Zeitschrift: Études de Lettres : revue de la Faculté des lettres de l'Université de Lausanne

Herausgeber: Université de Lausanne, Faculté des lettres

Band: 9 (1976)

Heft: 1

Artikel: Théories linguistiques et formalisation

Autor: Jolivet, Remi

DOI: https://doi.org/10.5169/seals-870916

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 07.01.2026

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

II

THÉORIES LINGUISTIQUES ET FORMALISATION

Dépassant le structuralisme, la grammaire générative transformationnelle a permis à la linguistique générale d'accomplir un pas décisif dans sa constitution comme « science » rigoureuse : celui de la formalisation.

Qui parcourerait un ouvrage récent de linguistique se réclamant des conceptions de Chomsky, s'il se laisse impressionner par les arborescences et l'implacable déroulement des déductions-dérivations, n'hésiterait sans doute pas trop à souscrire à ce jugement... pourtant entièrement contestable.

Contestable avant tout parce que les grammaires formalisées, au sens strict, sont les grammaires génératives qui ne font pas intervenir l'opération de transformation et dont les fondements sont à chercher dans le courant structuraliste américain connu sous le nom de *distributionnalisme*. Seules des grammaires établies sur cette base ont été effectivement complètement formalisées. Il reste vrai néanmoins que c'est à Noam Chomsky que l'on doit, pour une bonne part, cette formalisation. Mais, par elle-même, elle ne fait pas sortir la linguistique de ce cadre structuraliste.

Or la question que nous voudrions aborder est précisément là. La formalisation, comme son nom l'indique, est une forme particulière donnée à une théorie. Elle variera donc d'une théorie à une autre si celles-ci ne sont pas équivalentes. Ceci revient à dire que le caractère formalisé d'une théorie linguistique n'a pas — en principe — à entrer en ligne de compte lorsqu'on s'efforce de juger de l'adéquation de cette théorie aux faits dont elle prétend rendre compte. Il est donc parfaitement légitime de discuter une théorie linguistique sur la validité, ou plus exactement sur l'intérêt, des points de vue qu'elle retient pour définir son objet — langue ou langage — et sur sa cohérence interne. La formalisation peut, certes, faciliter la mise en évidence de cette cohérence mais elle ne la crée pas.

Rappelant ce que signifie — au sens strict — « formaliser » et sans critiquer dans leurs fondements les conceptions développées par

Chomsky, on voudrait simplement montrer ici qu'en tout état de cause la formalisation ne saurait donc garantir, en elle-même, l'adéquation, l'excellence ou la « vérité » d'une théorie linguistique.

1. *La formalisation*

1.1. Formaliser une théorie, ou la réduire à un *système formel,* cela a un sens bien précis. Il convient d'abord de ne pas confrondre formalisation et *symbolisation*, même si celle-ci est une étape quasi obligatoire de celle-là.

Devant un corpus, fini ou infini, réel ou fictif, la première tâche du linguiste, à quelque niveau qu'il travaille (phonétique, phonologie, signes, etc.), consiste à l'analyser en unités. Ce travail suppose la capacité de décider de ce qui est identique et de ce qui est différent. L'analyse n'est donc possible que grâce à une première opération d'abstraction car, du point de vue substantiel ou concret, deux événements, phoniques ou non, ne peuvent être absolument identiques, et s'il n'y a que des différences, sans constantes, il n'y a plus de science possible. Une fois les unités dégagées, on peut les remplacer par n'importe quel signe (lettre, groupement de lettres, chiffres, etc.) qui — par convention — les symbolisera. Mais il est bien clair que ce n'est pas cette symbolisation qui représente une abstraction mais l'opération même qui a permis le dégagement des unités.

De même, si, par une nouvelle opération d'abstraction, des unités reconnues préalablement comme différentes sont regroupées en une même classe (passage du « son » au phonème, du lexique aux « parties du discours », etc.) parce qu'elles présentent certaines caractéristiques communes, on pourra également symboliser cette classe par un moyen analogue. Là encore la symbolisation en elle-même ne représente aucun « progrès ».

Dans les deux cas la symbolisation est une notation, un codage qu'on peut vouloir construire le plus économiquement possible.

Ces remarques valent aussi pour la *visualisation,* au sens où l'on parle, notamment en pédagogie, de visualisation, par exemple, des rapports syntaxiques par un arbre. Il s'agit d'une forme de symbolisation plus complexe car elle intervient au niveau des relations qui peuvent être de divers types, entre éléments d'un certain ensemble (syntagme, phrase, etc.).

Les procédés, plus ou moins efficaces, de symbolisation ou de visualisation ne doivent donc pas faire illusion. L'important est ailleurs: dans les définitions que la théorie donne des éléments qui sont ensuite symbolisés.

1.2. La formalisation ne doit pas être confondue non plus avec l'*axiomatisation*, bien que l'axiomatisation formelle soit une étape nécessaire vers la formalisation.

Axiomatiser une théorie c'est exhiber les définitions de ses termes primitifs et les propositions que l'on se donne au départ, sans les démontrer, concernant ces termes primitifs.

L'axiomatique est souvent dite *matérielle* s'il y est encore fait référence, ne serait-ce que par l'emploi des mots du langage courant, aux caractéristiques concrètes, « matérielles » pourrait-on dire, des termes primitifs de la théorie. Définitions et propositions primitives revêtent alors un caractère plus ou moins net d'« évidences intuitives ». Mais du même coup, alors que les définitions devraient quasiment « créer » (par abstraction) les objets de la théorie, il risque de se maintenir « une zone plus ou moins floue de significations préalables implicites, auxquelles nous risquons de nous référer, sans le remarquer, dans la suite des démonstrations » [1].

La tentative d'axiomatisation de la géométrie proposée par Euclide est souvent présentée comme une axiomatisation matérielle [2].

En linguistique, diverses théories ont reçu au moins un début d'axiomatisation matérielle. L'essai le plus ancien est sans doute celui de Bloomfield [3] qui définit des termes tels que « énoncé », « communauté linguistique », « langue » et dont le premier postulat, considéré ailleurs comme le postulat fondamental de la linguistique [4] est le suivant :

> « Dans certaines communautés des énoncés distincts sont semblables ou partiellement semblables. »

Hjelmslev développe un corps de cent huit définitions [5], interreliées, dont deux n'en présupposent aucune autre (« analyse » et « opération ») tandis que deux autres sont présupposées dix-sept fois (« classe » et « fonction »). Interviennent également dans ces définitions un certain nombre de termes non définis, antérieurs à la théorie : « description », « objet », « dépendance », « homogénéité », « présence », etc.

Enfin, dans certains travaux récents, on s'efforce d'axiomatiser la linguistique fonctionnelle [6] en donnant des définitions de termes tels que « fonctionnel », « double articulation », « langage », « système sémiotique » etc. et en énonçant certaines propositions les concernant, par exemple, en notant en italiques les termes qui reçoivent une définition :

> « All *features* in *semiotic sets* are *functional.* »

Dans ces exemples l'axiomatisation porte sur les points de vue et les concepts théoriques et descriptifs de théories linguistiques générales. Ce ne sera plus du tout le cas dans les exemples que nous examinerons plus loin. Une grammaire formelle n'est pas la formalisation d'une théorie linguistique mais de la description d'une langue que permet de donner une théorie linguistique.

Par une nouvelle abstraction on passe de l'axiomatisation matérielle à l'axiomatisation *formelle*. Les termes primitifs sont alors des symboles vidés de toute signification intuitive (mais non pas vraiment vides de «sens» puisque chaque symbole s'oppose aux autres symboles utilisés). Ces symboles ne pourront acquérir une signification que dans une *interprétation* de la théorie. Si une même théorie reçoit plusieurs interprétations, un même symbole aura une signification différente dans chaque interprétation. L'axiomatisation formelle aura permis de mettre en évidence l'isomorphisme de ces interprétations.

1.3. Par rapport à l'axiomatisation formelle, la *formalisation* représente une étape supplémentaire dans le processus d'abstraction. Elle consiste à :

> « faire maintenant pour les règles logiques *selon lesquelles* on raisonne, ce qu'on avait fait précédemment pour les postulats *sur lesquels* on raisonne : les énoncer expressément, et en totalité. » [7]

Au terme de ce processus d'abstraction on a formalisé — au sens strict — une théorie. On l'a réduite à un système formel. « Réduire » est bien le verbe qui convient car, à chaque étape du processus d'abstraction, une part du « concret », de l'intuitif présent dans la théorie non formalisée en a été éliminée [8]. Au point qu'il est en principe possible de construire un système formel de façon complètement arbitraire, indépendamment de toute théorie naïve (c'est-à-dire non formalisée)[9].

