



**HAL**  
open science

## Conditions d'adéquation des théories linguistiques

Gabriel G. Bès

► **To cite this version:**

Gabriel G. Bès. Conditions d'adéquation des théories linguistiques. Cahiers du Centre Interdisciplinaire des Sciences du Langage - Université Toulouse Le Mirail, 1980, 2, pp.85-95. hal-01064409

**HAL Id: hal-01064409**

**<https://hal.science/hal-01064409>**

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.

Conditions d'adéquation des théories linguistiques

par Gabriel G. Bès\*

Un observateur externe, sans être naturellement trop pessimiste devrait caractériser la linguistique descriptive d'aujourd'hui par les traits suivants :

a) Pluralité de résultats en concurrence, dans le double sens d'existence d'une pluralité d'écoles et de sous-écoles qui proposent des descriptions linguistiques différentes et de l'existence d'une pluralité de résultats descriptifs possibles pour une même question à l'intérieur d'une même école (absence d'unicité dans les résultats descriptifs), sans qu'il existe la possibilité de choisir non arbitrairement parmi eux.

b) Absence de communication entre écoles différentes.

c) Absence d'un processus d'interaction entre élaborations théoriques dans le cadre de la linguistique descriptive et traitement de problèmes linguistiques concrets dans le très vaste domaine des langues humaines.

d) Absence d'un processus d'interaction entre élaboration théoriques dans le cadre de la linguistique descriptive et élaboration de modèles psycholinguistiques d'utilisation et d'acquisition du langage.

\* Professeur à l'Université de Clermont II

Certes, il existe une linguistique appliquée et une psycholinguistique ; (c) et (d) voudraient souligner, cependant, en durcissant quelque peu la situation, l'absence d'un processus d'interaction profonde de ces deux "branches" avec la linguistique descriptive : on n'a pas l'habitude de remettre explicitement en cause les hypothèses descriptives pour les adapter à la solution des problèmes appliqués ou pour les intégrer à des modèles psycholinguistiques, tout en souhaitant de conserver pour elles les caractéristiques de généralité et de formalisation d'une approche théorique.

Les quatre traits précédents ( (a) à (d) ) ne sont pas sans rapport, ce sont, plutôt, des manifestations d'un même constat : la linguistique descriptive depuis Saussure jusqu'à et y compris Chomsky s'est, sur le plan méthodologique, profondément enrichie, mais elle s'est désempirisée. Enrichie, parce qu'elle a gagné sur le plan de l'explicitation, de la rigueur et de la cohérence dans la formulation de ses outils conceptuels. Mais désempirisée, parce qu'il devient extraordinairement difficile, sans introduire un trop grand quota d'arbitraire, de tester par rapport à des données externes suffisamment fiables la vérité ou la fausseté des propositions descriptives.

L'objectif de cette communication ici résumée est double : d'une part, illustrer l'observation précédente dans le cadre de la grammaire générative ; d'autre part, esquisser un cadre conceptuel différent de celui couramment pratiqué pour aborder le langage. C'est dans ce cadre-ci qu'opère le C.I.L.N. (Groupe de recherches sur la condensation de l'information en langue naturelle) de l'Université de Clermont II, certains résultats de ses recherches sont présentés dans la communication "Description phonologique et codification économique du langage".

En grammaire générative, on affirme que la linguistique est une

science empirique : ses hypothèses sont vraies ou fausses par rapport à des données externes. Le point de vue ici avancé est que le cadre conceptuel effectivement utilisé en grammaire générative pour aborder la question de la pluralité des résultats contredit, sinon entièrement, au moins largement, cette affirmation.

Les notions de simplicité et de généralisation sont les deux outils principaux dont la grammaire générative dispose pour choisir entre hypothèses descriptives (grammaire universelle et grammaires particulières) en concurrence et qui portent sur un même domaine. On veut souligner ici leur faiblesse et leur incapacité à donner une solution à la situation de désempirisation des études sur le langage observée au départ.

Une pluralité de simplicités ont été utilisées en grammaire générative. Elles se classent en deux grandes rubriques : simplicités formelles, non susceptibles d'être vraies ou fausses en relation à des données externes, et simplicités empiriques, susceptibles d'être testées.

