

# Pourquoi y-a-t-il de la variation plutôt que rien ?

---

Bernard Laks

Institut Universitaire de France

Université Paris Ouest Nanterre La Défense, Modyco UMR 7114

Everyone knows that language is variable. Two individuals of the same generation and locality, speaking precisely the same dialect and moving in the same social circles, are never absolutely at one in their speech habits. Sapir (1921, 147)

## 1 Approches de la variation : Perspective historique

Dans l'histoire récente des sciences du langage, la publication de Weinreich, Labov et Herzog (1968) a marqué un tournant historique. Pour beaucoup de jeunes linguistes du moment, ce texte de référence a constitué l'acte de naissance de la sociolinguistique moderne et décidé de leur carrière intellectuelle. Certes, en 1968, le terme «sociolinguistique» a déjà quelques années d'âge<sup>1</sup>. Currie l'avait forgé dans les années cinquante pour rendre compte, entre autre, de la relation entre langues et castes en Inde (Currie 1952). Si au cours des années soixante, sociolinguistique et sociologie du langage se construisent progressivement autour de grandes conférences internationales qui marquent le champ<sup>2</sup>, c'est véritablement en 1968 que la perspective sociolinguistique s'impose en réunissant de façon originale la dialectologie structurale, spécialement urbaine, les travaux sur les contacts de langues et les mélanges d'une part, la linguistique historique de l'autre<sup>3</sup>.

Cette convergence n'a rien d'inattendue. Elle avait été annoncée un demi-siècle plutôt par Meillet qui, dans sa leçon inaugurale au Collège de France (Meillet 1921), plaidait que toutes les dimensions du changement linguistique (dans l'espace géographique, social, stylistique ou historique) ont une source unique, le caractère d'institution sociale de la langue. Il s'ensuit rappelait-il que tous les phénomènes de mélange, d'évolution, de morcellement, en un mot tout ce qui selon Whitney (1875) fait la vie du langage, ont toujours une source externe. C'est pourquoi les linguistes,

<sup>1</sup> Currie (1952) présente le terme sociolinguistique comme une création lexicale personnelle. Il en revendique fortement la paternité en 1981 (Currie 1981).

<sup>2</sup> Cf. Koerner (1991) pour une analyse historique (*caveat* la référence initiale à Currie y est mal datée)

<sup>3</sup> S'agissant des influences sur la sociolinguistique Cf. pour la dialectologie structurale Weinreich (1954), pour la dialectologie urbaine Labov (1966), pour les contacts de langues et les mélanges Weinreich (1951, 1953), pour la linguistique historique Lehmann (1962).

aussi longtemps qu'ils se placent uniquement du point de vue interne, peuvent certes décrire ces phénomènes mais uniquement pour ce qu'ils sont : ils échouent définitivement à les expliquer<sup>4</sup>. Pour atteindre une explication causale soulignait encore Meillet, il est nécessaire d'adopter un point de vue externe et social. Rapporter les faits linguistiques à leur cause, et ainsi les expliquer, impose alors de reconnaître le rôle moteur que joue la variation dans la phénoménologie linguistique. Cette variation est liée à la dimension intrinsèquement sociale du langage<sup>5</sup>. Telle fût la leçon de Meillet reprise un demi-siècle plus tard par Weinreich. Je me propose de montrer ici que cette leçon fût globalement ignorée par la linguistique du 20<sup>ème</sup> siècle mais qu'elle revient actuellement au premier plan avec une acuité renouvelée.

En 1968, dans leur manifeste pour une analyse empirique du changement linguistique, Weinreich et ses deux doctorants<sup>6</sup>, mettent donc la variation au cœur des phénomènes linguistiques. Pour ce faire ils conjointement plusieurs problématiques. Réfléchissant aux différents modèles du changement linguistique proposés depuis les néogrammairiens, ils discutent notamment les phénomènes d'évolution graduelle ou abrupte, le rôle comparé des lois phonétiques et de la dispersion lexicale, l'existence de résidus et d'exceptions à la mutation, etc.<sup>7</sup> Ceci les conduit à une analyse des modèles de la transmission intergénérationnelle comme source du changement. Les propositions de Paul (1909), comme celles de Chomsky et Halle (1968) sont critiquées à l'aune de la variabilité interne des grammaires et de l'inscription sociale de la langue dans les communautés linguistiques et les groupes de pairs où se forge l'identité linguistique. Comme souligné au troisième point de leur conclusion, si toute variation et toute hétérogénéité internes à une langue donnée ne conduisent pas nécessairement à un changement dans cette langue, tout processus de changement découle nécessairement d'une hétérogénéité et d'une variation internes, socialement repérées, évaluées et promues. Les langues sont ainsi vues comme des systèmes instables, ouverts, plastiques, déformables et poreux. Liées au contact, les notions de mélange, de mixte, d'interlangue, qui depuis Schuchardt (1909, 1922) au moins, posent la créolisation au principe d'évolution de toute langue, s'en trouvent réaffirmées.

Au total, c'est dans le caractère intrinsèquement social de la langue, dans l'intimité du lien entre langue et communauté linguistique socialement qualifiée que Weinreich, Labov et Herzog

<sup>4</sup> On appréciera à sa juste valeur la condamnation par Meillet de la grammaire interne fondée sur les catégories formelles de la seule logique. « L'ancienne grammaire générale est tombée dans un juste décri parce qu'elle n'était qu'une application maladroite de la logique formelle à la linguistique où les catégories logiques n'ont rien à faire » (Meillet 1921, 15)

<sup>5</sup> « Le seul élément variable auquel on puisse recourir pour rendre compte du changement linguistique est le changement social dont les variations du langage ne sont que les conséquences parfois immédiates et directes, et le plus souvent médiates et indirectes » Meillet (1921, 17)

<sup>6</sup> William Labov et Marvin Herzog ont tous les deux soutenu un PhD à Columbia sous la direction d'Uriel Weinreich en 1964.