En pratique, sauf pour l'étude abstraite des systèmes formels, on ne construit pas «gratuitement» un système formel qu'il serait impossible d'interpréter, ainsi les systèmes formels que sont les grammaires formelles ont-ils pour interprétation les énoncés des langues à partir desquelles ils ont été construits.

2. *Quelques grammaires formelles*

2.1. Les grammaires formelles, ou au moins celles que nous envisagerons ici, sont en effet un type particulier de système formel, des

systèmes combinatoires, conçus pour une interprétation par une langue.

La conception d'ensemble est la suivante. Les phrases de la langue pour laquelle on construit la grammaire formelle sont assimilées aux propositions d'une théorie (théorèmes en mathématique, par exemple). Le système formel doit permettre de déduire — on dit plus souvent, en linguistique, « dériver » — les phrases de cette langue à partir des éléments fournis au départ au système formel et de règles de déduction.

Plus généralement, les grammaires formelles sont des mécanismes finis capables de produire ou de reconnaître un ensemble de phrases qui constitue, *par définition*, le *langage* de la grammaire.

Le processus de production vient d'être exposé. Le processus de reconnaissance est symétrique. Une phrase étant donnée, peut-elle être produite par la grammaire ? Appartient-elle à son langage ? Il faut établir la possibilité de remonter de cette phrase aux éléments constitutifs du système formel par les règles de celui-ci. Dans la pratique, les problèmes de la production et de la reconnaissance ne sont évidemment pas toujours identiques.

On notera donc que dire « Chomsky a formalisé la théorie linguistique » n'est vrai qu'en accordant un sens bien particulier au groupe de mots « théorie linguistique ». Chomsky a en réalité montré que la description distributionnaliste du système linguistique qui sous-tend les phrase possibles d'une langue pouvait être considérée — au moins partiellement — comme une interprétation d'un système formel. Mais la théorie linguistique de Chomsky, c'est-à-dire ses points de vue sur le langage, la notion de compétence, la distinction des structures superficielles et des structures profondes, l'universalité et l'innéité postulées pour ces dernières, etc., échappe absolument à toute formalisation. On pourrait même soutenir que les concepts théoriques généraux sont peut-être moins élaborés chez Chomsky qu'ils ne le sont chez Hjelmslev par exemple.

2.2. Le mécanisme d'une grammaire formelle comporte les éléments suivants :

— V : un ensemble de symboles qui constitue le *vocabulaire* de la grammaire, subdivisé en :

  — V_a : ensemble de symboles qui n'apparaissent jamais dans les énoncés, les phrases de la grammaire, mais qui servent à produire ces phrases et à leur assigner

une description. C'est le vocabulaire *auxiliaire* ou *non terminal.*

— V_t : ensemble des symboles des éléments minimaux constitutifs des énoncés (phonèmes, morphèmes, mots... ou une autre unité, choisie arbitrairement). C'est le vocabulaire *terminal* de la grammaire.

Il n'y a pas de symbole qui appartienne à la fois à V_a et à V_t Par ailleurs, dans la suite, nous ne distinguerons pas dans l'écriture la notation du symbole de celle de l'élément qu'il symbolise.

— R : un ensemble de *règles de production* qui permettront de produire les phrases grammaticales pour la grammaire.

Soit en effet V_t^* l'ensemble des séquences possibles avec V_t = {*voit, Pierre, Paul*}, il comporte les séquences '*Pierre voit Paul*' et '*Paul voit Pierre*' mais également les séquences '*voit*', '*Pierre Paul voit*', '*Pierre Pierre Paul voit Paul Pierre voit*', etc.

L'ensemble des règles de production n'autorise la formation que d'une partie des séquences de V_t^*, celles qui seront alors, par définition, *grammaticales* pour la grammaire considérée.

— P : un symbole spécial, appelé *axiome.* La règle de production qui contient ce symbole doit être appliquée la première dans toute dérivation aboutissant à la production d'une phrase. L'axiome de la grammaire c'est l'indication du symbole de départ.

Une grammaire formelle générative G c'est donc :

— un vocabulaire auxiliaire V_a

— un vocabulaire terminal V_t

— un ensemble de règles de production R

— un axiome P

ou encore : $G = \{V_a , V_t , R , P\}$

Dans ce qui suit, les constituants d'une grammaire générative seront toujours donnés dans cet ordre et on ne répétera pas $V_a = \{...\}$, $V_t = \{...\}$, $R = \{...\}$.

2.3. Imaginons une machine susceptible de passer par une succession d'états (notés : P, S1, S2, S3, ...) en émettant un symbole (*a, b, c,* etc.) à chaque changement d'état. Une grammaire formelle générative associée à ce type de machine pourrait avoir la forme suivante :

$$G_1 = \{P, S1, S2, S3, ..., S25\}, \{a, b, c, ..., y, z\}, \{P \rightarrow aS1) (S1 \rightarrow bS2) (S2 \rightarrow cS3) ... (S24 \rightarrow yS25) (S25 \rightarrow z)\}, \{P\}$$

La règle (S2→*c*S3) signifie que la machine passe de l'état S2 à l'état S3 en émettant le symbole *c*, ou encore, du point de vue de la grammaire, que S2 est remplacé par la séquence *c*S3 dont l'élément S3 sera à son tour remplacé par *d*S4, etc. D'où le nom de *règles de réécriture* donné à ces opérations.

La première règle à appliquer est (P→*a*S1) puisque P est l'axiome de la grammaire, la seule règle applicable ensuite à la séquence obtenue est évidemment (S1→*b*S2) et ainsi de suite. La grammaire G_1 produit donc la séquence :

abcdefghijklmnopqrstuvwxyz

Si à la place des symboles *a, b, c,* etc., le vocabulaire terminal et les occurrences de ses éléments dans les règles de production avaient été composés des mots d'une phrase (de vingt-six mots) c'est cette phrase qui aurait été produite. Bien entendu il n'y a pas de limite au nombre des règles de production et on peut très bien concevoir une grammaire formelle sur ce modèle capable de produire toutes les phrases de tous les volumes de la Bibliothèque Nationale.

Une grammaire formelle de ce type est appelée *grammaire à états finis* ou encore *grammaire de listes* puisqu'elle revient à faire la liste des phrases de la langue.

On peut construire des grammaires de ce type qui permettent de sortir du cadre d'un corpus fini et d'engendrer un nombre infini de phrases. Pour cela on introduit des *boucles* dans la grammaire, c'est-à-dire une ou des règles permettant le passage d'un état donné à lui-même avec émission de symbole à chaque passage. Ces symboles peuvent alors être répétés indéfiniment. Un mécanisme de ce type est dit *récursif*. Pour qu'une grammaire engendre un nombre infini de phrases il faut qu'elle possède au moins une règle récursive. La grammaire suivante produit un nombre infini de phrases grâce à la règle (P→*a*P) qui est récursive :

$$G_2 = \{P\}, \{a, .\}, \{(P \rightarrow .) (P \rightarrow aP)\}, \{P\}$$

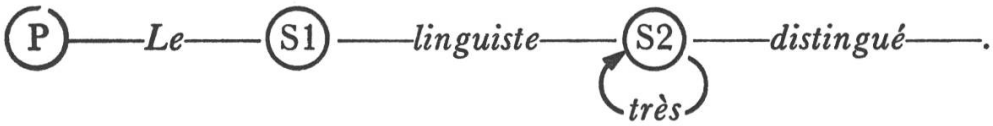
Cette grammaire produit les phrases suivantes :

.
a.
aa.
aaa.
aaaa. etc.

