



**HAL**  
open science

## Simplicités, empirisme et interprétation psychologique dans la grammaire générative.

Gabriel G. Bès

► **To cite this version:**

Gabriel G. Bès. Simplicités, empirisme et interprétation psychologique dans la grammaire générative.. Archives et documents de la société d'histoire et d'épistémologie des sciences du langage, 1981, 1, pp.1-84. hal-01064415

**HAL Id: hal-01064415**

**<https://hal.science/hal-01064415>**

Submitted on 16 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ARCHIVES ET DOCUMENTS

de la

SOCIÉTÉ D'HISTOIRE ET D'ÉPISTEMOLOGIE  
DES SCIENCES DU LANGAGE

N° 1

GABRIEL G. BÈS

*Simplicités, empirisme et interprétation psychologique  
dans la grammaire générative*

SAINT-CLOUD

S.H.E.S.L.

1981

ARCHIVES ET DOCUMENTS  
DE LA  
SOCIETE D'HISTOIRE ET D'EPISTEMOLOGIE  
DES SCIENCES DU LANGAGE  
(S.H.E.S.L.)

N°1

---

*Administration*, Th. OLIU- Ecole Normale Supérieure, Section d'Espagnol,  
Parc de Saint-Cloud, Grille d'honneur- 92211 Saint-Cloud.

*Rédaction*, S. AUROUX, J. AUTHIER, J. J. COURTINE, A. GRESILLON, Cl. HAROCHE, P. HENRY,  
D. MALDIDIER, F. SOUBLIN.

Toute correspondance ou manuscrit doit être adressée à Cl. HAROCHE, 85 rue d'Assas-  
Paris 75006- Les manuscrits reçus seront soumis au comité de lecture selon une  
procédure rapide. Lorsqu'un texte sera accepté, il sera demandé aux auteurs d'en  
fournir une version directement reproductible en offset.

Les "ARCHIVES ET DOCUMENTS DE LA S.H.E.S.L." sont diffusés aux seuls adhérents de la S.H.E.S.L.

Les raisons de leur lancement sont claires: la "Société d'Histoire et d'Epistémologie des Sciences du Langage" dispose déjà d'une revue bi-mensuelle (commercialisée) qui ne comporte qu'une centaine de pages par livraison.

Il nous a paru intéressant d'élargir les possibilités de publication et d'offrir aux membres de la S.H.E.S.L. des textes à diffusion rapide.

Les "ARCHIVES ET DOCUMENTS DE LA S.H.E.S.L." n'étaient pas commercialisés, et faisant l'objet d'une diffusion restreinte, on aura compris qu'il s'agit de "Working papers".

Cette solution présente un double avantage: ces documents devraient permettre à l'auteur de recevoir remarques et commentaires enrichissants, précisément dans la mesure où la S.H.E.S.L. est composée d'adhérents venant d'horizons très divers. Le lecteur, de son côté, pourra se trouver confronté à l'exposé de points de vue auxquels il peut ne pas être familiarisé. Fondamentalement, il s'agit ici de favoriser échanges et discussions, en un mot d'œuvrer au décloisonnement de disciplines universitaires qui ont parfois tendance à s'ignorer.

Nous souhaiterions susciter à côté des démarches plus ou moins traditionnelles et parfois circonspectes, des vues théoriques un peu audacieuses et stimulantes, en particulier en ce qui concerne l'Epistémologie, la Théorie et l'Histoire, et inciter ainsi les auteurs à prendre parfois le risque de sortir des garanties et des certitudes qu'offre un positivisme de bon aloi.

SIMPLICITES, EMPIRISME ET INTERPRETATION PSYCHOLOGIQUE  
DANS LA GRAMMAIRE GENERATIVE

G.G.BES

---

*La publication de ce document a été permise grace à une subvention de  
l'amicale des anciens élèves de l'Ecole Normale Supérieure de Saint-Cloud.*

## I Les questions posées

Plus de deux décennies après la publication des premiers écrits de Noam Chomsky sur la grammaire générative, une douzaine d'années après qu'elle eut été introduite en France(1) et alors que sa notation et sa terminologie se répandent jusque dans les textes scolaires, on peut et on doit se poser des questions sur ce mouvement linguistique qui apparaît aujourd'hui comme un des plus importants si l'on considère les signes externes du "pouvoir scientifique": nombre et importance des publications, répercussion sur les disciplines connexes, intérêt soulevé chez les jeunes chercheurs, appropriation des concepts-ou de la terminologie- par le milieu culturel, etc.. Il semble bien que les questions qui doivent être posées sont les suivantes: Est-ce que la grammaire générative a réellement avancé dans la connaissance du langage? Si oui, quelles sont les données nouvelles qu'elle a mises en lumière? Quelles sont celles qu'elle a expliquées? Bref, quel est cet ensemble d'acquis sur le langage que l'on pourrait considérer désormais comme suffisamment sûrs pour qu'on puisse penser qu'ils ne seront pas remis en doute très bientôt?

Si l'on veut éviter les erreurs souvent commises au moment de la publication des premiers écrits générativistes-rejet total ou adhésion non critique- il est impossible de répondre aux questions précédentes de manière simpliste ou globale: il est vraisemblable que le bilan final ne se fera pas par oui ou par non, mais que chaque aspect examiné exigera une réponse particulière et nuancée. Ce travail, centré sur la notion de simplicité en grammaire générative, s'inscrit dans cette perspective de fond de contribuer à dresser, de manière documentée et raisonnée, un bilan des apports de ce mouvement, qui déborde du linguistique sur le psychologique et le philosophique.

Le vocabulaire méthodologique employé ici sera très simple. Nous reconnaissons deux types de sciences: les sciences empiriques et les sciences formelles.

On dira qu'une proposition est une hypothèse empirique, ou plus simplement une hypothèse, si, par rapport à des données dont elle doit rendre compte de manière directe ou indirecte, on peut dire qu'elle est vraie ou fausse. On dira d'une science qu'elle est empirique si elle est composée d'hypothèses empiriques. En revanche une science formelle n'est pas composée d'hypothèses empiriques; elle ne possède pas des conditions d'adéquation par rapport à des données externes, les propositions qui l'intègrent se valident en fonction de critères internes.

Dans ce cadre de discussion, empirique ne veut donc dire ni "doué d'une existence perceptible", ni "constaté dans les faits mais non expliqué théoriquement", ni "ontologiquement réel", ni "formulé à la suite de l'observation du réel et par abstractions successives". Dans ce travail, empirique voudra simplement dire: proposition susceptible d'être testée par des données. On ne préjuge rien sur le fait que ces données soient perceptibles ou non, qu'elles soient susceptibles d'être produites et reproduites dans des conditions contrôlables ou non, qu'elles aient été observées avant de dégager l'hypothèse correspondante ou non. Toutes ces possibilités restent ouvertes dans la caractérisation ici adoptée d'empirique. (2).

Il sera parfois nécessaire d'introduire une distinction entre science entièrement empirique et science partiellement empirique. Une science entièrement empirique est celle qui est entièrement composée d'hypothèses empiriques, nul critère méthodologique autre que les exigences sur l'explicitation et la cohérence des hypothèses n'étant employé pour sélectionner celles-ci. Science partiellement empirique sera celle qui est composée d'hypothèses empiriques mais qui est complétée par des formulations méthodologiques générales, non empiriques, destinées à choisir entre plusieurs hypothèses possibles. Au cours du développement des sciences, on a attribué à la simplicité ce rôle d'arbitrage non empirique entre hypothèses en concurrence sur un même domaine.

L'interprétabilité psychologique d'une théorie ou ensemble d'hypothèses sur le langage concerne le problème posé par ce qu'on appelle couramment la réalité psychologique de la théorie ou son contenu psychologique. Ce problème peut utilement être posé à l'aide des notions de modèle de boîte noire et de modèle de boîte transparente. Sur un plan général, on dira qu'une théorie ou ensemble d'hypothèses empiriques, comprenant dans celles-ci les déductions particulières qu'on peut tirer des formulations générales, peut être associée soit au comportement d'un objet et, dans ce cas, on a un modèle de boîte noire, soit aux parties constitutives de l'objet, en général non observables, et à son comportement extérieur; dans ce deuxième cas, on a un modèle de boîte transparente. Dans le domaine du langage, on dira qu'une théorie ou hypothèse empirique est interprétable psychologiquement où, si l'on préfère qu'une théorie ou hypothèse possède une réalité psychologique si, non seulement les déductions faites à partir de la théorie ou de l'hypothèse sont associables à un comportement extérieur, mais si également, les parties constitutives de cette théorie ou de cette hypothèse sont elles-mêmes censées être associées aux parties constitutives internes de l'organisation mentale qui est responsable de ce comportement extérieur. (3).

Nous supposons que, sur un plan général, la notion de simplicité est une notion acceptée et connue intuitivement. Elle est susceptible, dans une formulation générale, d'être ramenée à une relation quantitative: A est plus simple que B si A possède moins d'éléments de type x que B. Etre plus simple signifie, par conséquent, être doué de moins d'éléments. Ce sont les entités A et B, ou le trait x, ou l'échelle par rapport à laquelle on détermine le plus et le moins, qui vont être modifiés dans chaque formulation particulière; mais, dans tous les cas, réapparaît le même schéma général.

Le statut de la simplicité n'est pas clair dans la méthodologie scientifique. Une simplicité formelle, d'explicitation difficile, a été parfois utilisée pour choisir entre plusieurs hypothèses empiriques possibles, d'où il résulte une situation quelque peu



paradoxe: dans l'organisation des hypothèses d'une science empirique apparaît un élément qui ne l'est pas(4). Mais une formulation de simplicité- ou plus simplement, une simplicité- n'a pas comme seule possibilité celle d'être formelle; en fait, rien ne s'oppose , en principe, à ce qu'une simplicité soit empirique. Dans ce cas, ce qui a été déclaré comme plus (ou moins) simple doit être associé à une donnée qui peut valider ou infirmer la proposition de simplicité. Par exemple, si l'on choisit entre hypothèses en concurrence par le biais de la simplicité, celle-ci sera une simplicité formelle si l'hypothèse choisie comme plus simple ne possède aucun type de vérification empirique susceptible de la distinguer de la moins simple; en revanche, la simplicité elle-même sera empirique si l'on exige de l'hypothèse choisie comme plus simple qu'elle possède certains traits d'adéquation externe, absents de la moins simple. Une simplicité empirique est, par conséquent une hypothèse à part entière , de laquelle on doit pouvoir dire si elle est vraie ou fausse. Par ailleurs, une simplicité empirique, dans le domaine du langage, peut ou non être interprétée psychologiquement. Elle aura une interprétation psychologique si elle est censée être associée à un mécanisme interne, mental, qui agit au cours des conduites linguistiques.

Le dilemme science formelle vs. science empirique est apparu dans la linguistique post-saussurienne, sans qu'il ait été ni clairement posé ni tranché.

En effet, la linguistique a été considérée parfois comme une science empirique parfois comme une science formelle. Saussure et Troubetzkoy la conçoivent comme science empirique de même que, par exemple, Martinet. En revanche, dans quelques secteurs du structuralisme américain, on cherche à valider des analyses linguistiques selon leur adéquation par rapport à une procédure analytique déterminée, fondée elle-même sur des définitions et des méthodes posées au préalable; dans cette optique, une description linguistique se valide moins par sa capacité de rendre compte des

données que par le fait d'avoir appliqué de manière cohérente un certain formalisme descriptif qui, lui-même, n'est pas empirique (5).

L'ambiguïté méthodologique de la linguistique apparaît dans toute son intensité avec Hjelmslev. Celui-ci soulève bien le problème du rapport de la théorie avec les données, mais il aboutit à une conclusion difficile à comprendre.

Il affirme (56b p.28-29 /56 p.14-15), en effet:

"1-La théorie elle-même ne dépend pas de l'expérience. Rien en elle n'indique si elle aura des applications en rapport avec des données de l'expérience. Elle ne postule nullement l'existence de ces données. Elle constitue ce que l'on peut appeler un pur système déductif, en ce sens que c'est la théorie à elle seule qui, à partir des prémisses qu'elle énonce, permet le calcul des possibilités qui en résultent.

2-Le théoricien sait d'expérience que certaines prémisses énoncées dans la théorie remplissent les conditions nécessaires pour que celle-ci soit applicable à certains faits".

La première de ces caractéristiques constitue, d'après lui, le caractère arbitraire de la théorie; la deuxième, son caractère approprié. Les données ne peuvent jamais "ni confirmer ni infirmer la validité de la théorie". Sa conclusion en est que "la question déjà soulevée du rapport entre théorie et réalité aura une réponse double: en vertu de son caractère arbitraire, la théorie est aréaliste, en vertu de son caractère adéquat, elle est réaliste [7]"

Si par "arbitraire" on comprend formel et empirique par "approprié", la conclusion de Hjelmslev enferme une contradiction impossible à dépasser: une théorie ne peut à la fois être vérifiable et non vérifiable par des données. Cependant, la contradiction pourrait être levée si l'on assignait à "applicabilité" non pas le sens de "confrontation avec les données pour tester les hypothèses" mais le sens de "détermination des possibilités de la théorie qui sont attestées sur le plan de l'expérience".

La détermination de l'applicabilité de la théorie ne serait que le constat de la réalisation substantielle dans le domaine des données de certaines possibilités prévues dans le calcul de la théorie linguistique(6). Mais le fait que d'autres possibilités prévues par la théorie n'apparaissent pas attestées ou qu'aucune possibilité prévue par la théorie le soit, n'infirmerait pas celle-ci. De même, les données qui attestent une des possibilités prévues par le calcul ne peuvent pas falsifier ou valider l'appareil de définitions et des procédés de la théorie, car des données sont décrites de manière "arbitraire" par celle-ci. Si cette interprétation est correcte, la glossématique de Hjelmslev est à classer parmi les sciences formelles: elle se présente dans un cadre méthodologique tel qu'il est impossible de décider si les définitions proposées par son appareil conceptuel sont vraies ou fausses.

Dans ce cadre d'ensemble, où l'alternative formalisme vs. empirisme n'était pas clairement tranchée, la notion de simplicité et son utilisation en linguistique ajoutent encore à la confusion. Comme la plupart des points importants de la méthodologie linguistique, elle a mérité en glossématique une attention particulière. Uldall semble considérer la simplicité comme un desideratum absolu, bien qu'il la subordonne à la cohérence et à l'exhaustivité (131 p. 20) et Hjelmslev la pose comme élément essentiel pour partager entre formulations concurrentes (56 p.18). La simplicité glossématique est nettement formelle: le recours à la simplicité est en dernière instance, dit Uldall (ibidem) un fait esthétique(7). Formelle aussi la simplicité de Hockett, proposée sous l'étiquette de "économie", bien que sa conception d'ensemble s'oriente vers une linguistique empirique (8).

Mais l'ambiguïté formalisme-empirisme, par rapport à la simplicité apparaît de manière particulièrement nette dans les écrits de Jakobson. Il insiste, d'une part, sur le fait qu'une description en traits binaires est, parce que plus simple, plus adéquate qu'une autre, qui n'utilise pas des traits binaires; il s'agit par conséquent d'une simplicité formelle.

Par ailleurs, il essaie de justifier la simplicité en fonction des processus psycho-physiologiques d'utilisation du langage; il incline ainsi vers une simplicité empirique(9).

Des déclarations très explicites de Noam Chomsky laissent penser qu'il a définitivement tranché sur les problèmes posés par la simplicité: d'une part, la simplicité, en grammaire générative, serait une simplicité empirique; par ailleurs, elle devrait être interprétable psychologiquement par rapport à la capacité humaine d'acquisition du langage. Des textes clairs témoignent de ces prises de position.

D'abord, la simplicité en grammaire générative est censée être empirique; non seulement elle est posée comme telle mais, aussi, la simplicité formelle est explicitement écartée:

"Encore une fois, nous disons de manière emphatique que le choix d'une mesure de simplicité est un problème empirique. Le problème de découvrir une mesure d'évaluation est très analogue à celui d'évaluer une constante en physique[...]. Il est aussi impossible de donner une justification A PRIORI d'une mesure d'évaluation particulière qu'une valeur particulière à une constante de gravitation (Chomsky et Halle 31p.108).

"Il est un autre sens selon lequel on parle de simplicité, élégance, etc..des théories (la physique de la relativité est plus élégante-simple- que la physique de Newton, l'astronomie de Copernic est plus simple que celle de Ptolomeo, etc..). Nous ne nions pas l'intérêt de cette notion, mais nous n'avons rien à dire sur elle. C'est un concept qui n'appartient ni à la linguistique ni à aucune discipline linguistique particulière mais plutôt à l'épistémologie générale ou à la philosophie des sciences" (Chomsky et Halle 31 p.109-110).

Par ailleurs, la mesure de simplicité est une hypothèse associée à la capacité d'acquisition du langage: elle est par

conséquent, selon notre terminologie, interprétable psychologiquement:

"Toutes ces théories ont, dès lors, besoin d'être complétées par une mesure d'évaluation, si l'on veut rendre compte de l'acquisition linguistique et justifier la possibilité de choisir des grammaires particulières; je continuerai d'admettre provisoirement- comme je l'ai fait jusqu'à présent- que c'est là un fait empirique concernant la théorie linguistique générale" (20b p.58/20 p.37).

"Supposons en d'autres termes que nous considérons un modèle d'acquisition linguistique comme un dispositif d'input-output, qui détermine à l'output une grammaire générative particulière, étant donné à l'input certaines données linguistiques primaires; une mesure de simplicité, conjointe à une spécification de (I)-(IV), constitue une hypothèse touchant la nature d'un tel dispositif. Le choix d'une mesure de simplicité est donc un problème empirique, ayant des conséquences empiriques" (20b p.59-60/20 p.38).

"Un trait essentiel de notre théorie du langage c'est qu'elle comprend une mesure d'évaluation qui permet l'assignation de valeurs aux grammaires en concurrence. C'est sur la base de cette mesure d'évaluation qu'un enfant qui apprend une langue choisit une des grammaires (dont il y a en principe, une grande infinité) qui sont compatibles avec le corpus assez restreint de données linguistiques auquel il avait été exposé. La grammaire qu'un enfant construit en apprenant sa langue native sera par conséquent toujours celle qui est classée plus haut en termes de cette mesure d'évaluation" (32p.251).

La lecture quelque peu attentive d'autres textes de grammaire générative et l'observation de l'utilisation de la notion de simplicité dans la pratique montre que l'accord de ces déclarations liminaires avec l'utilisation concrète de la notion de simplicité ne s'est pas toujours réalisé.

Dans le cadre de l'interrogation générale sur les progrès éventuellement accomplis par la grammaire générative(10), nous nous proposons d'aborder le problème de la simplicité en fonction des questions suivantes:

1.-Quelles sont la ou les notions de simplicité qui ont cours en grammaire générative? En particulier, la ou les simplicités proposées, sont-elles empiriques ou formelles? Sur quelles entités portent-elles?

2.-La ou les notions de simplicité empirique, sont-elles interprétables psychologiquement?

3.-Au-delà des déclarations de principe, est-ce que des hypothèses explicites de simplicité empirique et /ou de simplicité interprétable psychologiquement ont été formulées? Si oui, ont-elles été validées ou infirmées?

4.-Est-ce que le traitement du problème de la simplicité a été avancé par rapport à la situation antérieure, c'est à dire par rapport à la situation héritée par la grammaire générative?

Les définitions adoptées ci-dessus de formel et d'empirique ne laissent pas la possibilité d'envisager une troisième possibilité. Celle-ci cependant doit être introduite pour pouvoir rendre intelligibles certains textes, mais elle portera non sur le rapport formulation-données, mais sur le type d'entité considérée comme donnée. En effet, il est parfois impossible, face à un texte, de déterminer si une entité quelconque peut jouer ou non le rôle de donnée par rapport à une proposition déterminée; dans ce cas, on parlera de propositions éventuellement empiriques. Nous disons qu'une proposition (en particulier, une formulation de simplicité) est éventuellement empirique, si, acceptant le caractère de donnée d'une entité quelconque dont le statut est douteux, la proposition en question était empirique.

Désormais, pour faire référence à la simplicité, on utilisera MS (=mesure de simplicité) et trois formes verbales: définir; caractériser, à l'indicatif et caractériser, au conditionnel. La première sera utilisée dans le cadre d'une simplicité formelle, la deuxième, dans celui d'une simplicité empirique et la troisième, pour les cas d'une simplicité éventuellement empirique. La formule "MS définit A comme plus simple que B", indique qu'aucune condition de vérification externe ne sera exigée de l'entité A, sélectionnée comme plus simple; la formule "MS . caractérise A comme plus simple que B", indique que l'entité A, sélectionnée comme plus simple, devra être mise en rapport avec des données susceptibles de tester l'hypothèse sur la simplicité; la formule "MS caractériserait A comme plus simple que B" sera réservée pour les cas douteux, éventuellement empiriques, où l'entité A est mise en rapport avec des données externes, mais dont le statut est confus, le fait de savoir si ces données sont susceptibles ou non de tester une hypothèse n'étant pas clair. Les notions de définir et de caractériser (à l'indicatif et au conditionnel) peuvent s'appliquer aussi bien en termes absolus qu'en termes graduels, c'est à dire qu'une entité A quelconque peut être soit plus simple que B, soit plus simple que B à un degré déterminé; on convient donc que les termes utilisés peuvent recevoir les deux interprétations.