<sup>7</sup> Pour une mise au point détaillée plus récente Cf. Labov (1981, 1994, 2001)

(1968) voient la source première et le moteur du changement linguistique. La communauté linguistique rappellent-ils, est une organisation sociale concrète. Elle est donc, *ex definitio*, profondément hétérogène, divisée, hiérarchisée, structurée par des dynamiques sociales antagoniques. La variation et l'hétérogénéité linguistiques d'une part, la variation et l'hétérogénéité sociales de l'autre, ne sont alors que les deux aspects du même réel social. C'est ainsi parce qu'il n'existe jamais de communauté homogène parfaitement stable qu'il n'existe jamais de langue homogène parfaitement invariante et stable.

Ce lien intime entre langue et structure sociale, revendiqué par tous les Maîtres de la linguistique moderne du 20<sup>ème</sup> siècle, est saisi dans cette analyse sous le prisme du changement, mais Weinreich, Labov et Herzog en tirent des conclusions linguistiques de portée très générale qui vont profondément marquer le paysage théorique en donnant naissance à la (socio)linguistique variationniste. Si elle veut avancer vers la compréhension des phénomènes linguistiques au-delà d'une simple description phénoménologique écrivent-ils, la linguistique doit rompre le lien entre structure et homogénéité, organisation systémique et invariance. Toute société, toute culture et toute organisation humaine laisse apercevoir des différentiations internes fortes, des hiérarchies structurées, des hétérogénéités plus ou moins conflictuelles, et c'est l'essence même du social que d'organiser ces différences dans des systèmes dynamiques en constante évolution<sup>8</sup>. L'hétérogénéité et la variabilité sociales ne sauraient alors être conçues comme des dimensions parasites, surajoutées ou anomalies. Elles constituent les dimensions même du social. Comme je le montrerai ici, il s'ensuit que la variation et l'hétérogénéité doivent être situées au cœur même des systèmes linguistiques dont elles constituent le principe organisateur et fonctionnel.

La linguistique variationniste place ainsi la variation au cœur même du modèle linguistique et rejette par là même les approches invariantes dans la catégorie des grammaires homogènes. Une grammaire homogène et invariante n'étant rien d'autre qu'un système grammatical normé par le social même qu'il prétend ignorer. Ainsi, si toute langue est hétérogène et variable, la grammaire qui prétend la décrire et la modéliser doit l'être aussi. Mais plus, il faut voir dans la systémicité de la variation et dans l'organisation de l'hétérogénéité la source même de ce qui fait structure et système dans les langues. C'est en effet l'existence de modes et de régimes différents d'interlocution qui conduit à les grammaticaliser et les systématiser comme nous le verrons ci-dessous. Le Maître de Genève ne disait pas autre chose lorsqu'il déclarait que la langue comme système est ce qui advient lorsque des formes diverses de parole sont simultanément aperçues par la même conscience

<sup>8</sup> Sans entrer ici dans un débat post hégélien sur histoire, ses dynamiques et ses rationalités, tel par exemple que Fukuyama (1992) a cru pouvoir le synthétiser, on contrastera cette conception dynamique du social avec l'approche statique, idéale et parfaitement stérilisée que Chomsky (1965, 12) croit pouvoir se donner à des fins purement opératoires en se dotant d'un locuteur-auditeur idéal (Cf. *infra*). Pour une critique de cette prétention à s'extraire de l'histoire, des sociétés, de leurs dynamiques et de l'*hubris* même du social, Cf. Bourdieu (1997).

collective. Consécutivement, la linguistique de la parole (variable) apparaît pour Saussure comme condition préliminaire à toute linguistique de la langue. (Saussure 1916, 2759 ; 2001, 83)<sup>9</sup>.

Cinquante ans plus tard, l'idéalisme cartésien objectera avec Chomsky (1966) que pour être cognitivement instanciées, les structures linguistiques et la grammaire elle-même, sont par leur nature formelle nécessairement homogènes, monotones et invariantes, et que ceci tient au caractère immanent de la pensée et de la logique qui est à son principe. Cette position, qui est celle des grammairiens de l'homogénéité et de la statique linguistique depuis bien avant Port Royal<sup>10</sup> scotomise toutes les dimensions sociales de la grammaire. Chomsky par exemple, qui instrumentalise la compétence linguistique des sujets parlants, tient pour nulle toute l'analyse des compétences humaines comme des compétences socialisées. Il est ainsi porté à tenir pour non signifiant ce que Bourdieu (1980) a nommé « Théorie de la Pratique ». Pour le sociologue au contraire, les compétences sont le produit de l'intériorisation d'une extériorité toujours socialement qualifiée. Elles sont mises en acte, dans des contextes sociaux précis, comme des processus d'extériorisation d'une intériorité elle-même socialisée. Cette théorie de l'habitus, dont l'habitus linguistique et les compétences communicatives pratiques qui lui sont associées ne constituent qu'une instance particulière, rend compte de la place centrale qu'occupent, au cœur des structures linguistiques, la variation et l'hétérogénéité socialement motivées et contrôlées. Elle entre en résonance avec les travaux les plus récents du courant dit de la « cognition culturelle » (Tomasello 1999, 2008b) et avec ses extensions en linguistique contemporaine<sup>11</sup>. J'y reviendrai de façon détaillée ci-dessous.

Cette approche pose que dans une communauté linguistique réelle, et donc socialement structurée, l'aptitude à communiquer repose nécessairement sur une aptitude cognitive à maîtriser l'hétérogénéité et la variation linguistiques, maîtrise qui opère, comme l'habitus, à propos et en situation. En fait, bien plus que les grammaires et les structures linguistiques elles-mêmes, c'est cette capacité culturelle et cognitive à maîtriser hétérogénéité qui est transmise entre les générations par un processus de reproduction culturelle et sociale (Maniglier 2008). Dans cette perspective, la question du changement inter générationnel posée par Weinreich, Herzog et Labov trouve une explication transparente.

Rapporter ainsi hétérogénéité et variation linguistique aux processus sociaux et cognitifs qui régissent « la vie des signes au sein de la vie sociale » Saussure (2001, 34), permet, parce qu'on les rapporte à leur fonction, d'en comprendre l'ampleur et les limites. En effet, précisément parce

<sup>9</sup> Sur ces questions Cf. Laks (2011b), Béguelin (1990), Bouquet (1997).