Il y a plusieurs types de simplicité formelle. En effet, une mesure de simplicité formelle porte sur la comparaison de grammaires universelles, une autre sur des grammaires particulières d'une même langue naturelle déterminées par une même grammaire universelle ; une troisième a été utilisée pour choisir entre grammaires particulières d'une même langue naturelle mais déterminées par des grammaires universelles différentes. Elles présentent les désavantages de toute simplicité formelle, désavantages discutés au moins depuis l'utilisation d'une simplicité formelle en glossématique : elle désempirise la linguistique en soumettant les résultats descriptifs à des "règles de jeu" par rapport auxquelles il est impossible de se mettre d'accord sans arbitraire.

Plusieurs programmes de simplicités empiriques ont été envisagés en grammaire générative. Celle qui, effectivement et explicitement proposée, a mérité d'être évoquée à plusieurs reprises et est susceptible d'être testée, porte sur la comparaison du nombre de traits distinctifs binaires utilisés pour spécifier les règles de redondance lexicale et les matrices des formants. Elle est testable en relation avec la distinction entre absences systématiques et absences accidentelles parmi les séquences phoniques non attestées dans une langue : la grammaire particulière choisie par la mesure de simplicité devrait pouvoir rendre compte de cette distinction. Cette mesure de simplicité, dès qu'elle est utilisée pour traiter des situations qui échappent au caractère très artificiel de quelques exemples isolés, échoue dans des conditions telles qu'on doit considérer qu'elle est, certes, empirique, mais qu'elle est fausse.

La mesure de simplicité a été aussi utilisée en rapport avec la notion de généralisation, notion qui sera examinée ci-dessous. Dans ce cadre particulier, on emploie le raisonnement suivant :

On se donne deux langages différents (ensemble d'énoncés linguistiques et de jugements métalinguistiques portant sur eux), soit Lang-1 et Lang-2. On pense que la grammaire particulière n° 1 (soit GP-1) qui rend compte du Lang-1 possède une généralisation, ce qui n'est pas le cas de GP-2, qui rend compte du Lang-2 ; GP-1 et GP-2 sont déterminées par la même grammaire universelle. On pose une certaine mesure de simplicité, fondée sur le décompte de symboles utilisés dans les grammaires particulières. Si on constate que GP-1 est plus simple que GP-2, la grammaire universelle est validée : elle est susceptible d'exprimer une généralisation ; dans le cas contraire, c'est-à-dire si, en fonction de la mesure de simplicité choisie, GP-1 n'est pas plus simple que GP-2, on conclut que la GU n'est pas susceptible d'exprimer une généralisation et elle doit être modifiée. Ce cadre conceptuel est sous-jacent à la "théorie de la marque".

Cette utilisation de la simplicité semble particulièrement dangereuse car elle risque de gêner fortement le caractère empirique de toutes les hypothèses descriptives : en effet, le caractère adéquat de celles-ci est partiellement conditionné par le fait d'exprimer avec moins de symboles (d'un certain type) les données linguistiques qui sont censées présenter certaines caractéristiques particulières. C'est comme si, dans la définition même des objectifs de la linguistique, on incorporait une "règle de jeu", la mesure de simplicité, dont le caractère naturel n'a jamais été prouvé et qui, dans le schéma conceptuel choisi, reste non testable. Elle présente, par ailleurs, un autre inconvénient grave : en fonction de quels critères peut-on décider qu'une grammaire particulière doit posséder une généralisation? Ceci nous amène à aborder cette notion, second des outils conceptuels utilisés en grammaire générative pour aborder la question de la pluralité de résultats en concurrence sur un même domaine.

En termes généraux, on a une généralisation lorsque plusieurs propositions particulières peuvent être <sup>par</sup> remplacées ou déduites d'une seule proposition plus générale. On voudrait que les propositions particulières portent sur des données qui soient "d'un même type". On aurait par conséquent le schéma suivant : si les données linguistiques  $d_1, d_2, \dots, d_n$  appartiennent à une même catégorie, la GP correspondante doit les spécifier par une même règle ou ensemble de règles. Si tel est le cas, la GP et la GU qui la détermine expriment une généralisation et sont adéquates.

