elle se distingue de toutes les autres classes de langages figurant dans le diagramme ci-dessus. Cette remarque a son importance. En particulier nous pouvons la rapprocher des observations faites à la fin du chapitre III lorsque nous commentons le caractère apparemment général de la formulation du théorème d'intersection avec les langages réguliers que proposait Langendoen.

8.3. Conséquences méthodologiques

Les conséquences de l'étude mathématique de la classe $\mathcal{C}$ sont multiples. Elles permettent d'abord de répondre à la question de la spécificité des grammaires transformationnelles telle que nous la posions au début de ce chapitre. S'il est possible, par certains artifices formels, de ramener la forme apparente des règles transformationnelles à celle de simples règles de réécriture, nous voyons que cette réduction

s'arrête précisément aux apparences: les langages transformationnels dans l'acception la plus générale du terme sont des objets essentiellement différents des langages de constituants. Ils n'acceptent même pas les critères classificatoires les plus élémentaires de ceux-ci. Cette circonstance nous amène à nous interroger une fois de plus sur la pertinence d'une typologie des grammaires fondée sur la typologie des langages, c'est à dire déterminée à partir de la notion de capacité générative faible. C'est en effet là le principe qui guide la majorité des études de grammaire formelle, et nous en illustrons à l'instant la force avec la citation de Gross et Lentin donnée deux pages plus haut. Les propriétés des grammaires transformationnelles localement filtrantes à base indépendantes du contexte remettent en cause ce principe, non seulement parce qu'elles constituent un contre-exemple au mode de classification des langages formels, mais surtout parcequ'elles représentent la meilleure approximation mathématique des grammaires utilisées par les linguistes. En effet les langages de la classe $\mathcal{C}$ n'auraient été que le produit d'une exploration des grammaires formelles seulement destinée à satisfaire la curiosité des mathématiciens, leur étude n'aurait pas pu conduire à des considérations aussi critiques à l'égard de la problématique traditionnelle de la théorie mathématique des grammaires. Aussi bien nous avons déjà rencontré des langages formels, au chapitre IV, qui eux non plus ne s'inséraient pas dans les hiérarchies connues et nous n'avons pas pris prétexte de leur existence pour formuler de tel-

les réserves. Mais dans le cas des grammaires localement filtrantes il s'agit de systèmes linguistiquement motivés déterminés, comme les grammaires formelles de la hiérarchie de Chomsky, à partir de systèmes grammaticaux existants. D'ailleurs le rapprochement entre les langages de la classe et les langages bornés s'arrête à la seule incompatibilité avec la relation d'inclusion définissant la hiérarchie de Chomsky, car si on dispose d'un théorème de caractérisation pour ces deux classes de langages⁽¹²⁾, seuls les langages de la classe $\mathcal{C}$ se voient associer une grammaire.

Mais si la capacité générative faible n'est pas une propriété acceptable pour classer les grammaires quel critère retenir? Peut-être l'efficacité d'énumération pourrait-elle être un de ces critères. En particulier parmi les grammaires transformationnelles filtrantes, il devrait être possible de définir des degrés d'efficacité d'énumération. Nous avons vu au chapitre VI que les fonctions cycliques des grammaires transformationnelles pouvaient formaliser cette notion. Cependant comme Cobham (1965) le fait remarquer, dans une discussion sur la complexité de calcul en général et indépendamment de toute application à la linguistique, la hiérarchie de Ritchie (1963) s'appuie sur des caractéristiques des machines de Turing (longueur de bande utilisée, nombre d'étapes du calcul) qui ne sont pas nécessairement les plus appropriées pour définir un concept satisfaisant de complexité. Ceci parce que la notion d'étape de calcul est trop générale et peut correspondre à une opération très complexe: effacer

une bande entière, remettre la tête de lecture en position initiale, etc.... Cobham suggère de tenir compte d'autres aspects des machines de Turing comme, par exemple, les arrangements possibles des bandes, ou même le type de notation adoptée pour représenter les entiers naturels. Cependant quelles que soient les caractéristiques retenues pour spécifier la complexité d'une grammaire transformationnelle par la place de sa fonction cyclique dans les diverses hiérarchies une difficulté reste en suspens: ce type de critère permet plutôt d'évaluer les différents usages d'un modèle donné que de distinguer des modèles d'une façon essentielle. Mesurer la complexité d'une grammaire par sa fonction cyclique revient à estimer la complexité de la "performance du linguiste" qui utilise les grammaires filtrantes plutôt que d'estimer la complexité intrinsèque de la grammaire.

Au moins la difficulté de ces problèmes permet-elle de s'interroger sur l'affirmation, un peu catégorique selon nous, de Schützenberger (1978, p. 6) pour qui le modèle transformationnel "ne possède pas une structure assez riche pour bénéficier de raisonnements mathématiques excédant ceux des "Maths. Modernes". En admettant que ce jugement concerne également les résultats exposés ici, peut-on attribuer cet état de choses, comme le fait Schützenberger, "à des causes internes propres au langage lui-même"? En exprimant nos réticences devant ces propos nous ne pensons pas seulement au dangereux "réalisme" que sous-tend cette dernière thèse mais à des faits beaucoup plus précis. Les deux modèles mathé-

matiques construits par Peters et Ritchie ne forment qu'un certain nombre de propriétés des grammaires transformationnelles et une partie importante des opérations impliquées dans ces mécanismes sont à la fois définies avec suffisamment de précision et de rigueur et d'un usage assez systématique pour mériter une étude abstraite. Apparemment la présence et le rôle des symboles de frontière et des filtres n'ont jamais été supposés aussi important par les linguistes avant que n'en soit faite l'étude mathématique, et pourtant nous avons constaté à quel point ces dispositifs, introduits de façon ad hoc pour résoudre des cas marginaux de dysfonctionnement des grammaires, jouaient un rôle crucial dans le fonctionnement global des grammaires transformationnelles. De même, on pensait a priori, que les effacements étaient incompatibles avec la récursivité⁽¹³⁾ ou que le principe de récupération des effacements entraînait la récursivité, mais ces deux hypothèses, pourtant raisonnable se sont révélées inexactes. Toujours dans le même ordre d'idée, la règle de "Tree-Pruning" (la convention de réduction de Peters et Ritchie) est une opération qu'utilisent tous les transformationnalistes travaillant dans le modèle d'"Aspects" ou dans ses diverses variantes, mais qui pouvait penser que ces effacements "paradigmatiques" jouaient un rôle aussi important, sinon plus important, que les effacements "syntagmatiques" sur les éléments terminaux. On est donc en droit de se demander si la pauvreté mathématique du modèle transformationnel n'est pas plutôt la conséquence d'un manque de dialectique

que la conséquence d'une détermination indépassable de l'objet langage dont la complexité naturelle le mettrait hors de portée de toute mise en forme. Le problème ne se réduit d'ailleurs pas à la seule découverte de propriétés mathématiquement pertinentes des grammaires transformationnelles. S'il est probable que ces dernières ont de nombreuses caractéristiques qui demeurent enfouies sous les détours complexes d'une pratique encore mal éclaircie, d'autres ne semblent attendre que le bon vouloir des mathématiciens pour voir leur statut théorique clairement analysé. Peters et Ritchie (1973b) entrevoient déjà les "contraintes de structure superficielle" de Perlmutter comme restrictions pouvant modifier la nature des langages transformationnels. De la même façon tout dispositif intervenant au cours des dérivations semble un candidat possible pour l'exploration mathématique de la seule capacité générative faible et bien que, faute de compétence, nous ne puissions faire de suggestions précises nous croyons que des procédés tels que les règles de réajustement, par exemple, pourraient jouer un rôle en ce sens⁽¹⁴⁾.

On est donc à même de croire que les quelques études mathématiques des grammaires transformationnelles dont on dispose actuellement n'épuisent pas toutes les propriétés abstraites de ces systèmes et que la richesse des modèles mathématiques qui seront proposés dans le futur sera proportionnelle à la fidélité de la représentation des caractéristiques linguistiques. Mais les obstacles à la construction d'une théorie mathématique des grammaires génératives élabo-

réelles ne résident pas tous dans la difficulté à déterminer des propriétés empiriques mathématiquement significatives. Il n'est pas dit que les catégories des méthodes mathématiques usuellement adoptées par la grammaire formelle suffisent pour mener cette entreprise à son terme. En particulier nous avons vu que les concepts utilisés pour classer les grammaires, l'équivalence forte et l'équivalence faible, forment un canevas qui semble limiter la recherche. Aussi est-ce à une tentative originale de raffinement de cette distinction que nous consacrerons notre dernier chapitre.

NOTES DU CHAPITRE VIII

(1) Conformément aux conventions habituelles nous distinguons la dérivation directe, notée $\overset{*}{\Rightarrow}$, de la dérivation au sens large notée $\Rightarrow$.

(2) Sans entrer dans les détails de réalisation de l'auxiliaire, des affixes verbaux, etc....