<sup>10</sup> Cf. Arnauld et Lancelot (1660) et *infra*.

<sup>11</sup> Comme on le verra, la cognition culturelle est étroitement liée aux grammaires de constructions, aux grammaires exemplaristes et plus généralement aux grammaires fondées sur l'usage. Cf. Tomasello (2008a), Goldberg (2006) ; Barlow et Kemmer (2000).

qu'elle est et qu'elle fait système, la variation linguistique n'est jamais erratique. Elle est au contraire toujours contrainte et organisée. L'hétérogénéité structurale est ainsi grammaticalisée par le jeu de deux dynamiques opposées : les nécessités de la compréhension intracommunautaire limitent l'hétérogénéité et la différentiation, tandis que l'existence d'une organisation sociale différentiée limite l'homogénéité linguistique et l'invariance structurale. C'est le jeu et l'équilibre, toujours précaire, de ces deux dynamiques opposées qui structurent la variation et motivent le changement linguistique.

Observer la variation dans sa systématичité et rendre compte de l'hétérogénéité comme étant structurée impose évidemment d'adopter une méthodologie adéquate. On sait en effet que décontextualisée, l'observation détruit la systématичité des phénomènes variables et les fait paraître comme étant alléatoires<sup>12</sup>. Observer les faits linguistiques hors de l'écosystème social qui les conditionne détruit en effet tout ce que la pratique doit précisément à son caractère pratique<sup>13</sup>. C'est la raison pour laquelle l'analyse de la variation systématique commence nécessairement par une réflexion critique sur les observables.

## 2 Datum et exemplum, deux approches des données

Les Sciences du Langage, comme toutes les sciences, doivent se doter d'un fondement empirique clair. C'est dire que la construction des observables linguistiques y joue nécessairement un rôle central. Pour cela, et contre ce qu'il nomme l'« 'opinion commune », il faut avec Bachelard (1938, 14-16) rappeler qu'il n'y a pas de donnée scientifique qui ne résulte d'un questionnement théorique et d'une construction épistémologique précise : c'est le point de vue qui crée l'objet<sup>14</sup> et, comme il le souligne encore, « les instruments ne sont que des théories matérialisées. Il en sort des phénomènes qui portent de toutes parts la marque théorique » (*ibidem*.) C'est à l'aune de cette épistémologie critique qu'il convient d'éprouver la notion même d'observable linguistique, de fait de langue, ou encore de donnée langagière. La notion ne va pas de soi, et sur la scène linguistique contemporaine, deux grands types d'approches peuvent à son propos être contrastées, selon qu'elles prennent en compte la variation ou la détruisent.

---

<sup>12</sup> Sur le paradoxe de l'observateur Cf. Labov (1975, 1976). Sur le marché linguistique et les effets d'observation Cf. Encrev  (1976, 1982)

<sup>13</sup> Cf. Bourdieu (1994), Bourdieu et alii (1993)

<sup>14</sup> « Et quoi qu'on dise, dans la vie scientifique, les problèmes ne se posent pas d'eux-m mes. C'est précis ment ce sens du probl me qui donne la marque du v ritable esprit scientifique. Pour un esprit scientifique, toute connaissance est une r ponse   une question. S'il n'y a pas eu de question, il ne peut y avoir connaissance scientifique. Rien ne va de soi. Rien n'est donn . Tout est construit» Bachelard (1938, 14).

La linguistique des usages (Barlow et Kemmer 2000), dont la linguistique variationniste fait partie<sup>15</sup>, regarde les usages linguistiques attestés comme les faits dont il faut rendre compte dans une systématique grammaticale qui intègre donc nécessairement l'hétérogénéité et la variation. Faisant écho à la règle durkheimienne, elle traite les faits linguistiques et sociaux comme des choses et rejette l'introspection et ses catégorisations préconçues<sup>16</sup> au profit d'une recollection des productions linguistiques attestées. Une telle méthode d'investigation rencontre nécessairement l'hétérogénéité et la variation qu'elle a vocation à considérer chacune dans sa systématичité et son caractère proprement structural.

A l'opposé, la conception grammairienne, illustrée par la Grammaire Générative sous ses différents avatars (Minimalisme, Comparatisme universaliste, Biolinguistique) construit sa systématique sur la base d'exemples ou de jugements de grammaticalité, en faisant abstraction de ce qui s'observe dans les communautés linguistiques réelles. Toute variation et toute hétérogénéité sont alors rapportées à des dimensions extérieures à la langue et expulsées de l'observable ainsi construit. Considérée comme une dimension de pure performance, toute systématичité et toute pertinence grammaticale est déniée à la variation. Dans l'abstraction revendiquée d'une conception cartésienne de la langue et de la cognition, « L'objet premier de la théorie linguistique [reste] un locuteur-auditeur idéal, appartenant à une communauté linguistique complètement homogène, qui connaît parfaitement sa langue et qui lorsqu'il applique en une performance effective sa connaissance de la langue, n'est pas affecté par des conditions grammaticalement non pertinentes ». Cette nécessité de l'abstraction que Chomsky (1965, 12) revendique comme condition même de la rationalité scientifique<sup>17</sup>, a pour pendant le déni de toute linguistique de corpus<sup>18</sup>. Pour lui, l'empirisme en matière linguistique n'aurait pas beaucoup plus d'intérêt que la recollection des lépidoptères n'en a pour la pour la science en général<sup>19</sup>.

---

<sup>15</sup> Avec Langacker (1987), Tomasello (2008c) propose le terme plus théorique de linguistique cognitive fonctionnelle pour désigner les approches basées sur l'observations des usages qui ne tiennent pas pour acquises les catégories grammaticales préconstruites de la tradition grammaticale occidentale, mais visent à reconstruire l'analyse linguistique cognitive à partir de la structure exemplariste, statistique et probabiliste, des usages. J'utilisera ici le terme plus usité de linguistique des usages.