L'entité A, dans le cadre de la problématique ici abordée, est une entité proposée en linguistique, très généralement une grammaire ou une ou plusieurs règles d'une grammaire. On utilisera par conséquent grammaire aussi bien pour faire référence à une grammaire universelle ou particulière complète qu'à un sous-ensemble d'une ou de plusieurs règles ou entités d'une grammaire. On note grammaire universelle par GU et grammaire particulière par GP. On dira qu'une GP est déterminée par une GU si GP respecte les restrictions et le schéma général de l'organisation définis par GU. Par représentation nous comprenons cet objet théorique noté symboliquement et qui est destiné à rendre compte d'un énoncé: les arbres ou parenthétisations étiquetées constituent typiquement des représentations. Une grammaire spécifique (ou

engendre) un ensemble infini de représentations. Les données-  
sauf indication contraire- sont constituées par les énoncés d'une  
part, et d'autre part, par les jugements métalinguistiques  
intuitifs portés sur eux (d'ambiguïté, de grammaticalité, de  
paraphrase, etc...); on les placera sous la rubrique langage.  
Nous supposons qu'existe la possibilité de déterminer l'identité  
entre grammaires (particulières ou universelles), entre représen-  
tations et entre données. En fait, en ce qui concerne les objets  
théoriques (grammaires et représentations), le présupposé sur la  
possibilité ou non de les identifier va au-delà des possibilités  
actuelles de la théorie linguistique: une partie de la discussion  
entre sous-écoles rivales est centrée sur le problème de détermi-  
ner les identités ou différences (substantielle, notationnelle .  
etc..), entre hypothèses alternatives (11).

Pour notre discussion ce point est relativement secondaire car  
nous utiliserons les identifications des entités rapprochées qui  
ont été explicitement ou implicitement faites par celui qui  
propose une MS.

Nous acceptons également qu'existe la possibilité de  
déterminer l'équivalence entre grammaires particulières et entre  
grammaires: universelles. On dira que deux GP sont entièrement  
équivalentes, si elles-mêmes et les représentations qu'elles  
spécifient sont différentes et si ces représentations peuvent  
rendre compte, dans les mêmes conditions d'adéquation externe  
relatives à chaque énoncé et aux jugements métalinguistiques  
portés sur lui, de tout langage appartenant à la langue parti-  
culière en question. Par ailleurs, on dira que deux GP sont  
équivalentes par rapport à un langage x, si elles-mêmes et les  
représentations qu'elles spécifient sont différentes et si ces  
représentations peuvent rendre compte d'un même langage x, appar-  
tenant à la langue particulière en question, dans les mêmes  
conditions d'adéquation externe relatives à chaque énoncé et aux  
jugements métalinguistiques portés sur lui. Etre entièrement  
équivalent et être équivalent par rapport à un langage x, ne sont  
donc pas la même notion: si deux GP qui sont entièrement équiva-  
lentes doivent être équivalentes par rapport à un langage x, le  
contraire n'est pas vrai. Dans les formulations proposées par la



suite, on trouvera deux situations d'équivalence entre GP:

- (a) les GP considérées sont soit entièrement équivalentes, soit équivalentes par rapport à un langage x; ce sont les cas où les textes d'où les formulations sont tirés exigent une équivalence entre les GP mais ils ne permettent pas de trancher entre le type d'équivalence exigé pour que la MS soit susceptible d'être appliquée (formulations I, III, IV, et V); on peut penser que dans tous ces cas, on souhaite appliquer la MS à des GP entièrement équivalentes, bien que des illustrations ne portent que sur des GP équivalentes par rapport à un langage déterminé;
- (b) les GP considérées ne sont pas entièrement équivalentes, mais équivalentes par rapport à un langage déterminé (formulations VII et VIII).

L'équivalence au niveau des GU doit se définir par le biais de l'équivalence des GP qu'elles déterminent. Plusieurs cas sont à prévoir. Etant donné deux GU, chacune de ces GU peut déterminer, pour toute langue particulière, une GP qui est soit entièrement équivalente à la GP déterminée par l'autre GU, soit équivalente par rapport à un langage x. Ou bien, les GU peuvent déterminer, pour certaines langues naturelles, des GP entièrement équivalentes et, pour d'autres langues naturelles, des GP équivalentes par rapport à un langage x. Dans les formulations qui suivent, dans un seul cas (formulation II) il a été possible de dégager une situation d'équivalence claire entre deux GU: ce sont des GU qui, chacune, détermine, pour toute langue particulière, une GP qui est entièrement équivalente à la GP déterminée par l'autre GU. On dira, dans ce cas, que les GU sont entièrement équivalentes, et on s'abstiendra de proposer une terminologie pour les autres cas possibles d'équivalence entre GU.

Au moyen de lettres capitales on distinguera des grammaires universelles différentes: GU-A et GU-B sont des grammaires universelles différentes. En revanche, on utilisera des numéros pour distinguer les grammaires particulières et les langages: GP-1 et GP-2, Lang-1 et Lang-2 indiquent, respectivement, des grammaires particulières et des langages différents.

Il n'est pas toujours aisé de discuter les écrits de la grammaire générative, un certain type de vocabulaire- celui, en particulier, qui porte sur les notions générales- étant souvent employé sans rigueur excessive; comme, par ailleurs, il est impossible de discuter un texte peu explicite ou ambigu, il a été nécessaire de reformuler les textes originaux en propositions censées être claires et susceptibles d'être discutées. Le risque de déformation de la pensée originale est certain; pour le réduire autant que possible, chaque reformulation est suivie des textes originaux qui ont permis la reformulation et/ou des références bibliographiques précises.

Au § 2, qui suit, seront groupées les formulations qui comportent une simplicité formelle et dont la présentation originale n'était pas liée à la notion de généralisation. Le § 3 sera consacré aux simplicités mises en rapport avec la notion de généralisation; on y explicitera des simplicités formelles et des simplicités éventuellement empiriques. Le § 4 présentera les formulations comportant une simplicité empirique. Au § 5 on discutera la possibilité d'interprétation psychologique de la notion de simplicité et au § 6 on tirera le bilan d'ensemble, par rapport aux questions posées dans ce premier paragraphe et en fonction de l'argumentation développée aux § 2 à 5.

## II Simplicités formelles.

Il est possible de cerner au moins trois MS formelles qui diffèrent entre elles en fonction des entités sur lesquelles porte la comparaison et/ou des conséquences que chaque MS entraîne pour la théorie linguistique. Particulièrement importantes sont les formulations du § 2-1 et du § 2-2, car les MS qui y sont présentées sont directement impliquées dans le problème délicat du choix entre GU concurrentes.

### II-1 Formulation I

Soit GU-A et GU-B, qui déterminent, respectivement, GP-1 et GP-2, GP-1 et GP-2 sont soit entièrement équivalentes, soit

équivalentes par rapport à un Lang-1.

MS définit GP-1 comme plus simple que GP-2.

Par conséquent, GU-A doit être préférée à GU-B.

L'essentiel de la formulation précédente est l'utilisation de la MS pour choisir entre GU concurrentes: on doit préférer celle des GU qui détermine la GP plus simple susceptible de rendre compte d'un langage donné. Cette formulation se dégage, par exemple du texte suivant:

"La preuve la plus décisive de l'inadéquation d'une théorie linguistique consiste à montrer qu'elle ne peut absolument pas s'appliquer à une langue naturelle. Une preuve plus faible, mais parfaitement suffisante, de son inadéquation, serait de montrer que la théorie ne s'applique qu'avec lourdeur, c'est à dire que toute grammaire construite selon cette théorie est fort complexe, forgée ad hoc, et non révélatrice, que certains procédés très simples de description des phrases grammaticales ne peuvent s'accommoder des formes de la grammaire qui leur correspondent, et que certaines propriétés formelles fondamentales des langues naturelles ne peuvent être utilisées pour simplifier les grammaires[...].

La seule façon de tester l'adéquation de notre appareil est d'essayer de l'appliquer directement à la description des phrases anglaises. Dès que nous considérons une phrase un peu complexe et, en particulier, lorsque nous tentons de définir un ordre entre les règles qui produisent ces phrases, nous nous heurtons à de nombreuses difficultés[...]. je me bornerai à esquisser quelque cas simples où des progrès considérables peuvent être réalisés sur les grammaires de la forme (29b p.39-40 /29 p.34-35).

Le passage qui précède annonce les grands traits d'un raisonnement qui va se concrétiser tout au long du chapitre V

des Structures Syntaxiques et qui réapparaît aussi ailleurs: la GU entièrement composée de règles syntagmatiques est inadéquate parce que les GP qu'elle détermine- présentée sous forme de règles de l'anglais- sont compliquées; par contre une GU qui incorpore des transformations est adéquate parce qu'elle détermine des GP qui sont susceptibles de rendre compte des mêmes données et qui sont plus simples. A la suite de la discussion de plusieurs cas concrets, la conclusion d'ensemble qui en est tirée s'intègre, elle aussi, dans cette formulation I:

"Nous venons de voir qu'une grande variété de phénomènes apparemment distincts se mettent en place tout naturellement lorsque nous adoptons le point de vue de l'analyse transformationnelle et que, par conséquent, la grammaire de l'anglais devient beaucoup plus simple et ordonnée. C'est là l'exigence première que toute conception de la structure linguistique (c'est à dire tout modèle pour la forme des grammaires) doit satisfaire. Je pense que ces considérations justifient amplement l'affirmation énoncée plus haut, à savoir que les conceptions syntagmatiques sont fondamentalement inadéquates et que la théorie de la structure linguistique doit être développée dans le sens du présent exposé de l'analyse transformationnelle" (29b p.75/29 p.68-69).

Le schéma de raisonnement qui préside à la Formulation I réapparaît souvent dans les écrits de grammaire générative. Dans 20b p.127/20 p.87, Chomsky affirme que le lexique doit être séparé des règles syntagmatiques parce que "la grammaire peut être simplifiée de façon importante si [les propriétés idiosyncratiques des formants] sont exclues des règles de réécriture et enregistrées dans les entrées lexicales où elles vont à leur place la plus naturelle". Et il ajoute par la suite: "chaque fois que l'on essaiera de construire une grammaire précise, il se révélera rapidement que bien des formants ont des caractéristiques grammaticales uniques, ou presque, de sorte que la simplification de la grammaire qui peut être ainsi réalisée sera certainement substantielle".

On compare donc GP-1 et GP-2 que l'on suppose rendre compte des mêmes données et qui sont respectivement déterminées par GU-A et par GU-B. GU-A est une grammaire universelle avec lexique; GU-B sous-catégorise les formants dans les règles syntagmatiques et ne possède pas de lexique. Puisque GP-1 est plus simple, la solution proposée dans le cadre de GU-A est adoptée sur un plan général. On constate donc qu'une solution dans le domaine de la GU est choisie en fonction d'une simplicité formelle portant sur les GP (12).

Le même raisonnement apparaît appliqué au domaine des faits phoniques. Par exemple, Halle (51 p.91) choisit GU-A (avec règles ordonnées) au détriment de GU-B (avec règles non ordonnées) de la manière suivante. GP-1 et GP-2, déterminées respectivement par GU-A et GU-B, sont présentées comme étant toutes deux susceptibles de rendre compte de la formation du pluriel en anglais. Halle observe que "les longueurs relatives des deux ensembles de propositions (c'est à dire, des règles non formalisées) reflètent leur simplicité relative. Par conséquent, il est obligatoire d'introduire l'ordre dans cette circonstance si nous désirons satisfaire notre critère de simplicité". Le résultat en est qu'on choisit GU-A, GP-1 ayant été définie comme plus simple que GP-2 (13).

## II-2 Formulation II.

Soit GU-A et GU-B, qui sont entièrement équivalentes.  
MS définit GU-A comme étant plus simple que GU-B.  
Par conséquent, GU-A doit être préférée à GU-B.

La simplicité formelle porte cette fois non sur les GP mais sur les GU qui les déterminent. La présentation la plus claire en est faite par Postal (14). La particularité qui apparaît explicitement dans la formulation de Postal est le recours à des opérations ensemblistes: Posséder moins d'éléments est concrétisé comme A étant inclus dans B, le complément de A dans B n'étant pas vide; cela implique que les éléments x sur lesquels porte la quantification de MS sont des éléments qualitativement différents:

"Je ne voudrais me présenter comme un expert en philosophie de la science ni suggérer que les querelles théoriques en linguistique sont susceptibles en général de trouver leur solution dans la philosophie de la science. Cependant certains points sont relativement clairs. Deux théories différentes du même domaine étant données, il est possible de faire un choix clair entre elles si certaines relations logiques sont valables entre ces théories. En particulier, si l'outillage théorique ("theoretical machinery") d'une théorie est inclus dans celui de la deuxième et si celui-ci possède, en plus, un certain outillage théorique additionnel, alors toutes les autres choses étant identiques, la première théorie, conceptuellement plus restreinte, doit être choisie" (107 p.135).

Les formulations I et II qui précèdent proposent exactement le contraire de ce qui est nié avec force par ailleurs; non seulement la simplicité formelle est rejetée (cf. ci-dessus) mais explicitement on affirme que la MS utilisée dans le cadre de la grammaire générative ne doit s'employer pour comparer ni des GU différentes ni des GP déterminées par des GU différentes:

"Il apparaît également que des mesures d'évaluation comme celles dont on a discuté dans la littérature consacrée à la grammaire générative ne peuvent être employées pour comparer différentes théories de la grammaire; il est complètement absurde de comparer à l'aide d'une telle mesure une grammaire appartenant à une classe de grammaires proposées avec une grammaire appartenant à une autre classe [...]. C'est évidemment une question fondamentale que ce dernier problème et l'on devrait également la régler en se fondant, dans la mesure du possible, sur des bases empiriques d'adéquation descriptive et explicative. Mais ce n'est pas la question qui est mise en jeu lorsqu'on emploie une mesure d'évaluation pour tenter d'atteindre l'adéquation

explicative. A titre d'illustration concrète, examinons si les règles d'une grammaire doivent être non ordonnées (appelons cela la théorie linguistique Tn) ou ordonnées d'une manière spécifique (théorie To). A priori, on ne connaît aucun sens absolu de la "simplicité" ou de "l'élégance", développé dans la théorie linguistique ou dans l'épistémologie générale, qui permette de comparer Tn et To. Il est donc absurde de soutenir qu'en un sens "absolu" Tn soit "plus simple" que To ou inversement. Il est aisé d'inventer un concept général de "simplicité" qui donnera l'avantage à Tn sur To ou bien à To sur Tn; ce concept n'aura aucune justification connue, dans aucun de ces deux cas" (20b p.60/20p. 38-39).

### II-3 Formulation III.

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 et GP-2 sont soit entièrement équivalentes soit équivalentes par rapport à un Lang-1.

MS définit GP-1 comme plus simple que GP-2.

Par conséquent, GP-1 doit être préférée à GP-2.

Cette formulation-ci, à la différence des précédentes, ne s'applique qu'à des GP déterminées par une même GU; elle résume les textes qui suivent, empruntés aux Structures Syntaxiques:

"Une exigence encore plus modeste serait qu'étant donné un corpus et deux grammaires G1 et G2, la théorie dise laquelle des deux est la meilleure grammaire de la langue d'où le corpus est extrait. Dans ce cas nous pourrions dire que la théorie nous fournit une procédure d'évaluation des grammaires" (29b p.57 /29 p.51).

"Nous pensons qu'il serait déraisonnable d'attendre d'une théorie linguistique qu'elle fournisse plus qu'une procédure pratique d'évaluation des grammaires" (29b p.58/29 p.52)

"Dans chaque cas, notre seul souci a été de déterminer la complexité de la grammaire et nous avons cherché à montrer

que l'analyse proposée était manifestement plus simple que les autres analyses possibles que nous avons rejetées" (29b p.92 / 29 p.83).

Ce n'est pas seulement dans les Structures Syntaxiques que cette formulation est proposée (15). Dans Aspects..., certaines des solutions partielles proposées au problème de l'insertion lexicale sont fondées sur le même type d'argumentation. Par exemple le fait que le verbe et l'adjectif possèdent des traits de sélection contextuels qui sont déterminés par les traits de sélection du nom est justifié en fonction d'une MS encadrée dans la Formulation III:

"Bref, en décidant que l'analyse des Verbes en symboles complexes fera l'objet d'un choix indépendant et que les Noms seront sélectionnés suivant une règle fondée sur les Verbes, on est amené à compliquer considérablement la grammaire. Les problèmes sont multipliés quand on tient compte des règles sélectionnelles indépendantes Nom-Adjectif. En raisonnant d'une manière analogue, nous pouvons exclure la possibilité que le Sujet choisisse le Verbe et le Verbe l'Objet" (20b p.160/ 20 p.115).

Peu de remarques doivent être faites sur les Formulations I à III. Dans les trois cas il s'agit d'une simplicité formelle qui conduit aux difficultés difficilement surmontables de toute simplicité formelle (cf. notes 7 et 9). Celui qui a proposé une hypothèse a une tendance marquée à la considérer comme étant plus simple et plus belle que l'hypothèse proposée par son collègue ou son concurrent. Par ailleurs, pour pouvoir appliquer les Formulations avec un minimum de rigueur, il serait indispensable, dans les cas des Formulations I et II, de résoudre au préalable le problème de la comparaison des théories qui, parce que différentes, utilisent des entités différentes bien que la même notation graphique soit utilisée. Est-ce qu'une entité "SN" d'une grammaire syntagmatique peut être comptée comme une entité "SN" d'une grammaire transformationnelle? Pour ne pas glisser dans l'erreur de quantifier des entités dont la caractérisation est différente, il faudrait "traduire" le système notational de chaque théorie à



une sorte de sur-notation commune et appliquer la MS sur le résultat de telles traductions. Dans l'état actuel des connaissances sur l'aspect formel des théories linguistiques, il est pratiquement impossible d'envisager cette possibilité.

La Formulation I apparaît, parmi celles qui ont déjà été citées, comme celle qui a laissé le plus de traces dans le développement de la théorie. Dans un certain sens cette formulation dérive du principe général de simplicité, souvent invoqué pour choisir entre théories concurrentes (cf. note 4). La simplicité a pour elle une certaine tradition d'élégance, de beauté classique, d'harmonie, etc...; il se peut qu'un jour on puisse expliquer l'origine psychique de ce sentiment (cf. ci-dessous § 6). Mais la Formulation I ne se laisse pas entièrement englober dans ce principe général. Ce qui est propre à la Formulation I, c'est que la MS ne s'applique qu'aux GP et non aux GU respectives, malgré le fait que la fonction de cette MS est de sélectionner la GU. C'est à dire qu'une simplicité formelle sélectionne finalement la GU mais de manière médiate, par le biais de son application aux GP. Nous ne reconnaissons aucun principe dans la méthodologie générale des sciences d'où on puisse déduire que la simplicité doit porter sur quelques-unes des hypothèses de deux ensembles d'hypothèses rendant compte d'un domaine déterminé et non sur ces ensembles considérés dans leur totalité. Par ailleurs, il faut observer que si l'on utilise la même MS intuitive qui permet de choisir entre GP déterminées par GU différentes, le résultat des choix sur le plan des GU est l'opposé de sur celui des GP. Il est vrai que souvent, dans les exemples où la Formulation I a été appliquée, il y a des cas où effectivement "on a comme le sentiment" qu'une GP-1 déterminée par GU-A est plus simple que GP-2 déterminée par GU-B. Mais le même "sentiment" réapparaît, en sens inverse, si l'on rapproche GU-A et GU-B. GU-A avec règles syntagmatiques et transformations apparaît comme moins simple que GU-B avec les seules règles syntagmatiques. Et il est impossible, sur un plan formel, de justifier cette préférence pour la simplicité d'un certain type d'hypothèses au détriment de la complexité des hypothèses d'un autre type (16).

### III Simplicité et généralisation.

La notion de généralisation est particulièrement importante en grammaire générative. Utilisée pour justifier le choix d'hypothèses concurrentes, en particulier, grammaires universelles, elle n'a pas cependant été définie avec précision. Certes, on entend par généralisation le fait d'englober dans une proposition plus générale le contenu de plusieurs propositions particulières, qui deviennent ainsi dérivables de la proposition plus générale. Mais cette idée intuitive est rebelle à une formulation plus précise; en particulier, on s'efforce de dégager les traits susceptibles de distinguer une généralisation significative d'une généralisation qui ne l'est pas (17).

Il ne rentre pas dans le cadre de cette discussion de préciser la notion de généralisation. On ne l'abordera donc qu'en ce qui concerne ses rapports avec la notion de simplicité. Sans poser le problème de savoir ce qu'est une généralisation (celles dont on fera état sont censées être "significatives") nous envisageons deux possibilités:

1) Les généralisations sont des données ; elles sont par conséquent susceptibles de tester une hypothèse;

2) Les généralisations sont le résultat d'une définition; elles appartiennent à l'appareil conceptuel d'une théorie et ne sont pas susceptibles de tester une hypothèse.

Vu leur statut douteux, il est impossible, même dans le cadre de la possibilité (1), d'accepter les généralisations comme des données à part entière; pour cette raison, on utilisera caractériserait pour les MS censées être testées par rapport à des généralisations.

#### III-1 Formulation IV

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 et GP-2 sont soit entièrement équivalentes, soit équivalentes par rapport à un Lang-1.

MS définit GP-1 comme plus simple que GP-2.

Par conséquent, GP-1 possède une généralisation.

#### Formulation V

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 et GP-2 sont soit entièrement équivalentes, soit équivalentes par rapport à un Lang-1.

GP-1 possède une généralisation, ce qui n'est pas le cas de GP-2.

MS caractériserait GP-1 comme plus simple que GP-2.

Par conséquent, MS serait validée.

Si MS n'était pas susceptible de caractériser GP-1 comme plus simple que GP-2, MS serait infirmée et devrait être modifiée.