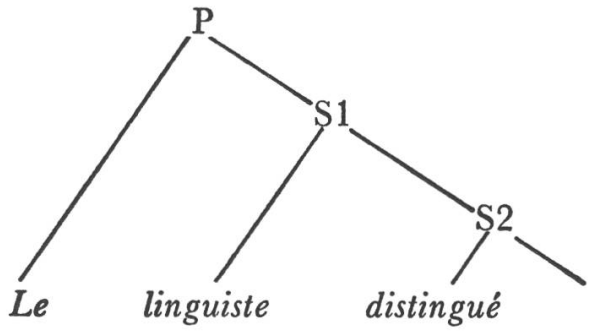
A la grammaire,

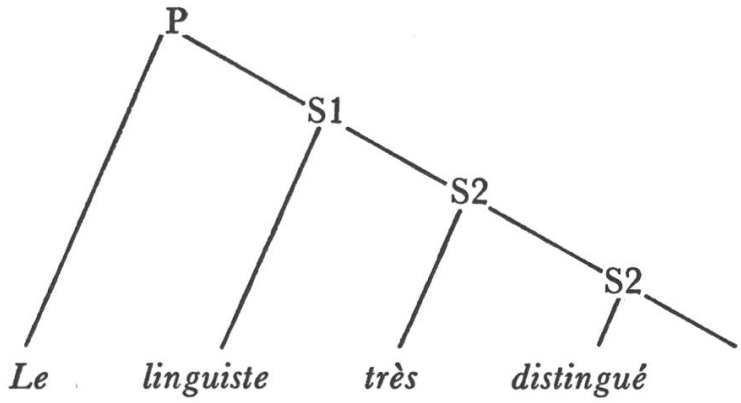
$G_3 =$ {P, S1, S2} , {*distingué, Le, linguiste, très, .*} ,
{(P→*Le*S1) (S1→*linguiste*S2) (S2→*très*S2) (S2→*distingué.*)} , {P}

récursive grâce à la règle (S2→*très*S2) et qui peut donc produire une infinité de phrases (*Le linguiste distingué. Le linguiste très distingué. Le linguiste très très distingué.*, etc.),
on peut associer le graphe :



Une arborescence peut visualiser la production de chaque phrase.





etc.

Du point de vue strictement formel ces grammaires sont pleinement satisfaisantes. Du point de vue linguistique ces grammaires ne conviennent pas parce qu'elles ne nous disent que fort peu de choses

sur la langue. La théorie linguistique qui est à la base de cette formalisation peut se résumer en une phrase : une langue est un ensemble, fini ou infini, de phrases constituées elles-mêmes d'une succession d'éléments appartenant à un vocabulaire. On voit les équivalences avec quelques concepts linguistiques : ensemble fini et corpus, ensemble infini et « créativité » du langage, succession et linéarité (ou syntagmatique), vocabulaire et paradigmatique. Tout cela, qui n'est pas faux, ne va néanmoins pas très loin dans l'analyse du fonctionnement des langues. On peut cependant ajouter que, du point de vue linguistique, cette grammaire a été améliorée par l'introduction de la notion de probabilité d'apparition d'un élément du vocabulaire en fonction des éléments apparus antérieurement. Sous cette forme la grammaire à états finis a connu des développements importants dans le cadre de la théorie de la communication [10]. Mais, même sous cette forme, cette grammaire ne peut absolument rien nous apprendre — directement — sur la hiérarchie syntaxique qui s'établit entre les éléments constitutifs d'une phrase.

Si nous exigeons — et cette exigence est proprement linguistique, elle n'a rien à voir avec la formalisation — qu'une grammaire formelle nous en dise plus, il faut qu'elle se fonde sur une théorie de la langue plus raffinée que la simple constatation de l'existence de l'axe paradigmatique et de l'axe syntagmatique. Ceci veut dire que la formalisation en elle-même ne rend pas compte du système de la langue, formaliser n'est pas découvrir ce système. Ce que l'on formalise ce n'est pas le système linguistique directement mais le système linguistique tel qu'une théorie linguistique permet de le décrire.

2.4. Il existe des grammaires formelles qui rendent compte non seulement de la succession des éléments dans une phrase mais aussi de la hiérarchie qui s'établit entre ces éléments, c'est-à-dire de la structuration de la phrase. Ces grammaires sont appelées *grammaires de phrases structurées.*

C'est le structuralisme américain qui a fourni la théorie linguistique de la structure des phrases permettant une formalisation d'un système linguistique. Au niveau syntaxique, la procédure analytique développée par cette théorie est l'analyse en *constituants immédiats.*

2.4.1. L'analyse en constituants immédiats décompose chaque phrase du corpus en éléments, puis chaque élément en éléments plus petits et ainsi de suite jusqu'à l'obtention d'éléments ultimes, indécomposables. On peut obtenir diverses représentations (visualisations) de la structure d'une phrase, de la hiérarchie qui organise ses divers com-

posants. Ainsi peut-on associer à l'analyse en constituants de la phrase *ce linguiste distingué prononce une conférence* trois représentations équivalentes.

— Représentation en « boîtes » de Charles Hockett [11] :

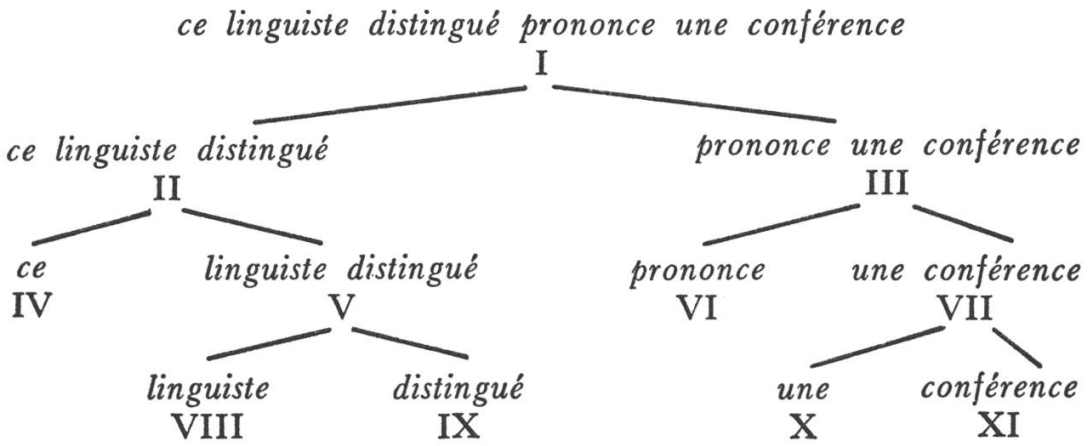
<table>
<tr><td colspan="6">ce linguiste distingué prononce une conférence
I</td></tr>
<tr><td colspan="3">ce linguiste distingué
II</td><td colspan="3">prononce une conférence
III</td></tr>
<tr><td rowspan="2">ce
IV</td><td colspan="2">linguiste distingué
V</td><td rowspan="2">prononce
VI</td><td colspan="2">une conférence
VII</td></tr>
<tr><td>linguiste
VIII</td><td>distingué
IX</td><td>une
X</td><td>conférence
XI</td></tr>
</table>

— Représentation par « parenthétisation » de Rulon Wells [12] :

(((*ce*) ((*linguiste*) (*distingué*)))
I II IV IV V VIII VIII IX IX V II

((*prononce*) ((*une*) (*conférence*))))
III VI VI VII X X XI XI VII III I

— Représentation « en arbre » :



Ces trois représentations, strictement équivalentes, rendent compte de la structure d'une phrase particulière. Pour généraliser on désigne chaque constituant d'une phrase par un terme général susceptible de s'appliquer aux constituants d'autres phrases partiellement ou totalement construites sur le même schéma syntaxique. Cela revient à dégager des « classes » syntaxiques abstraites sur la base d'une certaine communauté du comportement combinatoire ou fonctionnel des segments réalisant ces classes abstraites. Les représentations obtenues alors sont dites « étiquetées » :

<table>
<tr><td colspan="6">phrase
(P)</td></tr>
<tr><td colspan="3">syntagme nominal
(SN)</td><td colspan="3">syntagme verbal
(SV)</td></tr>
<tr><td rowspan="2">déterminant
(D)

Si l'on représentait de la même manière la phrase,

cette chanteuse ravissante connaît la musique

la parenté structurelle de cette phrase avec notre exemple apparaîtrait immédiatement.

La représentation « en arbre », plus commode et plus lisible, est très généralement adoptée ; on voit néanmoins qu'il existe d'autres procédés de visualisation qui sont strictement équivalents.

L'analyse en constituants immédiats permet donc :

— de dégager une structure hiérarchique pour une phrase,

— de comparer de ce point de vue des phrases différentes,

— de visualiser cette structure hiérarchique.