<sup>16</sup> « Qu'est-ce en effet qu'une chose? La chose s'oppose à l'idée comme ce que l'on connaît du dehors à ce que l'on connaît du dedans. Est chose tout objet de connaissance qui n'est pas naturellement compénétrable à l'intelligence, tout ce dont nous ne pouvons nous faire une notion adéquate par un simple procédé d'analyse mentale, tout ce que l'esprit ne peut arriver à comprendre qu'à condition de sortir de lui-même, par voie d'observations et d'expérimentations [...]. Traiter des faits d'un certain ordre comme des choses, [...] c'est en aborder l'étude en prenant pour principe qu'on ignore absolument ce qu'ils sont, et que leurs propriétés caractéristiques, comme les causes inconnues dont elles dépendent, ne peuvent être découvertes par l'introspection la plus attentive. » Durkheim (1927, préface à la seconde édition XIII)

<sup>17</sup> « La notion de langue est en elle-même un haut niveau d'abstraction. Les linguistes ont toujours à juste titre, procédé à une idéalisatіon : donnons nous disent-ils, l'idée d'une communauté linguistique homogène. C'est le seul moyen de procéder rationnellement. (...). Vous devez abstraire un objet, vous devez éliminer les facteurs non pertinents. Du moins si vous voulez en faire une étude non triviale. » Chomsky (1977, 74-75)

<sup>18</sup> Aarts (2000, 5) interroge Chomsky : 'What is your view of modern corpus linguistics?' la réponse de Noam Chomsky est sans appel : 'It doesn't exist.'

<sup>19</sup> « La sociolinguistique qui est censée naître de la sociologie et de la linguistique, ne tirera rien de la sociologie [...]. Vous pouvez aussi collectionner des papillons et faire beaucoup d'observations. Si vous aimez les papillons c'est très bien ; mais cette activité ne doit pas être

Comme je l'ai montré ailleurs (Laks 2008), au-delà d'une opposition épistémologique classique entre idéalisme et empirisme, l'opposition entre linguistique variationniste et linguistique générative réactive le clivage, également classique, entre les sciences du *datum* et les sciences de *l'exemplum*. Depuis leur plus lointaine origine en effet, la linguistique et la philologie voient s'opposer ces deux perspectives qui conditionnent le rapport spécifique aux données dont elles traitent. Afin d'éclairer le débat contemporain sur la place de la variation et de l'hétérogénéité dans les grammaires, il importe donc d'y revenir un instant.

## 2.1 Les sciences de *l'exemplum*

L'approche prescriptive ou grammaticale réfléchit à partir de (proto)types normés saisis comme des *exempla*. Dans la *techné* grammaticale grecque, chez Denys de Trace, ou dans son extension chez Appolonius Dyscole à une syntaxe déjà systématique, tout comme dans *l'Ars grammatica* latine de Varron à Donat jusqu'à Priscien, les données langagières sont toujours manipulées comme des exemples. Qu'ils illustrent des constructions ou des formations lexicales, ces exemples sont repris de prosateurs ou de poètes considérés comme classiques. Ils s'organisent en listes et forment un corpus pratiquement stable, transmis et repris de grammairien en grammairien<sup>20</sup>. On sait le lien profond entre grammaire et pédagogie, en Grèce comme à Rome. L'exemple grammatical s'impose alors comme instrument naturel pour les exercices de mémorisation et comme support des raisonnements inductifs. L'art grammatical consiste à découvrir les principes sous-jacents à ces listes d'exemples. Les notions de paradigme et de régularité systématique en découlent très logiquement. Ainsi, le lien entre l'exemple et la règle qu'il permet de découvrir est, dès l'origine de la grammaire classique, particulièrement fort<sup>21</sup>.

Dans cette science de *l'exemplum*, le savoir grammatical est donc construit sur la reprise emblématique de citations provenant de ceux qui, déjà, se voient qualifiés de « meilleurs auteurs ». Mais l'enseignement de la grammaire et de la rhétorique est également une *propédeutique*. Il se trouve étroitement lié, et ce sera encore plus fortement le cas au Moyen-âge et à la période Classique, à l'enseignement de la logique et à l'apprentissage d'une pensée droite, de la même façon que l'orthographe consiste dans l'apprentissage d'une écriture (et d'une lecture) droites. Par droites,

---

confondue avec la recherche rationnelle. [...] La lutte contre l'idéalisation est la lutte contre la rationalité ; elle signifie n'ayons pas de travail intellectuel significatif. » Chomsky (1977, 74-75)

<sup>20</sup> Sur la naissance de la théorie grammaticale et les reprises successives des traités de différents auteurs Cf. Baratin, Desbordes, Hoffman et Pierrot (1981)

<sup>21</sup> « Les exemples, reconnus comme tels, permettent de faire l'économie de la règle qu'ils symbolisent » Holtz (1981, 109). Je dois à l'amitié de Madeleine Keller de nombreuses références sur les grammairiens anciens, ainsi que sur les travaux de recherche du groupe *Ars Grammatica* qu'elle anime. Cf. *Grammatica* (2005), également Keller (2009)

il faut entendre réglées c'est-à-dire conformes à des règles immanentes que le grammairien et le philosophe découvrent en contemplant leurs corpus d'exemples, règles qui constituent le fondement même de la pensée vraie<sup>22</sup>

Dans cette approche proprement grammaticale, les faits de langue ne sont donc pas consistants en eux-mêmes. Ils n'interviennent que comme de simples supports pour une réflexion abstraite hypothético-inductrice. Les corpus d'exemples permettent ainsi de construire par idéalisation et abstraction une *theoria* de la langue, θεωρία que la pensée grecque conçoit précisément comme une façon de voir, de contempler les choses pour, par le raisonnement inductif, s'approcher de leur essence même et participer ainsi de la Connaissance. L'exemple est donc spéculatif par définition même, il donne à voir parce qu'il appelle la contemplation d'où surgit par illumination la *theoria*. Véhicule sur le chemin de la connaissance pour le Maître, il est aussi chemin de vérité pour l'élève qu'il fait progresser vers la maîtrise d'une pensée droite par l'imitation<sup>23</sup>. C'est une école et, comme telle, sa vertu consiste à dépasser l'imitation première pour, par intégration mentale, aboutir à l'édification (*ædificatio*). L'exemple ne vaut donc pas seulement par sa valeur exemplaire pour l'imitation, mais également par sa fonction de moteur inférentiel<sup>24</sup>. Comme l'a bien vu Valéry (1941), l'exemple n'est jamais d'usage direct, il vaut pour le paradigme et conduit à la règle qu'il dénomme et permet ainsi de mémoriser : « *Quia nominor leo* ne veut pas dire "parce que je me nomme lion", mais "je suis une règle de grammaire" ».