Ces deux formulations sont nettement différentes, l'une comportant une MS formelle, l'autre une MS éventuellement empirique. Or, malgré cette différence fondamentale, on doit les dégager l'une et l'autre, au moins partiellement, d'un même texte qui, par conséquent, est ambigu; le voici divisé en sous-textes par des lettres intercalées ((b) à (f)); groupés avec d'autres citations sur la même question, ils permettront, par la suite, de faire des citations plus analytiques:

a) "Le trait systématique utilisé toujours, ouvertement ou non, pour choisir parmi des descriptions alternatives, c'est le degré de généralisation obtenu. On a une généralisation lorsque des assertions différentes à propos d'éléments différents peuvent être remplacées par les mêmes assertions ou par des assertions similaires. Il est par conséquent naturel, et il n'est pas difficile, de développer des systèmes de notation qui, convertissant les considérations de complexité en considérations de longueur, conduisent à une définition partielle de complexité" (Chomsky 30 p.242).

b) "Dans l'établissement d'une mesure d'évaluation des grammaires, le problème principal est de déterminer parmi les généralisations concernant une langue, celles qui ont une portée ("which generalizations about a language are significant ones") et il faut choisir la mesure d'évaluation de façon à donner l'avantage à ces dernières".

c) "Nous avons une généralisation lorsqu'un ensemble de règles concernant des éléments distincts peut être remplacé par une règle unique (ou, plus généralement, par des règles partiellement identiques) concernant l'ensemble tout entier, ou encore lorsqu'il est possible de montrer qu'une "classe naturelle" d'éléments passe par un certain processus ou un ensemble de processus analogues".

d) "Ainsi le choix d'une mesure d'évaluation constitue une décision sur la nature des "processus analogues" et des "classes naturelles"-en bref, sur la nature des généralisations importantes ("in short, what are significant generalizations")".

e) "Le problème est de construire une procédure assignant à une grammaire une mesure numérique de valeur sur la base du degré de généralisation linguistique significative auquel elle parvient".

f) "La mesure numérique évidente que l'on peut appliquer à une grammaire est la longueur, définie sur la base du nombre de symboles. Mais si ce doit être une mesure douée de sens, il est nécessaire de construire les notations et de restreindre la forme des règles de telle façon que les considérations importantes de complexité et de généralité soient converties en considérations de longueur- cela afin que les généralisations réelles abrègent la grammaire et que les pseudo-généralisations ne le fassent pas. Ce sont ainsi les conventions de notation adoptées pour présenter une grammaire qui définissent la "généralisation importante" ("significant generalization"), si l'on admet que la mesure d'évaluation est la longueur" (Chomsky 20b p.63-64 / 20 p.42).

g) "Il (le système de notation) a été proposé comme une explication de la notion "généralisation linguistique significative"; le degré de généralisation linguistique significative atteint par une grammaire-sa "simplicité", dans le sens technique du terme- est mesuré par le nombre de symboles qui apparaissent dans le schéma sous-jacent qui se développe en cette grammaire par l'utilisation du système de notation" (Chomsky et Halle 32 p.36).

Dans la Formulation IV, la MS est utilisée pour définir, dans le sens formel, l'existence d'une généralisation. Les textes qui permettent cette interprétation sont a in fine, c et, en particulier, d, f in fine, et g. Nous comprenons que GP-1 possède une généralisation, à la différence de GP-2, qui n'en a pas, si GP-1 est plus simple que GP-2: plusieurs règles de GP-2 sont remplacées par une seule règle de GP-1. Selon cette interprétation, "posséder une GP une généralisation" est le résultat d'une définition (18). La Formulation IV "complète" en quelque sorte la Formulation III; la différence en est que, d'après la Formulation IV, la GP-1 qui est définie comme étant plus simple est, par là même définie comme possédant une généralisation. La MS de la Formulation IV est posée comme un principe général, valable non pas pour quelques règles mais pour toute la grammaire, par Jacobs et Rosenbaum (19):

"Par exemple, une grammaire où on a doublé une certaine quantité d'information pour décrire un phénomène particulier est moins générale qu'une grammaire pour le même phénomène où cette information est présentée une seule fois" (63 p.224).

La Formulation V se dégage des textes b et e ci-dessus présentés. A la différence de la Formulation IV, la Formulation V serait empirique si la généralisation était une donnée à part entière. En effet, la lecture stricte de ces textes induit à interpréter que la MS devrait être modifiée jusqu'à ce qu'elle soit capable de s'adapter aux généralisations rencontrées, dans ce

sens que si une GP possédait une généralisation, elle devrait être sanctionnée comme plus simple; si elle ne l'était pas, ce serait la MS qu'il faudrait corriger. C'est à cause de ce schéma susceptible de tester une MS par rapport à une généralisation que celle-ci serait éventuellement empirique.

La Formulation V a été utilisée d'une manière plutôt marginale(20). En revanche, la formulation qui suit, dans le § III-2, apparemment proche de la V, mais en fait très différente a été employée en grammaire générative de manière non accidentelle.

### III-2 Formulation VI

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 rend compte du Lang-1 et GP-2 du Lang-2.

GP-1 rend compte des données telles qu'elle doit posséder une généralisation, ce qui n'est pas le cas pour GP-2.

MS définit GP-1 comme plus simple que GP-2.

Par conséquent, GU-A doit être considérée comme une grammaire universelle susceptible d'exprimer une généralisation.

Si MS n'était pas susceptible de définir GP-1 comme plus simple que GP-2, GU-A devrait être modifiée.

Dans cette Formulation VI, comme dans la précédente Formulation V, la MS est mise en rapport avec des GP dont l'une devrait posséder une généralisation et l'autre non, étant l'une et l'autre déterminées par une même GU. Mais l'analogie entre les deux formulations s'arrête ici. En effet, dans la Formulation VI- mais non dans la V- GP-1 et GP-2 rendent compte des données différentes. Par ailleurs, dans la Formulation VI, ce n'est pas la MS qui devrait être mise en cause si elle n'était pas susceptible de définir comme plus simple GP-1, censée posséder une généralisation, mais la GU qui a déterminé GP-1. De manière analogue, si la MS définissait GP-1 comme plus simple, on ne devrait pas pour autant en conclure que la MS serait validée; dans ce cas, ce serait la GU qui a déterminé GP-1 qui le serait. Les exemples qui suivent, empruntés à Chomsky et Halle (32 p.400 et suivantes) permettent d'illustrer cette formulation et ses conséquences (cf. également le texte (f) de la citation précédente empruntée à Chomsky 20b p.64 /20 p.42).

Soient les paires de règles qui suivent:

(a -1)  $i \longrightarrow u$

(a -2)  $i \longrightarrow \dot{i}$

(b -1)  $t \longrightarrow s$

(b -2)  $t \longrightarrow \theta$

(c -1)  $k \longrightarrow \overset{v}{c} \quad / \quad - \begin{bmatrix} -\text{cons} \\ -\text{post} \end{bmatrix}$

(c -2)  $\overset{v}{c} \longrightarrow *k \quad / \quad - \begin{bmatrix} -\text{cons} \\ +\text{post} \end{bmatrix}$

Chomsky et Halle affirment que dans ce cas (et d'autres analogues qui y sont présentés) l'exemple 1 comporte plus de traits distinctifs ou le même nombre de traits distinctifs que l'exemple 2 correspondant bien que, dans chaque paire, l'exemple 1 soit "plus attendu dans une grammaire que le cas (ii) (= exemple 2) et devrait par conséquent être "plus simple" en termes d'une mesure significative empirique d'évaluation" (32 p.401). Et plus loin ils ajoutent:

"Tous ces exemples, et beaucoup d'autres, marquent la nécessité d'élargir la théorie pour intégrer les effets du contenu intrinsèque des traits et pour distinguer les cas "attendus" ou "naturels" des règles et des configurations de symboles des autres qui sont non attendus et non naturels. Dans le sens linguistiquement significatif de la notion "complexité", une règle qui sonorise les voyelles ne devrait pas contribuer à la complexité de la grammaire mais une règle qui les assourdit devrait le faire, tandis que dans le cas des obstruantes la décision opposée serait requise" (32 p.402).

Dans la suite de leur raisonnement, les conséquences tirées des cas précédents sont claires: la GU universelle est modifiée et on y introduit la notion de marque afin que, dans les cas nécessaires, les GP qu'elle détermine soient définies comme plus simples.

Cette manière de raisonner implique qu'on dispose d'une MS "évidente", qui n'est plus considérée comme hypothèse empirique, l'interrogation ne portant pas sur elle mais sur la GU. C'est effectivement l'idée que l'on trouve explicitée dans d'autres textes:

"La mesure numérique évidente que l'on peut appliquer à une grammaire est la longueur, définie sur la base du nombre de symboles" (Chomsky 20b p.64/ 20 p.42).

"Probablement il est mieux de considérer une mesure d'évaluation comme un procédé pour assigner un entier à une grammaire, un entier plus petit correspondant à une valeur plus haute" (Chomsky et Halle 31 p.108).

"La "valeur" d'une suite de règles est la réciproque du nombre de symboles dans la représentation minimale" (Chomsky et Halle 32 p.334).

Le schéma de raisonnement qui se dégage, schéma extrêmement important, car il est à la base de la justification de la grammaire universelle, est donc le suivant:

1) Deux GP (GP-1 et GP-2) sont données; elles rendent compte des données différentes mais sont déterminées par la même GU-A.

2) On sait qu'une de ces grammaires doit posséder une généralisation (ou, avec une terminologie encore plus métaphorique, qu'elle est plus "naturelle" ou "attendue"), ce qui n'est pas le cas pour l'autre.

3) On dispose d'une MS postulée et admise.

4) La GU qui détermine GP-1 et GP-2 est une GU adéquate- ou, en d'autres termes, est une GU susceptible d'exprimer une généralisation- si la GP censée posséder une généralisation est définie comme plus simple par la MS; dans le cas contraire, GU n'est pas adéquate et elle doit être modifiée.

Ce schéma de raisonnement a également été utilisé dans le domaine syntaxique, par exemple dans Aspects (p.128-134/p.90-95) à propos des règles de sous-catégorisation. La Formulation VI y est concrètement utilisée de la manière suivante (les numéros entre parenthèses, sauf indication contraire, correspondant aux numéros utilisés dans Aspects pour individualiser les règles ou les



conventions):

GP-1: Règle (34)

GP-2: Règle (35)

GU-A: Convention de notation des traits de sous-catégorisation. Les traits de sous-catégorisation sont explicités et à droite de la règle qui opère sur (+V) et dans les indications du contexte de (+V).

MS : Non entièrement explicitée mais fondée sur le nombre de symboles utilisés dans GP.

On sait : GP-1 doit posséder une généralisation, qui est absente de GP-2.

On constate: par application de MS, GP-1 n'est pas définie comme plus simple que GP-2.

Conclusion: GU-A doit être modifiée.

Le même texte présente la Formulation VI utilisée en sens contraire, c'est à dire, dans le sens de la confirmation de la GU: GP-1: Règles (40) et (42); la règle (42) est équivalente à la règle (34) de l'exemple précédent.

GP-2: Règle (40-bis) (non explicitée dans le texte; devrait être comparée à la règle (40)) et règle (35), comparée à (42).

GU-B: Convention de notation des traits de sous-catégorisation (convention (36)). Les traits ne sont explicités que dans les indications de contexte de la règle qui opère sur (+V).

MS: Non entièrement explicitée mais fondée sur le nombre de symboles utilisés dans GP et identique à la MS de la règle précédente.

On sait: GP-1 doit posséder une généralisation, qui est absente de GP-2.

On constate: par application de MS, GP-1 est définie comme plus simple que GP-2.

Conclusion: GU-B est adéquate ou, plus exactement, plus adéquate que GU-A (21).

La MS de la formulation VI est formelle (22). Elle l'est non pas parce que les notions de "posséder une généralisation" ou "être naturel", etc., qui doivent qualifier les grammaires particulières sont peu sûres, non objectives, etc. Le statut douteux de

ces notions ne devrait permettre que de classer la MS comme éventuellement empirique. Mais la MS de la Formulation VI est formelle parce qu'elle ne peut être testée par aucune donnée; même si les notions qui qualifient les grammaires particulières accédaient, dans des formulations ultérieures, à la catégorie de données, la MS continuerait à être formelle. Dans le schéma de raisonnement où elle prend place, la MS est acceptée comme un a priori qui n'est pas à discuter; c'est en revanche la GU qui est discutée à partir de cet a priori non empirique.

Cette conclusion est par conséquent différente de celle qui a été dégagée par Botha, bien que nous arrivions tous deux à considérer la MS de la Formulation VI comme n'étant pas empirique. En effet, Botha, après avoir examiné une série d'exemples empruntés à Chomsky et Halle - et qui rentrent tous dans la Formulation VI conclut à ce que (14 p.186) "Chomsky et Halle n'appuient pas leurs affirmations sur le caractère naturel ou attendu des catégories phonologiques, des règles et des systèmes phonologiques, avec des faits dont la fiabilité et la pertinence puissent être contrôlés. Par conséquent, il n'est pas possible de prétendre que ces affirmations démontrent la nature empirique de leur conception de valeur" (déterminée, rappelons-le, par la MS). Or, cette constatation ne suffirait, dans notre cadre conceptuel, qu'à considérer la MS comme éventuellement empirique. Si nous la considérons comme formelle, c'est parce que, comme il vient d'être expliqué, elle est posée a priori et c'est à la GU de s'y adapter.

La divergence avec l'interprétation de la MS par Botha vient de ceci: la MS effectivement employée par Chomsky et Halle dans les exemples cités par Botha n'est pas la MS dont il présente les traits quelques pages auparavant, traits qui traduisent les déclarations de Chomsky et Halle sur la MS mais non l'utilisation qu'ils en font. En effet, Botha (14 p.178) indique que "la fonction de la mesure d'évaluation (=MS) est de sélectionner la grammaire évaluée comme plus haute, pour un langage, parmi les alternatives qui ont été construites par application de la théorie linguistique générale aux données linguistiques". Cette MS sélectionne donc des hypothèses en concurrence qui doivent rendre compte des mêmes données et cette exigence sur l'identité des données dont les GP

doivent rendre compte est explicite dans les textes de Chomsky et Halle (cf. 32 p.168 et, plus particulièrement, p.251 et 331) cités par Botha; sa formulation de p.178 est par conséquent en accord avec les textes originaux. Mais dans tous les exemples qu'il étudie (ibidem p.183-186) la MS a été appliquée par les linguistes américains à des GP qui rendent compte des données différentes. Par ailleurs, dans cette application de la MS- qui relève de la Formulation VI- ce n'est pas la MS qui est en jeu; c'est la GU qui doit être modifiée afin que la MS -inchangée- puisse définir comme plus simple les GP qu'elle détermine et qui sont censées être plus "naturelles" ou posséder une généralisation(23).

La MS formelle de la Formulation VI introduit une singularité très importante parmi les MS formelles; les MS formelles des Formulations I à IV définissent comme plus simples des entités qui sont entièrement équivalentes ou équivalentes par rapport à un langage, que ces entités soient des GU ou des GP. Dans tous ces cas, la MS s'applique selon la tradition du "ceteris paribus", plus ou moins explicite, dans le recours traditionnel à la simplicité, c'est à dire, comme proposition non empirique venant à discriminer, dans une science empirique, entre plusieurs hypothèses concurrentes portant sur un même domaine. Or, dans la Formulation VI, la MS joue un rôle infiniment plus important et qui, en fait, fait subir des distorsions au caractère empirique de toutes les hypothèses de la théorie. En effet, si l'on accepte la Formulation VI, il faut accepter que l'on ne puisse plus formuler les objectifs couramment annoncés en grammaire générative; il faudra expliquer que la linguistique est, certes, la discipline qui se propose de rendre compte de la compétence et de la faculté du langage mais qu'elle doit faire ceci par la formulation d'une GU telle qu'elle détermine des GP définies dans certaines conditions comme plus simples que d'autres GP par une certaine MS postulée a priori. La MS devient ainsi une "règle de jeu" qui est, d'une part, impossible à justifier et qui, d'autre part, compromet le caractère empirique du reste des hypothèses.

La formulation VI répond à une lecture stricte des utilisations concrètes de la MS dans un certain nombre de raisonnements clefs qui ont été faits en grammaire générative pour

justifier telle ou telle solution incorporée à la GU. On peut se demander si une lecture différente d'autres textes, qui ont des rapports avec la formulation VI mais qui en diffèrent profondément, ne serait pas susceptible de faire mieux comprendre, sinon les utilisations effectives, du moins les intentions des linguistes américains sur la simplicité.

Dans Aspects (20b p.64-66 / 20 p.43-44) on trouve un exemple factice comportant les éléments qui suivent:

GP-1 (rend compte du Lang-1): Règle (15) possédant 4 symboles.

GP-2 (rend compte du Lang-2): Règle (15-bis), non explicitée dans le texte, possédant 20 symboles.

GU-A: Règles syntagmatiques avec certaines conventions de notation (utilisation de parenthèses).

MS: non entièrement explicitée mais fondée sur le nombre de symboles utilisés dans les GP.

Lang-1: item(16).

Lang-2: item(17).

Lang-1 et Lang-2 sont différents mais possèdent le même nombre (vingt) de symboles.

On constate: a) GP-1 est spécifiée comme plus simple que GP-2.

b) Le nombre de symboles de GP-2 est identique au nombre de symboles du Lang-2.

On conclut: Nombre de symboles de GP-2 moins Nombre de symboles de GP-1 = degré de généralisation de GP-1. Cette conclusion est explicitée dans le texte qui suit:

La notation parenthétique a dans ce cas la signification suivante: elle dit que la différence entre quatre et vingt symboles est une mesure du degré de généralisation linguistique (significative-dans-l'original) qui est atteint dans une langue ayant pour le syntagme Auxiliaire les formes données dans la liste(16) (=Lang-1), par opposition à une langue ayant par exemple, pour représenter cette catégorie, les formes de la liste (17) (=Lang-2)".

[...] De ce fait, l'adéquation (l'adoption, dans l'original) des conventions de notations usuelles revient à poser que, sous-jacente à l'ensemble des formes de la liste (16) (=Lang-1), il y a une généralisation linguistique (significative dans l'original) et que ce n'est pas vrai pour l'ensemble des formes de (17) (=Lang-2).

Chomsky souligne que le fait que GP-1 possède une généralisation par rapport à GP-2 est le résultat de l'adoption d'une certaine GU-A caractérisée par certaines conventions de notation; une GU caractérisée autrement serait susceptible de déterminer la GP rendant compte du Lang-2 comme plus simple que la GP qui rend compte du Lang-1; on dirait dans ce cas que celle -là mais non celle-ci possède une généralisation:

"L'adoption des conventions de notation revient à formuler l'hypothèse empirique que les régularités du type illustré en (16) (=Lang-1) sont celles que l'on trouve dans les langues naturelles et sont du type attendu par les enfants apprenant une langue, et qu'au contraire, les régularités cycliques du type illustré en (17) (=Lang-2), encore que prises abstraitement, elles soient parfaitement authentiques, ne sont pas caractéristiques des langues naturelles, ne sont pas du type que les enfants rechercheront intuitivement dans le matériel linguistique, et pour le sujet apprenant la langue sont bien plus difficiles à employer ou à reconstruire à partir de données dispersées. Ce que l'on pose, dès lors, c'est qu'à partir d'exemples dispersés, tirés de (16), le sujet apprenant la langue construira la règle (15), engendrant l'ensemble entier avec ses interprétations sémantiques, alors qu'à partir d'exemples dispersés qui pourraient être regroupés par une règle cyclique, il n'incorporera pas cette "généralisation" à sa grammaire[...]. Il serait aisé de suggérer une convention de notation différente qui abrègerait la liste (17) en une règle plus courte que le liste (16), et, ce faisant, de formuler une hypothèse de travail empirique différente sur ce qui constitue une généralisation linguistique".

Si, suivant ces suggestions de Chomsky, on modifie la GU, l'exemple précédent devient:

GP-3 (rend compte du Lang-1, équivalente à GP-1): non-explicitée dans le texte.

GP-4 (rend compte du Lang-2, équivalente à GP-2): non explicitée dans le texte.

GU-B: non explicitée dans le texte; conventions de notation différentes de celles de GU-A et qui "favorisent" les régularités cycliques.

MS: non entièrement explicitée mais fondée sur le nombre de symboles utilisés dans les GP. Identique à celle de l'exemple précédent.

On constate: GP-4 est spécifiée comme plus simple que GP-3.

On conclut: GP-4 possède une généralisation.

Le texte précédemment transcrit de Chomsky et celui qui suit insistent sur ceci: la conclusion selon laquelle GP-1 (qui rend compte du Lang-1) possède une généralisation et non pas GP-4 (qui rend compte du Lang-2) relève d'une hypothèse empirique sur GU-A et sur GU-B:

"Pour rapprocher la théorie linguistique de l'adéquation explicative, nous pouvons essayer de raffiner la mesure d'évaluation des grammaires ou de resserrer les contraintes formelles pesant sur elles, de façon qu'il devienne plus difficile de découvrir une hypothèse de valeur élevée qui soit compatible avec les données linguistiques primaires. Il n'est pas douteux que les théories actuelles de la grammaire ont besoin d'être modifiées de ces deux manières à la fois, la seconde étant en général la plus prometteuse" [...]. De cette façon, la théorie linguistique pourra se rapprocher de l'adéquation explicative et contribuer à l'étude des processus mentaux et de la capacité intellectuelle de l'homme - plus précisément, à la détermination des capacités qui rendent possible l'apprentissage linguistique, eu égard aux limitations empiriquement constatées de temps et de données" (20b p. 64 /20 p. 46-47).

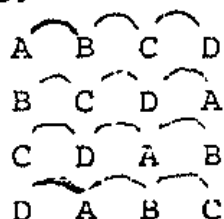
Dans quel sens peut-on dire que la conclusion selon laquelle GP-1 - et non pas GP-2- possède une généralisation est une hypothèse empirique? Quel type de données peut tester cette hypothèse? Le texte cité ouvre la porte à deux types d'observation

a) Les régularités illustrées en (16) (=Lang-1) sont celles que l'on trouve dans les langues naturelles, contrairement aux régularités cycliques illustrées dans (17) (=Lang-2).

b) Les régularités illustrées en (16) sont celles qui déterminent un comportement particulier au cours de l'acquisition

Il faut souligner que les observations de type (a) sont difficilement acceptables comme données: comment peut-on déterminer l'existence ou non d'une régularité dans un langage?

Pour illustrer le problème, soit l'exemple factice du Lang-2 présenté par Chomsky. Il obéit presque entièrement au schéma suivant:



Or, on peut affirmer qu'on trouve ce type de langage cyclique dans les langues où l'ordre des syntagmes n'est pas fixe, par exemple l'espagnol, où les énoncés qui suivent sont grammaticaux:

la niña come un chocolate en el campo  
 come un chocolate en el campo la niña  
 un chocolate en el campo la niña come  
 en el campo la niña come un chocolate  
 ("la fille mange un chocolat à la campagne").