Toutes ces possiblités renvoient à la *description* d'un énoncé. Mais l'analyse en constituants immédiats autorise aussi une interprétation en terme de *production* d'un énoncé. Il suffit de partir du constituant le plus général (phrase) et de le réécrire sous la forme des éléments qui le constituent (par exemple : syntagme nominal et syntagme verbal) et ainsi de suite jusqu'aux éléments ultimes, lexicaux. La procédure analytique de segmentations successives d'une séquence est en fait analogue à la procédure productrice, générative, des règles de réécriture.

2.4.2. Ces remarques faites il est assez facile de formaliser la description d'un corpus d'énoncés sur la base de l'analyse en constituants immédiats. La grammaire formelle comprendra :

— un vocabulaire auxiliaire V_a contenant les symboles des classes syntaxiques abstraites ;

— un vocabulaire terminal V_t qui contient les éléments ultimes des phrases, en pratique les symboles des unités lexicales ou grammaticales « concrètes » ;

— un ensemble R de règles grâce auxquelles on pourra réécrire un constituant en ses éléments composants, jusqu'aux éléments ultimes (appartenant à V_t). Dans la partie gauche de la règle figure le constituant à réécrire et dans la partie droite les éléments par lesquels il convient de le réécrire, séparés par le signe +, partie gauche et partie droite sont séparées par le signe →. Le sens intuitif de la règle P→SN+SV, par exemple, est le suivant :

  — du point de vue du mécanisme formel, P doit être remplacé (réécrit) par la suite SN+SV,

— du point de vue linguistique SN+SV « est un » P: la succession d'un syntagme nominal et d'un syntagme verbal constitue une phrase. De ce fait même, SN et SV, dans cet ordre, entretiennent entre eux une certaine relation ;

— un axiome P qui constitue la partie gauche de la première règle à appliquer, puisque cette partie gauche ne pourrait, elle, être produite par aucune règle.

La grammaire formelle produisant notre exemple :

ce linguiste distingué prononce une conférence

serait la suivante :

G_4 = {P, SN, SV, GN, D, A, N, V} , {*ce, distingué, conférence, linguiste, une, prononce*} , {(P→SN+SV) (SN→D+GN) (SN→D+N) (SV→V+SN) (GN→N+A) (D→*ce/une*) (N→*linguiste/conférence*) (A→*distingué*) (V→*prononce*)}, {P}

(le signe / se lit « ou » ; la notation (D→*ce/une*) est donc une notation abrégée pour les deux règles (D→*ce*) et (D→*une*).)

En appliquant ces règles dans un certain ordre on pourra obtenir la phrase

ce linguiste distingué prononce une conférence

et associer à cette phrase une visualisation des rapports hiérarchiques définis par la grammaire formelle (par exemple l'arbre présenté ci-dessus).

On vérifiera qu'il est également possible d'obtenir d'autres phrases à partir de cette même grammaire :

ce conférence prononce une linguiste distingué
une conférence prononce ce linguiste distingué
ce conférence distingué prononce une linguiste
une conférence distingué prononce ce linguiste

Le *langage* de la grammaire G_4 comporte outre ces quatre phrases et la phrase de notre exemple, cinq phrases où *distingué* apparaît dans les deux SN et cinq phrases où *distingué* n'apparaît pas car les deux SN ont été réécrits par (SN→D+N). Toutes ces phrases sont *grammaticales pour G_4* du point de vue simplement formel. Le langage formel de G_4 n'a donc pas d'interprétation en français puisque nombre de ces phrases sont certainement non grammaticales du point de vue linguistique. En fait il manque à G_4 la possibilité de tenir compte de certaines incompatibilités (« accord en genre » ; une *conférence* peut, en un certain sens, *prononcer* quelque chose mais pas

un *linguiste* !). Il existe divers procédés permettant de remédier à ces inconvénients, le plus satisfaisant, pour le linguiste, étant l'utilisation de règles de réécriture qui tiennent compte du contexte d'apparition d'un élément pour sa réécriture (cf. ci-après).

2.4.3. Il existe en effet plusieurs variétés de ces grammaires formelles. Elles ont fait l'objet d'études approfondies, notamment de la part de Noam Chomsky [13]. Les critères pertinents pour l'analyse et le classement de ces grammaires sont les suivants :

— *longueur* (définie par le nombre d'éléments de V) de la partie gauche et de la partie droite des règles ;

et surtout :

— *conditions d'application des règles.* On distingue en effet des *règles contextuelles* et des *règles non contextuelles.*

Les règles (SN→D+N) et (SN→D+GN) sont non contextuelles. Toute occurrence de SN peut toujours être réécrite D+N ou D+GN. Par contre, la règle ($a+b+c \rightarrow a+d+e+f+c$) est une règle contextuelle : le symbole b est réécrit $d+e+f$ lorsqu'il est précédé immédiatement de a et suivi immédiatement de c.

La règle ($V+N \rightarrow V_{tr}+N$) est également contextuelle, elle pourrait se lire : « verbe » se réécrit « verbe transitif » s'il est suivi d'un « nom ». Une autre notation possible pour ces deux règles serait la suivante :

($b \rightarrow d+e+f / a__c$) (où / se lit ici : « dans le contexte »)
($V \rightarrow V_{tr}$ / __N)

En généralisant la forme des règles de réécriture dans l'une des formules suivantes :

$A \rightarrow X/Y__Z$ ou $YAZ \rightarrow YXZ$

il y a quatre types de règles possibles dont trois sont contextuels et un non contextuel :

$A \rightarrow X/Y__Z$	ou	$YAZ \rightarrow YXZ$	avec Y et Z non nuls
$A \rightarrow X/Y__$	ou	$YA \rightarrow YX$	avec Y non nul, Z nul
$A \rightarrow X/__Z$	ou	$AZ \rightarrow XZ$	avec Y nul, Z non nul
		$A \rightarrow X$	avec Y et Z nuls

Il est clair, intuitivement, qu'une grammaire formelle dépourvue de règles contextuelles risque fort d'être inadéquate du point de vue linguistique. La démonstration de cette intuition a souvent été faite [14]. Elle s'inscrit dans le cadre plus général de l'étude des rapports entre grammaires ou langages formels et langues naturelles.

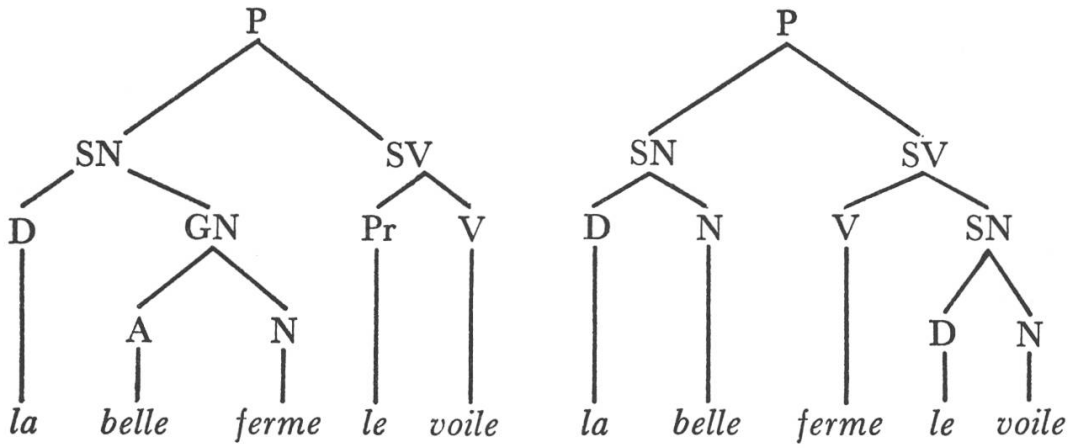
3. *Grammaires formelles et langues naturelles*

3.1. Soit la phrase :

la belle ferme le voile

elle est ambiguë, le syntagme nominal sujet pouvant être soit *la belle* soit *la belle ferme*. Ce type d'ambiguïté est fréquent dans les langues naturelles où il s'explique à la fois par des faits morphologiques et par le statut du niveau syntaxique et du niveau lexical, largement indépendants l'un de l'autre. Une grammaire formelle adéquate pour une langue naturelle devra donc rendre compte de ces faits, c'est-à-dire produire deux (ou plus de deux) séquences identiques mais en attribuant à chacune une structuration différente, ce qui apparaîtra dans la visualisation. La grammaire suivante serait, de ce point de vue, satisfaisante pour la phrase prise comme exemple.