## 2.2 La linguistique de *l'exemplum* aujourd'hui : la Grammaire Générative

La Grammaire Générative s'inscrit dans cette tradition. En effet, dès lors qu'elle récuse en droit les données variables de l'usage et qu'elle choisit de fonder son raisonnement sur des corpus d'exemples construits à l'aide du jugement de grammaticalité, elle est conduite à épouser un cadre épistémologique que Chomsky (1966) théorisera comme celui de la linguistique cartésienne. Dans le dégagement progressif de ce cadre théorique, la question des données linguistiques fût cruciale. Pour la linguistique des années 1950, le corpus structuraliste était explicitement compris comme un

<sup>22</sup> « Art de penser et non point art de bien penser, parce qu'un art a toujours pour tâche de donner des règles; que les règles définissent toujours une action correcte et qu'il n'y a pas plus d'art de mal penser qu'il n'y a de règles pour peindre mal. La pensée incorrecte est une pensée sans règle; et une règle qui ne serait « point bonne » ne saurait en aucune manière être considérée comme une véritable règle. » Foucault (1967, 7)

<sup>23</sup> On aura reconnu le triptyque *Contenplatio, imitatio, illuminatio* qui dans la mystique chrétienne, par exemple chez Saint Bonaventure, conduit à l'*ædificatio*. On pense également au rôle joué, au-delà de l'œuvre de Thomas à Kempis elle-même, par *L'imitation de Jésus-Christ*

<sup>24</sup> "Ce n'est plus l'imitation qui est requise de l'apprenant par production et confrontation des modèles, mais l'intellection. On lui demande de comprendre la rationalité du système de la grammaire, les "causes" comme disent les maîtres. Sanctius, Sanchez de Las Brozas : *De causis linguae latinae* : "Qui a jamais dit: *Ego amo deum et deus amatur a me?* Dira-t-on aussi : *facio orationem do tibi damnum* et bien d'autres tournures du même genre. C'est auprès des meilleurs auteurs qu'il faut apprendre le latin et non chez les grammairiens. La grammaire n'apprend pas à parler latin, mais elle renvoie la langue latine à un art de sorte qu'ensuite, par imitation du latin, on puisse parler" Chevalier (2007,155)

compendium d'usages attestés. C'est précisément par sa critique acérée que Chomsky rompt avec Harris. En récusant les données de performance et en rompant avec les linguistiques du *datum*, la Grammaire Générative se rattache dès lors explicitement aux linguistiques de *l'exemplum*. La critique du modèle syntagmatique (Chomsky 1957) et l'instrumentalisation du théorème Gold (1967) concourent à définir une approche linguistique et cognitive nouvelle (Chomsky 1965, 1968) pour laquelle les données dont il faut rendre compte ne sont plus constituées par des productions spontanées de locuteurs en situation de communication, mais par le produit d'un examen introspectif de constructions décontextualisées manipulées comme des *exempla* (Cf. *Supra* note 16 la critique de l'introspection par Durkheim) C'est donc très logiquement que Chomsky (1995) finira par poser comme objet de la Grammaire Générative, non le langage observable externe mais le langage interne de la pensée, retrouvant ainsi la voie ouverte par Port Royal<sup>25</sup>.

La Grammaire Générative se présente alors comme une théorie des états mentaux qui se propose de modéliser les fonctions cognitives propres à la capacité langagière de l'espèce humaine. Son objet n'est donc plus l'analyse des faits d'usage (performance) mais bien la (re)construction de la logique des processus mentaux impliqués dans l'organisation syntaxique des phrases (compétence). Comme l'a noté Milner (1989), pour caractériser ces processus linguistico-cognitifs, la Grammaire Générative tâche à circonscrire le possible, et surtout l'impossible, de la langue et des langues. Le jugement de grammaticalité est alors l'instrument de cette partition et il constitue le véritable observable dont la théorie linguistique doit rendre compte. Par le jugement de grammaticalité se construit un corpus d'exemples du fonctionnement formellement correct (ou incorrect) de la capacité cognitive particulière à une langue. La formalisation de cette grammaire particulière (GP), spécifiquement paramétrée, permet à son tour d'inférer et de formaliser les principes de la grammaire universelle (GU) qui caractérise la faculté de langage propre à l'espèce. C'est en ce sens que la Grammaire Générative est une théorie cartésienne de l'esprit dont l'objet n'est en rien le système de communication interindividuelle des sujets sociaux, mais le langage (universel) de la pensée humaine.

Le locuteur-auditeur idéal et abstrait, pourvoyeur de jugements et d'exemples de grammaticalité n'est donc en rien un sujet social. Encrevé (1986) l'a rappelé avec force, le locuteur chomskyen est sourd et muet, il ne communique pas et n'a aucune relation interindividuelle. Il est pris dans un

---

<sup>25</sup> "We are concerned, then with states of language faculty, which we understand to be some array of cognitive traits and capacities, a particular component of the human mind/brain. The language faculty has an initial state, genetically determined; in the normal course of development it passes through a series of states in early childhood, reaching a relatively stable steady state that undergoes little subsequent change (...). We call the theory of the state attained its grammar and the theory of the initial state universal grammar. (...) When we say that Jones has the language L, we now mean that Jones's language faculty is in the state L [...] To distinguish this concept of language from others, let us refer to it as I-language, where I is to suggest 'internal', 'individual', and 'intentional'. The explanatory model outlined deals specifically with language acquisition under the idealized conditions of an homogeneous speech community. (...) The (acquisition) process is (viewed) as if it were instantaneous. Chomsky (1995, 18-19).