Mais il paraît difficile que le fait d'attester l'existence de ce langage cyclique valide empiriquement la GU-B et la MS de l'exemple factice précédent. Le texte de Chomsky doit donc être compris comme exigeant, pour pouvoir conclure à l'existence d'une régularité dans un langage, à la fois l'observation d'un type caractérisé de distribution des éléments dans les suites du langage et aussi l'observation d'un comportement spécial

de la part des usagers par rapport à ces suites au cours de l'acquisition , en particulier, la fait qu'ils soient capables de maîtriser tout le langage à partir des "exemples dispersés". La formulation qui se dégage de cet exemple factice ainsi interprétée est donc la suivante:

Formulation VI bis

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 rend compte du Lang-1 et GP-2 du Lang-2.

Les êtres humains dans le processus d'acquisition ont un comportement différencié par rapport au Lang-1 et au Lang-2, la maîtrise du Lang-1 étant susceptible d'être acquise à partir des exemples dispersés, ce qui n'est pas le cas du Lang-2.

MS caractérise GP-1 comme plus simple que GP-2.

Par conséquent, GU-A et MS sont des hypothèses validées sur le plan explicatif et sont susceptibles de rendre compte de l'acquisition du langage.

Si MS ne caractérisait pas GP-1 comme plus simple que GP-2, il faudrait:

(a) modifier la MS et/ou (b) modifier la GU et présenter une autre destinée comme la première à s'intégrer avec une MS dans un modèle d'acquisition ou (c) modifier la GU et présenter une autre qui opérerait sans MS ou (d) présenter un modèle tout à fait différent de l'acquisition du langage.

Dans la Formulation VI bis, MS et GU seraient des hypothèses empiriques; elles auraient par ailleurs une interprétation psychologique dans le domaine de l'acquisition du langage, ceci étant en rapport avec cela. En effet, la donnée clef qui rend la Formulation VI bis empirique, est l'observation du comportement dans le processus d'acquisition: c'est cette adéquation externe, accompagnée de la suppression du caractère a priori accordée à la MS , qui permet de parler de formulation empirique. Or, le constat qu'on est amené à faire à propos de la Formulation VI bis est le suivant: nous ne connaissons aucun cas où la formulation ait été appliquée à une situation attestée. Une pré-condition indispensable pour qu'elle soit appliquée serait de définir les restrictions dans les rapports que



doivent entretenir les données différentes dont GP-1 et GP-2 rendent respectivement compte; le même problème apparaît aussi pour appliquer la Formulation VI et, autant qu'on sache, n'a jamais été posé. Mais par ailleurs, il existe une question plus générale et plus fondamentale: dans le cadre heuristico-méthodologique que la grammaire générative s'est donné pour envisager le problème de l'acquisition du langage, la Formulation VI bis et le schéma de recherche qui en découle ne peuvent pas être appliqués. On présentera les fondements de cette dernière affirmation au § 5.

#### IV Simplicité et empirisme.

Les MS <sup>formelles</sup> ou éventuellement empiriques n'épuisent pas les MS de la grammaire générative. Des essais d'utilisation de MS empiriques ont eu lieu et ils seront examinés ici (Formulations VII et VIII). On donnera d'abord, de là première, une formulation générale; celle-ci apparaît appliquée selon des modalités différentes, qui seront présentées par la suite comme des types différentes de la Formulation VII. La présentation de ces types sera précédée d'un classement des données linguistiques qui permettra de les comprendre. Dans la discussion de chaque formulation, on essayera de dégager si elle est suffisamment explicite pour être appliquée et si, une fois appliquée et en tant qu'hypothèse empirique, elle a été validée ou infirmée. On présentera finalement la formulation VIII.

##### IV-1 Formulation VII

Soit GU-A, qui détermine GP-1 et GP-2.

GP-1 et GP-2 ne sont pas entièrement équivalentes mais elles sont équivalentes par rapport à un Lang-1.

GP-1 rend compte aussi du Lang-2, ce que ne fait pas GP-2.

MS caractérise GP-1 comme plus simple que GP-2.

Par conséquent, GU-A et MS, déterminant GP-1, qui est susceptible de projeter du Lang-1 au Lang-2, constituent un mécanisme de projection adéquat.

Si MS ne caractérisait pas GP-1 comme plus simple que GP-2, MS et/ou GU-A seraient infirmées et devraient être modifiées.

Cette formulation VII est fondée sur les textes qui suivent:

"Etant donné une telle théorie, nous pouvons essayer de construire des grammaires pour les langues existantes et nous pouvons déterminer si les grammaires les plus simples que nous pouvons trouver (c'est à dire les grammaires que la théorie générale nous contraint de choisir) satisfont ou non aux conditions externes d'adéquation. Nous continuerons à réviser notre notion de simplicité et notre caractérisation de la forme des grammaires jusqu'à ce que les grammaires choisies par la théorie satisfassent effectivement aux conditions externes" (29b p.60/ 29 p.54).

"[...]

Nous essaierons de montrer que les grammaires les plus simples satisfont à certaines conditions externes d'adéquation alors que les grammaires plus complexes, qui impliquent des décisions différentes relatives à l'attribution de phrases au noyau, etc., ne les remplissent pas" (29b p.61/29 p.55).

"D'abord, nous essayons d'enrichir la structure de la théorie linguistique de manière à restreindre la classe des grammaires qui sont compatibles avec les données présentées- en d'autres termes, nous essayons de faire l'affirmation universelle plus forte et légitime sur la structure du langage. Deuxièmement, nous essayons de construire une procédure d'évaluation susceptible de sélectionner une parmi la pluralité de grammaires permises par la théorie linguistique proposée et compatibles avec les données présentées" (31 p.107).

"Pour construire un modèle raisonnable de l'acquisition, il est nécessaire de réduire la classe des grammaires à la fois réalisables et compatibles avec les données linguistiques primaires, jusqu'au point où nous pourrions faire usage, pour choisir entre elles d'une mesure formelle d'évaluation (20b p.56/20 p.35).

Ces citations, et d'autres que l'on pourrait ajouter confirment les points essentiels de la Formulation VII: les GP comparées doivent rendre compte d'un même ensemble de données et être déterminées par la même GU. Par ailleurs, ces citations sont englobées dans la problématique de l'adéquation explicative de la théorie linguistique, problématique qui coïncide, dans l'optique chomskyenne, avec celle de l'acquisition du langage. C'est d'une part ce cadre-là et d'autre part, les utilisations concrètes de cette MS, qui autorisent à introduire les autres traits propres à la Formulation VII et notamment, la capacité projective de la MS opérant sur des GP déterminées par une GU donnée (24). En effet la MS et la GU constituent une "machine de projection", celle-ci pouvant se définir comme une machine qui reçoit comme entrée des données  $\underline{d}$  et qui produit comme sortie une GP susceptible de rendre compte de  $\underline{d}$  et, en plus, d'autres données  $\underline{d}'$ ; les données  $\underline{d}$  et  $\underline{d}'$  sont incluses dans une classe  $\underline{D}$  et l'intersection de  $\underline{d}$  et  $\underline{d}'$  est nulle. Les données  $\underline{d}$  et  $\underline{d}'$  seront respectivement nommées données originales et données projectives. La MS et la GU sont une machine de projection car elles permettent d'obtenir une GP susceptible de rendre compte des données qui leur sont présentées (données originales, le Lang-1) et, en plus, d'autres données (données projectives, le Lang-2).

MS et GU établissent donc, comme toute machine de projection, un rapport entre les données originales et les données projectives. Une manière de caractériser les différents doublets des MS et des GU ( $MS_1 \widehat{GU}_1 \dots MS_n \widehat{GU}_n$ ) est de classer les types des données originales et le type des données projectives entre lesquelles chaque  $MS_i \widehat{GU}_i$  établit un rapport. Ceci exige un classement des données dont une GP doit rendre compte.

Nous distinguons deux types de données linguistiques: les données observables et les données profondes. Données observables sont celles qui se révèlent dans les signifiants des énoncés et peuvent par conséquent être captées à partir des indices perceptibles. Tout ce qui a trait aux "sons", au sens le plus large, relève des données observables: les unités phoniques elles-mêmes, mais aussi leur linéarité, l'agencement des unités

phoniques en signifiants, la linéarité des signifiants, etc. Les données observables n'excluent pas les considérations sur le signifié, pourvu que celles-ci soient rattachables à des différences perceptibles dans le signifiant. Les jugements de grammaticalité, d'être identique ou différent à l'audition, de déviation morphologique ou syntaxique en surface, sont autant des données observables. Les données profondes sont celles qui ne sont pas susceptibles d'être captées par un processus perceptif ou par l'observation des signifiants des énoncés. Les jugements d'ambiguïté syntaxique, d'identité de paraphrase et/ou des rapports syntaxiques non manifestés en surface, sont des données profondes. Les données profondes, en lignes générales, sont celles qui ne respectent pas la consigne structuraliste selon laquelle à toute différence de sens devrait correspondre une différence de son. Ceci est typiquement le cas des ambiguïtés syntaxiques, car aucune distinction dans le signifiant ne correspond à la distinction de signifié propre à l'ambiguïté. Les données profondes ont une relation avec les données observables correspondantes; néanmoins et par la définition de la notion même de données profondes, l'observation des données observables qui véhiculent les données profondes, ne permet pas de déterminer l'existence de celles-ci. Par exemple, dans "Il m'a expédié le message de la pharmacie", rien, dans le signifiant de la pharmacie ou dans le reste de l'énoncé, ne permet de conclure à l'ambiguïté de l'énoncé et, en particulier, à l'ambiguïté des rapports syntaxiques de de la pharmacie. Mais cette donnée profonde qu'est l'ambiguïté syntaxique, a, cependant, une relation avec le signifiant qui la véhicule, c'est à dire avec il m'a expédié le message de la pharmacie. On nommera "donnée observable correspondante" la donnée observable qui est en relation avec une donnée profonde. On établit la convention que l'expression données profondes désigne et les données profondes et les données observables correspondantes.

La catégorisation en données observables et en données profondes se superpose avec la catégorisation en données originales et données projectives.

On a, par conséquent, les quatre types de données qui suivent:

- A: originales et observables
- B: originales et profondes
- C: projectives et observables
- D: projectives et profondes

La projection doit donc s'effectuer entre données de la catégorie A ou B et données de la catégorie C ou D. Or, la catégorie D peut se subdiviser en deux sous-catégories. En effet, on peut distinguer les deux situations suivantes:

a) Situation D-1 ou de coïncidence des données observables.

Les données observables correspondantes (c'est à dire les données observables qui sont en relation avec les données profondes) sont les mêmes données observables que celles de données originales à partir desquelles la projection s'est effectuée.

b) Situation D-2 ou de non-coïncidence des données observables.

Les données observables correspondantes ne sont pas les mêmes données observables que celles des données originales à partir desquelles la projection s'est effectuée.

On a par conséquent:

	données originales		Données observables correspondantes	Données profondes projectives
Situation D-1	W	projectées en	W	Z
Situation D-2	W	projectées en	Y	Z

Il est possible, à partir de ce classement des données, de caractériser les différents doublets de MS et de GU en tant que mécanismes projecteurs. Chaque MS GU sera caractérisé par le type de données originales et le type de données projectives qu'il met en rapport. Or, nous avons pu repérer, dans les écrits de la grammaire générative, au moins trois types de projection différents où mention explicite est faite du rôle de la MS:

Type I : De données A (originales et observables) à données C (projectives et observables).

Type II: De données A (originales et observables) à données D-1 (projectives profondes et coïncidence des données observables)

Type III: De données A (originales et observables) à données D-2 (projectives profondes et non coïncidence des données observables).

Chaque type de projection définit une modalité différente de la Formulation VII. Ces trois modalités seront successivement examinées par la suite.

#### IV-1-1 Formulation VII-Type I

Le Type de la formulation VII apparaît en relation avec les variations morphologiques des matrices phoniques d'un même formant. Les données A (observables) sont constituées par les variations morphophonologiques régulières constatées dans certains contextes et pour certains items lexicaux; les données C (projectives) sont constituées par d'autres variations morphophonologiques qui apparaissent dans d'autres contextes, ces nouvelles variations et/ou ces nouveaux contextes n'ayant pas été constatés parmi les données A. Le texte de Chomsky (21 p.65) qui suit, illustre cette possibilité:

"La grammaire parviendrait au niveau encore plus haut d'adéquation explicative[...] si la théorie linguistique qui lui est associée était susceptible de pourvoir une structure pour les règles phonologiques et une mesure d'évaluation satisfaisant la condition suivante: l'ensemble de règles évalué comme plus haut et de la forme appropriée, ayant été sélectionné pour engendrer un ensemble d'items où les variantes de "telegraph" sont exclues, serait l'ensemble de règles qui en fait prédit cette variation contextuelle de "telegraph"."

Nous ne connaissons aucune formulation explicite de MS et de GU où cette possibilité ait été concrétisée. Elle est plus un programme de formulation empirique qu'une formulation explicite;

impossible donc de contrôler si, dans un certain nombre de cas, une hypothèse quelconque relevant de cette formulation se vérifie ou non.

Le Type I de la Formulation VII apparaît aussi ailleurs, mais explicité dans une formulation suffisamment claire pour être testée. Il est appliqué au problème de séquences phoniques non attestées mais possibles. En anglais, par exemple, brick, blik et bnik sont, respectivement: (a) une séquence phonique attestée dans un corpus de l'anglais; (b) une séquence non attestée mais possible; blik est le signifiant d'un mot qui n'a pas de signifié mais qu'un anglophone admettrait comme séquence possible, susceptible d'appartenir à l'anglais et de recevoir un sens; (c) une séquence non attestée mais impossible, bnik n'étant pas accepté par un anglophone comme le signifiant d'un mot possible dans sa langue. Par définition donc, aussi bien les séquences non attestées possibles, ou absences accidentelles, que les séquences non attestées impossibles ou absences systématiques, ne peuvent pas appartenir aux données observables: ceci pose le problème de projeter la distinction entre absences accidentelles et absences systématiques à partir des séquences attestées.

Le doublet MS GU, suffisamment explicite et qui a été proposé en grammaire générative pour résoudre le problème ainsi posé, présente les caractéristiques suivantes:

GU: (a) le signifiant des formants du lexique est défini par une matrice de traits distinctifs binaires.

(b) Les règles de redondance lexicale incorporent aux traits des matrices lexicales des traits redondants; ces règles ont, d'une part, certaines capacités formelles, et d'autre part, opèrent sur des traits distinctifs binaires.

MS :GP-1 est plus simple que GP-2 si le nombre des traits distinctifs binaires utilisés aussi bien dans la spécification du lexique que dans les règles de redondance est moins élevé dans GP-1 que dans GP-2.

Le mécanisme de projection, dans cette hypothèse de doublet MS  $\overline{GU}$  (désormais MS GU-A), est constitué par la présence ou absence d'une règle de redondance. Par exemple, soit les règles

(empruntées à Chomsky et Halle ,31 p.101)

- (1) Segment consonantique → Liquide  
dans le contexte: ~~≠~~ Occlusif- voyelle
- (2) Segment consonantique → /r/  
dans le contexte: ~~≠~~ /b-ik/

La première règle exprime le fait qu'entre consonne occlusive initiale et voyelle il n'y a en anglais que /r/ ou /l/; elle permet d'épargner, dans le lexique, la spécification de "liquide" du deuxième élément phonique, pour tous les formants tels que craft, blend, prize, etc. La deuxième règle exprime le fait qu'entre b et ik on ne trouve que /r/ dans le lexique anglais; par là même, elle exclut /blik/ parmi les séquences non attestées mais possibles. Les deux règles sont adéquates par rapport aux données originales et observables, mais c'est la seule règle (1) qui établit la projection de manière adéquate n'excluant pas le cas du possible blik, ce qui, en revanche, fait la règle (2). C'est donc la présence ou l'absence d'une règle de redondance qui conditionne la projection des séquences non attestées possibles.

Les règles de redondance opèrent conjointement avec un lexique, et un même langage peut être spécifié modifiant soit le lexique soit les règles de redondance, soit les deux, les GP qu'on obtiendrait modifiant ces éléments pouvant être équivalentes par rapport à un langage déterminé. Par exemple, dans le cas de l'exemple suivant, on aurait:

GP-1:

- (a) Lexique sans indication de "liquide" dans le contexte: consonne-voyelle.
- (b) Règle de redondance (1).

GP-2:

- (a) Lexique avec indication de liquide pour tous les formants dans le contexte: consonne-voyelle.
- (b) Aucune règle de redondance équivalente à la (1).

GP-1 et GP-2 sont équivalentes par rapport à un langage constitué de séquences douées de signifié.



GP-3 :

(a) Lexique avec indication des traits spécifiques de /r/ dans la matrice de brick.

(b) aucune règle de redondance équivalente à la règle (2).

GP-4:

(a) Lexique sans indication des traits spécifiques de /r/ dans la matrice de brick.

(b) Règle de redondance (2).

GP-3 et GP-4 sont équivalentes par rapport à un langage constitué de séquences douées de signifié.

Les situations illustrées par les paires GP-1 et GP-2, GP-3 et GP-4, constituent un cas typique où la MS devrait opérer. Si on applique celle qui a été proposée dans MS  $\widehat{\text{GU-A}}$ , on observe que GP-1 et GP-3 sont correctement caractérisées comme plus simples que, respectivement, GP-2 et GP-4. En effet, dans le cas de GP-1 et de GP-2, le nombre des traits que l'on doit utiliser pour expliciter la règle (1) est moindre que le nombre des traits épargnés dans le lexique; la MS oblige donc à incorporer la règle, ce qui, comme il a été remarqué, est adéquat. En revanche, dans le cas de GP-3 et de GP-4 le nombre des traits que l'on doit utiliser pour expliciter la règle (2) est plus élevé que le seul trait épargné dans la matrice phonique du formatif brick: par conséquent la règle de redondance (2) ne doit pas être incorporée.

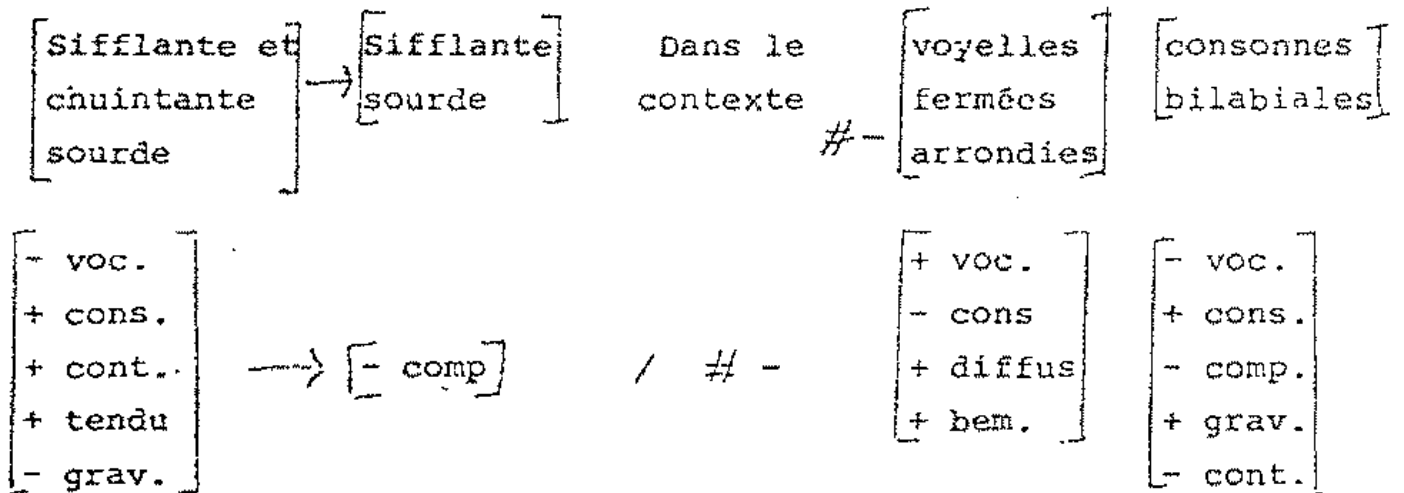
La MS du doublet MS  $\widehat{\text{GU-A}}$  est donc empirique: la donnée qui servira à la vérifier c'est la possibilité de distinguer les absences accidentelles des absences systématiques, parmi les séquences qui n'ont pas été observées.

Nous n'avons pas connaissance d'un processus actif de recherche destiné à proposer des grammaires particulières que la MS du doublet MS  $\widehat{\text{GU-A}}$  devrait caractériser comme plus simple. Exception faite de quelques exemples simplifiés à l'extrême, nous n'avons pas connaissance d'autres cas où cette MS ait été appliquée (25). Mais le fait qu'elle soit suffisamment explicite et

empirique permet de la tester.

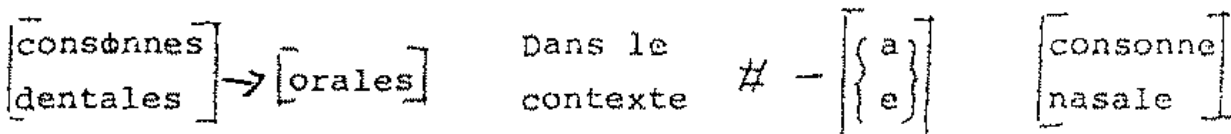
Il est relativement facile de trouver des contre exemples qui infirment la capacité projective de MS<sup>1</sup>GU-A. En voici trois, dont le premier est emprunté au français et les deux autres à l'espagnol.