(3) Ceci, évidemment, si on accepte l'argumentation de Peters présentée au chapitre VI établissant qu'un automate linéairement borné ne dispose pas d'assez de bande pour accéder aux structures profondes des phrases engendrées par une grammaire transformationnelle.

(4) Ces deux règles de base peuvent alors être appliquées indépendamment pour produire (5). Ces complications inutiles amènent Kuroda à s'écarter du respect strict du principe de Katz et Postal, tout en en respectant l'esprit. Pour lui, les transformations sont un moyen de raffiner le sens.

(5) Nous ne croyons pas déformer les propos de Kuroda en les interprétant ainsi.

/

(6) Les conséquences épistémologiques de ce résultat sont nombreuses et importantes, mentionnons-en deux: 1°) Contrairement à ce qui est souvent affirmé dans la littérature les grammaires génératives ne sont pas des mécanismes neutres par rapport à la production et à la reconnaissance. Le théorème de Peters et Ritchie montre que ces deux opérations peuvent mettre en jeu des procédures essentiellement différentes. 2°) Les procédures d'analyse syntaxique du structuralisme américain n'opérant que sur des éléments attestés ne peuvent atteindre un degré de complexité supérieur à celui des grammaires indépendantes du contexte.

(7) Kimball utilise le même concept mathématique de grammaire

transformationnelle que Peters et Ritchie; sa démonstration procède principalement par la construction d'homomorphismes et par l'exploitation des propriétés des langages de Dyck, ce qui nous la fait qualifier d'algébrique. Peters et Ritchie, eux, construisent une machine de Turing simulant le comportement de la grammaire transformationnelle.

(8) Fonction récursive élémentaire est ici entendue au sens de Kalmar. Voir Yasuhara (1971) p. 119 et sqq. pour une définition précise et l'étude de quelques propriétés de cette classe de fonctions. Informellement la classe des fonctions élémentaires est la plus petite classe contenant les fonction successeur, addition, soustraction limitée et qui est fermée pour la composition, la transformation explicite, l'addition limitée et la multiplication limitée. La classe des fonctions récursives élémentaires est donc une sous-classe de celle des fonctions primitives récursives. C'est même une sous-classe propre. Kleene (1952, p.285) donne l'exemple de la fonction b^a définie par $0^a=1$, $b^a = a^{(b^a)}$ comme exemple de fonction primitive récursive non élémentaire.

(9) Post-cyclique n'est pas le terme approprié. Les grammaires localement filtrantes font plutôt intervenir une "post-cyclicité relative". Celles-ci contrairement aux grammaires non-filtrantes, produisent des structures de surfaces au terme de chaque cycle transformationnel qui peuvent être sujettes à l'application de nouveaux cycles. C'est en ce sens qu'on peut dire que des transformations s'appliquent sur des structures de surfaces.

(10) Les langages prédictiblement énumérables sont les langages finis plus les domaines d'arrivée des fonctions récursives élémentaires injectives, ce qui est le cas de L_2 . Cette caractérisation des langages prédictiblement énumérables est due à Marsh et Ritchie.

(11) Cette propriété, plutôt inattendue se démontre très simplement. Soit L un langage récursivement énumérable non récursif

il existe une grammaire transformationnelle filtrante à base indépendante du contexte engendrant L . Soit G cette grammaire. Remplaçons dans G toutes les occurrences de $\#$ par un symbole quelconque a , ne figurant pas dans le vocabulaire de G . Nous obtenons ainsi un langage L_1 sur $V \cup \{a\}$ engendré par G_1 non filtrante (donc $L_1 \notin \mathcal{C}$). Or $L = L_1 \cap V^*$. V^* est régulier, donc récursif. Puisque la classe des langages récursifs est close sous l'intersection L_1 n'est pas récursif. Voir Peters et Ritchie (1973b, pp. 185-6) pour la démonstration détaillée.

(12) Faut-il voir une simple coïncidence dans le fait qu'à chaque fois que nous avons rencontré un théorème de caractérisation des langages, les classe de ces langages étaient incomparables avec les classes de langages caractérisés par leur grammaire.

(13) A la lumière des résultats exposés dans ce chapitre on peut reconsidérer l'affirmation de Putnam selon laquelle tout ensemble récursivement énumérable peut être engendré par une grammaire transformationnelle, puisque si les grammaires localement filtrantes engendrent certains ensembles non récursifs, elles n'engendrent pas pour autant tous les ensembles récursivement énumérables.

(14) Langendoen (1975) à partir d'une argumentation en partie mathématique déduit que les règles de réajustement sont phonologiques plutôt que syntaxiques. Nous ne voyons pas la portée du problème, seules deux choses étant indiscutables: 1°) Les règles de réajustement portent sur des constituants et des variables de constituant. 2°) Ces règles sont indispensables pour engendrer des phrases grammaticales.

CHAPITRE IX

UNE METHODE TOPOLOGIQUE

DE

CLASSIFICATION DES GRAMMAIRES

de l'informatique

9. L'intérêt porté par les théoriciens des "computer sciences" aux premiers travaux de Chomsky a eu pour conséquence positive de faire avancer très rapidement la connaissance des propriétés mathématiques des grammaires. Cependant cet intérêt pour la théorie des grammaires formelles trouve sa contrepartie dans l'orientation que les mathématiciens lui ont imposée. Les liens étroits unissant la théorie des automates à la théorie des grammaires formelles d'une part, la plus grande familiarité de la théorie des automates d'autre part et la typologie des langages formels induite par la hiérarchie des grammaires enfin, ont conduit les mathématiciens à étudier principalement les propriétés formelles des langages. Le problème général devenant de caractériser algébriquement la classe $\mathcal{L}(\mathcal{A})$ des langages acceptés par une classe d'automates donnée $\mathcal{A}$, ou inversement de rechercher si on peut associer une classe particulière d'automates à une classe de langages satisfaisant certaines propriétés algébriques données. Bref, la théorie des grammaires formelles est devenue théorie des langages formels^(I). Si à cette évolution des intérêts et de la terminologie correspondait une mutation réelle de l'approche mathématique, alors il faudrait se réjouir d'une telle transformation. Elle aurait pour conséquence de rendre envisageable, sinon possible, l'étude des propriétés mathématiques des langues naturelles en les comparant aux langages formels car ces propriétés ne seraient plus celles induites par des grammaires

$\mathcal{L}(\mathcal{A})$

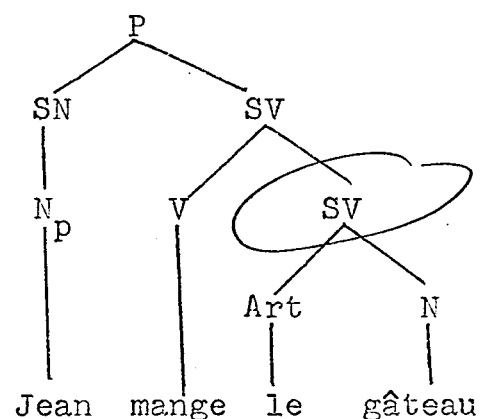
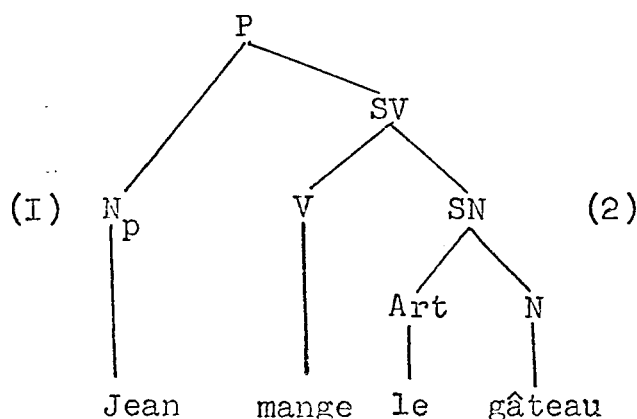
génératives dont le choix demeure toujours en partie arbitraire, mais des propriétés intrinsèques d'ensembles de séquences. La seule propriété mathématique des langages naturels connaissable à ce jour, et encore sous toutes les réserves exprimées au cours des chapitres III, V et VI, étant leur localisation dans une hiérarchie grossière des langages formels⁽²⁾.