Un des problèmes centraux de ce schéma de raisonnement est d'accepter, en tant que donnée, le classement de plusieurs énoncés linguistiques différents comme appartenant, à certain niveau d'abstraction, à une même catégorie. Une série d'auteurs, en grammaire générative, l'ont fait. La variabilité des résultats et les disputes qui ont vu le jour déconseillent fortement ce choix méthodologique. Le schéma interne à la discussion de Lakoff ("Pronominalization, Negation, and the Analysis of Adverbs", dans R.A.Jacobs et P.S.

Rosenbaum Readings in English Transformational Grammar, Waltham, Mass.,..., Ginn and Co., 1970, p. 145-165) avec Chomsky ("Deep Structure, Surface Structure and Semantic Interpretation" dans Studies on Semantics in Generative Grammar. The Hague, ..., Mouton, 1972, p. 90-91) à propos des énoncés qui suivent, témoigne de l'énorme difficulté de faire porter les jugements métalinguistiques sur leur classement:

- (1) John said that Bill finked on Max, but I didn't believe it.
- (2) Goldwater won in the West, but it didn't happen in the East.
- (3) Goldwater won in 1964, but it won't happen in 1968.
- (4) Irving refused the peanut butter sandwich, but it wouldn't have happened with a bagel.
- (5) Having the police after him bothered John, but it would never bother Mary.
- (6) Noon found Harry making love to Zelda.
- (7) Noon found Harry making love to Zelda at 12 o'clock.\*

L'argumentation de Lakoff, dans sa première formulation, peut se résumer ainsi : (1), (2) et (3) appartiennent à une même classe; la description qu'on en donne doit exprimer ce fait. Ce à quoi Chomsky répond : selon les critères employés, des énoncés du type du (4) appartiennent aussi à la même classe et il est impossible de les représenter selon l'organisation proposée pour (1), (2) et (3). Réplique de Lakoff : (4) et (5) appartiennent à une même classe, à laquelle (3) et (6) n'appartiennent pas, ces énoncés-ci présentent un comportement particulier par rapport au complément circonstanciel de temps : voir l'agrammaticalité de (7).

Il paraît clair que le centre méthodologique de cette discussion est la constitution de classes d'énoncés qui sont différents mais qui, à un certain niveau d'abstraction (pronominalisation de "it", négation, compléments

circonstanciels) présenteraient des éléments communs. Ceci dépend, d'une part, des critères qu'on se donne sur l'identité entre objets comparés et, d'autre part, de l'univers d'énoncés qu'on considère. Il n'est pas sans intérêt de remarquer la rhétorique "alluviale" de ce type de discussion: Lakoff présente d'abord les énoncés (1), (2) et (3), Chomsky ajoute des énoncés du type (4) par la suite, Lakoff rajoute encore (5), (6) et (7). Si on poursuivait la discussion, en ajoutant :

(8) Noon found Harry making love to Zelda in the garden.

la question de la constitution des classes d'énoncés deviendrait encore plus confuse. En effet, (8), à la différence de (7), n'est pas agrammatical, le circonstanciel de lieu ayant un comportement différent du circonstanciel de temps. Or, ceci devrait amener, pour les mêmes raisons qui ont conduit à séparer (4) et (5) de (3), à distinguer (2) et (3) : on détruit ainsi l'homogénéité de la classe (1), (2), (3) et, par là, on ruine le point de départ de la discussion, c'est-à-dire le fait d'admettre comme acquise l'homogénéité de la classe formée par (1), (2) et (3).

Un autre choix méthodologique peut se faire par rapport à la notion de généralisation. L'explicitation de cette notion peut se fonder sur celle de projection. En termes simples, on a une projection lorsque les règles d'une GP sont exigées par un sous-langage (ensemble justificatif) d'un langage A et cette même GP rend non seulement compte de ce sous-langage mais, aussi, d'un autre sous-langage (ensemble projectif) du même langage A. Dans ce cas, la GP proposée pour rendre compte du sous-langage justificatif peut projeter du sous-langage justificatif au sous-langage projectif.