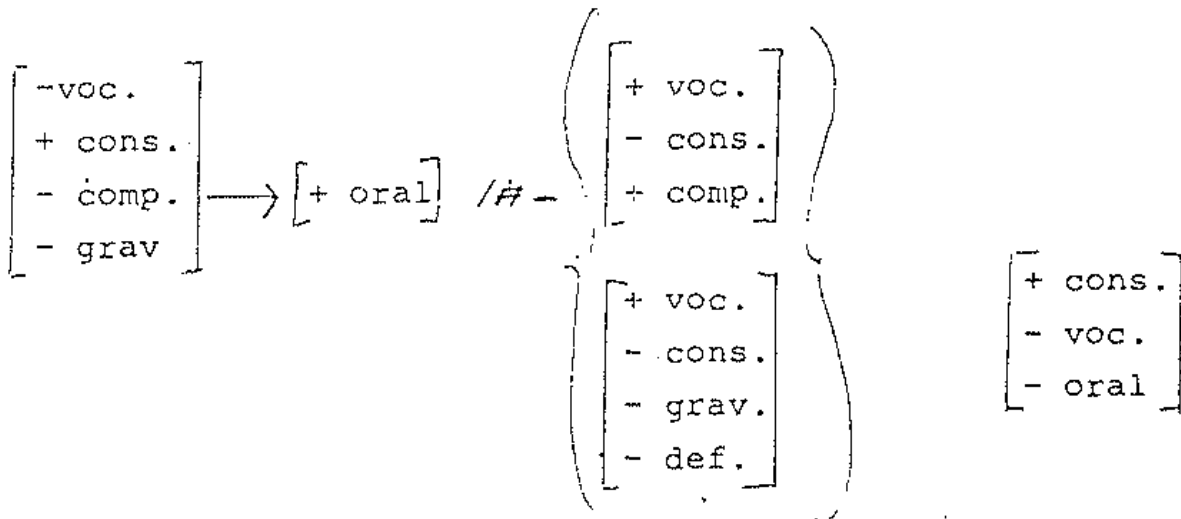
Il se trouve que le lexique du français est faiblement représenté dans le contexte chu- ou chou-; en revanche, les initiales su- et sou- abondent. Cette observation permet de formuler la règle de redondance suivante, dont nous donnons deux formulations, l'une plus technique que l'autre:



Cette règle comporte 15 traits; dans le dictionnaire consulté, le lexique du français ne comporte pas moins de 190 items où le trait qui distingue /s/ de /s<sup>v</sup>/ pourrait ne pas être indiqué; elle devrait par conséquent, selon MS GU-A, être incorporée. Or, cette règle classerait, par ex. les séquences chouquette, choumer, chumer, etc. parmi les non possibles, ce qui est inadéquat (26).

Le lexique espagnol présente peu de n initiales suivis de a ou de e et d'une consonne nasale. En revanche, on compte environ une centaine d'items avec t ou d dans cette position. On peut donc formuler la règle de redondance suivante:

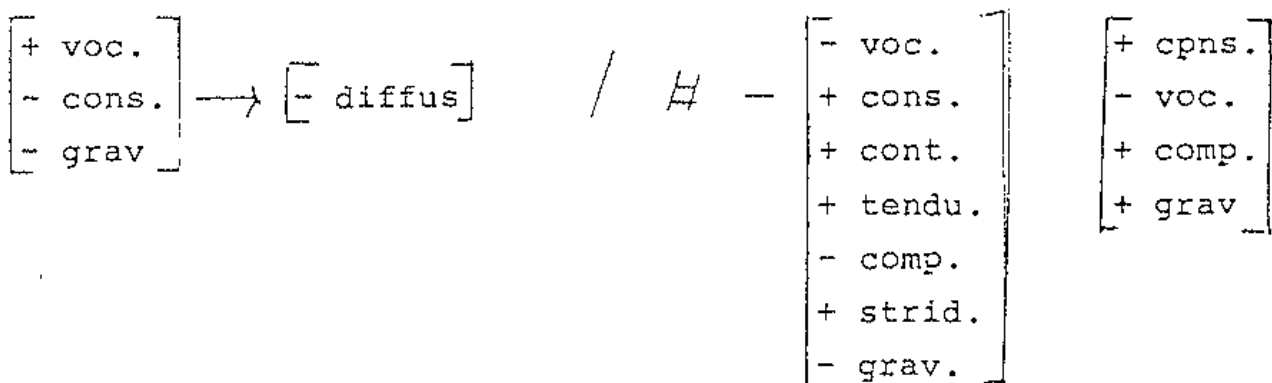
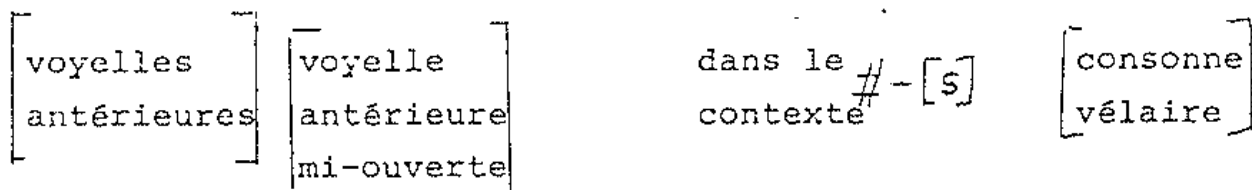




Cette règle, qui comporte 15 traits, devrait être incorporée. Cependant dans les données projectives, elle classe nanzarina, nentro, nentáculo, nemplado, etc., comme des absences systématiques, ce qui est inadéquat.

Le lexique espagnol présente peu de séquences initiales is ou his; en revanche la séquence es initiale et précédant une consonne vélaire est abondamment représentée (190 entrées).

La règle de redondance suivante, qui comporte 15 traits devrait par conséquent être incorporée:



Elle conduirait, cependant, à des résultats inadéquats, car des séquences telles que iscala, iscabeche seraient classées dans les données projectives comme des absences systématiques.

MS<sup>GU</sup>-A est la seule hypothèse explicite que nous connaissons dans le cadre de la Formulation VII-Type I. Dans la mesure où les seuls cas connus menant à des résultats corrects ne sont que des exemples largement artificiels, et comme dans les trois cas approchant de la complexité de deux langues naturelles elle s'est révélée fautive, on peut conclure que MS<sup>GU</sup>-A est une hypothèse qui est, certes, empirique mais, pour l'instant, complètement infirmée (27).

IV-1-2 Formulation VII-Type II. Le Type II de la Formulation VII- de données A, originales et observables, à des données D-1, données projectives, profondes qui présentent une coïncidence avec des données observables- apparaît utilisé dans certains traitements de l'ambiguïté et, plus généralement, dans des discussions sur le rapport entre structure de surface et structure profonde. Dans (29b p.97-98 /29 p.88-89), Chomsky remarque que l'énoncé (111) The shooting of the hunters peut être compris comme possédant le schéma syntaxique de (112 i) The growling of lions ou celui de (112 ii) The raising of flowers. Et il ajoute:

" En termes transformationnels, cependant, l'explication est claire et automatique. Une analyse précise de l'anglais montre que nous pouvons simplifier la grammaire en rejetant des syntagmes du type (111) et (112) hors du noyau et en les réintroduisant par une transformation.

Si l'on procède de cette manière, l'énoncé (111) The shooting of the hunters aura deux origines transformationnelles distinctes; il aura une représentation ambiguë au niveau transformationnel".

L'observation transcrite peut se reformuler dans les termes suivants: si l'on utilise une GU avec une composante de transformations, on peut vérifier que la grammaire plus simple qui rend compte de la grammaticalité de (111), (112-i) et (112-ii), ces suites étant considérées parmi les données originales et

observables, élimine ces énoncés de l'ensemble des énoncés noyaux et les spécifie au moyen des transformations. Or, ce faisant, elle assigne une double structure au signifiant (111), rendant ainsi compte de son ambiguïté. C'est-à-dire que la GP caractérisée comme plus simple rend compte non seulement des données originales et observables portant sur la grammaticalité des suites (111), (112i) et (112 ii) mais, aussi, de l'ambiguïté de (111), c'est à dire des données profondes dont la partie observable coïncide avec les données originales et observables. La GP déterminée par cette GU et caractérisée comme plus simple par la MS, est susceptible de projeter des données A, portant sur la grammaticalité des suites (111), (112 i) et (112 ii) à des données D-1, portant sur l'ambiguïté de (111); elle ajoute donc de la " profondeur" aux données observables.

Dans 21 (p.66-67) réapparaît un raisonnement analogue au cours de la discussion de (3) John is easy to please et de (4) John is eager to please. Chomsky observe que le niveau d'adéquation explicative serait atteint si la théorie était susceptible de sélectionner une grammaire adéquate sur le plan descriptif étant donné les énoncés (3), (4) et d'autres suites qui ont un rapport avec eux. L'adéquation descriptive coïncidant largement avec la notion de données profondes, on peut légitimement supposer que l'on demande à la GP de rendre compte du niveau profond des énoncés (3) et (4) lorsqu'on ne dispose comme données originales que de la partie observables de ces mêmes énoncés:

"Il y a de bonnes raisons de penser que dans les cas de (3) et de (4), la théorie de la grammaire transformationnelle peut s'approcher du niveau de l'adéquation explicative et présenter une explication partielle de l'intuition linguistique du locuteur. C'est à dire que la grammaire qui assigne les descriptions structurales correctes possède des généralisations qui ne sont pas exprimées par les grammaires qui ne pourvoient pas les descriptions structurales correctes et est par conséquent évaluée comme plus haute (c'est à dire, plus simple) dans un sens que l'on peut

apparemment préciser sans beaucoup de difficulté" (21p. 67).

Mis à part les textes qui précèdent, largement programmiques, on ne connaît aucune formulation explicité de MS, dans le cadre de la Formulation VII-Type II, permettant d'opérer conjointement avec l'une des GU proposées en grammaire générative. Il a donc été impossible de dégager un doublet MS GU suffisamment clair- comme il était le cas dans la Formulation précédente- pour être testé.

#### IV-1-3 Formulation VII-Type III.

Le Type III de la Formulation VII- de données A, originales et observables, à des données D-2, projectives profondes sans coïncidence des données observables- est suggéré dans le texte qui suit:

"Supposons que nous avons une grammaire transformationnelle de l'anglais construite pour engendrer de la manière la plus économique l'éventail complet de constructions adjectives à l'exclusion de comparatives. On peut montrer (renvoi à C.S. Smith, "A class of Complex Modifiers in English", dans Language 37 (1961), p.342-365) qu'une large variété de constructions comportant des comparatifs seront engendrées automatiquement par la grammaire avec la disposition correcte des ambiguïtés, cas particuliers et "exceptions" apparentes, si l'on ajoute à la grammaire au point approprié dans la séquence des règles ordonnées, la transformation généralisée qui forme la construction comparative la plus simple(...) C'est par conséquent un cas intéressant où l'on peut dire que la théorie générale de la grammaire transformationnelle offre une explication pour un amoncellement de données de surface assez désordonnées" (21 p.68).

Le texte qui précède s'interprète de la manière suivante: à partir des données originales et observables (des suites qui correspondent aux constructions adjectives et à la comparative dans sa forme plus simple) la GU et la MS proposées dans le cadre

de la grammaire générative sont susceptibles de proposer une GP telle qu'elle doit rendre compte des données précédentes et, en plus d'autres données profondes -comme les ambiguïtés- et ceci aussi bien pour les données A (constructions adjectives et la forme la plus simple de la comparative) que pour les autres données projectives (les autres comparatives que les plus simples). L'interprétation que Fodor, Bever et Garrett (39b) donnent du modèle d'acquisition du langage présenté par Chomsky dans Aspects (cf. 20 p. 30-31/ 20b p. 50-51), modèle qu'ils jugent inadéquat, rentre dans les termes de ce type III de la Formulation VII.

Hormis la suggestion du texte précédent, on ne connaît aucune formulation explicite de GU et de MS susceptible d'être testée.

#### 4-2 Formulation VIII

Soit GU-A, qui détermine GP-1, et GU-B, qui détermine GP-2.

GP-1 et GP-2 ne sont pas entièrement équivalentes mais elles sont équivalentes par rapport à un Lang-1.

GP-1 rend compte du Lang -2; ce qui n'est pas le cas de GP-2.

MS caractérise GP-1 comme plus simple que GP-2.

Par conséquent, GU-A doit être préférée à GU-B, et GU-A et MS sont validées GP-1 étant susceptible de projeter du Lang-1 au Lang-2.

Si MS ne caractérisait pas GP-1 comme plus simple que GP-2, MS et/ou GU-A seraient infirmées et devraient être modifiées.

Cette Formulation VIII est, comme la précédente, empirique, mais à la différence de la Formulation VII, les GP comparées par MS sont déterminées par des GU différentes. Elle a été appliquée par Halle dans le domaine phonique, plus précisément dans les règles accentuelles de l'anglais (49 et 46)<sup>28</sup>. La présentation qu'on peut en dégager est la suivante :

GU-A : les règles sont organisées selon un ordre disjoint.

GU-B : règles non ordonnées.

GP-1 : item (6) dans une version non formalisée et item (9) dans une version formalisée.

GP-2 : item (3) dans une version non formalisée et item (4) dans une version formalisée.

- Lang-1 : liste des mots présentés dans l'item (2), qui comporte 23 mots.
- Lang-2 : (a) mots du type venom, avec voyelle accentuée non tendue, qui reçoivent l'accent principal dans l'avant-dernière position (item (12), qui comporte 6 mots).
- (b) monosyllabes avec voyelle non tendue, qui deviennent accentuées.
- (c) mots du type Suez, qui reçoivent l'accent principal dans la dernière syllabe avec voyelle non-tendue (item (13), qui comporte 6 mots).
- MS : exprimée dans l'item (11) par le terme "prolixité" ; elle porte sur le nombre de symboles.

La projection de type (c) dans le Lang-2 ne peut être atteinte que si l'on ajoute à GP-1 (constituée par l'item (6) ou (9), une "règle d'exception", non formulée dans le texte, mais qui, par les commentaires qu'on y trouve, profite de l'ordre disjoint de GP-1 : si un mot est marqué comme exception à une des règles qui sont abrégées dans l'item (9), on lui applique automatiquement la règle suivante. On a par conséquent : GP-1 bis : GP-1 (=item (6) ou (9)) plus règle d'exception.

Aussi bien GP-1 (et GP-1 bis) que GP-2, rendent compte du Lang-1, c'est-à-dire qu'elles placent l'accent principal dans les mots anglais du Lang-1 en tenant compte des caractéristiques phoniques des segments qui composent leurs signifiants et en particulier, du caractère tendu ou non des voyelles dans les successions des voyelles et des consonnes <sup>29</sup>. Trois positions ont été prévues à partir de la dernière syllabe ; elles apparaissent dans les exemples qui suivent, empruntés au Lang-1 :

Posit. 3	Posit. 2		Posit. 1
A <u>me</u> rica	Ariz <u>o</u> na	Wis <u>o</u> nsin	Tenn <u>es</u> see
al <u>u</u> min <u>u</u> m	arb <u>o</u> ret <u>u</u> m	carb <u>o</u> ru <u>u</u> nd <u>u</u> m	sup <u>re</u> me

Pour pouvoir considérer que GU-A avec MS sont vérifiées, il faudrait ne pas pouvoir accumuler les contre-exemples. Or, la variété des contre-exemples est si grande, ils sont si faciles à trouver, parce que



nombreux, que la MS et/ou la GU proposée apparaissent comme complètement infirmées. Et il paraît impossible de considérer ces contre-exemples comme des "exceptions", quelle que soit la notion d'exception qu'on se donne. On peut en dégager trois types :

(A) Contre-exemples à la capacité générative commune à GP-1 (et à GP-1 bis) et à GP-2. Ce sont les cas où l'accent principal apparaît placé en dehors des situations attestées dans le Lang-1. Toutes les possibilités combinatoires de contre-exemples possibles sont attestées pour les trois dernières positions ; en plus, on trouve des contre-exemples dans des mots qui placent l'accent principal en dehors des trois dernières positions prévues :

(A-1) Posit. 4 attestée, alors que posit. 3 ou 2 ou 1 prévues :

dedicatory                      degeneracy

(A-2) Posit. 3 attestée, alors que posit. 2 prévue :

concretely                      flippancy

(A-3) Posit. 3 attestée, alors que posit. 1 prévue (une règle d'antéposition, que l'on peut ajouter aussi bien à GP-1 (et à GP-1bis) qu'à GP-2, rend compte de ces cas) :

concretize                      concupine

(A-4) Posit. 2 attestée, alors que Posit. 3 prévue :

flotilla                      coition

(A-5) Posit. 2 attestée, alors que posit. 1 prévue :

fellow                      female

(A-6) Posit. 1 attestée, alors que posit. 3 prévue :

convalesce                      clarinet

(A-7) Posit. 1 attestée, alors que posit. 2 prévue :

drawee                      convulse

(B) Contre-exemples à la capacité générative particulière à GP-1 (sans règle d'exception). Ce sont les cas où l'accent principal apparaît placé en dehors des situations attestées dans les formes (a) du Lang-2 (mots du type venom) :

canal                      cadet

(C) Contre-exemples à la capacité générative particulière à GP-1bis. Ce sont les cas où, à la différence des situations attestées dans les formes (c)

du Lang-2, la règle d'exception ne produit pas des résultats adéquats :

Forme attestée :	Devrait être selon GP-1 et GP-2 :	Devrait être selo. GP-1 bis :
<u>f</u> eline	<u>f</u> eline	<u>f</u> eline
glu <u>o</u> se	glu <u>o</u> se	glu <u>o</u> se
defi <u>ci</u> ency	defi <u>ci</u> ency	defi <u>ci</u> ency
copu <u>l</u> ative	copu <u>l</u> ative	copu <u>l</u> ative
cre <u>m</u> atory	cre <u>m</u> atory	cre <u>m</u> atory
con <u>v</u> alesce	con <u>v</u> alesce	con <u>v</u> alesce

## 5 L'interprétabilité psychologique des formulations empiriques.

L'absence d'un processus de formulations empiriques explicites sur la simplicité et le fait que les seules qu'il a été possible de tester soient infirmées, est à rapprocher d'une autre constatation : ceux qui, dans le cadre de la psycholinguistique générative ont abordé le problème de l'acquisition du langage, n'ont pas incorporé une MS explicite et empirique aux propositions qu'ils ont faites pour expliquer le processus d'acquisition. Cette omission est d'autant plus remarquable que dans la vision chomskyenne du problème -dont ils se réclament- la simplicité est une pièce essentielle<sup>30</sup>. Ces deux constatations, bien que différentes, sont en rapport entre elles.

On a pu constater au § 4 que les MS empiriques, aussi bien celles qui ont été explicitées que celles qui sont suggérées, sont destinées à être intégrées dans un mécanisme projectif. Or, il y en a plusieurs types de projection possibles ; ci-dessus on n'a présenté que les modalités imaginées en fonction d'une typologie posée au départ : rien n'empêche d'imaginer beaucoup d'autres modalités différentes. Celles-ci peuvent être toutes empiriques, dans ce sens qu'il sera toujours possible de vérifier si la GP caractérisée comme plus simple est susceptible ou non de rendre compte d'un certain type de données projectives à partir d'un certain type de données originales. Il se pose donc un problème méthodologique capital : laquelle de ces modalités de projection choisir et pourquoi<sup>31</sup> ?

Nous croyons que seules deux réponses possibles existent : ou bien il faut choisir arbitrairement ou bien il faut admettre que la solution choisie doit être interprétable psychologiquement et devenir, par conséquent, une hypothèse empirique dans l'acquisition du langage. Le choix arbitraire implique que la linguistique ne sera qu'une discipline partiellement empirique : des "règles de jeu" non motivées, devraient être introduites dans son cadre méthodologique pour restreindre la multiplicité des possibilités descriptives. On cherchera, p. ex., le doublet MS<sup>GU</sup> susceptible de projeter des données originales et observables à des données projectives et profondes sans coïncidence des données observables, ou bien des données originales et observables qui ne supposent pas l'utilisation de symboles recursifs à des données projectives et observables qui supposent un nombre X de degrés d'autoenchaînement, etc. L'imagination est fertile et peut

suggérer beaucoup de possibilités : On risque, certes, de s'approcher encore plus du cas limite vers lequel on glisse doucement déjà, où chaque linguiste proposera le type de projection qui lui convient. Il y aurait alors autant de linguistiques que de linguistes ; toutes seraient empiriques mais avec des règles de jeu différentes.

L'interprétabilité psychologique constitue une solution possible au problème du choix du type de projection ; elle implique de replacer le problème des descriptions linguistiques dans le cadre des hypothèses sur les conduites linguistiques, dans ce cas particulier, les conduites linguistiques d'acquisition. Une machine de projection peut, en effet, être une hypothèse empirique interprétable psychologiquement dans le cadre de l'acquisition du langage, car elle peut résoudre le problème de l'acquisition, à partir d'un langage fini, des structures finies permettant de rendre compte d'un langage infini. Dans ce cas, le type de projection deviendrait à son tour une hypothèse empirique : à tel stade du développement de l'acquisition, on devrait chercher à présenter des hypothèses permettant de projeter tel type de données à tel autre type, car on aurait constaté que ces données sont effectivement et respectivement les données d'entrée pour l'enfant (les données originales) et les données sur lesquelles il possède une connaissance intuitive (les données projectives).

Pourquoi, malgré la prétention à l'interprétabilité psychologique des hypothèses sur la GU, et, plus particulièrement, dans le cadre de ce travail, de la MS, rien à notre connaissance n'a été fait dans un schéma de recherche visant à expliciter, tester, valider ou infirmer des hypothèses sur la MS, empiriques et interprétables psychologiquement<sup>32?</sup> Cette question porte aussi bien sur la Formulation VI bis du § 3.1 que sur les différents types de la Formulation VII et sur la Formulation VIII au § 4.