Cependant il n'est pas dit que la "théorie des langages formels" soit véritablement une théorie des langages formels. Nous avons observé au chapitre IV qu'une propriété comme l'ambiguïté inhérente ne pouvait être considérée comme une propriété des langages. De même, à la note (5) de ce chapitre, nous avons noté que le théorème de caractérisation des langages indépendants du contexte n'était pas un véritable théorème de caractérisation puisque l'identification du langage (satisfaisant les propriétés du théorème) comme indépendant du contexte ne pouvait être faite que par le recours à sa grammaire. Il en irait de même pour un grand nombre de propriétés algébriques des langages formels comme celles de clôture sous la substitution, l'intersection, etc.... La théorie des langages formels propose donc des propriétés algébriques qui généralement dépendent des grammaires. L'inconvénient est que ces propriétés, tout en étant grammaticales, ne sont pas très intéressantes pour la connaissance mathématique des grammaires elles-mêmes, car elles n'ont trait qu'à une caractéristique, la capacité générative faible, que l'on n'a aucune raison de considérer comme déterminante sinon qu'elle est plus simple à étudier

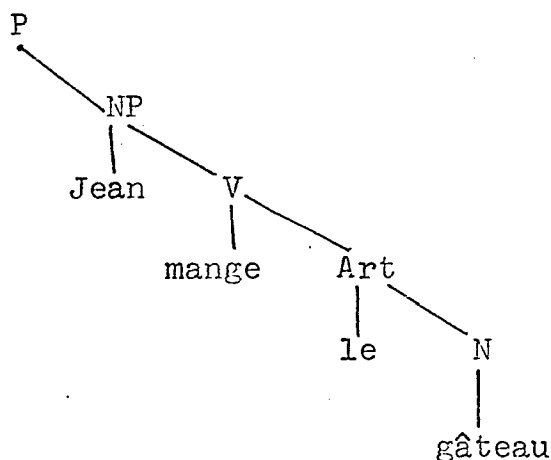
et qu'elle intéresse en premier chef les théoriciens des langages formels qui disposent des outils algébriques nécessaires à son étude. Evidemment on objectera que connaître la grammaire, c'est connaître la nature des objets qu'elle engendre. Mais on pourrait souhaiter dans la perspective d'un renouvellement orienté vers l'application à la grammaire générative, que la grammaire formelle cesse de s'intéresser exclusivement aux propriétés algébriques pour considérer également les propriétés géométriques des grammaires.

9.I. La similitude structurale des grammaires

Comme nous l'avons dit, c'est la distinction entre capacité générative faible et capacité générative forte qui est en partie à l'origine de l'impasse dans laquelle se trouvent les rapports entre grammaire formelle et grammaire générative. Or cette distinction présente d'autres inconvénients de nature plus proprement empirique. Considérons les deux descriptions suivantes de la même phrase "Jean mange le gâteau":



elles sont quasiment identiques. L'unique différence étant que les règles de réécriture qui engendrent le premier arbre pourraient comprendre une production de la forme $P \rightarrow \begin{Bmatrix} \text{SN} \\ \text{N}_p \end{Bmatrix} \text{SV}$ tandis que le second arbre est engendré par des règles où le choix facultatif entre les différentes réalisations du syntagme nominal est introduit après l'expansion obligatoire de P en SN, $P \rightarrow \text{SN SV}$, $\text{SN} \rightarrow \begin{Bmatrix} \text{N}_p \\ \text{Art N} \end{Bmatrix}$. Il n'y a donc là qu'une différence notationnelle, les informations présentées par les deux indicateurs syntagmatiques étant par ailleurs de même nature. On ne pourrait pas en dire autant de la description



si on devait la comparer à l'un des indicateurs syntagmatiques précédents. Or la théorie des grammaires formelles, comme la théorie des grammaires génératives ne propose aucun moyen permettant d'opposer la différence entre (3) et (1) ou (2) à la similitude de (1) et (2).

Pour tenter de résoudre ce problème, Kuroda (1976)⁽³⁾ a développé des méthodes topologiques de classification des grammaires. Ces méthodes conçues pour résoudre le problème d'une

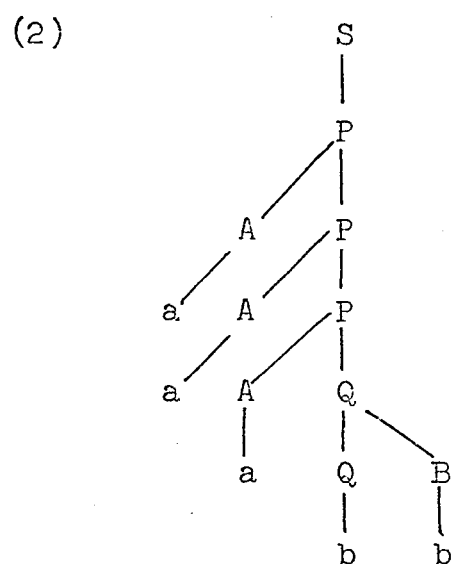
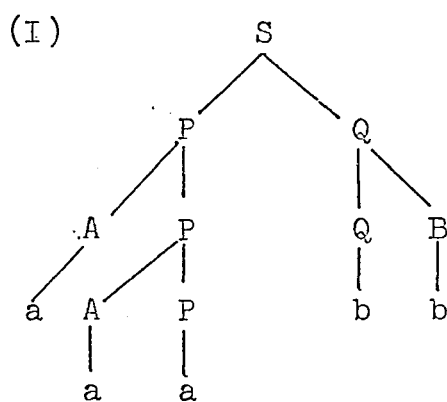
caractérisation de la similitude structurale traitent nécessairement, étant donné les outils mathématiques utilisés et le problème étudié, des propriétés géométriques des indicateurs syntagmatiques. Donc, en traitant ce problème particulier, elle propose une direction d'étude alternative à la théorie des grammaires formelles qui lui permet de sortir ces ornières tracées par la tradition algébrique. Un des intérêts du point de vue adopté par Kuroda est qu'il permet de formuler un certain nombre de problèmes classiques de la grammaire formelle dans des termes entièrement nouveaux. Avant d'examiner plus en détail les conséquences de l'étude topologique des grammaires de constituants, nous exposerons brièvement quelques unes des notions techniques employées dans cette étude ainsi que les principaux résultats démontrés par Kuroda.

9.2. Une méthode topologique de classification des grammaires ⁽⁴⁾

9.2.I. Les grammaires suivantes G_1 et G_2 engendrent toutes deux le langage $a^n b^m$:

G_1 :	$S \longrightarrow PQ$	G_2 :	$S \longrightarrow P$
	$P \longrightarrow AP$		$P \longrightarrow AP$
	$P \longrightarrow a$		$P \longrightarrow AQ$
	$Q \longrightarrow QB$		$Q \longrightarrow QB$
	$Q \longrightarrow b$		$Q \longrightarrow b$
	$A \longrightarrow a$		$A \longrightarrow a$
	$B \longrightarrow b$		$B \longrightarrow b$

Elles attribuent respectivement à la séquence $a^2 b^3$, par exemple, les arbres suivants:



Les grammaires G'_1 et G'_2 engendrent également le langage $a^n b^m$:

G'_1 :

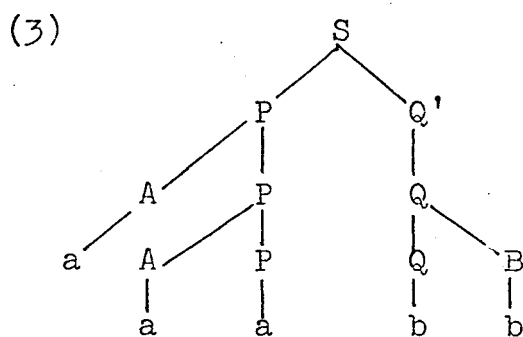
- $S \longrightarrow PQ'$
- $Q' \longrightarrow Q$
- $P \longrightarrow AP$
- $P \longrightarrow a$
- $Q \longrightarrow QB$
- $Q \longrightarrow B$
- $A \longrightarrow a$
- $B \longrightarrow a$

G'_2 :

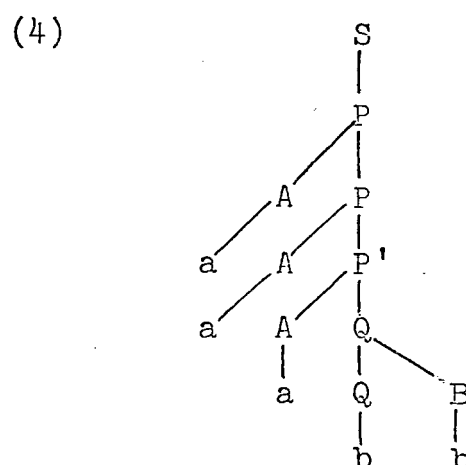
- $S \longrightarrow P$
- $S \longrightarrow P'$
- $P \longrightarrow AP$
- $P \longrightarrow AP'$
- $P' \longrightarrow AQ$
- $Q \longrightarrow QB$
- $Q \longrightarrow b$
- $A \longrightarrow a$
- $B \longrightarrow b$

en attribuant à la séquence a^3b^2 les arbres (3) et (4).

On voit que (I) et (3) d'une part et (2) et (4) d'autre part sont quasiment identiques et qu'on ne peut en dire de même pour (I) et (2) et (3) et (4).