G_5 = {P, SN, SV, D, N, V, A, Pr} , {*belle, ferme, la, le, voile*} , {(P→SN+SV) (SN→D+N) (SN→D+GN) (GN→A+N) (SV→V+SN) (SV→Pr+V) (D→*la/le*) (N→*belle/voile/ferme*) (V→*ferme/voile*) (A→*belle*) (Pr→*le*)} , {P}



Il existe néanmoins d'autres cas d'ambiguïté dont les grammaires formelles présentées ici ne peuvent rendre compte. Ainsi dans les deux phrases suivantes :

cette critique de Chomsky ne vaut rien (« génitif objectif » vs. « subjectif »)

il pèse cent vingt livres (*livre* : objet ou unité de poids)

La notion de transformation permet de traiter le premier cas. Deux phrases distinctes en structure profonde, c'est-à-dire produites

par une grammaire générative sans règles de transformation, peuvent aboutir, en structure superficielle, donc après l'application de transformations, à une même succession d'éléments, dotée d'une même structure hiérarchique (syntaxique). En simplifiant à l'extrême on aurait en structure profonde d'une part *X critique Chomsky* et d'autre part *Chomsky critique X*. Enfin, la partie sémantique de la grammaire générative et transformationnelle, au moins dans la théorie « standard », permettrait de rendre compte de l'ambiguïté de *livre* dans la deuxième phrase [15].

3.2. Une autre caractéristique importante des langues naturelles est, on l'a vu, le caractère illimité de leur « langage » (au sens formel), c'est-à-dire le nombre infini des phrases qu'elles permettent de produire. Ceci parce que la longueur des phrases n'est pas structurellement limitée puisqu'il existe des éléments susceptibles d'être indéfiniment répétés :

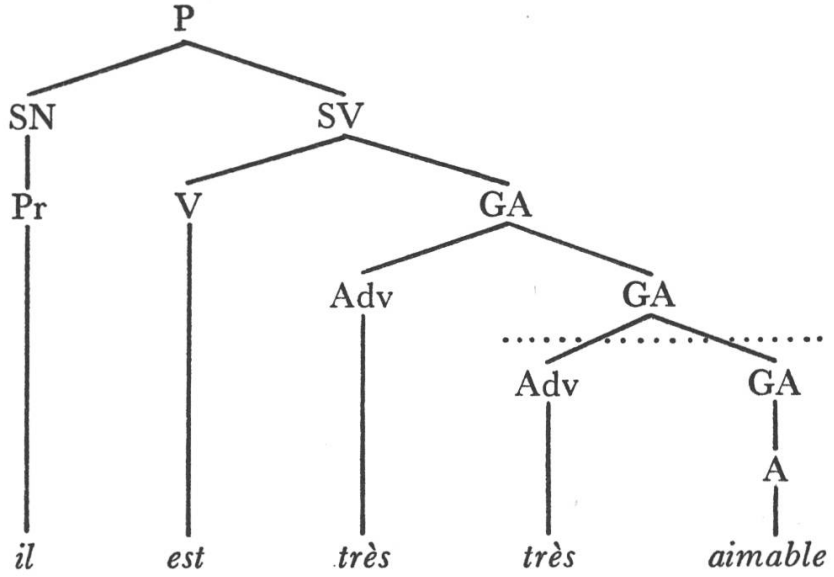
il est très très (...) aimable
Jean est un garçon aimable, gentil, serviable (...)
Jean que Marc (...) connaît connaît Matthieu

Une grammaire formelle adéquate devra donc comporter des mécanismes récursifs. Si un élément ou une séquence d'éléments figure comme partie gauche et dans la partie droite d'une même règle (ou dans la partie droite d'une règle qui n'a pas encore été appliquée), cet élément ou cette séquence d'éléments est récursif et peut être répété.

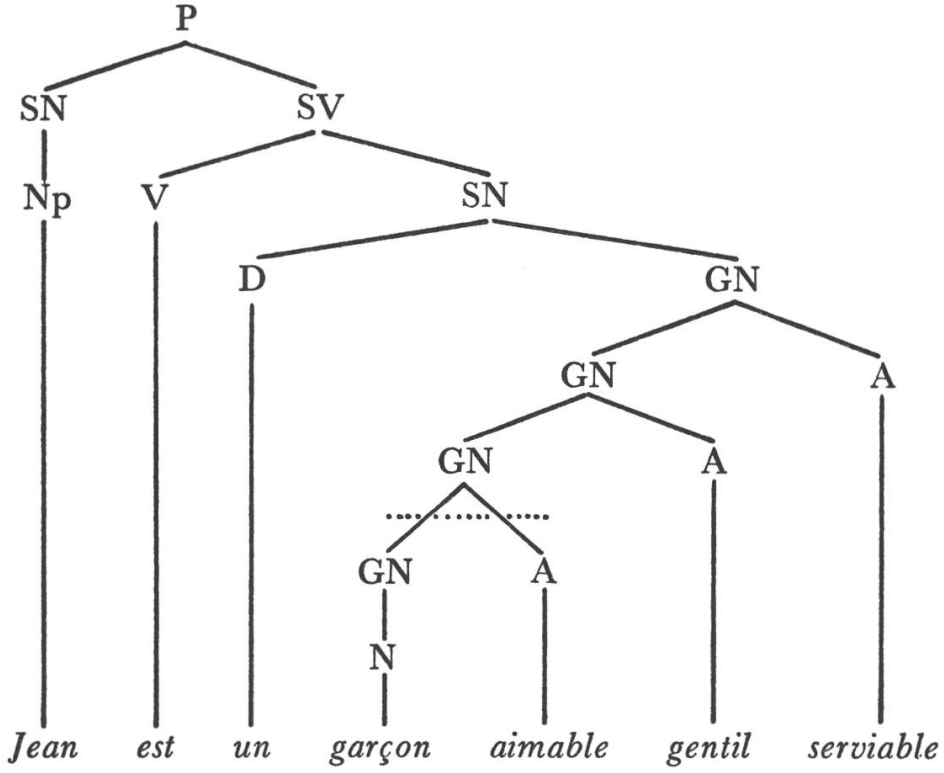
La grammaire G_6 ci-dessous permettrait de produire l'infinité des phrases constituées sur le modèle des trois exemples précédents. Elle produirait également bien d'autres phrases, grammaticales ou non du point de vue du français.

G_6 = {P, SN, SV, Pr, V, GA, Adv, A, Np, D, GN, N, PR, Prel}
{*il, un, aimable, gentil, serviable, Jean, Marc, Matthieu, garçon, est, connaît, que, très*}
{(P→SN+SV) (SN→Pr) (SN→Np) (SN→Np+PR) (SN→D+GN) (SV→V+GA) (SV→V+SN) (GA→Adv+GA) (GA→A) (GN→GN+A) (GN→N) (PR→Prel+SN+V) (Pr→*il*) (Np→*Jean/Marc/Matthieu/...*) (V→*est/connaît*) (D→*un*) (Adv→*très*) (A→*aimable/gentil/serviable/...*) (N→*garçon*) (Prel→*que*)}
{P}

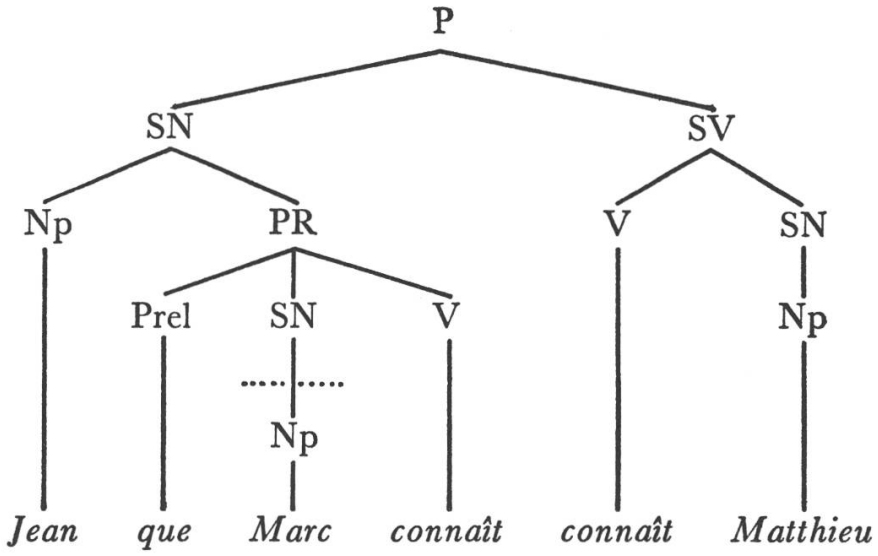
La règle (GA→Adv+GA) permet d'obtenir la phrase :



Avec la règle (GN→GN+A) on peut produire :



Enfin, les deux règles (SN→Np+PR) et (PR→Prel+SN+V) donneraient :

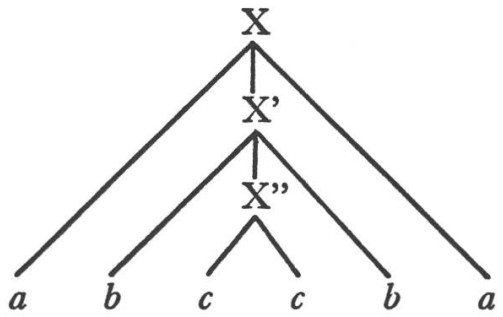


On notera toutefois que la structure des relations syntaxiques entre éléments coordonnés (*très* dans le premier exemple et les adjectifs dans le second) n'est pas rendue de façon linguistiquement satisfaisante puisque la grammaire formelle introduit une hiérarchie là où elle n'existe pas normalement dans la langue ; comparez :

un grand grand homme = personnage tout à fait illustre,
un grand grand homme = personnage illustre et de haute taille.