Il est très probable que la réponse à cette question soit liée au cadre heuristique que s'est donné la grammaire générative. Dans ce cadre, la problématique de l'acquisition est volontairement simplifiée et défigurée : on suppose deux points immédiats ; dans l'un, le sujet n'a rien intériorisé de sa compétence particulière, dans l'autre, il a tout développé. La GU et la MS sont posées comme des pièces essentielles du mécanisme de projection permettant l'accomplissement de ce processus instantané<sup>33</sup>. Comme il a été remarqué ci-dessus, on peut certes,

dans ce schéma, prétendre à l'empirisme de la GU et de la MS ; on ne peut en dire autant de leur interprétabilité psychologique : rien, dans ce schéma de recherche ne permet ni de vérifier ni d'invalider une hypothèse sur l'interprétabilité psychologique de la GU et de la MS. En effet, d'une part, ces hypothèses, dans le meilleur des cas, sont testées dans un domaine très différent de celui dans lequel elles sont censées être interprétables psychologiquement. D'autre part, dans le cadre heuristique fixé, il est par définition impossible d'invalider leur interprétabilité psychologique : comment peut-on tester que la projection proposée par une formulation particulière sur la GU et la MS est la correcte -dans le sens "est celle qui est interprétable psychologiquement"- si, par définition, il n'existe pas de situation effective permettant d'observer les données susceptibles de la valider ou de l'infirmer ? En particulier, il est par définition exclu d'avoir accès aux données originales et aux données projectives qui, dans les différentes étapes de l'acquisition, seraient susceptibles de tester les mécanismes de projection. Il est à remarquer, par ailleurs, que les projections proposées sont peu plausibles. P. ex., au § 4, on a testé des hypothèses empiriques censées être susceptibles de projeter la distinction entre absences systématiques et absences accidentelles et de projeter la place de l'accent à partir des traits phoniques dans la succession des voyelles et des consonnes. Le fait qu'elles soient empiriquement invalidées rend très académique une discussion sur leur interprétabilité psychologique. On remarquera, cependant, le faible degré de plausibilité qu'aurait, au départ, une tentative dans ce sens. Si l'on accepte qu'un enfant connaît tôt les séquences phoniques possibles et que, en revanche, le lexique est justement une des zones de sa compétence qu'il met le plus de temps à acquérir, on conviendra qu'un mécanisme de projection qui dépend si étroitement de la présence de tel ou tel signifiant dans le lexique, paraît particulièrement mal placé pour projeter la distinction entre absences systématiques et absences accidentelles à partir des données originales effectivement attestées vers des données projectives effectivement attestées. Soit, par ailleurs, dans cette même ligne de raisonnement, la projection du placement de l'accent principal en anglais (cf. § 4.3). Est-il plausible de penser qu'un petit anglophone va projeter la place de l'accent d'un corpus des données originales où apparaissent des mots tels que antidisestablishmentarian,

endocrinological, epistemological, arboratum, carbonatum, suicidal, developmental, anthropomorphisms (mots qui apparaissent parmi les 23 présentés dans le Lang-I de données originales) à un langage projectif, où on trouve veron, medal, villa (cités parmi les 6 du Lang.II) ou bien, butter, butcher, nothing, silly ?

Puisqu'il est impossible de tester l'interprétabilité psychologique d'un mécanisme de projection présenté dans le cadre conceptuel d'aujourd'hui, une autre pièce de la construction chomskyenne sort singulièrement affaiblie, à savoir ses affirmations, prétendument empiriques, sur le caractère inné de NS et de GU<sup>34</sup>. Ceci résulte de deux types d'observations.

Il existe aujourd'hui une pluralité de modèles de GU ; ceux-ci sont trop puissants, dans la mesure où ils parvent "tout faire"<sup>35</sup> ; ils sont, par conséquent, entièrement équivalents. On peut en conclure qu'ils vont se distinguer non pas par le fait que l'un ne présente pas des projections tandis que l'autre les présente, mais parce qu'ils présentent des projections différentes<sup>36</sup>. Et il est impossible, dans le cadre conceptuel posé par Chomsky, de choisir de manière empirique entre les différents types de projection. Même si on prétendait restreindre les GU à un point tel qu'elles détermineraient des GP susceptibles de faire tout et seulement ce qui est requis par les langues naturelles, il n'y a pas lieu à penser que cette exigence va éliminer l'existence de différents types de projection entre différents sous-langages d'un même langage naturel. Or, chacun de ces modèles pourrait prétendre au statut d'universel et d'inné, ce qui pose, dans le cadre conceptuel adopté, une situation dont on voit mal l'issue possible.

Par ailleurs, le raisonnement sur le caractère inné de NS et de GU, est fondé sur les considérations suivantes : (a) à partir des données originales finies et restreintes, un être humain est susceptible d'intérioriser une compétence portant sur des données projectives, infiniment plus larges que les données originales, et qui peuvent être considérées infinies ; un mécanisme de projection est donc nécessaire pour expliquer ce processus ; (b) étant donné qu'au stade final on trouve dans les GP et dans les représentations qu'elle spécifie, des structures "complexes", "sous-jacentes", "abstraites", on doit conclure à ce que ces structures sont déterminées par une organisation innée -GU et NS-, qui déterminent les

GP - car ces structures, de par leur "complexité", "sous-jacente" ou "abstraction", ne peuvent pas avoir été acquises à partir des données originales.

Or, ce raisonnement, dans le cadre conceptuel où il est présenté, aboutit à une proposition non empirique, car elle ne peut pas être fausse. En effet si l'on ne considère que deux points extrêmes, où dans l'un on n'a rien intériorisé et dans l'autre tout ; si, dans les données originales on écarte et/ ou on minimise les apports des structures cognitives autres que celles du langage ; si on étudie une compétence neutre par rapport aux conduites de production et d'interprétation, il paraît évident que ce qu'on trouve au stade final et qui apparaît comme étant impossible d'avoir été appréhendé à partir des données originales, ne peut pas ne pas avoir été déterminé par quelque chose, ce quelque chose ne pouvant être qu'une organisation que l'on pose au moment initial comme étant innée. Un mécanisme de projection est, certes, nécessaire, comme il a été souligné par Chomsky, pour rendre compte de l'acquisition que ce mécanisme soit constitué d'une GI et d'une MS innées, c'est une autre question, question qui, dans l'état actuel de la recherche, ne pourra pas être empirique tant qu'elle restera posée dans le cadre heuristique de recherche adopté par Chomsky<sup>37</sup>.

## 6 Bilan et perspectives

6-1 Dans les §§ 2, 3 et 4 ont été présentées une pluralité de MS utilisées en grammaire générative. Elles résultent d'une analyse qui ne peut pas être exhaustive de tous les écrits linguistiques inspirés par la grammaire générative, mais qui se veut représentative d'un certain nombre de textes clés. Certes, l'ambiguïté de la notion de simplicité avait déjà été remarquée, mais la presque dizaine de formulations différentes qu'il est possible d'expliciter à partir des exemples concrets, témoignent des très grandes faiblesses du cadre conceptuel d'ensemble<sup>38</sup>. Et il paraît impossible, dans l'état actuel des formulations, de réduire cette trop grande diversité. Il reste cependant à élucider complètement la nature des rapports de la simplicité avec la notion de généralisation, et ceci à l'intérieur de la problématique posée par la projection des structures. D'ores et déjà il paraît établi, par la discussion précédente, que la MS utilisée dans la Formulation VI n'est pas empirique et, ce qui est plus grave, qu'elle hypothèque lourdement le caractère empirique de toute proposition sur la composition d'une GU. Les Formulations VI-bis, VII et VIII posent, en revanche, le problème autrement, et il n'est pas exclu qu'elles puissent être rapprochées, à condition de fixer des contraintes aux données originales et projectives, et aux données différentes qui interviennent dans la comparaison des GP rapprochées dans la Formulation VI-bis.

Le § 4 montre que les MS empiriques explicites qui ont été proposées sont très facilement susceptibles d'être invalidées. Le § 5 établit qu'il devient impossible de donner une interprétation psychologique aux MS empiriques susceptibles d'être formulées dans le cadre actuel. Or, au-delà du constat d'absence de MS explicites, empiriques, et validées, il est possible de formuler l'hypothèse (empirique) suivante :

Aucune MS empirique qui puisse être proposée dans le cadre méthodologique actuellement utilisé en grammaire générative n'est susceptible d'être conjointement : (a) explicite ; (b) incorporable à des hypothèses sur l'acquisition du langage théoriquement cohérentes avec les autres hypothèses linguistiques qui portent sur la description de la compétence et qui sont spécifiques de la grammaire générative ; (c) testable dans le domaine psychologique des conduites d'acquisition du langage, si celles-ci sont observées sans l'"idéalisation" qui les fausse ; (d) vérifiée dans un certain nombre des cas pour qu'un consensus se dégage à sa



faveur. Cette hypothèse implique de faire la prévision forte suivante : même si une hypothèse empirique explicite était formulée, testée et validée dans le cadre méthodologique actuel de la grammaire générative, cette hypothèse sur la MS ne serait pas validée par des observations susceptibles d'être faites dans le domaine de l'acquisition du langage, si ces observations sont conduites sur des données proches des conditions réelles et non sur des données résultant d'une "idéalisation", heuristique ou autre. Ceux qui croient que la ou les MS envisagées en grammaire générative ont des possibilités d'être interprétables psychologiquement devraient considérer que cette prévision est susceptible d'être invalidée : il leur suffirait de présenter la MS avec les caractères précédemment signalés.

La non-unicité des résultats descriptifs hante la Linguistique depuis longtemps. On peut accepter, avec Bach<sup>1</sup>, que l'année 1957 a été une année capitale et charnière pour la linguistique : d'une part, ont été publiées quelques grandes récapitulations de la linguistique structurale, comme le recueil de Joos ; d'autre part, les Structures Syntactiques de Chomsky, dont on justifiait en partie les nouvelles propositions parce qu'on attendait qu'elles soient susceptibles de trancher, de manière fondée, entre solutions concurrentes<sup>39</sup>. En effet, le recueil de Joos finit par la discussion d'un problème classique en linguistique structurale : l'assignation des allophones aux phonèmes dans une langue donnée, en l'occurrence l'espagnol, problème qui a été à la source de non-unicités multiples dans la linguistique post-saussurienne et qui, sur un plan général, n'a pas été résolu de manière satisfaisante. Aujourd'hui la grammaire générative ayant, certes, enrichi la connaissance du langage, se trouve confrontée à un problème analogue : son cadre conceptuel n'est pas suffisamment riche pour pouvoir donner la possibilité de choisir de manière empirique entre solutions concurrentes. L'absence d'interprétabilité psychologique des hypothèses entraîne la nécessité de postuler des règles de jeu permettant de se donner un "critère" pour choisir entre plusieurs hypothèses possibles. Ces règles de jeu peuvent se multiplier facilement, d'où l'éclatement en écoles et sous-écoles rivales, et, ce qui paraît plus grave, l'absence de communication entre elles.

Cependant, la réponse à la question N°4 du § 1 (Est-ce que le traitement du problème de la simplicité a avancé par rapport à la situa-

tion antérieure ?) doit être, semble-t-il, affirmative. Deux points, qui se dégagent de la discussion précédente, justifiant un bilan finalement positif. Grâce à Chomsky, on doit admettre :

(a) La possibilité que la notion de simplicité soit interprétable psychologiquement.

(b) La possibilité que le mécanisme de projection qui doit être incorporé aux hypothèses sur le langage comprenne une MS.

6-2 Il semble bien qu'il y ait des raisons pour justifier l'incorporation des considérations de simplicité à un mécanisme de projection. En effet, un système linguistique doit être prévu pour rendre compte de deux caractères des langues naturelles qui semblent bien leur être inhérente :

(a) des principes généraux, aussi bien inter-langues qu'intra-langues organisent tout système linguistique : ce seraient les universaux de la GU et les généralisations des GP recherchés par la grammaire générative ;

(b) des cas particuliers, des exceptions, des applications partielles d'un principe général, des distinctions graduelles et non polaires, apparaissent partout dès qu'une description est tant soit peu poussée à un certain degré de détail.

On peut déduire, de ce double constat, que le formalisme permettant d'exprimer une description linguistique, quel qu'il soit, doit être préparé pour satisfaire aussi bien (a) que (b). Par ailleurs, étant donné que la structure linguistique doit être finie, le vocabulaire du formalisme qui l'exprime doit l'être aussi. Ceci implique que tout principe général exprimé par le biais de formules où le vocabulaire utilisé est défini en compréhension peut être exprimé par des formules finies équivalentes, où le vocabulaire utilisé est défini en extension. Ceci entraîne que, dans les langues naturelles, les cas réductibles à des principes généraux sont susceptibles d'être formulées de deux manières : selon le formalisme qui lui est propre (formulation en compréhension des classes intervenantes) et aussi, de manière équivalente, selon le formalisme propre aux cas de type (b) (formulation en extension des

classes intervenantes). Si ce raisonnement est exact, il y a des raisons pour croire que nécessairement, étant donné leurs caractères propres, toute formalisation des langues naturelles se doit d'envisager le problème du choix entre alternatives équivalentes : les conditions pour justifier l'incorporation nécessaire des considérations sur la simplicité semblent donc exister<sup>40</sup>.

Mais un autre argument peut encore être donné pour rendre plausible l'incorporation de considérations de simplicité à la théorie du langage.

Les données extérieures auxquelles l'enfant est confronté, celles qui ne relèvent pas de son organisation mentale et qui varient de langue à langue, qu'elles soient réductibles ou non à des principes généraux, se présentent sous la forme de cas particuliers : les différents types d'énoncés, les phonèmes, les morphèmes, les constructions syntaxiques, les traits sémantiques sont, pour ainsi dire, concrétisés sous la forme d'exemplaires individuels dans la parole effective. Il est imaginable que, au moins à un certain stade du développement de l'acquisition, ces exemplaires soient incorporés en tant que tels ou par le biais d'abstractions qui restent très près des cas concrets. Ce serait par la suite, lorsque le stockage et l'organisation des cas généralisables par le biais du formalisme adapté aux cas particuliers deviennent trop complexes et coûteux en relation au langage traité, que des considérations de simplicité obligeraient à réorganiser leur structuration sous la forme de principes généraux. Dans cette optique, l'acquisition du langage serait envisagée comme un processus de réorganisation de structures, conditionné par un certain idéal de "rentabilité". Celle-ci résulterait d'une relation entre le langage que l'enfant doit traiter et plusieurs coûts : coûts de stockage du système linguistique et des instructions d'utilisation de ce système et coût de fonctionnement de ces instructions dans la production et l'interprétation des énoncés<sup>41</sup>. Ces coûts seraient précisés par plusieurs mesures, qui ne porteraient pas sur des éléments d'un même type et qui seraient vraisemblablement conflictuelles, en ce sens que la réduction d'un type de coût (c'est-à-dire, la "simplicité" d'un des facteurs) devrait entraîner l'augmentation du coût des autres. Des conditions d'équilibration dans la recherche d'une réduction des

différents types de coût devraient donc exister ; elles ne détermineraient pas la simplicité du système d'ensemble, mais un équilibre entre plusieurs simplicités différentes. Chaque réorganisation des structures, intervenant dans les processus d'acquisition, déterminerait un effet double : d'une part, diminuer relativement, par rapport au langage traité, les différents coûts de ces structures ; d'autre part, élargir le langage susceptible d'être traité par la nouvelle réorganisation.

Si la notion de simplicité était interprétable psychologiquement et validée, les perspectives qui s'ouvriraient seraient intéressantes : il serait possible d'avancer une explication à ce recours à une simplicité formelle dont on a fait usage dans les sciences empiriques. En effet, un ensemble de connaissances, sur n'importe quel domaine, doit être structuré et codifié pour être stocké à long terme. Est-ce qu'il est inimaginable de penser que cette structuration et cette codification peuvent être conditionnées par des considérations de coût et d'équilibration entre coûts, considérations qui seraient formellement analogues, voire identiques, à celles qui opèrent sur le langage humain ? Le langage marquerait ainsi de son empreinte les autres structures cognitives.

Un long chemin reste à faire avant que les spéculations précédentes commencent à se concrétiser sous la forme d'hypothèses explicites et testables. Cependant, on peut risquer l'hypothèse (toujours empirique) suivante :

Si un modèle général et explicatif de l'acquisition du langage présente conjointement les caractéristiques suivantes :

(a) il incorpore un modèle explicite de description du système linguistique ;

(b) il est explicite et testable par rapport à des conduites d'acquisition observées dans des conditions effectivement attestées et proches des situations réelles ;

(c) il se révèle vérifié dans un nombre suffisant de cas pour qu'un consensus se dégage à sa faveur ; il doit incorporer une ou plusieurs MS empiriques et explicites.

Sera-t-il plus facile de vérifier cette hypothèse que d'infirmer la précédente ?

Des présentations partielles de ce travail ont été faites, en janvier-février 1975, à la Faculté de Psychologie et des Sciences de l'Education de l'Université de Genève et au Département des Lettres Modernes de l'Université de Simon Fraser (Burnby, Canada) et, en mars 1977, au C.I.R.A.L., de l'Université de Paris III. Cette rédaction profite largement des discussions avec A. Hurtado, E.W. Roberts, J.P. Bourge, J.P. Dumas, C. Matzkin et D. Hadjadj, à qui elle doit, en plus, de très utiles observations stylistiques. Je remercie R. Zuber de m'avoir indiqué l'existence du travail de Sober (cf. 124b).

1 - Non sans quelque arbitraire on peut situer vers 1967 l'introduction et diffusion en France de la grammaire générative : après la date de parution de Grunig 43 et avant celle de Ruwet 111.

2 - Les "sciences formelles" et les "sciences empiriques" de ce travail correspondent aux "sciences formelles" et aux "sciences sur des faits" ("factuales", dans l'original) de Bunge (17 p. 38 et suiv.) Par conséquent l'expression "hypothèse empirique", dans ce travail ne doit pas être comprise dans le sens de Bunge (17 p.284). Sur le statut des hypothèses, en particulier dans le domaine linguistique, cf. Botha 14 p.174-176, Botha 13 et Botha 15. Ce travail a été conduit indépendamment des analyses minutieuses de Botha, en particulier de 14, où on aborde des sujets analogues. En lignes générales, on présente dans cette contribution un matériel complémentaire de celui de Botha et/ou qui a été interprété de manière différente ; on signalera ci-dessous des convergences et des divergences.

3 - Sur les notions de modèle de boîte noire et de boîte transparente, cf. Bunge 18.

4 - Sur la simplicité utilisée pour choisir entre hypothèses empiriques en concurrence, cf. Bochenski 12 p. 229-230, Copi 33b p. 383-384 ; Botha 13 p. 47-48 et les renvois qui y sont faits et, tout particulièrement, Popper 105 p. 128-136 et Bunge 18 p. 129-184. Celui-ci, d'une part, met en relief les aspects multiples d'un système hypothétique qui peuvent être plus ou moins simples et, d'autre part, montre comment la simplicité favorise seulement quelques uns des principes généraux par rapport auxquels il est possible d'évaluer des hypothèses. Le recours à la simplicité dans les

autres domaines (cf. Derwing 35 p. 243-245), scientifiques est moins "évident", pose plus des problèmes et est <sup>plus</sup> controversé, que ne le croient certains linguistes (comme, p. ex, Lakoff 77, p. 147, où ses considérations sur le rasoir d'Occam sont à opposer à celles de Bunge 13 p. 184 ; Lamb 79 p. 55 et Dell 34 p. 163). Ci-dessus dans le texte, par simplicité formelle, nous comprenons ce que Popper classe parmi les concepts esthétique et pragmatique de la simplicité et non pas la simplicité en tant que principe qui favorise la testabilité des hypothèses (cf. Popper, ibidem p. 129-134 et Bunge, ibidem, p. 138-150).

5 - On constate, chez Saussure (cf. 113 ) l'effort pour mettre en rapport la langue avec la "conscience" des usagers, de même que chez Troubetzkoy (129 en partie. p. 1-15, p. 31-41, p. 66-68). La notion de "fonction" est celle qui rend empirique la conception linguistique de Martinet (cf. en particulier, 85 p. 36-108). Le résultat normal de cette conception empirique devrait être la possibilité de tester les conceptions linguistiques par rapport aux faits externes ("conscience" des sujets, fonction de communication, etc). En revanche, chez Bloch 11, la vérification des résultats descriptifs se fait de manière prioritaire par rapport aux procédés descriptifs.

6 - Sur la notion de "calcul" et, sur celle de forme, considérée indépendamment des données où elle se manifeste. cf. Hjelmslev 56 p. 87, p. 96-106, p. 114-115, Rès 6 et Rès 7 chap. IV.

7 - L'analyse de la simplicité proposée en glossématique conduit vite à la critique dont il paraît fatalement paisible toute simplicité formelle : un objet présente plusieurs aspects différents et il est possible de le considérer plus ou moins simple selon l'aspect dont on tient compte. Cette question a déjà été abordée par Uldall 131 p. 23 ; cf. aussi Martinet 86 ; Spang-Hanssen 125 ; Fischer-Jørgensen 37 et Siertsema 117 b p. 63-66. Il ne semble pas que Lamb ait tenu compte de cette discussion en reformulant le principe de simplicité de Hjelmslev (cf. 73 p. 551). Sa reformulation prouve, encore une fois, qu'une théorie linguistique

présente des aspects différents dont chacun est candidat à être préféré à l'autre dans une hiérarchie d'application d'une simplicité formelle. Le critère qu'il propose (cf. 79 p. 554 et suiv.) n'en échappe pas : la note de p. 555 nous informe -sans justification empirique- que, pour faire le calcul de la simplicité, il ne considère que les lignes de son système de représentation, ce qui lui permet d'affirmer que la description de la fig. 9 est plus simple (et plus générale) que celle de la fig. 8 ; or, s'il prenait en considération les noeuds et non les lignes, le résultat serait le contraire.

8 - Cf. Hockett 57 p. 110-111 qui le présente cependant avec beaucoup de précautions (sur l'orientation empirique de Hockett, cf. en particulier 58). La quête de la symétrie des systèmes en linguistique structurale est également en rapport avec une simplicité formelle (cf. Hockett 57 p. 109-110 et Twaddell 130).

9 - Cf. Jakobson et alii 67 p. 45-47, 66 p. 7, 65 p. 510-522. Il a été montré que les avantages de la simplicité formelle sont loin d'être aussi évidents que Jakobson le voudrait : la notion "trait distinctif" présente plusieurs aspects différents qui déterminent des simplicités différentes (cf. Bès 8) ; par ailleurs, la simplicité formelle proposée n'est pas susceptible de résoudre le problème de la non unicité des descriptions (cf. Bès 7, p. 228-241).