G₁

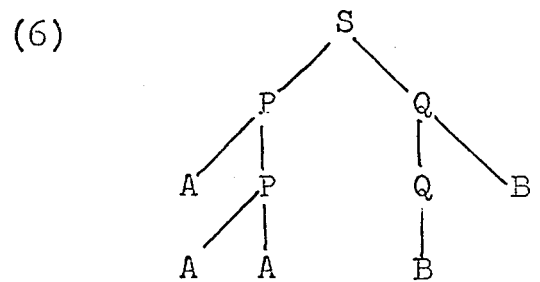
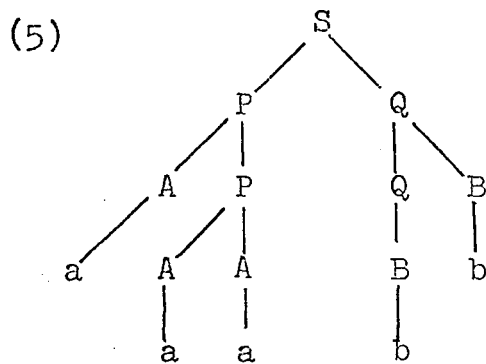


On peut rendre compte de la similarité des arbres de G_1 et de G'_1 et de celle des arbres de G_2 et G'_2 en introduisant la topologie du pré-ordre sur les langages d'arbres K_1, K'_1, K_2, K'_2 engendrés par ces grammaires. Pour définir cette topologie considérons la relation $<_o$ sur K définie de la manière suivante : deux arbres τ et σ vérifient $\tau <_o \sigma$ si et seulement si (τ) est un sous-arbre de (σ) où (τ) et (σ) sont les arbres effeuillés (privés de leurs symboles terminaux) obtenus à partir de τ et de σ . La relation de pré-ordre $<_o$ induit une topologie sur K , la topologie dont les ouverts sont engendrés par les ensembles $V(T) = \{ \tau / \tau \in K \text{ et } T <_o \tau \}$. Soit $\tau_1^{n,m}$ l'élément de K_1 associé à la séquence $a^n b^m$, la base des ouverts pour K est alors $V(\tau_1^{m,n}) = \{ \tau_1^{x,y} / m \leq x \text{ et } n \leq y \}$. Pour K_2 la base de la topologie est formée des ensembles $V(\tau_2^{m,n}) = \{ \tau_2^{\bar{m},y} / n \leq y \}$. La fonction injective f de K_1 dans K_2 définie par $f(\tau_1^{m,n}) = \tau_2^{m,n}$ n'est pas continue car les images par f des plus petits voisinages des éléments de K contiennent strictement les plus petits voisinages des éléments de K .

nages des éléments correspondant de K_2 . Ceci traduit le fait que les voisinages de $\tau_2^{m,n}$ ne contiennent que des arbres dont les branches se ramifient vers la droite. Par contre, g définie de K_1 à K'_1 par $g(\tau) = \tau'$ est continue de même que son inverse g^{-1} ; c'est donc un homéomorphisme. K_2 et K'_2 sont également homéomorphiques par la fonction définie similairement à g .

En première approximation, cette méthode est satisfaisante car elle classe les quatre grammaires en deux classes disjointes de grammaires homéomorphes. De façon intuitive, nous pouvons dire qu'elle regroupe d'un côté les grammaires d'"aspect" indépendant du contexte (G_1 et G'_1) et de l'autre les grammaires d'"aspect" linéaire à droite, puisque pour G_2 et G'_2 , une fois commencé l'engendrement des b , l'engendrement des a n'est plus possible. Cependant cette méthode n'est pas assez générale pour définir une classification intéressante des grammaires. En effet considérons $\bar{G}_1$ obtenue à partir de G_1 en remplaçant $P \longrightarrow a$ et $Q \longrightarrow b$ par les règles $P \longrightarrow A$ et $Q \longrightarrow B$ et $\bar{G}_2$ obtenue à partir de G_2 en remplaçant $Q \longrightarrow b$ par $Q \longrightarrow B$.

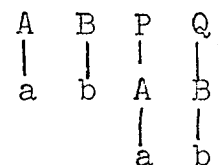
La topologie du préordre pour les langages d'arbre $\bar{K}_1$ et $\bar{K}_2$ associée à ces deux grammaires est alors la topologie discrète puisque tout arbre effeuillé a pour noeuds terminaux des noeuds stériles (c'est à dire des symboles pré-terminaux de la grammaire). Par exemple, l'arbre associé à a^3b^2 dans $\bar{G}$ est (5) et l'arbre effeuillé correspondant est (6)



On voit que $V(\bar{\tau}_1^{3,2})$ engendre l'ouvert $\{\bar{\tau}_1^{3,2}\}$ et qu'il en sera de même pour tout $\bar{\tau}_1^{m,n}$, K_1 et K_2 étant munis chacun de la topologie discrète, ils sont homéomorphes, bien que leurs arbres ne soient pas similaires.

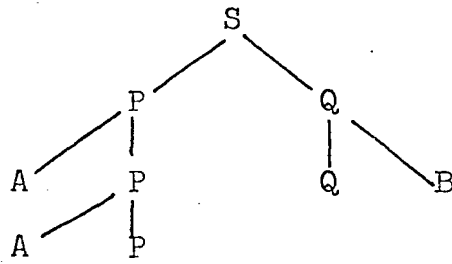
9.2.2. Pour remédier à cette difficulté et obtenir une classification plus satisfaisante, on peut alors associer une classe de topologie à un langage plutôt qu'une topologie. On procède de la manière suivante.

Plutôt que de définir les ouverts à partir des arbres effeuillés, on les définit à partir des arbres émondés. L'émondage d'une branche consiste en l'effacement de tous les noeuds de cette branche, sauf sa racine. L'émondage d'un arbre consiste en l'émondage des branches d'un ensemble d'émondage Π fixé. Par exemple, si l'ensemble d'émondage est Π :



L'arbre (5) devient alors l'arbre (7) $(\bar{\tau}_1^{3,2})_\Pi$.

(7)



Un ensemble d'émondage détermine lui aussi une topologie du préordre sur un langage d'arbres K . La relation de préordre étant définie de la manière suivante: $\tau <_{\Pi} \sigma$ si et seulement si $(\tau)_{\Pi}$ est un sous-arbre de $(\sigma)_{\Pi}$. L'ensemble $V_{\Pi}(\tau) = \{\sigma / \tau <_{\Pi} \sigma\}$ étant le plus petit voisinage de τ dans la topologie $T_{\Pi}(K)$.

Grâce à ce procédé, on peut rendre les langages $L(G_1)$ et $L(\bar{G}_1)$ homéomorphes. En effet, $T_0(\bar{K}_1)$ étant discrète, $f(\tau) = \tau$ n'était pas continue de $T_0(K_1)$ dans $T_0(\bar{K}_1)$ mais on voit immédiatement que f est continue de $T_0(K_1)$ dans $T_{\Pi}(\bar{K}_1)$. Cependant il n'existe pas de choix approprié d'un ensemble d'émondage pour K_2 qui rendrait f continue de K_1 à K_2 . L'intuition de similitude entre les arbres de G_1 et de $\bar{G}_1$ est donc traduite par l'existence d'un ensemble d'émondage pour $\bar{K}_1$ qui le rende homéomorphe à K_1 .

On peut donc définir une classe de topologie pour un langage d'arbre K en faisant varier Π sur tous les ensembles finis de branches de K . Mais alors, lorsqu'on considère la classe de topologie ainsi définie pour un langage d'arbres la question générale suivante se pose. Etant donné deux langages d'arbres

K_1 et K_2 et une application f de K_1 et de K_2 , existe-t-il pour Π_1 quelconque de K_1 un ensemble d'émondage Π_2 de K_2 tel que f soit continue de $T_{\Pi_1}(K_1)$ dans $T_{\Pi_2}(K_2)$? Cette question est compliquée, entre autres parce que pour un langage K il ne suffit pas que $\Pi \subset \Pi'$ pour garantir la continuité de l'identité de $T_{\Pi}(K)$ dans $T_{\Pi'}(K)$. Il est cependant possible de montrer que pour tout Π il existe une extension Π' de Π rendant l'identité continue de $T_{\Pi}(K)$ dans $T_{\Pi'}(K)$. (Kuroda, 1976, Th.I, p.224).