Ce défaut est pratiquement inévitable lorsque les mécanismes récursifs sont confiés à une grammaire générative. En introduisant l'opération de transformation, Chomsky parvient à traiter ces phénomènes de façon généralement plus satisfaisante [16]. Mais après l'introduction de cette opération extrêmement puissante la grammaire n'est plus un système formel.

3.3. Les divers types de structure hiérarchique entre éléments d'une phrase peuvent être étudiés de façon formelle. Ainsi la possibilité, au moins théorique, d'imbriquer des relatives *(Jean que Marc que Luc estime connaît écrit)* peut être, en simplifiant, généralisée sous la forme suivante où la même lettre désigne des éléments en rapport de dépendance syntaxique (de sujet à verbe dans l'exemple linguistique) :



Formellement il s'agit d'un langage « en miroir » : une séquence (*abc*) suivie de sa répétition dans l'ordre inverse (*cba*). On peut montrer, ce qui est évident sur la figure, que la grammaire susceptible de produire ce type de langage avec cette structuration devra comporter une règle dont la partie droite est de longueur supérieure à 2 : les deux éléments en rapport de dépendance avec, imbriqué entre les deux, la source d'une nouvelle séquence. Dans G_6 ci-dessus c'était le cas de la règle (PR→Prel+SN+V) dont la partie droite est de longueur 3.

Il est possible ainsi, pour chaque type de langage, de définir les types de grammaire susceptibles de les produire. Ceci du point de vue formel. L'application linguistique est immédiate. Si l'on observe, dans une langue naturelle, un certain type de dépendance, déjà étudié formellement, on sait du même coup quelle grammaire formelle pourra produire les phrases de cette langue.

On a pu ainsi démontrer que les grammaires dont toutes les règles sont non contextuelles ne permettent pas de produire de façon linguistiquement satisfaisante toutes les phrases d'une langue naturelle. Seules les grammaires formelles possédant des règles contextuelles pourraient donc être retenues pour formaliser la description d'un système linguistique, lorsque cette description est sous-tendue par les principes linguistiques généraux du distributionnalisme et de l'analyse en constituants immédiats.

3.4. En fait on a pu aller plus loin et montrer que les grammaires formalisant la description d'une langue à partir de l'analyse en constituants immédiats sont linguistiquement inadéquates. Ce qui revient à établir l'insuffisance de la théorie linguistique distributionnelle... mais sûrement pas, du même coup, de toutes les théories linguistiques se réclamant du « structuralisme », le rapport entre ce concept (?) et les théories qui s'en réclament n'étant pas du même ordre que le rapport entre une théorie et son application. L'examen de quelques-unes des limitations des grammaires génératives non transformationnelles est, à cet égard, assez instructif.

On a déjà noté que, dans le cas d'éléments coordonnés, les grammaires génératives ne sont pas linguistiquement satisfaisantes : elles introduisent « trop de structure ». Ceci est en réalité la conséquence de la conception chomskyenne de la grammaire comme un mécanisme (système formel) *fini* susceptible de produire un nombre infini de phrases. Si la finitude du mécanisme n'était pas exigée pour des raisons théoriques (limitation des capacités mémorielles, de la « compétence »), on pourrait concevoir une grammaire générative com-

prenant, à la place d'éléments récursifs, un nombre infini de règles. Cette grammaire serait linguistiquement correcte du point de vue descriptif. A la place des deux règles de G_6 ci-dessus (GA→Adv+GA) et (GN→GN+A) on aurait respectivement : (GA→A) (GA→Adv+A) (GA→Adv+Adv+A) (GA→Adv+Adv+Adv+A), etc. et (GN→N) (GN→N+A) (GN→N+A+A) (GN→N+A+A+A), etc.

Sur ce point on pourrait donc soutenir que l'analyse en constituants immédiats n'est pas vraiment mise en défaut mais que l'inadéquation provient d'une exigence supplémentaire, posée au stade de la formalisation, celle du caractère fini du mécanisme génératif. En fait, même dans le cadre transformationnel, certains types de coordination continuent à poser des problèmes difficiles [17]. Ceci tient sans doute à la place tout à fait particulière qu'occupe la coordination parmi les procédés de formation des phrases. C'est la notion même de coordination qu'il faudrait sans doute d'abord approfondir pour parvenir à des solutions plus satisfaisantes.

La deuxième limitation importante des grammaires génératives est l'impossibilité dans laquelle elles se trouvent de rendre compte de certains cas d'ambiguïté. On en a vu deux exemples typiques (cf. 3.1.) et la façon dont la grammaire transformationnelle prétend les traiter : dans un cas (génitif objectif vs. subjectif) de façon « syntaxique » et dans l'autre *(livre* unité de poids ou objet) grâce à la composante « sémantique » de la grammaire. Mais la limite entre syntaxe et sémantique est en fait posée de façon arbitraire. Il est plus simple de voir dans ces deux cas d'inadéquation la conséquence du fait que l'analyse distributionnelle ne tient pas compte du sens. C'est bien clair dans le deuxième cas. Mais l'explication vaut aussi dans le premier : la préposition *de* peut avoir plusieurs sens (cp. *la maison de pierre* et *la maison de Pierre*, *le train de nuit* et *le train de Paris,* etc.) et, dans certains contextes, il peut y avoir ambiguïté. Si *de* est considéré comme un marqueur de fonction, on dira qu'il y a plusieurs fonctions *de*, compatibles d'ailleurs dans l'énoncé: *la maison de poupée de ma fille*. La même analyse vaudrait pour le latin où deux génitifs l'un objectif et l'autre subjectif peuvent déterminer un même noyau :

> *tametsi pro veteribus Helvetiorum injuriis populi Romani ab his poenas bello repetisset* [18]

Postuler l'existence d'une structure profonde et d'une structure superficielle reliées par des règles transformationnelles est donc une solution possible pour résoudre ces difficultés, mais ce n'est ni la seule solution ni la plus simple dès lors qu'on accepte de considérer les uni-

tés linguistiques, phonèmes mis à part, comme dotées d'une forme et d'un sens.

Mais la limitation la plus importante de l'analyse en constituants immédiats réside dans le fait qu'elle ne permet pas de s'affranchir du caractère linéaire de l'énoncé. Elle ne peut donc pas analyser de façon linguistiquement satisfaisante une unité linguistique unique lorsque son signifiant est divisé en fractions non contiguës dans la chaîne. Or cette situation se rencontre fréquemment dans les langues. C'est le cas par exemple de la négation *ne ... pas* en français, la présence de *ne*, au moins dans un certain registre, impliquant le plus souvent celle de *pas* et réciproquement, lors même que ces deux segments ne sont pas contigus. Il existe bien entendu divers procédés permettant de produire ces unités d'une façon formelle dans une grammaire générative mais ce sont des artifices techniques sans fondement ni valeur linguistiques. Cette limitation est absolument fondamentale et tous les linguistes qui ont travaillé dans le cadre théorique fourni par le distributionnalisme ont tenté de résoudre le problème. Qu'il s'agisse de Zellig S. Harris qui traite de façon intuitivement satisfaisante la question des unités de type *ne ... pas* sur la base de la dépendance complète des fragments non contigus : l'un n'apparaissant pas sans l'autre il s'agit d'une seule unité [19]. Qu'il s'agisse encore de Rulon Wells qui, s'attaquant à des cas autrement ardus (ex. un sujet séparé du verbe par une incise) est conduit à réintroduire le sens pour définir l'unité du constituant [20]. Ces solutions sont intuitivement acceptables du point de vue descriptif. Mais elles ne sont pas formalisables dans le cadre d'une grammaire de production, générative, car elles ne permettent pas d'établir directement, sur la séquence linéaire, une hiérarchie linguistiquement satisfaisante entre les éléments de la phrase [21].