Cette caractérisation de la notion de généralisation ne présente pas les inconvénients de la caractérisation précédente : on vérifie la GP proposée à l'aide des jugements de grammaticalité, d'ambiguïté, etc., mais il n'est

pas nécessaire d'avoir recours à des jugements portant sur l'homogénéité éventuelle des énoncés considérés. La caractérisation évoquée a été utilisée dans des discussions en grammaire générative. On comprend qu'une GU-A (grammaire universelle A) et les GP qu'elle détermine sont considérées comme susceptibles d'exprimer une généralisation si les GP sont susceptibles de projeter d'un sous-langage justificatif à un sous-langage projectif ; cette GU-A est, par conséquent, adéquate. Comme cette argumentation est utilisée dans un cadre de comparaison de GU différentes, la conclusion se complète ainsi : ... alors que GU-B et les GP qu'elle détermine n'expriment pas, elles, une généralisation (c'est-à-dire qu'elles ne sont pas susceptibles de projeter des sous-langages justificatifs aux sous-langages projectifs correspondants) et, par conséquent, ne sont pas adéquates.

Si ce schéma ne présente pas les inconvénients du précédent, il en présente d'autres. En effet, les utilisations effectives qui en ont été faites sont souvent peu convaincantes parce que les ensembles justificatifs et projectifs considérés sont très réduits : il suffit de modifier l'ensemble de référence -c'est-à-dire, le langage constitué par le sous-langage justificatif et le sous-langage projectif- pour que les GP déterminées par les GU comparées organisent les projections de manière différente ; il n'est donc pas possible de conclure qu'une GU-A et les GP qu'elle détermine peuvent exprimer une généralisation alors que GU-B et les GP correspondantes ne peuvent pas le faire ; on doit constater que chacun des systèmes descriptifs comparés exprime des généralisations différentes. Le problème de la non-unicité des résultats reste donc entier.

La notion de projection semble particulièrement adéquate pour tester les systèmes descriptifs : elle peut devenir, à un niveau intéressant de formalisation, une réponse à la question de l'acquisition du langage. Elle est

donc susceptible d'être interprétée psychologiquement. Par conséquent, il est légitime de se demander si on peut corriger ses défauts actuels.

On pourrait penser que, pour ce faire, il suffirait d'augmenter les exigences méthodologiques : on peut, par exemple, exiger que l'ensemble de référence soit beaucoup plus vaste que ceux considérés aujourd'hui, que les descriptions respectives soient complètement explicitées, que les jugements métalinguistiques soient beaucoup plus fiables que ceux considérés aujourd'hui. Toutes ces précautions seraient sans doute salutaires ; elles ne peuvent pas, cependant, résoudre le problème de fond.

La notion de projection est psychologiquement interprétable dans le cadre de l'acquisition du langage. Malheureusement, Chomsky a posé le problème dans un cadre heuristico-méthodologique tel qu'il est impossible d'avoir accès à des données du processus d'acquisition pour tester les modèles de description linguistique. On sait, en effet, qu'il se donne une situation volontairement idéalisée où on ne considère que deux points : dans l'un, l'enfant n'a pas constitué sa compétence et dans l'autre, il l'a complètement intériorisée. Il est donc par définition impossible, dans ce cadre, d'avoir accès à l'observation du processus d'acquisition pour fixer des contraintes aux sous-langages justificatif et projectif qui seraient susceptibles d'être attestés au cours de l'acquisition. Même en multipliant les exigences méthodologiques, la question du type de projection qu'on se propose d'obtenir reste entière et, par là, l'existence d'une pluralité de résultats descriptifs, parmi lesquels il est impossible de choisir sans arbitraire.

Pour pallier les défauts actuels de la notion de généralisation, il semble bien qu'il soit nécessaire de considérer qu'elle doit être attestée, en droit, dans le domaine de l'acquisition du langage. Ceci implique que les

hypothèses descriptives de la linguistique n'ont aucune priorité logique par rapport à celles qui portent sur les autres éléments d'un modèle plus complexe : dans cette vision du problème, le linguistique est un élément en interaction avec d'autres. On peut le distinguer des autres et, si on le souhaite pour des raisons heuristiques, l'étudier, jusqu'à un certain point, en l'isolant, mais, en dernière instance, la validation d'une hypothèse descriptive sera médiate, par le biais de la validation d'un modèle plus complexe, où les éléments linguistiques sont traités en interaction avec des éléments d'un autre type. C'est dans un tel cadre que les généralisations seraient susceptibles d'être testées, en fonction de la question de fond suivante : est-ce qu'elles sont susceptibles de s'intégrer à des modèles explicatifs portant sur l'acquisition du langage, modèles qui devraient être testés par rapport à des données dont l'idéalisation devrait être la plus réduite possible et, en tout cas, sans que cette idéalisation rende banale la vérification des hypothèses proposées ?