La classe $\mathcal{B}$ des ensembles d'émondage tels que l'identité de $T_{\Pi}(K)$ dans $T_{\Pi'}(K)$ est continue, avec $\Pi \subset \Pi'$, est appelé système inductif d'émondage. Si pour une fonction f de K dans K_2 on peut trouver pour tout Π_1 de K_1 un Π_2 de K_2 tel que f soit continue de $T_{\Pi_1}(K_1)$ dans $T_{\Pi_2}(K_2)$, alors f est dite structurellement fortement continue. La continuité structurellement forte est évidemment une condition très forte et les langages d'arbres la satisfaisant présentent de très grandes similitudes. Mais le fait, que pour un langage d'arbres K et un ensemble d'émondage quelconque Π , on puisse toujours trouver une extension Π' de Π (avec $\Pi \subset \Pi'$) telle que l'identité soit continue de $T_{\Pi}(K)$ dans $T_{\Pi'}(K)$ amène à se poser une nouvelle question : existe-t-il un moyen de définir une classe de topologies ("plus compliquées") telle que l'identité soit continue pour tout $\Pi' \supset \Pi$ sans avoir à chercher une extension convenable Π'' de Π' ? La réponse est positive mais il est inutile, pour les

besoins de notre exposé, d'en donner la formulation qui requièrerait toute une série de définitions préliminaires. Contentons-nous de noter qu'il est possible de définir sur des langages d'arbres des hiérarchies de topologies, 1°) permettant de rendre compte de la similitude des langages d'arbres en général et de la similitude des grammaires en particulier, 2°) évitant les difficultés et les limites de topologies ad hoc comme celle du simple préordre, 3°) évitant d'imposer des contraintes trop fortes sur la similitude des langages d'arbres pour que les hiérarchies de topologies définies sur ceux-ci puissent autoriser la continuité.

Ces remarques étant faites, nous pouvons mentionner, de façon informelle, quelques unes des conséquences les plus significatives de l'étude de Kuroda.

9.3. Un renouvellement de la théorie des grammaires formelles

Signalons d'abord que la majorité des concepts mentionnés précédemment pour les langages d'arbres demeurent valables pour les langages de séquences qui leur sont associés. Pour chaque ensemble d'émondage Π de K , on peut définir la topologie $T_{\Pi}(L)$ sur L comme la topologie la plus fine telle que l'application $\tau \mapsto \tau/\tau/$ (où $\tau/$ désigne la séquence terminale de τ) soit continue de $T_{\Pi}(K)$ dans $T_{\Pi}(L)$. Il est possible par conséquent de donner une classification topologique des langages formels qui soit différente des classifications algébriques habituelles et qui ne découlent pas de la forme des

productions des grammaires. Cette caractérisation des langages formels ne se distingue pas des classifications algébriques seulement par le fait qu'elles soient indépendantes des grammaires mais elles se traduisent également par une différence notable de leurs propriétés. Par exemple, la substitution, pour les langages indépendants du contexte, ne préserve pas les homéomorphismes structuraux.

Ensuite, les hiérarchies de grammaires de Chomsky ne correspondent pas à la classification des grammaires structurellement homéomorphes. Ainsi le langage en miroir $\{x\tilde{x}\}$ (où x est engendré par une grammaire régulière) est homéomorphe au langage $\{x\}$ par l'application $x\tilde{x} \longrightarrow x$. D'une façon plus générale se présente la question des modifications que l'on peut opérer sur les règles de grammaire tout en laissant leur structure topologique invariée (donc tout en conservant quasiment l'équivalence forte au sens de Chomsky). Il vaut la peine d'observer que cette question se pose également pour les règles transformationnelles. La méthode topologique de classification des grammaires permet donc, en un certain sens, d'unifier deux modes de dérivation qui dans le point de vue classique était complètement distincts.

Enfin notons que même la notion d'ambiguïté, dans le cadre topologique, reçoit une formulation nouvelle. Le degré d'ambiguïté d'une grammaire formée par l'union de deux grammaires est égale à la somme des degrés d'ambiguïté de chacune

des deux grammaires unies si ces trois grammaires sont faiblement équivalentes. Mais "dans ce cas, l'accroissement du degré d'ambiguïté algébrique ne provoque aucun changement dans la structure topologique du langage de séquences; l'ambiguïté introduite par la grammaire "union" n'est pas structurellement "essentielle". Toutefois en général l'ambiguïté peut être "essentielle". Par exemple, une séquence peut être associée à différents arbres de manière telle que la fonction canonique $t \rightarrow /t/$ n'est localement homéomorphe à aucun de ces arbres". (Kuroda, 1976, pp. 317-18).

Si le problème de l'ambiguïté peut être abordé en des termes nouveaux, c'est bien parce qu'ici l'ambiguïté d'une séquence d'un langage formel a un sens nouveau. Elle dépend véritablement de la façon dont un arbre structure une séquence terminale et non plus de la façon dont une grammaire la dérive.

Nous ne pouvons évidemment pas préjuger de l'usage que pourra faire la linguistique mathématique de cette nouvelle façon de concevoir la grammaire formelle. Cependant nous ferons observer que toutes les difficultés liées à l'usage linguistique de la théorie mathématique des grammaires que nous avons pu relever au cours des chapitres précédents disparaissent dans ce cadre nouveau. C'est là une conséquence directe de l'abandon de la capacité générative faible comme objet privilégié d'étude. Cette réorientation ne signifie pas pour

autant, comme nous venons de le voir, que les langages sont oubliés. Au contraire, en les délaissant provisoirement, on se donne les moyens de mieux les connaître, ou plutôt de les connaître pour eux-mêmes. En fait, bien plus que l'introduction de méthodes topologiques, nous pensons que la véritable originalité de l'étude de Kuroda consiste en l'usage de la notion de langage d'arbres. Cet objet intermédiaire, qui est différent d'une grammaire (un langage d'arbres n'est pas nécessairement engendré par une grammaire) et qui est plus qu'un langage permet enfin à la théorie mathématique des grammaires de s'intéresser à d'autres problèmes que celui de l'appartenance d'une séquence à l'ensemble hypothétique des phrases grammaticales. Permettre d'étudier les propriétés mathématiques des grammaires sans référence à la "grammaticalité" n'est pas le moindre de ses avantages.

NOTES DU CHAPITRE IX

- (1) Comme signe de cette évolution on peut noter qu'à l'heure actuelle, le terme "langage formel" apparaît presque exclusivement dans la littérature.
- (2) Plus exactement on devrait dire que la seule chose que l'on sache c'est où ne se trouvent pas les langues naturelles.
- (3) Cette section est un résumé succinct de l'étude de Kuroda. Pour plus de précisions, de détails et de clarté, on se réfèrera à cet article.
- (4) Un langage d'arbres est un ensemble d'arbres dont les racines ont même "nom".

CHAPITRE X

CONCLUSION

10. Au cours de l'introduction, nous avons rappelé le point de vue de Hockett sur les rapports de la linguistique et des mathématiques. L'auteur y dissociait nettement la démarche empirique de la démarche mathématique. La dernière ne doit et ne peut apparaître qu'une fois la première menée à terme et se limite par conséquent à une mise en forme élégante des conclusions de l'investigation empirique. D'ailleurs, pour Hockett, l'unique nécessité de la mathématisation semble être d'enjoliver les descriptions linguistiques d'un symbolisme rutilant. Nous ne répèterons pas qu'une telle position nous semble indéfendable mais ce qui nous paraît mériter quelques éclaircissements, ce sont les termes dans lesquels Hockett pose le problème général des rapports entre mathématique et linguistique, à savoir: peut-on ou non appliquer les mathématiques à la linguistique? Selon nous, qu'il s'agisse de linguistique ou d'autre chose, la question de l'opportunité du recours aux mathématiques ne peut être posée dans les termes de la possibilité de leur application.

L'origine de cette problématique s'explique facilement. Les sciences humaines, d'une part, sont des sciences jeunes qui rencontrent d'énormes difficultés dans leurs essais de construction de méthodes objectives. D'autre part, ce faisant, elles sont constamment préoccupées par des considérations épistémologiques. Aussi le succès avec lequel les sciences exactes ont employé les mathématiques a-t-il pour elles valeur d'exemple et de référence. Convaincus que la mathématisation est la voie

royale pour accéder à une connaissance digne de ce nom, les sciences humaines sont perpétuellement tentées d'"appliquer" les mathématiques à leurs domaines au lieu de s'interroger sur la possibilité de les appliquer. Or ce qui justifie le recours aux mathématiques c'est uniquement la nature des problèmes que peuvent lui poser les sciences empiriques. Dans certains cas, ce recours peut même être nécessaire. Par exemple le problème de la décidabilité de la grammaire transformationnelle exposé au cours du chapitre VI ne pouvait être traité que mathématiquement. Ceci évidemment parce que la décidabilité est une propriété mathématique, mais aussi, d'une façon plus générale, parce qu'il s'agit d'un problème où figurent des raisonnements faisant intervenir l'infini. On peut d'ailleurs affirmer que cette dernière caractéristique est une condition sine qua non de l'usage des mathématiques. Partant il ne faut plus s'étonner du fait que la théorie des grammaires génératives doive être nécessairement complétée d'une théorie mathématique des grammaires. En effet, dès l'instant où se pose le problème de la justification de la théorie des grammaires génératives surviennent des difficultés épistémologiques et empiriques liées à la définition du langage comme ensemble infini d'expressions. C'est d'ailleurs, nous l'avons vu, en analysant le théorème d'inadéquation des grammaires régulières, dans ce cadre de réflexion qu'est née la théorie des grammaires formelles.

Deux conditions préliminaires doivent donc être satisfai-

tes pour justifier l'usage des mathématiques : 1°) que la recherche se heurte à un obstacle, 2°) que la nature de cet obstacle relève d'un traitement mathématique.