Ici aussi il est donc difficile de faire la part de ce qui dans cette limitation provient d'un défaut de l'analyse en constituants et de ce qui provient d'une nouvelle définition, indépendante, de la notion même de grammaire comme

> « mécanisme fini capable d'engendrer un ensemble infini de phrases grammaticales, et de leur associer automatiquement à chacune une description structurale. » [22]

Quoi qu'il en soit, on peut tirer deux conclusions générales importantes de l'étude de ces limitations :

— une description linguistique qui ne fait pas une place à la sémantique ne peut être entièrement adéquate. Toute théorie linguistique générale devra donc traiter le problème du sens ;

— une description linguistique qui ne considère comme pertinentes que des relations entre segments contigus, donc des relations marquées par la position relative des constituants de la phrase, ne peut être entièrement adéquate. Toute théorie linguistique générale devra donc se doter de concepts théoriques et de procédures descriptives permettant de s'affranchir de la linéarité du signifiant.

La théorie linguistique de Chomsky et les grammaires transformationnelles qui en découlent satisfont ces deux exigences. La théorie fonctionnelle d'André Martinet et les descriptions linguistiques qui en découlent satisfont aussi ces deux exigences.

4. *Linguistique et formalisation*

4.1. On a vu que la grammaire formelle la plus élémentaire (grammaire à états finis) était capable de produire un nombre illimité de phrases. Si l'on fait du caractère formalisé ou formalisable d'une théorie linguistique (ou de la description d'un système linguistique qu'elle autorise) son seul critère de validité, alors la théorie linguistique qui sous-tend ce type de grammaire élémentaire (une langue est un ensemble de phrases constituées d'une succession d'éléments choisis dans un vocabulaire) est valide. Soit. Mais cette théorie est-elle suffisante ? On accordera que c'est le droit du linguiste de chercher plus loin et de requérir d'une grammaire formelle qu'elle mette en évidence d'autres caractéristiques du langage. Ce sont ces interventions du linguiste, qui n'ont rien à voir avec la formalisation en elle-même, que nous rappellerons ici pour commencer.

C'est le linguiste qui décrète que la grammaire formelle doit pouvoir produire une infinité de phrases et non pas seulement les phrases attestées dans un corpus fini, aussi étendu que l'on voudra le prendre. Du point de vue formel cette exigence n'est pas difficile à satisfaire: il suffit d'introduire dans la grammaire au moins une règle récursive, ou une infinité de règles. Mais cette facilité n'est qu'apparente car elle se heurte aux autres exigences posées par le linguiste.

L'une de ces exigences est que la grammaire formelle attribue aux phrases qu'elle produit une hiérarchie déterminée indépendamment de et antérieurement à la formalisation. On a vu que cette exigence amenait à rejeter l'utilisation de règles récursives.

Mais si l'on réclame en outre d'une grammaire qu'elle soit un mécanisme fini, alors la deuxième solution possible pour produire un nombre illimité de phrases doit également être rejetée et la première exigence ne peut plus être satisfaite. Les grammaires formelles de type génératif sont donc inadéquates, mais, comme dans le cas des gram-

maires à états finis, pour des raisons strictement linguistiques sans rapport avec la formalisation en elle-même.

C'est le linguiste enfin qui décide des phrases que devra produire la grammaire formelle et de celles qu'elle devra exclure. D'où l'ambiguïté du concept de *grammaticalité*. Est grammatical, au sens formel, l'ensemble des phrases qu'une grammaire formelle est susceptible de produire. Mais si ces phrases doivent avoir une interprétation pour les phrases d'une langue naturelle donnée, la notion de grammaticalité retrouve un sens plus intuitif et plus familier. Seront grammaticales les phrases attestées dans un corpus ou acceptées par une fraction importante de la communauté linguistique considérée. On retrouve tous les problèmes familiers aux linguistes de variétés des usages, de norme, etc.

4.2. Supposons maintenant toutes ces difficultés résolues. Même si nous disposions d'une grammaire du français entièrement formalisée de façon linguistiquement satisfaisante, on se rendrait alors compte que la formalisation, si elle a incontestablement un grand intérêt scientifique, a aussi ses limites, pratiques et théoriques.

Limitation pratique d'abord car, en linguistique comme en logique,

> « Le progrès vers la rigueur se paie d'un alourdissement presque insupportable de l'exposé. » [23]

Ou encore, selon l'opinion de deux spécialistes soviétiques :

> « Les grammaires formelles ne contribuent pas à élever le degré de clarté des descriptions linguistiques et ne conviennent pas à une utilisation directe par l'homme. » [24]

Ceci a des conséquences, notamment pédagogiques. L'enseignement du français, par exemple, et la recherche en linguistique ne sont pas sans rapports mais ne sont pas non plus la même chose. De même que partir des fondements selon une démarche scientifique rigoureuse et unificatrice n'est peut-être pas la meilleure façon d'enseigner les mathématiques aux enfants, de même l'utilisation directe de grammaires « scientifiques », toujours très complexes, qu'elles soient ou non formalisées, n'est-elle certainement pas la meilleure façon d'améliorer le maniement de la langue.

4.3. Si l'on considère à présent les systèmes formels en eux-mêmes, on a pu mettre en évidence certaines limitations internes, inhérentes à ces systèmes et irréductibles. Signaler rapidement quelques-unes de ces

limitations, de façon tout à fait intuitive, aura sans doute le mérite de troubler certaines conceptions encore répandues et trop idéalistes de la « rigueur scientifique » [25]. Ainsi, dans un système formel assez puissant, tels que ceux utilisés pour formaliser les mathématiques ou la logique, il existe des *propositions indécidables,* c'est-à-dire qu'on ne peut, dans le système formel, ni les dériver ni dériver leur négation. Il est même impossible, dans la plupart des systèmes formels, d'établir une procédure générale permettant de décider si une proposition donnée est ou non dérivable dans le système formel considéré. Enfin, la plupart des systèmes formels admettent des modèles, des interprétations qui ne sont pas isomorphes. Pour en revenir aux grammaires formelles, il existe un certain nombre de propriétés indécidables, c'est-à-dire qu'on ne peut établir au terme d'une procédure générale de calcul valable pour toutes les grammaires formelles d'un type donné. Ces propriétés indécidables sont d'autant plus nombreuses que la grammaire est plus compliquée, donc plus satisfaisante linguistiquement.

4.4. Ces critiques et ces limitations n'autorisent absolument pas une quelconque condamnation de l'attitude formalisante en linguistique. Qu'il s'agisse de la formalisation de la théorie ou des descriptions fournies par la théorie il y a, au contraire, beaucoup à attendre de ces recherches. Deux théories formalisées deviennent plus aisément comparables, il devient possible d'établir éventuellement leur équivalence, ou, si elle ne sont pas équivalentes, de choisir entre elles en s'efforçant, par exemple, d'introduire une mesure de leur simplicité relative, etc. La formalisation des descriptions linguistiques permettrait également un progrès de la typologie linguistique et, pourquoi pas, la mise en évidence progressive, par l'accumulation d'études, de quasi-universaux de fait. C'est un peu la perspective de Chomsky pour qui la formalisation au niveau « profond » a une importance capitale en ce qu'elle seule, en vidant les langues diverses de ce qui fait précisément leur diversité, permettrait de découvrir la structure des mécanisme innés et universels postulés.

Mais qu'une théorie soit ou non formalisée n'implique pas qu'on puisse esquiver les questions suivantes, si l'on désire en évaluer la validité :

— la théorie proposée vaut-elle effectivement pour tous les objets qu'elle définit comme siens (systèmes sémiologiques, langues naturelles, etc.) ?

— les hypothèses sur lesquelles repose la théorie sont-elles explicitées ? Sinon, quelles sont-elles ? Quelles sont les données empiriques antérieures à la théorisation qui y prennent statut d'hypothèses ?