solipsisme radical, sans horizon dialogique quelconque. En définitive, GU et GP sont gagées sur un corpus d'exemples sans rapport direct avec les usages observables dans une communauté linguistique précise. On comprend mieux alors le déni affiché par Chomsky pour la recollection des faits et pour l'enquête linguistique, centrales dans les linguistiques du *datum*. Récusant la base taxinomique de la pensée scientifique, il s'appuie sur une épistémologie des sciences uniquement hypothéticodéductive. Il le souligne avec mépris, la recollection des données d'usage, pas plus que celle des papillons, ne permet de fonder une pratique de recherche réellement scientifique (Cf. Supra note 19)

## 2.3 Les sciences du *datum*

En matière de linguistique et de philologie, la description des usages est aussi ancienne que la perspective grammaticale que je viens de rappeler. Si la linguistique de *l'exemplum* ne s'intéresse aux faits de langue que comme support inférentiel, la linguistique du *datum* se construit comme une observation minutieuse et une description des usages attestés dans leur diversité, leur hétérogénéité et leur variation. La notion d'usage l'emporte alors sur l'exemplarité des données observées.

Chez Cicéron, Horace et les grands rhéteurs<sup>26</sup>, comme chez Quintillien à leur suite, l'usage est toujours premier<sup>27</sup>. Pour ces linguistiques du *datum*, la tâche première du grammairien consiste ainsi à dresser des listes de faits langagiers que l'on peut plus justement appeler des corpus d'usages observés. Ce sont ces recollections qui fondent les régularités linguistiques car, comme le stipule Meigret (1542) à l'aube de l'histoire du français : "Les règles sont dressées sur l'usage et façon de parler lesquels ont toute puissance, autorité et liberté". Les règles ne s'infèrent donc pas des exemples comme des principes immanents, elles se déduisent « de la commune observance qui comme une loi nous les a tacitement ordonnées »<sup>28</sup>. On est comme on le voit très proche de ce qui sera théorisé quelques siècles plus tard sous le vocable de « communauté linguistique ». Ce qui, on l'a vu *supra* s'articule chez Saussure comme une conscience collective supportant un savoir linguistique commun partagé.

---

<sup>26</sup> « Licuit semperque licebit signatum praesente nota producere nomen. Multa renascentur quae iam cedidere, cadentque quae nunc sunt in honore uocabula, si uolet usus, quem penes arbitrium est et ius et norma loquendi. [On a toujours eu, on aura toujours la liberté de mettre en circulation un mot marqué au coin de l'année. Beaucoup renaîtront, qui ont aujourd'hui disparu, beaucoup tomberont, qui sont actuellement en honneur, si l'exige l'usage, ce maître absolu, légitime, régulier de la langue].» Horace (457/1944) Tr. Richard

<sup>27</sup> « Consuetudo uero certissima loquendi magistra, utendumque plane sermone ut nummo, cui publica forma est. Omnia tamen haec exigunt acre iudicium, analogia praecipue, quam proxime ex Graeco transferentes in Latinum proportionem uocauerunt [Quant à l'usage, c'est le maître le plus sûr, et l'on doit se servir hardiment du langage établi, comme de la monnaie marquée au coin de l'État. Mais tout cela exige beaucoup de discernement, surtout l'analogie, mot tiré du grec, et auquel correspond dans notre langue celui de rapport]» Quintilianus (1842, I, 6, 3) Tr. Nisard

<sup>28</sup> « Je confesse que cela serait raisonnable, si les règles qu'on fait de grammaire, commandaient à l'usage : vu qu'au contraire les règles sont dressées sur l'usage et façon de parler » Meigret (1542, 46). Sur ces questions Cf. Glatigny (1982)

Il reste que si le grammairien du *datum* est un collecteur de faits d'usage dont il tente de rendre raison<sup>29</sup>, les oppositions théoriques restent fortes. Entre contemporains les choix sont tranchés, car si la recollection des usages construit des corpus de référence, ces corpus, exhibent implicitement ou explicitement des principes de construction et de clôture forts différents. Qui donc commande sur l'usage et quelle est la valeur prescriptive de cet usage ? Telles sont les questions posées et le champ de la linguistique descriptive des usages aux 16<sup>ème</sup> et 17<sup>ème</sup> siècle est parcouru de conceptions opposées du *datum* linguistique. Au moment où se forgent le français moderne et ses premières grammaires, le débat oppose ainsi les normativistes prescripteurs aux descriptivistes constateurs. Avec Meigret, Ramus (1562) donne par exemple tous priviléges au peuple, c'est-à-dire à l'entièreté de la communauté des parleurs, pour régenter la langue<sup>30</sup>, tandis que Vaugelas restreint ce pouvoir quasi judiciaire à quelques élites choisies<sup>31</sup>.

On voit ainsi que les options normatives et prescriptives, tout comme les options descriptives sont également compatibles avec les linguistiques du *datum*. Ce qui fait le partage reste les principes de construction et de clôture des corpus et la qualité explicitement normative que le grammairien confère aux données. Si dans le cas des linguistiques de *l'exemplum* la question ne fait pas débat, ou presque, tout exemple construit par jugement qualitatif ou recueil sélectif pouvant toujours être vu comme une intériorisation pratique de la norme visée (Cf. Bourdieu 1982), pour les linguistiques du *datum* au contraire, prescription et description constituent des choix antithétiques. C'est en définitive non seulement le type de l'enquête, mais surtout la théorie de l'enquête linguistique mise en œuvre qui fait le départage. Soit l'enquête vise à dégager le bon usage en en sélectionnant au sein d'une communauté linguistique donnée les représentants légitimes, soit elle vise à décrire l'objet linguistique actif dans cette communauté. Dans le premier cas le corpus exhibe les formes correctes qu'il donne pour patron, et stigmatise incidemment les solécismes les plus vulgaires<sup>32</sup>, dans le second il décrit des usages dans leur hétérogénéité et leur variabilité attestées.