La discussion précédente a conduit à proposer l'empirisation des hypothèses linguistiques par le biais de leur intégration à des modèles sur l'acquisition. Mutatis mutandis, les mêmes propositions peuvent être faites dans d'autres domaines, en particulier, dans celui de l'intégration des hypothèses descriptives à des modèles d'utilisation du langage (de production et d'interprétation). Sur un plan plus général, le cadre conceptuel envisagé pour tester les hypothèses descriptives consiste à ne jamais traiter le système linguistique sous-jacent comme une pièce isolable en droit, qui aurait un champ d'observations propres par rapport auquel il serait testable. Le système linguistique, dans cette optique, doit être traité comme un élément complexe en interaction avec d'autres, la description de cette interaction peut être testée, mais pas le système considéré isolément. La linguistique ne devrait plus avoir "pour unique et véritable objet la langue envisagée en elle-même et pour elle-même", comme le voulait Saussure, avec juste raison à son époque.

Sur le plan heuristique, peu importe, à la limite, le domaine choisi au départ, et, en particulier, que ce domaine soit classé sous la rubrique de psycholinguistique ou de linguistique appliquée. Dans tous les cas, il s'agit, d'une part, de se donner un domaine d'exigences externes (acquisition du langage, utilisation -production et/ou interprétation- du langage, enseignement de la langue maternelle, enseignement d'une langue seconde, utilisation des langues naturelles en documentation, etc.), dans lequel il sera possible d'observer les données de manière fiable et, d'autre part, de proposer des hypothèses descriptives sur le système linguistique en interaction avec d'autres types d'hypothèses, cet ensemble d'hypothèses devant être contrôlé par rapport au champ d'observation choisi. Si le champ d'observation relève de la linguistique appliquée, le contrôle final des hypothèses proposées sera obtenu par la possibilité d'améliorer une praxis particulière dans le domaine du langage.

Voilà, brièvement résumée, l'analyse de la situation des études sur le langage qui a conduit à l'organisation du C.I.L.N. La pluralité d'écoles et de résultats descriptifs en linguistique ne serait, selon cette analyse, que la conséquence d'un processus de désempirisation de la linguistique, processus qui ne peut pas être renversé dans le cadre conceptuel couramment pratiqué. Il faut donc insérer d'emblée les hypothèses linguistiques dans un système plus vaste, au risque probable de dégager, au départ, une pluralité d'hypothèses qui vont différer selon le champ d'adéquation externe choisi ; cette diversité devrait se réduire progressivement avec l'approfondissement des questions abordées dans chaque domaine.

Le champ choisi par le C.I.L.N. est celui de la condensation de l'information en langue naturelle : il s'agit, à différents niveaux de formalisation, d'étudier et de traiter cette caractéristique si étonnante des textes

en langue naturelle. d'associer un certain quantum d'information à un signifiant qui est notoirement compressible, sans pourtant que quelque chose qu'on peut considérer comme l'information "essentielle" -"essentielle" selon les propos visés par chaque type de texte- en pâtisse trop. On espère, ce faisant, dégager des traits généraux portant sur l'organisation et le conditionnement mutuel de différents types de structures dans un texte en langue naturelle et contribuer à améliorer dans le domaine du langage plusieurs pratiques centrées sur la condensation de l'information.

#### Note

Le texte précédent est un résumé de la communication présentée aux Journées de Linguistique, organisées par le Centre Interdisciplinaire des Sciences du Langage de l'Université de Toulouse - Le Mirail (mars-avril 1979); le matériel complémentaire peut être trouvé dans G.G. Bès "Les conduites simulées ; fonctions de communication et de représentation", dans J. Piaget, J.P. Bronckart et P. Mounoud (Dir.) La Psychologie. Paris, Gallimard, La Pléiade, sous presse, et dans G. G. Bès "Simplicités, empirisme et interprétation psychologique en grammaire générative". Les travaux du C.I.L.N. sont publiés dans Condenser ; on trouvera notamment, dans son premier numéro (février 1980), M. Bellot et alii, "La problématique de la contraction de texte" (p. 13-44), G.G. Bès et P.M. Fauchère, "Le système documentaire Vercingétorix I" (p. 57-94) ; cf. ici même le résumé de la communication "Description phonologique et codification économique du langage".