— la théorie étant donnée, deux linguistes travaillant indépendamment sur les mêmes données parviennent-ils à des résultats identiques ?

— Peut-on envisager une vérification expérimentale (statistique, psycholinguistique, etc.) de certaines hypothèses ou thèses de la théorie, ou encore de la valeur de son application ? (Cf. les recherches — non concluantes — pour prouver la réalité psycholinguistique de l'opération de transformation [26].)

Bien que le domaine très vaste des recherches formelles sur les langages et les grammaires ait été volontairement limité ici aux grammaires génératives non transformationnelles, on espère n'avoir pas trop faussé l'esprit de ces recherches lorsqu'elles s'appliquent aux langues naturelles [27]. On voudrait surtout avoir montré :

— que parler de « théorie linguistique formalisée » est ambigu. Jusqu'à présent ce ne sont pas des théories linguistiques au sens courant du terme « théorie » qui ont été formalisées mais des descriptions de langues selon ces théories ;

— que la formalisation n'est ni le seul, ni sans doute le principal aspect qu'il faille considérer dans l'évaluation d'une théorie linguistique.

Remi JOLIVET.

NOTES

[1] Robert Blanché, *L'axiomatique*, collection SUP « Initiation philosophique », Presses Universitaires de France, Paris, 5e éd. 1970, pp. 55-56.

[2] Stephen C. Kleene, *Logique mathématique,* collection U, Armand Colin, Paris, 1971, p. 199.

[3] Leonard Bloomfield, « A Set of Poslulates for the Science of Language », *Language,* 2, 1926, pp. 153-164, repris dans : Martin Joos, *Readings in Linguistics I, The Development of Descriptive Linguistics in America, 1925-56,* The University of Chicago Press, Chicago & Londres, 1966, pp. 26-31; traduit dans André Jacob, *Genèse de la pensée linguistique,* Armand Colin, Paris, 1973, pp. 184-196.

[4] Leonard Bloomfield, *Le langage,* Payot, Paris, 1970, pp. 77, 137.

[5] Louis Hjelmslev, *Prolégomènes à une théorie du langage,* traduction Una Canger, Minuit, Paris, 1971, pp. 44, 51, 161-169.

[6] Jan W. F. Mulder, *Sets and Relations in Phonology. An Axiomatic Approach to the Description of Speech,* Clarendon Press, Oxford, 1968 ; Jan W. F. Mulder, Sándor G. J. Hervey, « Postulates for Axiomatic Functionalism », *Actes du Premier Colloque de Linguistique Fonctionnelle* (Gröningue, 1974 ; à paraître); Sándor G. J. Hervey, « Axiomatic-Deductive Linguistics », *Actes du Premier Colloque de Linguistique Fonctionnelle* (ed. cit.); Jan W. F. Mulder, « Syntax in Axiomatic Functionalist Linguistics », *Actes du Deuxième Colloque de Linguistique Fonctionnelle* (Clermont-Ferrand, 1975 ; à paraître); Sándor G. J. Hervey, « Semantics in Axiomatic Functionalist Linguistics », *Actes du Deuxième Colloque de Linguistique Fonctionnelle* (ed. cit.); Gwendoline Soutar, « Written 'Language' as a Semiotic System : an Axiomatic Functionalism Point of View », *Actes du Deuxième Colloque de Linguistique Fonctionnelle* (ed. cit.).

[7] Robert Blanché, op. cit., p. 59.

[8] Sur tout ceci, cf. Ferdinand Gonseth, *Les mathématiques et la réalité ; Essai sur la méthode axiomatique,* Librairie Albert Blanchard, Paris, 2e éd. 1974, notamment pp. 75-93, 226-243.

[9] Roger Martin, *Logique contemporaine et formalisation,* Presses Universitaires de France, Paris, 1964, pp. 22, 55 ; voir aussi la discussion sur ce qu'est *un symbole* aux pages 22 à 32.

[10] Claude E. Shannon & Warren Weaver, *The Mathematical Theory of Communication,* University of Illinois Press, Urbana, Chicago & Londres, 1972 (traduction française parue : Retz-CEPL, Paris, 1974).

[11] Charles F. Hockett, *A Course in Modern Linguistics,* Macmillan, New York, 1958.

[12] Rulon S. Wells, « Constituants immédiats », *Langages,* 20, 1970 (1° éd., en anglais, 1947), pp. 61-100.

[13] Noam Chomsky, *Structures syntaxiques,* Seuil, Paris, 1969 ; Noam Chomsky, « Trois modèles de description linguistique », *Langages,* 9, 1968, pp. 51-76 ; Noam Chomsky & George A. Miller, *L'analyse formelle des langues naturelles,* Gauthier-Villars, Paris ; Mouton, Paris & La Haye, 1968.

[14] Paul M. Postal, « Limitations of Phrase Structure Grammars », *The Structure of Language,* Prentice-Hall, Englewood Cliffs, New Jersey, 1964, pp. 137-151.

[15] Jerrold J. Katz & Jerry A. Fodor, « Structure d'une théorie sémantique avec applications au français », *Cahiers de Lexicologie,* 1966-II & 1967-I, pp. 39-72, 47-66.

[16] Noam *Chomsky,* « La notion de règle de grammaire », *Langages,* 4, 1966, pp. 81-104 ; Noam Chomsky, *Aspects de la théorie syntaxique,* Seuil, Paris, 1971, pp. 183-188.

[17] Noam Chomsky, *Aspects de la théorie syntaxique,* ed. cit., p. 188, n. 11.

[18] « Bien qu'il eût vengé d'anciens outrages des Helvètes au peuple romain », César, *De bello Gallico,* 1, 30, 1, cité par O. Riemann, *Syntaxe latine,* Klincksieck, Paris, 1890, p. 93. Les exemples français paraissent peu « naturels » : *la critique de Chomsky de cet auteur.*

[19] Zellig S. Harris, « La structure distributionnelle », *Langages,* 20, 1970, p. 30.

[20] Rulon S. Wells, op. cit., pp. 86-88.

[21] Pour une idée de la démonstration, cf. Nicolas Ruwet, *Introduction à la grammaire générative,* Plon, Paris, 1968, pp. 154-157.

[22] Nicolas Ruwet, op. cit., pp. 55-56.

[23] Roger Martin, op. cit., p. 20.

[24] A. V. Gladkij & I. A. Mel'cuk, *Eléments de linguistique mathématique,* Dunod, Paris, 1972, p. 17.

[25] Jean Ladrière, « Les limitations internes des formalismes et leur signification philosophique », *Dialectica,* 14, 1960, pp. 299-328 ; Jean Ladrière, « Limites de la formalisation », in Jean Piaget, *Logique et connaissance scientifique,* Encyclopédie de la Pléiade, Gallimard, Paris, 1967, pp. 312-333.

[26] « Psycholinguistique et grammaire générative », *Langages,* 16, 1969 ; Jacques Mehler & Georges Noizet, *Textes pour une psycholinguistique,* Mouton, Paris & La Haye, 1974, pp. 487-563.

[27] Pour des développements plus importants sur les grammaires formelles, on consultera, outre les ouvrages de Noam Chomsky & George A. Miller cité à la note 13 et de A. V. Gladkij & I. A. Mel'cuk cité à la note 24, les ouvrages ou articles suivants, classés approximativement par difficulté croissante :

Jean-Pierre Desclés, « Linguistique et formalisation. La logique et le nombre », in Bernard Pottier, *Langage,* Dictionnaire du Savoir moderne, CEPL, Paris, 1973, repris dans *Comprendre la linguistique,* Marabout-Université, Verviers, 1975, pp. 143-173.

Gilles Fauconnier, « La grammaire générative. La description d'un mécanisme », in Bernard Pottier, op. cit., pp. 213-260.

Georges Van Hout, *Franc-Math, Essai pédagogique sur les structures grammaticales du français moderne, Introduction mathématique,* Didier, Paris, 1973, pp. 298-337.

Michel Hughes, *Initiation mathématique aux grammaires formelles,* Larousse, Paris, 1972.

A. V. Gladkij, *Leçons de linguistique mathématique,* Documents de linguistique quantitative nº 5 & 6, Dunod, Paris, 1970.

Maurice Gross & André Lentin, *Notions sur les grammaires formelles,* Gauthier-Villars, Paris, 1970.

On consultera aussi les deux numéros (34 et 35) de la revue *Mathématiques et sciences humaines,* Gauthier-Villars, 1971, consacrés à la formalisation en linguistique.