Certes, toute donnée linguistique est toujours naturelle dès lors qu'elle est produite par un sujet parlant, y compris lorsqu'elle est construite comme exemple par des linguistes, où suscitée chez un locuteur par appel à un jugement de grammaticalité. Il est seulement nécessaire de savoir

---

<sup>29</sup> Comme le note justement Glatigny (1982, 104) « Alors que Palsgrave prescrit, que Sylvius déduit, [Meigret] explique. »

<sup>30</sup> « Le peuple est souverain de sa langue et la tient comme un fief de franc alleu, et n'en doit reconnaissance a aucun seigneur. L'école de cette doctrine n'est point es auditoires des professeurs hébreux, grecs et latins en l'Université de Paris; elle est au Louvre, au Palais, aux Halles, en Greve, à la place Maubert » Ramus (1562)

<sup>31</sup> « De ce grand Principe, que le bon usage est le Maître de notre langue, il s'ensuit que ceux-là se trompent, que en donnent toute la juridiction au peuple [...], C'est la façon de parler de la plus saine partie de la Cour, conformément à la façon d'écrire de la plus saine partie des Auteurs du temps. Toutefois quelque avantage que nous donnions à la Cour, elle n'est pas suffisante toute seule de servir de règle, il faut que la Cour et les bons Auteurs y concourent, et ce n'est que de cette conformité qui se trouve entre les deux, que l'Usage s'établit. » Vaugelas (1647/1934)

<sup>32</sup> On se souviendra que les traités de bon usage se complètent utilement de prescriptions « ne dites pas, mais dites » sur le mode où *l'appendix probi* fût adjoint à *Instituta Artium*

comment et à quelles fins elle a été produite. Ainsi, répondant à Milner qui posait le primat du seul jugement de grammaticalité, Bourdieu réaffirme le caractère de *datum* de toute production linguistique, y compris les erreurs où les phrases impossibles forgées par les grammairiens à des fins argumentatives, dès lors que l'on prend en compte leur contexte social et pragmatique de production<sup>33</sup>. Toute donnée est pour lui le produit d'une observation située ou d'un questionnement : d'une enquête. Il est donc toujours nécessaire d'expliciter la situation d'enquête dans laquelle elle a été recueillie car ce n'est que dans la mise en relation de cette donnée, fut-elle fautive ou absurde, à ses conditions sociales de production et de réception que l'on peut en construire une analyse explicative (*Cf. Supra* la position de Meillet). Contrairement à la grammaire de *l'exemplum* qui hypostasie le jugement de grammaticalité, «La sociologie (ou, si l'on veut, la sociolinguistique comme branche de la sociologie) [...] s'accommode de toutes les formes d'acceptabilité; son *datum*, c'est la relativité absolue de l'acceptable et non l'absolutisation d'une forme particulière d'acceptabilité. [...] , (Bourdieu in Bourdieu et alii 1977, 45). Il s'ensuit que toute (socio)linguistique du *datum* est en définitive gagée sur la théorie de l'enquête qui est implicitement ou explicitement à son principe. Ainsi, comme l'a montré Encrevé (1976, 1982), c'est par exemple en partant d'une déconstruction du paradoxe de l'observateur et en visant un état de langue vernaculaire que Labov (1966, 1976, 1979) construit le *datum* particulier sur lequel s'érige sa (socio)linguistique.

## 2.4 La linguistique du *datum* aujourd'hui : le structuralisme

Tout comme la linguistique de *l'exemplum*, celle du *datum* trouve son actualité dans les approches contemporaines. Comme je l'ai montré ailleurs (Laks 2008), le structuralisme, qu'il soit européen ou américain, fonde son épistémologie sur la recollection et le classement systématique des données observées. Comme toutes les grandes sciences modernes depuis Linné (1735) et Buffon (1749) jusqu'à Lamarck (1809) et Darwin (1859) c'est une science taxinomique. Dans son expression formelle, que celle-ci prenne la forme d'un modèle structural (Troubetzkoy 1939, Hockett 1942), fonctionnel (Martinet 1962) ou transformationnel (Harris 1951), il s'agit toujours de fonder une systématique sur la taxinomie raisonnée d'un corpus de faits linguistiques. Ce corpus est le produit d'un récolement, donc d'une enquête.

Dans cette enquête, la notion de communauté linguistique est centrale. Contrairement à l'approche chomskyste qui réduit la communauté linguistique à un point unique virtuel dont le

---

<sup>33</sup> « La sociologie se donne pour première tâche le recensement des formes d'acceptabilité, c'est-à-dire de la relation entre une phrase et les situations où elle est acceptable. [...] Donc, premier travail du sociologue : recenser. Ensuite, il reste à faire la science des conditions de production de la phrase et des conditions de son acceptabilité. » Bourdieu in Bourdieu et alii 1977, 45).

locuteur-auditeur idéal constitue en quelque sorte l'éponyme, l'enquête structuraliste opérationnalise la notion de relation de communication. Celle-ci qui permet de circonscrire la communauté linguistique comme le réseau des interlocuteurs potentiels. En effet, pour les structuralistes, depuis Bloomfield au moins, l'objet premier de l'enquête linguistique n'est pas la langue, mais la communauté linguistique : la langue est ce qui s'observe au sein d'une communauté linguistique donnée comme support de la communication entre les individus. La nature sociale et culturelle de la langue, au sens de Whitney et Saussure, fonde donc la recollection des observables linguistiques dans le périmètre de l'ensemble des locuteurs qui usent du même code. Imaginez dit Bloomfield (1933, 46) qu'à chaque fois qu'un locuteur adresse une occurrence linguistique à un autre locuteur, on trace une ligne entre ces deux sujets sociaux. Au bout d'un certain temps, ce graphe virtuel, par la densité de ses zones noires, grises et blanches, permettrait à la fois de cerner la communauté linguistique et la langue qui l'organise en même temps qu'il ferait apercevoir les densités relatives de communication et d'échange. La notion de réseau de communication est comme on le voit consubstantielle à cette approche.