La première condition est fréquemment négligée. Nous y verrons la raison de l'intérêt secondaire de tant de travaux, réalisés pour la plupart dans le cadre de la linguistique structuraliste, où on a pu "appliquer" l'algèbre de Boole à la phonologie, la théorie des espaces métriques à la morphologie ou encore la théorie des ensembles à la syntaxe^(I). L'absence de motivations concrètes conduit nécessairement ces études à des impasses, puisqu'elles se singularisent par le fait qu'elles n'ont pas de véritables objets d'étude. En ce sens, bien qu'à un degré moindre, l'application de la théorie de l'information présente les mêmes limites. Découvrir le taux de redondance moyen des suites de phonèmes d'une langue naturelle procède d'une démarche certainement plus fine que celles à peine citées, mais faute d'être inscrite dans une perspective de recherche nettement définie, cette découverte reste sans lendemain. Sur ce dernier point, la grammaire formelle se distingue encore une fois des "applications" des mathématiques à la linguistique, car à la pertinence de son emploi est associé un effet qu'on ne peut raisonnablement juger comme secondaire: chaque fois que la théorie des grammaires formelles permet de résoudre un problème posé par la grammaire générative, elle en soulève un autre. Par exemple, une fois démontré l'inadéquation des grammaires à états finis, la même question se pose pour les

grammaires indépendantes du contexte. Ou encore, une fois découvertes les conditions garantissant la décidabilité des grammaires transformationnelles, survient la question de la localisation des langages transformationnels dans la hiérarchie des langages formels. Retenir ce critère de la multiplication et de la ramification des problèmes comme le symptôme d'une mathématisation opportune semblera sans doute peu rigoureux. Il est vrai qu'il relève de l'intuition mais on admettra qu'il a pour lui tout le poids de la confirmation qu'apporterait une statistique sur l'usage des mathématiques dans l'ensemble des sciences exactes. La seconde condition préliminaire justifiant l'usage des mathématiques est plus complexe à analyser. Relativement à l'échantillon réduit des problèmes que nous avons exposés dans les pages précédentes, on peut distinguer trois situations.

a) Le problème rencontré est, au "codage disciplinaire" près, identique à un problème déjà traité par les mathématiques. C'est le cas de la décidabilité des grammaires transformationnelles. (A un degré différent, c'est aussi celui de la reformulation de la loi de Zipf par Mandelbrot qui se trouve être analogue dans sa forme à la loi des gaz parfaits.)

b) La nature du problème le prédispose à être intégré dans une théorie mathématique déjà constituée et disposant des méthodes nécessaires à son traitement. C'est le cas de l'étude de la similitude structurale des grammaires de constituants. Celle-ci pouvant être menée à terme parcequ'elle s'appuie sur

des concepts, tel que la connexité d'un graphe, et de structures, telle que la topologie du pré-ordre, dont les propriétés mathématiques sont bien connues.

c) L'originalité du problème est telle qu'on ne dispose d'aucune méthode mathématique pour l'aborder. C'est le cas de la démonstration de l'inadéquation descriptive des grammaires à états finis. Ce qui prédisposait le problème à recevoir un traitement mathématique, c'est qu'il impliquait, comme nous l'avons dit plus haut, des raisonnements où figure l'infini.

Malgré leur différence, ces trois types de situations ont en commun de proposer un matériau conceptuel qui les désigne spécialement comme relevant d'une investigation mathématique. Evidemment les deux conditions que nous proposons pour sélectionner les questions linguistiques candidates à un traitement mathématique ne sont que des conditions nécessaires car une fois satisfaits ces réquisits préliminaires, tout reste à faire. En particulier dans le cas c) la construction d'outils mathématiques originaux est une entreprise qui n'est pas dépourvue de risques comme nous l'avons vu au cours du chapitre V.

Sans doute s'étonnera-t-on de la banalité de cette première conclusion sur la caractérisation des problèmes linguistiques mathématisables. Il arrive pourtant qu'on perde de vue, en "appliquant" les mathématiques, ces deux conditions. C'est pourquoi nous avons cherché à en repréciser le sens à travers l'analyse détaillée de quelques exemples.

Nous proposerons par contre une conclusion moins classique pour expliquer l'origine du succès limité de l'étude des propriétés mathématiques des grammaires. Car il faut bien parler de limite quand on songe au nombre restreint de problèmes linguistiques traités par la grammaire formelle. Généralement les mathématiciens et les épistémologues sont d'accord pour attribuer la responsabilité de telles limites aux insuffisances des sciences empiriques. Les mathématiques, selon ce point de vue, seraient douées d'une neutralité de contenu les dégageant de toute responsabilité quant à l'inutilité de leur usage. Et il est entendu que si la théorie mathématique d'un phénomène linguistique ne procure aucun résultat intéressant, ce ne peut être que la conséquence d'une formulation impropre des questions linguistiques.

Nous voudrions nuancer cette interprétation en répétant que si la connaissance mathématique des grammaires demeure aujourd'hui encore superficielle ce n'est pas nécessairement parce que leurs propriétés sont soit triviales, soit inintéressantes. Au cours du chapitre IV, nous avons montré que l'incompatibilité entre la théorie mathématique de l'ambiguïté et l'ambiguïté linguistique provenait en partie de l'équivocité de la notion de langage formel ambigu. De même l'analyse des théorèmes de Peters et Ritchie sur la décidabilité des grammaires transformationnelles révèle quels a priori guident implicitement l'étude mathématique des grammaires lorsqu'elle se contente de considérer les restrictions sur la forme des pro-

ductions comme le seul moyen de contrôle de la capacité générative faible. Cette carence de la grammaire formelle n'est qu'une manifestation particulière d'un problème que nous avons rencontré à maintes reprises. Il s'agit de la réduction des propriétés grammaticales aux propriétés des langages. La capacité générative faible, répétons-le, n'est pas la caractéristique essentielle d'une grammaire formelle. La confiance des spécialistes en cette assimilation est démentie par les résultats discutés aux chapitres IV et VIII (exemples des langages bornés et des langages transformationnels non filtrés). De plus les perspectives ouvertes par l'étude de Kuroda confirment tout l'intérêt qu'il peut y avoir à abandonner le point de vue exclusivement "langagier" si on veut entreprendre une étude mathématique des grammaires qui tienne compte des aspects linguistiquement significatifs.

Evidemment, cette dernière remarque dépend de ce que l'on veut dire par "linguistiquement significatif". De l'ensemble des questions dont nous avons discuté dans cette thèse il ressort que par "linguistiquement significatif" nous n'entendons pas, comme nous l'expliquons dans l'introduction, significatif pour le langage mais plutôt significatif pour les méthodes employées dans la description du langage. Cette distinction, qui justifie selon nous l'étude mathématique des grammaires génératives, pourra sembler trop générale. En particulier on peut se demander si les discussions sur les problèmes de décidabilité n'évacuent pas toute référence au contenu linguistique. En dehors

des arguments que nous avons proposés pour montrer la pertinence de telles études (le contrôle des notions empiriques) nous ajouterons que même au niveau de leur traitement mathématique ces problèmes conservent encore des aspects propres à leurs origines linguistiques. Autrement dit, exprimer une théorie grammaticale par le formalisme des machines de Turing, par exemple, ne lui fait pas perdre nécessairement son contenu. Ce phénomène est analogue à celui que Wang (1963) nous fait remarquer pour les mathématiques. Lorsqu'on représente différentes théories mathématiques dans le calcul des prédicats, ces théories semblent perdre leur contenu, mais on retrouve toujours celui-ci au niveau des méthodes de démonstration qui continuent d'exprimer la spécificité de chacune de ces théories.

Il reste peut-être à démontrer l'actualité d'un débat sur les rapports de la grammaire formelle et de la grammaire générative puisqu'un spécialiste comme Schützenberger (1978), par exemple, fait remonter leurs funérailles aux articles de Chomsky du "Handbook of Mathematical Psychology" et que les résultats les plus récents discutés dans ces pages ont près de dix ans. Nous pensons quant à nous que pour la grammaire générative, le recours à la mathématisation n'est pas facultatif. Nous en avons une confirmation récente avec les théories "rétro" de Gazdar {1981} qui propose de décrire tous les phénomènes linguistiques ayant motivé une analyse transformationnelle à l'aide de grammaires indépendantes du contexte {2}. Quand on

sait que pratiquement toutes les justifications avancées par Gazdar pour ce retour au passé sont d'ordre mathématique, on reconnaîtra, qu'en attendant de revenir aux grammaires à états finis, le travail critique de la grammaire formelle a encore quelques beaux jours devant soi. Le temps que les grammaires génératives disparaissent de la surface du globe...

c'est fait - Etienne Bickel,

NOTES DU CHAPITRE X

- (1) Il ne s'agit pas de fiction. Voir Marcus (1967), Brainerd (1971), Cooper (1963). D'ailleurs tout est permutable dans ces applications des mathématiques et on pourrait tout aussi bien (lorsque cela n'a pas été fait) appliquer la théorie des espaces métriques à la phonologie, etc...
- (2) Chomsky a actuellement une position presque analogue puisque le composant transformationnel est réduit à une seule transformation de mouvement non effaçante. Mais évidemment, contrairement à Gazdar, son intention explicite n'est pas de renier toutes les théories transformationnelles qu'il a pu proposer par le passé.