10 - L'étude des simplicités a été faite comme si les textes formaient un corpus unique, sans distinction internes, cette méthodologie étant justifiée par le type des citations couramment utilisées en grammaire générative : les textes se renvoient les uns aux autres sans faire état, en ce qui concerne la simplicité, des modifications et/ou des changements de critère. Il doit être cependant possible d'ordonner chronologiquement la pluralité des conceptions sur la simplicité, de telle manière que, à un moment donné, l'incohérence serait moins grande qu'en considérant synchroniquement les écrits ; cet aspect évolutif n'a pas été considéré ici.

11 - Cf. par exemple, la discussion portant sur la différence notationale et/ou substantielle entre sémantique interprétative et sémantique

généralive : Chomsky 22, Postal 108 p. 244-270, Katz 73 p. 397 et suiv. Les participants à cette discussion ne semblent pas avoir tenu compte du fait que l'incorporation d'une MS empirique, s'appliquant, comme toute MS, aux symboles utilisés pour noter une structure, peut déterminer qu'une différence "notationnelle" ait des répercussions empiriques.

12 - Ailleurs, dans Chomsky 20 on retrouve le même type de raisonnement, même si les expressions "simple" et "simplicité" n'y sont pas employées toujours : p. 164, en discutant l'élimination des redondances des entrées lexicales ; p. 170-173, pour justifier un certain type de description des éléments de flexion. Katz l'a utilisé pour justifier le modèle de sémantique interprétative (72 p. 250-257, 73 p. 408-412) ; le schéma du raisonnement sous-jacent a été présenté par Botha 13 p. 291-292).

13 - La Formulation I, appliquée au domaine phonique, réapparaît ailleurs : cf. dans Halle (50 p. 338) la présentation sur le sandhi en sanscrit, et ibidem (p. 338-340) la première partie des discussions sur la spécification des signifiants. Le même schéma est appliqué par Schane 114, à propos de l'élision des consonnes finales en français, bien que les expressions "simple" et "simplicité" n'y soient pas utilisées.

14 - La comparaison de Postal porte sur le modèle de la sémantique généralive comparé et préféré au modèle de la sémantique interprétative. Chez Chomsky, un schéma analogue -mais non identique- à celui de la Formulation II est utilisé pour justifier l'élimination de la "redondance" d'un modèle de GU et/ou pour favoriser son "homogénéité" et/ou sa "simplicité" : cf. dans 20 les discussions sur l'élimination des indicateurs de transformations (p. 134-135) et des transformations de permutation (p. 144) . On peut comprendre dans ces cas que le modèle de GU préféré comme plus "homogène" est, en même temps, plus restrictif, les GP qu'il détermine ne devant pas être par conséquent entièrement équivalentes à celles de l'autre modèle ; en revanche, Postal utilise la MS dans le cadre du "ceteris paribus". En raison de ces optiques différentes, ni les observations de Botha 13 p. 309-320, ni la critique (tacite) de Postal par Chomsky (26, p. 125-129) ne semblent pertinentes.



15 - La proposition de Halle (53 p. 190) sur les restrictions à introduire à la MS, n'étant pas appuyée sur la considération des données externes, est à classer dans ce schéma, de même que la discussion de Kucera (76) bien que, au départ, celui-ci présente une MS à classer dans la Formulation VII. Dans la Formulation III peuvent être classées également les MS discutées par Haiman (55) et chacune des deux MS étudiées par Yen (152 p. 58-68), mais il est difficile de comprendre à partir de quelle base, prêterment empirique, il peut par la suite (ibidem, p. 68-69) comparer les deux MS rapprochées. En fait, ce qu'il essaye de comparer ce sont deux modèles de GU.

16 - Il n'est pas surprenant de constater qu'une approche formelle de la simplicité permet à Katz (cf. note 17) et à Postal (cf. note 14) d'aboutir à des résultats opposés dans l'évaluation des mérites de simplicité des deux mêmes modèles en concurrence.

17 - Botha a mis en relief et la difficulté à déterminer le contenu de la notion de "généralisation significative" et son importance (cf. 14 chap. III ; 13 p. 276-280). De même, Peters (101 p. 186 et note 12) souligne que les intuitions portent sur les énoncés et non sur les généralisations (cf. aussi Kiparsky 74 p. 172 et Derwing 35 p. 136-141). Le statut des généralisations en tant que données susceptibles de tester les hypothèses a donc été explicitement mis en doute, ce qui justifie que la MS de la Formulation V soit éventuellement empirique. Dans Bès 5 ont été proposés des critères, fondés sur la notion de projection, permettant de distinguer les différents types de généralisation. Ces critères peuvent constituer des hypothèses empiriques et seraient susceptibles d'être interprétés psychologiquement, mais ils ne permettent pas, à eux seuls, de solutionner le problème des modèles en concurrence et ce pour les mêmes raisons qu'on discutera ci-dessous à propos des projections relevant de la MS. On n'a pas été plus explicite en grammaire générative sur la notion de régularité que sur celle de généralisation. De certains emplois des deux termes on peut conclure à ce que "régularité" est une

propriété relevant des langages et "généralisations" des grammaires : un langage qui possède une régularité doit être spécifié par une grammaire qui possède une généralisation. Je dois à Jean-François Dumas cette utile distinction, qui permet de comprendre un certain nombre de textes portant sur ces questions.

18 - Plusieurs traits des Formulations IV et V ont été inférés des contextes où apparaissent les citations. Le problème du contenu conceptuel de "généralisation" se double d'un problème terminologique. Il se peut que ce terme désigne à la fois une donnée et un concept caractérisé dans la théorie linguistique, comme, p. ex., c'est le cas de grammatical, qui parfois a été utilisé aussi bien dans le sens de "ce qui est spécifié par la grammaire" que dans celui de "ce qui est considéré comme non déviant par un usager". Dans ce cas, la généralisation, notion conceptuelle, devrait rendre compte de la généralisation, relevant des données (mais cf. note 17 in fine). Ceci n'est qu'une question terminologique ; ce cas de "généralisation-conceptuelle rendant compte de généralisation-donnée" rentre dans la Formulation V.

19 - Si la notion de classe naturelle en phonologie était considérée comme le résultat d'une définition, la MS qui y intervient serait une variante de celle de la Formulation IV : le texte de Halle 51 p. 90 peut être interprété dans ce sens (cf. aussi Harris 54 p. 26 et Postal 106 p. 73). En revanche, si elle était considérée comme une donnée (dans la note 22 on trouvera des références attestant l'utilisation de "classe naturelle" comme une donnée) elle appartiendrait à la Formulation V.

20 - On ne connaît aucun exemple où la Formulation V apparaisse strictement appliquée, mais elle l'est dans des cas assimilables à ceux qui y sont prévus : cf. Garcia 41, Saporta et Contreras 112 p. 10-12, Zirmar 133. Le trait commun de ces utilisations est le fait qu'une MS est testée selon sa capacité de sélectionner la GP considérée par ailleurs comme possédant une adéquation externe, celle-ci n'étant pas forcément une "généralisation" mais restant toujours très faiblement précisée.

21 - La Formulation VI, dans 20 p. 96-97 s'applique à nouveau et permet la justification de GU-C comportant les conventions (44) et (45).

22 - La Formulation VI s'inscrit dans un schéma de discussion de modèles où, contrairement aux opinions recueillies dans la note 17, on pose, au départ, que l'intuition sur la similitude de deux faits linguistiques, ou sur leur "complexité", "naturalité", etc., peut être considérée comme une donnée susceptible de tester des hypothèses : cf. Ross 110 p. 229 ; McCawley 89 p. 172, 88 ; Halle 53 p. 183-186 ; Bach 2 p. 76. La Formulation VI est utilisée par Halle 48 et par Halle 50 p. 334-338, Harms 54 p. 7-9 et Dell 34 p. 166-167 pour justifier les traits binaires en phonologie ; cf. aussi Harms ibidem p. 110-111 la discussion des propositions de McCawley sur les jonctures, Schane 115 p. 34, Contreras 33 et Bach 2b (Ces deux derniers étudiés par Sober 124b p. 106-107). Sur le plan syntaxique, le schéma de la Formulation VI est utilisé pour justifier l'introduction dans la GU des contraintes "universelles" qui ne sont pas censées être respectées dans toutes les GP mais qui, si elles ne l'étaient pas, rendraient les GP où elles ne sont pas respectées plus complexes : cf. Chomsky 20c p. 235-236, 24b p. 75-77, Chomsky et Lasnik 32b p. 437-437 ; un raisonnement analogue dans Chomsky 22b p. 176-177.

23 - Les mêmes observations sont valables pour les exemples rapportés par Botha 14 p. 69-75, p. 137-139 et le schéma de raisonnement qui s'y applique. Il faut souligner que, malgré des déclarations explicites de Chomsky et de Halle, ce n'est pas la MS qui est mise en cause dans les exemples où se concrétise l'utilisation de la Formulation VI. P.ex., dans 32 p. 361, on lit que "en formulant la méthathèse de cette manière, nous soulevons un problème à notre mesure d'"évaluation", et ailleurs (ibidem p. 400) que "jusqu'à un certain point, cette mesure donne les résultats désirés, mais dans beaucoup des cas, elle échoue complètement". Or, les suites de ces raisonnements montrent que ce n'est pas la MS qui est mise en cause -elle continue à être posée comme correcte- mais les caractéristiques de la GU : à la p. 361 plusieurs conventions pour noter les méthathèses sont proposées, et, à la suite des exemples de p. 460, la notion de marque est incorporée.

24 - Le fait que la MS intègre un mécanisme de projection a été remarqué par Peters 101.

25 - Contrairement à l'interprétation de Sober (124b p. 106-107), la MS de cette illustration ne porte pas sur la comparaison des deux règles de redondance (1) et (2) entre elles. Hormis l'exemple présenté ci-dessus, MS GU-A apparaît utilisé par Halle 50 p. 340-342 et par Chomsky 21 p. 64 dans des exemples qui sont aussi artificiels que celui rapporté plus haut ; allusion en est faite dans Chomsky et Halle 32 p. 381-382, dans Postal 106 p. 164, p. 178 et dans Bach 2 p. 246, sans toutefois que des exemples nouveaux y soient présentés. D'après Chomsky (21 p. 64, et cf. dans le même sens 20 p. 169-170) "Halle montre comment une adhésion cohérente au principe de minimiser les spécifications des traits dans la composante phonologique fournit un fondement pour la distinction entre les absences ("gap") accidentelles et celles qui ne le sont pas". Or, dans les deux travaux de Halle cités par Chomsky dans 21, c'est-à-dire dans 45 et 52, il a été impossible de retrouver des traces de MS GU-A et de la possibilité de distinguer entre absences accidentelles et absences systématiques (cf. en particulier 52, chap. I et II). La MS qui est présentée (ibidem p. 29-30) est à classer dans la Formulation III, les quelques observations qui y sont faites -p. ex. p. 29- pour essayer de la justifier sont très loin d'être suffisantes pour la rendre empirique. L'affirmation de Chomsky est donc excessive parce que, autant qu'on sache, elle est non fondée. Stanley (127) présente un formalisme différent de celui de Halle pour rendre compte des redondances lexicales mais, pour ce faire, il doit présupposer que la MS conduit à des résultats adéquats (127 p. 401, p. 428, p. 431), ce qui est faux, les contre-exemples pour le formalisme de Halle sont également valables pour celui de Stanley. Bever et Rosenbaum (10 p. 3-7) ont utilisé une MS dans un cadre analogue à celui de MS GU-A, mais appliqué aux formants lexicaux. L'exemple qui y est présenté est aussi artificiel que ceux utilisés en phonologie ; par ailleurs, ni la MS utilisée ni la définition du lexique (vocabulaire des traits utilisés p. ex) ne sont suffisamment explicitées pour qu'elles soient testables.

26 - Pour que l'hypothèse empirique portant sur MS GU-A soit suffisamment explicite il faudrait indiquer des critères précis pour cerner la notion "lexique d'une  $G^m$ ", selon qu'un item soit ou non incorporé, il sera possible, dans certains cas, d'incorporer une règle de redondance (cf. Kucera 76 p. 1115 note). Ce problème n'étant pas réglé, il a été nécessaire de le solutionner de manière arbitraire : les dictionnaires bilingues français-anglais (Paris, Larousse, 1960, collection Jupiter) et espagnol-allemand, manuel, (Berlin, Langenscheidt KG, 1962) ont été utilisés pour illustrer les lexiques du français et de l'espagnol ; ils présentent l'avantage, étant donné l'utilisation à laquelle ils sont destinés, de ne pas incorporer trop de régionalismes, technicismes, etc.

27 - Le fait de pouvoir présenter des contre-exemples et de pouvoir infirmer MS  $\widehat{GU-A}$  prouve, en même temps, que la MS de la Formulation VII est bien empirique, contrairement aux conclusions générales de Darwin (35 p. 246-247). De même Botha (14 p. 237-241) dans ses commentaires à une situation qui n'est pas entièrement identique à celle de MS GU-A, semble croire à l'impossibilité de principe de calculer la complexité dans le lexique et dans les règles de redondance. De deux auteurs qu'il cite en son appui, l'un (Harns 55 p. 609) affirme, sans preuves empiriques, que les traits du lexique ne peuvent être considérés à égalité avec les traits des règles morphologiques, et l'autre (Postal 106 p. 134-138), si l'on replace le texte transcrit dans le texte original, n'écarte nullement la possibilité de mettre en rapport la complexité du lexique avec la complexité des règles. On ne comprend pas quel est, a-priori, le problème qui se pose par rapport à une MS qui tient compte des traits distinctifs et dans le lexique et dans les règles de redondance : une vague notion de "plausibilité" ou autre ne saurait être un obstacle de principe à une MS parfaitement empirique -bien qu'invalidée- portant sur la quantification d'entités -les traits distinctifs binaires- qui sont utilisés aussi bien dans les matrices lexicales que dans les règles de redondance.

28 - 46 reproduit la première partie de 49 ; on ne considèrera donc que celui-ci. Il faut souligner que ce ne sont pas les propositions de

Chomsky et de Halle dans 32 celles qui sont discutées, mais la capacité projective attribuée à la MS et à la GU dans Halle 49 . Dans ce texte, il est explicitement indiqué qu'il s'agit de placer l'accent principal dans des mots (cf. p. 2-3) ; par ailleurs, l'item (2) présente des mots qui n'écartent ni préfixes, ni suffixes, ni la composition (p.ex. anti-disestablishmentarian, endocronological, epistemological, anthropomorphous) ce qui autorise à en faire autant dans les contre-exemples.

29 - On ne discutera que les projections du Lang-2 qui sont propres à la GP plus simple et non pas toutes les projections discutées dans Halle 49 : il y en a -p. ex. la gradation accentuelle dans des mots comme personify qui peuvent être projetées par GP-1 ou par GP-2 plus d'autres règles susceptibles de s'appliquer aussi bien à la sortie de GP-1 que de GP-2. Les transcriptions phonétiques et la place de l'accent principal permettant de caractériser les contre-exemples -dont seulement des spécimens de chaque type ont été indiqués- ont été empruntés au Every man's English Pronouncing Dictionary de Daniel Jones (11<sup>e</sup> ed., Londres et New York, 1956).

30 - Il est risqué d'avancer une affirmation portant sur l'inexistence de quelque chose dans un domaine aussi vaste que la psycho-linguistique générative. Disons qu'il a été impossible de trouver une MS explicite et empirique ou une discussion approfondie et détaillée de ce problème dans un certain nombre de travaux représentatifs, aussi bien dans ceux qui ont été inspirés par la grammaire générative que dans ceux qui ont pris une position critique. Hormis certaines allusions très générales dans McNeill 92 p. 34-36, 94 p. 61-62 (et cf. aussi l'opinion de McNeill citée par Derwing 35 P. 154n), Fodor 39 p. 114, p. 118-120, Braine 16 p. 171-172, p. 180, et cf. ci-dessus § 4-1-3) il a été impossible de retrouver une MS explicite incorporée aux hypothèses sur l'acquisition dans Slobin 120 ; elle n'y apparaît pas ni dans la présentation de Slobin 119, ni dans les contributions de Palermo 98, de Schelesinger 117, et d'Ervin Tripp 36 ; elle n'apparaît pas non plus dans Huxley et Ingram 62 (cf. Schelesinger 116, Bellugi 4 Mehler 95, Sinclair 118), dans Flores d'Arcais et Levelt

38 (cf. Huxley 61, Slobin 122, Riegel 109b) dans Lyons et Wales 82 (cf. McNeill 93, Klima et Bellugi 75), ni dans Smith et Miller 124 (cf. l'introduction, les commentaires de Slobin à McNeill 94, et Jenkins 70). Une MS en tant qu'hypothèse incorporée à un modèle d'acquisition est absente aussi des travaux tels que Jakubowicz 68 et Jakubowicz et alii 69 (cf. aussi Mehler 95 et Mehler et Noizet 96b) qui présentent un aperçu général des apports de la grammaire générative en psycholinguistique ; elle n'apparaît pas non plus dans des articles qui, incorporés à des recueils, ont fait le point des problèmes dans ce domaine (cf. McNeill 91, Slobin 121, Campbell et Wales 19) et dans la contribution de Bever (9), qui représente un tournant important en psycholinguistique générative. Elle est absente du recueil récent sur les rapports de la théorie linguistique avec la réalité psychologique (cf. Halle et alii 53b).

31 - Le problème se complique du fait que la notion de généralisation peut -et vraisemblablement, doit- se préciser en fonction de la possibilité de projection (cf. notes 17, 18 et 36). Par conséquent, il faut non seulement choisir entre différents types de projection, caractérisés par les données originales et projectives, mais, aussi, choisir entre projections obtenues par MS GU et projections obtenues par GU sans intervention de MS (tout au moins dans les formulations existantes, car il est probable que des considérations de simplicité relative doivent intervenir toujours ; cf. ci-dessous). Par ailleurs, si l'on admettait que tous les types de projection devraient être admis, il faudrait préciser quelles portions de chaque grammaire devraient être traitées par chaque type de projection.

32 - La distinction entre hypothèse empirique et hypothèse empirique interprétable psychologiquement correspond à celle présentée par Botha (14 p. 170-176) entre théorie linguistique non mentaliste et théorie linguistique mentaliste, celle-ci se présentant comme la description d'un mécanisme mental ; cf. dans le même sens, Peters 101 p. 174-175. Mais, contrairement à Botha et à Peters, il paraît impossible que l'on n'exige pas de la théorie linguistique d'être interprétable psychologiquement. Si l'on n'exige que la linguistique soit empirique et si le raisonnement qui suit est valable, elle ne sera plus qu'une discipline

partiellement empirique, avec des marges de variation trop grandes, qui devront être comblées par des décisions non-empiriques.

33 - cf. Chomsky 20 p. 32-37 et note 19 p. 202, Chomsky et Halle 32 p. 331, Chomsky 25 p. 15, p. 119-122, p. 148 et note de p. 245.

34 - cf. Chomsky 24 et 25, Les affirmations sur le caractère inné de MS et de GU sont faussement empiriques dans le champ d'observation des données linguistiques ; elles ne le seraient pas, en droit, si l'on se donnait la biologie comme domaine d'observation susceptible de tester l'hypothèse innéiste (ce point m'a été éclairé par une discussion avec J.-F. Dumas, qui, cependant, envisage l'hypothèse innéiste dans une perspective différente).

35 - Sur le trop grand pouvoir des modèles de GU, cf. Peters et Ritchie 102, 103 et 104, Levelt 81 v. II, chap. 5, Bach 2 chap. 9 à 11.

36 - Cf. dans Bès 5 des exemples de non unicité des projections déterminées par les généralisations significatives ; la variation des résultats est déterminée par l'utilisation de modèles de GU différentes et/ou par la considération de corpus différents.

37 - Dans Bès 5 ce raisonnement a été discuté avec plus de détail, en particulier, en contestant la neutralité de la compétence par rapport à la performance. La non considération des apports des structures cognitives autres que celles du langage dans le processus d'acquisition apparaît dans Chomsky 23, 20 chap. I, cf. Levelt 81 v. 3 p. 144-156.

38 - L'ambiguïté de la notion de simplicité en grammaire générative avait été, en termes généraux, remarquée, elle est implicitement signalée dans la réponse de Householder 59 et, explicitement, dans Garcia 41, Mathews 87 p. 276, Botha 13 p. 291, Mounin 97 p. 219-221 ; cf. aussi Pottier 109 p. 11 et Hagège 44 p. 180. Les huit formulations ici explicitées donnent cependant une mesure plus exacte de l'acuité du problème. Elles n'épuisent



cependant pas les possibilités existantes, au moins trois autres simplicités pouvant s'y ajouter : (1) la simplicité à laquelle fait allusion Chomsky dans 23 p. 14, destinée à délimiter des données ; (2) la simplicité portant sur les dérivations, suggérée par Chomsky et Halle (32 p. 63) ; (3) la simplicité de Kiparsky (74), dont la validation empirique devrait se faire dans des domaines (dialectologie, diachronie) qui débordent celui de la linguistique descriptive. Il est facile de déduire de cette diversité jusqu'à quel point ont été peu fructueuses les discussions sur la simplicité en grammaire générative, développées sans tenir compte de la riche variété des MS (cf. Chomsky 27, Chomsky et Halle 31, Halle 47 Householder 59, 60, Mathews 87). Dans ce contexte, le traitement par Sober (124b chap. 3) de la simplicité en grammaire générative comme un phénomène unitaire, où il <sup>ne</sup> voit, entre 1957 et 1968, qu'un "subtle change" (ibidem p. 89) et son essai de l'exprimer par la notion générale de simplicité qu'il propose, semblent voués à l'échec ; les illustrations qu'il présente relèvent, en effet, de trois formulations différentes : Formulation VI, exemples empruntés à Halle (124b p. 96-100), à Chomsky et Halle (124b p. 107-111), à Contreras (124b p. 114-118) et à Fach (124b p. 118-120) ; Formulation V, exemples sur les classes naturelles en phonologie, d'après Halle (124b p. 100-102) ; Formulation VII - Type I, exemple -nal interprété, cf. note 25- emprunté à Chomsky et Halle (124b p. 106-107). Il fait état, en plus, de la simplicité qui relève de la Formulation I (124b p. 94-95, p. 102-103).

39 - " Le livre de Chomsky -disait Lees (80 p. 378) peut bien fournir l'opportunité d'appliquer des mesures de simplicité explicites pour accorder la préférence à une forme de grammaire plutôt qu'à une autre [...] même s'il n'y a pas de science naturelle susceptible de prétendre à une procédure d'évaluation vraie, Chomsky croit que la linguistique peut très bien permettre ce degré de capacité théorique [...]. Pour rejeter la mauvaise alternative d'une paire de théories proposées, le mieux que le savant des sciences naturelles peut faire est de proposer un "expériment crucial" ; mais les linguistes peuvent faire un pas au-delà et formuler des critères rigoureux de l'excellence des grammaires" (ibidem p. 380-381).

40 - Ces considérations de simplicité n'auraient pas le caractère supplé-  
tif que Botha attribue à la simplicité chomskienne, son objectif n'étant  
pas de pallier une insuffisance dans la formulation de la GU (14 p. 179-  
180). Contrairement à Garcia (41 p. 869), des considérations de simpli-  
cité ne sont pas méthodologiquement superflues : "la simplicité -dit-elle-  
est innécessaire lorsqu'elle est en accord avec les intuitions des usa-  
gers et sans pertinence lorsqu'elle les contre-dit". Or, les intuitions  
sur le langage sont des données et la simplicité une hypothèse empirique  
qui les explique ; déclarer/<sup>les hypothèses/</sup>inutiles parce qu'on a accès aux données qui  
les testent, reviendrait à déclarer la physique inutile parce qu'on con-  
naît les mouvements des astres ou la science inutile parce qu'on peut  
observer l'univers.

41 - Ces suggestions, très largement à l'état de programme de travail  
supposent une vision des rapports de la compétence avec la performance  
différente de celle qui est habituelle ; cf. Bès 5, où on introduit la  
distinction, parmi les instructions d'utilisation, entre les règles  
(finies) et les procédures (infinies). Le système linguistique sous-  
jacent, les règles et les procédures se conditionnent réciproquement  
et constituent un triplet d'utilisation, lequel doit être vraisembla-  
blement différent pour auditeur et locuteur. Cette notion d'équili-  
bration entre simplicités conflictives s'apparente à celle d'économie, de  
Martinet, utilisée avec succès dans le cadre diachronique (cf. Martinet  
83 chap. 4 et 84 chap. 6 ; cf. aussi Derwing 35 p. 154n).

- 1- BACH, E. "Linguistique structurelle et philosophie des sciences", in Diogenes 51 (1965), p.117-136.
- 2- BACH, E. Syntactic theory. New York..., Holt, Rinehart and Winston, 1974.
- 2b- BACH, E. "Two Proposals concerning the simplicity metric in phonology", in Glossa, 2 (1969), p. 128-149.
- 3- BACH, E. et R.T.HARMS (eds). Universals in Linguistic Theory. New York..., Holt, Rinehart and Winston, 1968.
- 4- BELLUCI, U. "Simplification in Children's Language", in 62 p.95-119.
- 5- BES, C.C. "Les conduites simulées; fonctions de communication et de représentation", in J.Piaget, J.P.Bronckart et P.Mounoud (Dir.), La Psychologie. Paris, Gallimard, collection La Pléiade, sous presse.
- 6- BES, C.C. "Forme et substance", in A.Martinet (Dir.), La Linguistique; guide alphabétique. Paris, Denoël-Contier, 1969, p.117-124.
- 7- BES, C.C. Identités et différences dans les unités de deuxième articulation. Thèse. Université René Descartes-Paris V, 1972.
- 8- BES, C.C. "Trait distinctif", in Word, 23 (1967), p.37-46.
- 9- BEVER, T.C. "The cognitive basis for linguistic structures", in J.R.Hayes (Ed.) Cognition and the Development of Language. New York, Wiley, 1970, p.279-362.
- 10- BEVER, T.C. et P.S. ROSENBAUM. "Some Lexical Structures and Their empirical Validity", in 64 p.3-19.
- 11- BLOCH, B. "A Set of Postulates for Phonemic Analysis", in Language, 24 (1948), p.3-46.
- 12- BOCHENSKI, I.M. Los métodos actuales del pensamiento. Traduction de R.Drudis Baldrich. Madrid, Ediciones Rialp, 1957.
- 13- BOTHA, R.B. The Justification of Linguistic Hypotheses. The Hague..., Mouton, 1973.
- 14- BOTHA, R.B. Methodological Aspects of Transformational Generative Phonology. The Hague..., Mouton, 1971.
- 15- BOTHA, R.B. The Methodological Status of Grammatical Argumentation. The Hague..., Mouton, 1970.
- 16- BRAINE, M.D.S. "On Two Types of Models of the Internalization of Grammars", in 120 p.153-186.
- 17- BUNGE, M. La investigación científica. Traduction de M.Sacristan. Barcelona, Aries, 1969.
- 18- BUNGE, M. Teoría y realidad. Traduction de J.L.García Molina et J.Sempere. Barcelona, Ariel, 1972.
- 19- CAMPBELL, R. et R.WALES. "The Study of Language Acquisition", in J.Lyons (Ed.), New Horizons in Linguistics. Penguin Books, 1970, p.242-260.
- 20- CHOMSKY, N. Aspects of the Theory of Syntax. Cambridge, Mass., The M.I.T. Press, 1965.
- 20b. CHOMSKY, N. Aspects de la théorie syntaxique. Traduction de J.-C.Milner. Paris, Editions du seuil, 1971.
- 20c- CHOMSKY, N. "Conditions on Transformations", in S.Anderson et P.Kiparsky (Eds.), A Festschrift for Morris Halle. New York, Holt, Rinehart and Winston, 1973, p.232-286.

- 21- CHOMSKY, N. "Current Issues in Linguistic Theory", in 40 p.50-118.
- 22- CHOMSKY, N. "Deep structure, Surface structure, and Semantic Interpretation", in 28 p.62-119.
- 22b-CHOMSKY, N. "Etats initiaux et états stationnaires; l'approche linguistique", in M. Piattelli-Palmarini (Ed.), Théories du langage - Théories de l'apprentissage. Paris, Seuil, 1979, p.169-177.
- 23- CHOMSKY, N. "Explanatory Models in Linguistics", in E. Nagel, P. Suppes et A. Tarski (Eds.), Logic, Methodology and Philosophy of Science. Stanford, Cal., Stanford University Press, p.528-550.
- 24- CHOMSKY, N. Le Langage et la pensée. Traduction de L.-J. Calvet. Paris, Payot, s.d.
- 24b-CHOMSKY, N. "On Wh-movement", in P.W. Culicover, Th. Wasow et A. Akmajian, Formal Syntax. New York..., Academic Press, 1977, p.71-132.
- 25- CHOMSKY, N. Reflections on language. New York, Pantheon Books, 1975.
- 26- CHOMSKY, N. "Some Empirical issues in the Theory of Transformational Grammar", in 28 p.120-202.
- 27- CHOMSKY, N. "Some General Properties of Phonological Rules", in Language, 43 (1967), p.102-128.
- 28- CHOMSKY, N. Studies on Semantics in Generative Grammar. The Hague..., Mouton, 1972.
- 29- CHOMSKY, N. Syntactic Structures. The Hague, Mouton, 1957.
- 29b-CHOMSKY, N. Structures syntaxiques. Traduction de Michel Braudeau. Paris, Editions du Seuil, 1969.
- 30- CHOMSKY, N. "A Transformational Approach to Syntax", in 40, p.211-245.
- 31- CHOMSKY, N. et M. HALLE. "Some Controversial Questions in Phonological Theory", in Journal of Linguistics, 1 (1965), p.97-138.
- 32- CHOMSKY, N. et M. HALLE. The sound Pattern of English. New York, Harper and Row, 1968.
- 32b-CHOMSKY, N. et H. LASNIK. "Filters and controls", in Linguistic Inquiry, 8 (1977), p.425-504.
- 33- CONTRERAS, H. "Simplicity, Descriptive Adequacy, and Binary features", in Language, 45 (1), 1969, p.1-8.
- 33b-COPI, I.M. Introducción a la lógica. Traducción de N. Miguez. Buenos Aires, Eudeba, 1962.
- 34- DELL, P. Les règles et les sons. Paris, Hermann, 1973.
- 35- DERWING, B.L. Transformational Grammar as a theory of Language Acquisition. Cambridge, at the University Press, 1976.
- 35b-DUMAS, J.-P. La grammaire générative et le problème de l'innéisme. Thèse. Université de Clermont II, 1978.
- 36- ERVIN-TRIPP, S. "An Overview of theories of Grammatical Development", in 120 p.189-212.
- 37- FISCHER-JORGENSEN, E. "Remarques sur le principe de l'analyse phonémique", in Recherches Structurales, TOLC, V (1949), p.214-234.

- 38-FLORES D'ARCAIS, C.B. et W.J.M. LEVELT (Eds.) Advances in Psycholinguistics. Amsterdam... ,North-Holland, 1970.
- 39-PODOR, J.A. "How to learn to talk: Some Simple Ways", in 124 ,p.105-128.
- 39b-PODOR, J.A., T.C.BEVER et M.P.CARFETT. The Psychology of Language. New York, McCraw-Hill, 1974.
- 40-PODOR, J.A. et J.J. KATZ (Eds.) The structure of Language. Englewood Cliffs, New Jersey, Prentice-Hall, 1964.
- 41-CARCIA, E. "Auxiliaries and the Criterion of Simplicity", in Language, 43 (1967), p.853-870.
- 42-CROSS, M., M.HALLE et M.P.SCHUTZENBERGER (Eds.) The formal analysis of natural languages. Proceedings of the First International Conference. The Hague..., Mouton, 1973.
- 43-CRUNIC, B. "Les théories transformationnelles. Exposé critique", in La Linguistique, 2 (1965), p.1-24; 1(1969), p.31-101.
- 44-HAGEGE, C. La grammaire générative. Réflexions critiques. Paris, P.U.F., 1976.
- 45-HALLE, M. "On the Bases of Phonology", in 40 p.324-333.
- 46-HALLE, M. "Frontiers of Linguistic Theory" (Tirage à part photocopié).
- 47-HALLE, M. "How not to measure length of lexical representations and other matters", in Journal of Linguistics, 5 (1969), p.305-308.
- 48-HALLE, M. "Markedness", in Proceedings of the Sixth International Congress of Phonetic Sciences, p.61-72.
- 49-HALLE, M. "A New Organon, the Formalism of Phonological Rules", in 42 p.1-11.
- 50-HALLE, M. "Phonology in Generative Grammar", in 40, p.334-352.
- 51-HALLE, M. "On the Role of Simplicity in Linguistic Descriptions", in R.Jakobson (Ed.), Structure of Language and its Mathematical Aspects. Providence, R.I., American Mathematical Society, 1961, p.89-94.
- 52-HALLE, M. The sound Pattern of Russian. Cravenhage, Mouton, 1959.
- 53-HALLE, M. "Theoretical issues in phonology in the 1970's", in Proceedings of the 7th International Congress of Phonetic Sciences (1972), p.179-205.
- 53b-HALLE, M., J.BRESNAN, C.A.MILLER (Eds.) Linguistic Theory and Psychological Reality. Cambridge, Mass...., The M.I.T. Press, 1978.
- 54-HARMS, R.T. Introduction to phonological theory. Englewood Cliffs, New Jersey, Prentice-Hall, 1968.
- 55-HARMS, R.T. "The measurement of Phonological Economy", in Language, 42 (1966), p.602-611.
- 56-HJELMSLEV, L. Prolegomena to a theory of Language. Traduction de P.J.Whitfield. Madison, The University of Wisconsin Press, 1961.
- 56b-HJELMSLEV, L. Prolégomènes à une théorie du langage. Paris, Ed.de Minuit, 1963.
- 57-HOCKETT, Ch.F. A Course in Modern Linguistics. New York, MacMillan, 1958.
- 58-HOCKETT, Ch.F. "A note on structure", in 71 p.279-280.
- 59-HOUSEHOLDER, F.W. "Phonological Theory: A Brief Comment", in Journal of Linguistics, 2 (1966), p.99-100.

- 60- HOUSEHOLDER, F.W. "On some recent claims in phonological theory", in Journal of Linguistics, 1 (1955), p.13-34.
- 61- HUXLEY, R. "The development of the correct use of subject-personal pronouns in two children", in 38 p.141-165.
- 62- HUXLEY, R. et E. INGRAM (Eds.) Language Acquisition: Models and Methods. London, Academic Press, 1971.
- 63- JACOBS, R.A. et P.S. ROSENBAUM (Eds.) English Transformational Grammar. Waltham, Mass., Blaisdell, 1968.
- 64- JACOBS, R.A. et P.S. ROSENBAUM (Eds.) Readings in English Transformational Grammar. Waltham, Mass., Cinn and Co., 1970.
- 65- JAKOBSON, R. Selected Writings. Phonological Studies. Cravenhage, Mouton, 1962.
- 66- JAKOBSON, R., C.C.M. FANT et M. HALLE. Preliminaries to Speech Analysis. Cambridge, Mass., The M.I.T. Press, 1961.
- 67- JAKOBSON, R. et M. HALLE. Fundamentals of Language. The Hague, Mouton, 1956.
- 68- JAKUBOWICZ, C. "Recherches récentes en psycholinguistique", in L'année psychologique 70 (1970), N°1, p.247-293.
- 69- JAKUBOWICZ, C., J. SECUI et J.M. FIEDERPALVI. "L'utilisation des théories linguistiques en psychologie": l'apport de la grammaire générative", in Bulletin de Psychologie, 276(1968-1969), p.685-695.
- 70- JENKINS, J.J. "Reflections on the Conference", in 124 p.347-359.
- 71- JOOS, M. (Ed.) Readings in Linguistics. New York, American Council of Learned Societies 1957.
- 72- KATZ, J.J. "Interpretative semantics vs. generative semantics". in Foundations of Language, 6 (1970), p. 220-259.
- 73- KATZ, J.J. Semantic Theory. New York..., Harper and Row, 1972.
- 74- KIPARSKY, P. "Linguistic Universals and Linguistic Change". in 3 p. 171-210.
- 75- KLEVA, E.S. et BELINGI U. "Syntactic regularities in the speech of children". in 82 p. 183-219.
- 76- KLVČERA, H. "Distinctive Features, Simplicity and Descriptive Adequacy". in To Honor Roman Jakobson. The Hague, Mouton, 1966, p. 1114-1125.
- 77- LAKOFF, G. "Pronominalization, Negation and the Analysis of Adverbs". in 64 p. 145-165.
- 78- LAMB, S.M. "Epilegomena to a Theory of Language". in Románic Philology, 29 (1966), p. 531-573.
- 79- LAMB, S.M. "Prolegomena to a Theory of Phonology". in Language, 42 (1966) p. 536-573.
- 80- LEES, R.B. Compte-rendu de 29. in Language, 33 (1957), p. 375-408.
- 81- LEVELT, W.J.M. Formal Grammars in Linguistics and Psycholinguistics. 3v. The Hague..., Mouton, 1974.
- 82- LYONS, J. et R.J. VALES Eds. Psycholinguistic Papers. Edinburgh, Edinburgh University Press, 1960.
- 83- MARTINET, A. Economie des changements phonétiques. Berna, Francke, 1955.
- 84- MARTINET, A. Éléments de linguistique générale. Paris, A. Colin, 1970.

- 85- MARTINET, A. La linguistique synchronique. Paris, P.U.F., 1965.
- 86- MARTINET, A. "Au sujet des fondements de la théorie linguistique de Louis Hjelmslev". in BSL, 42 (1946), p. 19-42.
- 87- MATHEWS, P.H. "Some remarks on the Householder-Halle contro-versy". in Journal of Linguistics, 4 (1968), p. 275-283.
- 88- McCawley, J.D. "On the role of notation in generative phonology". in 42 p. 51-62.
- 89- McCawley, J.D. "Where Do Noun Phrases Come From ?" in 64 p.166-183.
- 90- McNEILL, D. The Acquisition of Language : The Study of Developmental Psycholinguistic. New York, Harper and Row, 1970.
- 91- McNEILL, D. "Are there specifically linguistic universals ?" in 128 p. 530-535.
- 92- McNEILL, D. "The Capacity for the Ontogenesis of Grammar". in 120 p. 17-40.
- 93- McNEILL, D. "The creation of language by children". in 82 p. 99-132.
- 94- McNEILL, D. "Developmental Psycholinguistic". in 124 p. 15-84.
- 95- MEHLER, J. "Introduction". in Psycholinguistique et grammaire générative, Langages, 16 (1969), p.3-15.
- 96- MEHLER, J. "Studies in Language and Thought Development". in 62 p. 201-229.
- 96b-MEHLER, J. et G. NOIZET. "Vers un modèle psycholinguistique du locuteur". in Textes pour une psycholinguistique. Paris..., Mouton, 1974, p. 7-22.
- 97- MOUNIN, G. La linguistique du XX<sup>e</sup> siècle. Paris, P.U.F., 1972.
- 98- PALERMO, D.S. "On Learning to Talk : Are Principles Derived from the Learning Laboratory Applicable ?". in 120 p. 41-62.
- 99- PARIENTE, J.C. "Grammaire générale et grammaire générative". in Actes de la recherche en sciences sociales, 5/6 (1975), p. 36-49.
- 100- PETERS, S. Ed. Goals of linguistic theory. Englewood Cliffs, New Jersey, Prentice-Hall, 1972.
- 101- PETERS, S. "The projection problem : How is a grammar to be selected". in 100 p. 171-187.
- 102- PETERS, P.S. et R.W. RITCHIE. "On the Generative Power of Transformational Grammars ". in Information Sciences, 6 (1973), p. 49-83.
- 103- PETERS, P.S. et R.W. RITCHIE. "A Note on the Universal Base Hypothesis". in Journal of Linguistics, 5 (1969), p. 150-152.
- 104-PETERS, P.S. et R.W. RITCHIE. "On restricting the Base Component of Transformational Grammars". in Information and Control, 18 (1971), p. 483-501.
- 105- POPPER, K.P. La lógica de la investigación científica. Traducción de V. Sanchez de Zavala. Madrid, Editorial Tecnos, 1973.
- 106- POSTAL, P.M. Aspects of Phonological Theory. New York, Harper and Row, 1968.

- 107- POSTAL, P.M. "The Best Theory". in 100 p. 131-170.
- 108- POSTAL, P.M. "On the Surface Verb "Remind". in C.J. Fillmore et D. Terence Langendoen Eds. Studies in Linguistic Semantics. New York... Holt, Rinehart and Winston, 1971, p. 181-270.
- 109- POTTIER, B. "La grammaire générative et la linguistique". in Travaux de linguistique et de littérature, Université de Strasbourg, VI (1968), p.7-26.
- 109b-RIEGEL, K.F. "Relational interpretation of the language acquisition process". in 38 p. 224-236.
- 110- ROSS, J.R. "On Declarative Sentences". in 64 p. 222-272.
- 111- RUWET, N. Introduction à la grammaire générative. Paris, Plon, 1968.
- 112- SAPORTA, S. et H. CONTRERAS. A Phonological Grammar of Spanish. Seattle, University of Washington Press, 1962.
- 113- SAUSSURE, F. de. Cours de linguistique générale. Paris, Payot, 1968.
- 114- SCHANE, S.S. "The formalization of exceptions in phonology". in 42 p. 63-70.
- 115- SCHANE S.A. Generative phonology. Englewood Cliffs, New Jersey, Prentice Hall, 1973.
- 116- SCHELESINGER, I.M. "Learning grammar : From Pivot to Realization Rule". in 62 p. 79-93.
- 117- SCHELESINGER, I.M. "Production of Utterances and Language Acquisition". in 120 p. 63-101.
- 117b-SIERTSEMA, B. A Study of Glossematics. 2è ed. The Hague, Martinus Nijhoff, 1965.
- 118- SINCLAIR, H. "Sensorimotor Actions Patterns as a Condition for the Acquisition of Syntax". in 62 p. 121-130.
- 119- SLOBIN, D.I. "Data for the Symposium". in 120 p. 3-14.
- 120- SLOBIN, D.I. Ed. The Ontogenesis of Grammar. New York..., Academic Press, 1971.
- 121- SLOBIN, D.I. "Seven Questions about Language Development". in P.C. Dodwell Ed. , New Horizons in Psychology. Penguin Education, 1972, p. 197-215.
- 122- SLOBIN, D.I. "Universals of grammatical development in children". in 38 p. 174-186.
- 123- SLOBIN, D.I. et C.A. FERGUSON Eds. Studies of Child Language Development. New York, Holt, Rinehart et Winston, 1972.
- 124- SMITH, F. et G.A. MILLER Eds. The Genesis of Language. Cambridge, Mass., The M.I.T. Press, 1966.
- 124b-SOBER, E. Simplicity. Oxford, Clarendon Press, 1975.
- 125- SPANG-HANSSSEN, H. "On the Simplicity of Descriptions". in Recherches Structurales, TCLC 5(1949), p. 61-71.



- 126- STAATS, A.W. "Linguistic-Mentalistic Theory versus an Explanatory S-R Learning Theory of Language Development". in 120 p. 103-150.
- 127- STANLEY, R. "Redundancy rules in phonology". in Language 43 (1967), p. 393-435.
- 128- STEINBERG, D.E et L.A. JAKOBOVITS Eds. Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology. Cambridge, At The University Press, 1971.
- 129-TROUBETZKOY, N.S. Principes de phonologie. Traduction de J. Cantineau. Paris, Klincksieck, 1957.
- 130- TWADDELL, W. FREEMAN. "Uses and Abuses of Symmetry". in Second Texas Conference on Problems of Linguistic Analysis in English. Austin, Texas, The University of Texas, 1962, p. 129-137.
- 131- ULDALL; H.J. Outline of glossematics. Part I : General Theory. Copenhagen, Nordisk Sprog-og Kulturforlag, 1957.
- 132- YEN, S.L. "Two Measures of Economy in Phonological Description" in Foundations of Language, 4 (1968), p. 58-69.
- 133- ZIMMER, K.E. "On the evaluation of alternative phonological descriptions". in Journal of Linguistics, 6 (1969), p. 89-98.