REFERENCES

- ARBIB, M. (1969) Theories of Abstract Automata . Englewood Cliffs, N.J. : Prentice-Hall
- BAR-HILLEL, Y. (1964) Language and Information. Selected Essays on their Theory and Application. Reading, Mass. : Addison-Wesley Publ. Co et Jerusalem, Israel: The Jerusalem Academic Press Ltd.
- BAR-HILLEL, Y., PERLES, M. et SHAMIR, E. (1961) "On formal properties of simple phrase structure grammars". Zeitschrift für Sprachwissenschaft und Kommunikationsforschung 14. Rééd. in Bar-Hillel (1964), pp.116-150.
- BAR-HILLEL, Y. et SHAMIR, E. (1960) "Finite state languages: formal representation and adequacy problems". Bulletin of the Research Council of Israel . Vol. 8F, N°3. Rééd. in Bar-Hillel (1964) pp.97-115.
- BENVENISTE, E. (1969) "Mécanismes de Transposition". Cahiers Ferdinand de Saussure, Genève, Droz, 25 (Mélanges H. Frei)
- BRAINERD, B. (1971) Introduction to the Mathematics of the Language Study, Elsevier. New-York.
- BRØNDAL, V. (1948) Les Parties du Discours , Copenhagen, Munksgaard.
- CHOMSKY, N. (1956) "Three models for the description of language". IRE Transactions on Information Theory IT-2. Trad. in Langages N°9 (1968)
- CHOMSKY, N. (1957) Syntactic Structures . La Haye : Mouton Trad. franç. Structures Syntaxiques (1968), Le Seuil.

- CHOMSKY, N. (1959) "On certain formal properties of grammars",
Information and Control , 2:2. pp.137-167.
- CHOMSKY, N. (1961) "On the notion " rule of grammar"" in
Jakobson (1961) Structure of Language and its Mathematical
Aspects , Providence, R.I.
- CHOMSKY, N. (1962) "Context-free grammars and pushdown storage",
Quart. Prog. Dept. N°65. M.I.T. Res. Lab. Elect. pp.187-194
- CHOMSKY, N. (1963) "Formal Properties of Grammars", Handbook
of Mathematical Psychology , Vol. 2 . Luce, Bush et Galanter
Eds. New-York: Wiley.
- CHOMSKY, N. (1964) Current Issues in Linguistic Theory , La
Haye, Mouton.
- CHOMSKY, N. (1965) Aspects of the Theory of Syntax , Cambridge,
Mass. The MIT Press. Trad. franç. Aspects de la Theory de
la Syntaxe (1971) Le Seuil .
- CHOMSKY, N. (1970) "Remarks on Nominalizations" in Readings in
English Transformational Grammar , R. Jacobs et P. Rosen-
baum, Eds. New-York, Ginn .
- CHOMSKY, N. (1975) The Logical Structure of Linguistic Theory
New-York, Plenum.
- CHOMSKY, N. (1979a) "Knowledge of Grammar" Conférence prononcée
à la Scuola Normale Superiore de Pise. Scuola Normale Supe-
riore de Pise;
- CHOMSKY, N. (1979b) "On Binding", ronéotypé à paraître dans
Linguistic Inquiry
- CHOMSKY, N. et MILLER, G. (1958) "Finite States Languages"
Inf. and Cont., 1:2, pp.91-112
- CHOMSKY, N. et SCHUTZENBERGER, M.P. (1963) "The algebraic theory
of context-free languages", Computer Programming and Formal
Systems, North Holland, Amsterdam.

COBHAM, A. (1965) "The intrinsic computational difficulty of functions", Logic, Methodology and Philosophy of Science. Proceeding of the 1964 Internat. Congr. North-Holland, Amsterdam, pp.24-30.

COOPER, W. (1963) Set Theory and Syntactic Description , Mouton La Haye.

DE CORNULIER, B. (1978) "Syllabe et suite de phonèmes en phonologie du Français" in Etudes de Phonologie Française, Editions du CNRS. pp. 31-69.

? *DAVUST*
DEKKER, J.C.E. (1953) "The constructivity of maximal dual ideals in certain boolean algebras ", Pacific Journal of Mathematics N°3, pp.73-101.

DE SAUSSURE, F. (1916) Cours de Linguistique Générale , Payot.

GABBAY, D. et MORAVCSIK, J. (1974) "Branching quantifiers English and Montague grammar", Theoretical Linguistics, Vol.I, N°1/2 pp.139-157.

GAZDAR, G. "Phrase Structure Grammar", à paraître dans G.K. Pullum & P. Jacobson, On the Nature of Syntactic Representation

GINSBURG, S. (1966) The Mathematical Theory of Context Free Languages, McGraw Hill, New-York.

GINSBURG, S. et PARTEE, B. (1969) "A mathematical model of transformational grammars", Inf. and Cont. XV, pp.297-334 .

GOLD, E. (1967) "Language identification in the limit", Inf. and Control 10, pp.441-74.

GRANGER, G. (1968) Essai d'une Philosophie du Style, A. Colin.

GROSS, M. (1972) Mathematical Models in Linguistics , Prentice Hall, Englewood Cliffs. N.J.

GROSS, M. (1975) Méthodes en Syntaxe , Hermann.

GROSS, M. et LENTIN, A. (1970) Notions sur les Grammaires Formelles, Gauthier-Villars. 2ème édition.

GRZEGORCZYK, A. (1953) "Some classes of recursive functions" Rozprawy Matematyczne, IV. Varsovie.

GUENTHNER, F. et HOEPELMAN, J.P. (1975) "A note on the representation of "Branching Quantifiers", Theoretical Linguistics, Vol.II, pp.285-89.

GUIBERT, J. (1974) La Parole, Compréhension et Synthèse par les Ordinateurs, PUF.

HAMBURGER, H. et WEXLER, N. (1973) "Identifiability of a class of transformational grammars", Approaches to Natural Language Hintikka et al., pp.153-166. Dordrecht, Reidel.

X HARRIS, Z. (1957) "Co-occurrence and transformations in linguistic structure", Language 33, pp.283-340.

HARROP, R. (1961) "On the recursivity of finite sets", Zeitschr. Math. Logik und Grundlagen d. Math. Vol. 7, pp.136-140.

HINTIKKA, J. (1974) "Quantifiers vs. Quantification Theory", Linguistic Inquiry, Vol.V, pp.153-178.

HOCKETT, C. (1954) "Two models of grammatical description", Linguistics Today, Word, 10, pp.210-33.

HOCKETT, C. (1967) Language, Mathematics and Linguistics, Mouton La Haye.

HOCKETT, C. (1968) The State of the Art, Mouton, La Haye.

HOPCROFT, J.E. et ULLMAN, J.D. (1969) Formal Languages and Their Relations to Automata. Reading. Mass : Addison-Wesley.

JACKENDOFF, R. (1974) Introduction to the $\bar{X}$ Convention, Mimeograph, Indiana University Club, Bloomington Ind.

JANSSEN, T. KOK, G. MEERTENS, L. (1976) "On restrictions of transformational grammars reducing the generative power", Foundations of Language, Vol. 12, pp. 111-118.

Josli + Levy
KATZ, J. et POSTAL, P. (1964) An Integrated Theory of Linguistic Descriptions, Cambridge, MIT Press.

KIMBALL, J. (1967) "Predicate definable over transformational derivations by intersection with regular languages", Inf. and Contr. XI, pp. 177-95.

KLEENE, S. (1952) Introduction to Metamathematics, Amsterdam, Van Nostrand.

KURODA, S. Y. (1969) "On attachment transformations", Modern Studies in English : Readings in Transformational Grammar, S. A. Schane et D. Reibel Eds. Englewood Cliffs. N. J.

KURODA, S. Y. (1973) "Généralisation de la notion d'équivalence des grammaires - Une méthode topologique", The Formal Analysis of Natural Languages. M. Gross, M. Halle et M. P. Schützenberger Eds. La Haye, Mouton, pp. 362-371.

KURODA, S. Y. (1976) "A topological study of phrase-structure languages", Inf. and Cont. 30:4, pp. 307-379

KURYLOWICZ, J. (1948) "Contribution à la théorie de la syllabe" Biuletyn Polskiego Towarzystwa Językoznawczego, VIII, pp. 80-114. Rééd. dans Essais Linguistiques (1973) Wilhelm Fink Verlag, Munich.

LAKOFF, G. (1970) "Linguistics and natural logic", Synthese Vol. 22, N° 1/2. Trad. franç. Linguistique et Logique Naturelle (1976) Klinksiack.

LANDWEBER, P. (1963) "Three theorems on phrase-structure grammars of type I", Inf. and Cont. 6:2, pp. 131-136.

- LANGENDOEN, D.T. (1975) "Finite-state parsing of phrase-structure languages and the status of readjustment rules in grammar", Linguistic Inquiry , 6, pp.533-54.
- LANGENDOEN, D.T. (1977) "Type-3 and Type-2 grammars", Amsterdam Studies in the Theory and History of Linguistic Science, Vol.II, Amsterdam, Benjamin. pp.159-171.
- LAPOINTE, S.G. (1977) "Recursiveness and deletion", Linguistic Analysis, Vol.3:3, pp.227-263.
- LEVELT, W. (1974) Formal Grammars in Linguistics and Psycholinguistics , Vol.2, La Haye , Mouton
- MARCUS, S. (1967) Introduction à la Linguistique Mathématique, Dunod.
- MILLER, G.A. (1962) "Some psychological studies of grammar", American Psychologist, 17, pp.748-762.
- MILLER, G.A. et CHOMSKY, N. (1963) "Finitary models of language users", Handbook of Mathematical Psychology , Vol.2, LUCE et al. eds. New-York, Wiley.
- MONTAGUE, R. (1974) Formal Philosophy - Selected Papers , Ed. et intr. R.H. Thomason. New-Haven, Yale Univ. Press.
- NAUR, P. (1960) Report on ALGOL-60, McGraw Hill, New-York.
- NERODE, A. (1958) "Linear automaton transformations", Proc. Amer. Math. Soc. , 9. pp.541-544.
- PARIKH, R.J. (1961) "Language generating devices", Quart. Prog. Rept., 60, MIT Res. Lab. Elect. pp.199-212.
- PERLMUTTER, D. (1970) Deep and Surface Structure Constraints in Syntax, New-York, Holt, Rinehart et Winston.
- PETERS, S. (1973) "On restricting deletion transformations" in The Formal Analysis of Language , M. Gross et al. eds. La Haye, Mouton , pp.372-384.

- PETERS,S. et RITCHIE,W. (1969) "Context sensitive immediate constituent analysis - Context-free languages revisited", Proc.of the 9th. ACM Symposium on Theory of Computing, I-8. Rééd. in Mathematical Systems Theory, 6 (1973) pp.324-333.
- PETERS,S.et RITCHIE,R. (1971) "On restricting the base component of transformational grammars" Inf. and Cont. 18, pp.483-501.
- PETERS,S.et RITCHIE,R (1973a) "On the generative power of transformational grammars. Inf. Sciences ,6, pp.49-83.
- PETERS,S. et RITCHIE,R. (1973b) "Non Filtering and local filtering transformational grammars. in Approaches to Natural Language , J. Hintikka et al. Eds.Dordrecht, Reidel.pp.180-93.
- POSTAL, P.M. (1964a) Constituent Structure. A Study of Contemporary Models of Syntactic Description (= Supplément à l'International Journal of American Linguistics, Vol.30), Bloomington, Indiana, et La Haye, Mouton.
- POSTAL,P.M. (1964b) "Limitations of Phrase Structure Grammars". in The Structure of Language : Readings in the Philosophy of Language , Ed. par J.A. Fodor et J.J. Katz, Englewood Cliffs, N.J. , Prentice-Hall.pp. 137-51.
- PUTNAM, H. (1961) "Some issues in the theory of grammar." In The Structure of Language and its Mathematical Aspects , R. Jakobson, Ed. Providence, Rééd. in Harman, G.:On Noam Chomsky: Critical Essays (1974) Anchor Books, pp.80-103. New-York.
- PUTNAM, H. (1974) "On the corroboration of theories" in The Philosophy of Karl Popper , Ed. P.Schlipp, La Salle. pp. 221-240.
- QUINE, W. (1953) From a Logical Point of View : 9 logico-philosophical essays, Cambridge , Harvard University Press.
- QUINE,W. (1974) The Philosophy of Logic, Englewood Cliffs, N.J. Prentice Hall.

REICH, P. (1969) "The finiteness of natural language", Language 45, pp. 831-843.

RITCHIE, R. (1963) "Classes of Predictably Functions" in Transactions of the American Mathematical Society, 106, pp. 139-176.

ROSS, J. R. (1967) Constraints on Variables in Syntax, Thèse de Doctorat, MIT. Indiana Linguistic Club.

RUWET, N. (1977) Introduction à la Grammaire Générative, Plon.

SCHUTZENBERGER, M. P. (1978) "Mathématiques et Linguistique", Séminaire sur les fondements des sciences, Cahiers de l'Université Louis Pasteur, Strasbourg, N°92, pp. 25-34.

SHANNON, C. ^{Carde} et WEAVER, W. (1949) The Mathematical Theory of Communication, University of Illinois Press, Urbana.

STANLEY, R. J. ^{Richard} Finite state representations of context-free languages³, M.I.T. Res. Lab. Electron. Quart. Progr. Rept., 76 pp. 276-279

WANG, H. (1963) "Mechanical Mathematics and Inferential Analysis" in Computer Programming and Formal Systems, North Holland, Amsterdam, pp. 1-24.

^{memorables} WELLS, R. ^{Wells} 1947⁴ Immediate Constituents⁴, Language 23, pp. 81-117.

WEXLER N. et HAMBURGER, H. (1973) "On the insufficiency of surface data for the learning of transformational languages", in Approaches to Natural Language, J. Hintikka et al. eds. Dordrecht, Reidel. pp. 167-179.

YASUHARA, A. (1971) Recursive Function Theory and its Logic New-York, Academic Press.

TABLE DES MATIERES

I. PROBLEMES ET LIMITES DE L'APPLICATION DE LA THEORIE DES GRAMMAIRES FORMELLES	I
I. INTRODUCTION	2
I.1. Mathématisation et inertie symbolique des langages humains	4
I.2. Le champ de la linguistique mathématique	10
I.3. Particularités de la grammaire formelle	17
Notes du chapitre I	22
II. UN CRITERE DISTINGUANT LA GRAMMAIRE FORMELLE DE LA GRAMMAIRE GENERATIVE.....	25
2.1. La nature mathématique de la grammaire génération.....	28
2.2. L'abstraction de la grammaire générative	42
2.3. La grammaire formelle est une syntaxe pure	46
Notes du chapitre II	50
III. LES FONCTIONS APPARENTES DE LA GRAMMAIRE FORMELLE ...	53
3.1. Grammaire et système formel	55
3.2. La théorie des grammaires formelles comme ins- trument de sélection des grammaires génératives	63
Notes du chapitre III	69
IV. AMBIGUITE MATHEMATIQUE ET AMBIGUITE LINGUISTIQUE	70
4.1. Grammaires ambiguës et langages ambigus	72
4.2. Le traitement mathématique de l'ambiguïté	82
4.3. Les difficultés méthodologiques d'une solution théorique	90
Notes du chapitre IV	98

II. LA THEORIE DES GRAMMAIRES FORMELLES COMME INSTRUMENT DE CONTROLE DE LA THEORIE DES GRAMMAIRES GENERATIVES	101
V. L'INADEQUATION DES GRAMMAIRES A ETATS FINIS.....	102
5.1. Sens et portée de la notion de "grammaticalité".	108
5.2. De la mise en forme d'une intuition à sa démonstration	117
Notes du chapitre III	130
VI. LA DECIDABILITE DES GRAMMAIRES TRANSFORMATIONNELLES ET LA CONJECTURE DE PUTNAM.....	133
6.1. La conjecture de Putnam	136
6.2. Le principe de récupération des effacements	143
6.3. Propriétés globales et propriétés locales	155
Notes du chapitre VI	171
VII. UN THEOREME DE LIMITATION POUR LE PROGRAMME DE CHOMSKY	174
7.1. Le programme de Chomsky	176
7.2. Un théorème de limitation	184
Notes du chapitre VII	189
III. PERSPECTIVES POUR LA THEORIE MATHEMATIQUE DES GRAMMAIRES	190
VIII. VERS UNE THEORIE MATHEMATIQUE DES LANGAGES TRANSFORMATIONNELS	191
8.1. Sur l'existence des langages transformationnels	193
8.2. Propriétés remarquables des grammaires locale- ment filtrantes.....	206
8.3. Conséquences méthodologiques	210
Notes du chapitre VIII	217

IX.	UNE METHODE TOPOLOGIQUE DE CLASSIFICATION DES GRAMMAIRES	220
	9.1. Similitude structurale des grammaires	223
	9.2. Une méthode topologique de classification des grammaires	225
	9.3. Un renouvellement de la théorie des grammaires formelles	232
	Notes du chapitre IX	236
X.	CONCLUSION	237
	Notes du chapitre X	247
	REFERENCES	248
	TABLE DES MATIERES	256