Certes, ces réseaux ne sont pas sociologisés. La communication y est encore vue comme symétrique et réciproque et les rapports sociaux de domination et de prestige, qui rendent toujours l'échange inégal, sont ignorés. Mais il suffit, comme le fit le premier Labov, d'ajouter à l'enquête linguistique une dimension sociologique (Labov 1966), ou mieux, comme le fit le second Labov<sup>34</sup>, de déconstruire la communauté linguistique unique et de prendre en compte l'asymétrie des relations sociales (Labov 1972), ou encore comme Milroy (Milroy et Milroy 1985) d'interroger la réticulation sociale des réseaux, pour dépasser l'enquête linguistique structuraliste et construire une forme d'enquête proprement sociolinguistique. Dans tous les cas, comme le souligne Hymes (1972, 43), "The natural unit for sociolinguistic taxonomy is not the language but the speech community"<sup>35</sup>

Ce qui fait le départ de cette sociolinguistique structurale avec la dialectologie classique, c'est précisément l'attention portée à l'organisation sociale interne de la communauté linguistique. Là où la dialectologie voyait la variation et le changement comme des signes de la désintégration de la communauté originelle et de sa langue, et plaçait en conséquence les NORMS<sup>36</sup> au cœur de son investigation linguistique, les sociolinguistiques modernes voient la communauté linguistique comme fondamentalement hétérogène, divisée, hiérarchisée et structurée. Elles en concluent que la variation sociolinguistique est bien au cœur des langues réellement vivantes, c'est-à-dire socialisées. Dans cette approche, la communauté linguistique ne présente jamais d'uniformité des usages et des

<sup>34</sup> Sur la distinction entre le premier et le second Labov, Cf. Encrevé (1976)

<sup>35</sup> Pour une analyse détaillée du concept de communauté linguistique Cf. Patrick (2002)

<sup>36</sup> Acronyme forgé par Chambers et Trudgill (1980) pour désigner les « Non-mobile, Old, Rural, Male Speakers » qui constituent les locuteurs de référence des dialectologues classiques.

pratiques linguistiques. Elles sont au contraire aussi stratifiées et diversifiées que la structure sociale elle-même. Les relations de communication sont alors réglées par un corps de normes sociales et linguistiques reconnues par tous qui évaluent et hiérarchisent ces pratiques différenciées de façon unique<sup>37</sup>. Dissensus des usages et des pratiques donc, et consensus des normes et des systèmes d'évaluation : on n'est pas loin du contrat durkheimien reformulé par Saussure comme la dimension sociale du trésor linguistique déposé dans le cerveau de chaque individu.

La communauté linguistique qui est au centre des linguistiques structurales, et plus largement de toutes les linguistiques du *datum*, est donc d'abord le lieu de l'hétérogénéité, de la différenciation et des oppositions, des antagonismes et des luttes. Tant qu'on la regarde au niveau des individus elle apparaît dispersée et variable, mais dès qu'on l'envisage comme une structure sociale, et les individus comme des agents sociaux impliqués dans ce treillis, elle apparaît comme un système fonctionnel de gestion et d'organisation de cette hétérogénéité, système dont la reconnaissance de normes communes, ce qui ne veut pas dire également partagées, est la clé de voute. Dans les termes de la sociologie de Bourdieu, l'hétérogénéité et la différenciation sociale des agents sont des principes de fragmentation, mais la reconnaissance légitime des principes sociaux de classement et d'organisation reconstitue l'unicité du social. C'est pourquoi les luttes sociales peuvent être regardées comme des luttes pour la définition légitime de ce qui s'impose légitimement à tous : normes, classes, hiérarchies. Ce qui au niveau individuel et interindividuel apparaît comme variation et hétérogénéité apparaît au niveau social, celui de la communauté linguistique, comme fonctionnel et organisé. C'est ainsi dans la communauté linguistique que l'acte, éminemment labile, du locuteur se constitue comme fait social fonctionnel, c'est-à-dire comme *datum* pour la (socio)linguistique<sup>38</sup>.

### 3 Datum et exemplum : le tournant des années 2000 :

Le paysage de la linguistique contemporaine reste informé par ces oppositions théoriques fondamentales. Dans les années soixante, la rupture avec le structuralisme et l'avènement de la linguistique cartésienne (Chomsky 1965, 1966) avaient durablement imposé la linguistique de *l'exemplum* qu'est la Grammaire Générationnelle comme paradigme linguistique dominant. Ceci avait conduit à une négation du caractère structurant de la communication interpersonnelle. En effet si comme on vient de le rappeler la Grammaire Générationnelle, comme toutes les linguistiques de

---

<sup>37</sup> "The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms. These norms may be observed in overt types of evaluative behavior, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage." Hymes (1972, 120)

<sup>38</sup> "A SpCom is defined in functionalist terms as a system of organized diversity held together by common norms and aspirations... Members of such a community typically vary with respect to certain beliefs and other aspects of behavior. Such variation, which seems irregular when observed at the level of the individual, nonetheless shows systematic regularities at the statistical level of social facts". Gumperz (1982, 24)

*l'exemplum*, substitue un sujet ontologique, au locuteur-auditeur socialisé et dénie ainsi toute pertinence linguistique au concept même de communauté linguistique, elle est nécessairement conduite à dénier tout caractère fonctionnel à la communication interpersonnelle.

Comme le rappelle avec force Chomsky dans ses plus récentes analyses, pour la linguistique cartésienne, le langage humain n'est ni défini ni mis en forme par la communication. L'approche par la fonction sociale ou communicative est récusée en droit comme en fait : ni d'un point de vue phylogénétique, et encore moins d'un point de vue ontogénétique, la langue ne se constitue pour des besoins fonctionnels, qu'ils soient communicatifs ou autres<sup>39</sup>. La langue chomskyenne se définit sur un terrain uniquement cognitif et mental, et dans une approche strictement individuelle<sup>40</sup>. Comme l'a bien noté Labov (1987), dans l'approche idéaliste de la Grammaire Générale, la langue est une propriété ontologique de l'individu en tant qu'il appartient à l'espèce homo alors que dans l'approche matérialiste, le langage est une propriété fonctionnelle des groupes sociaux formés par l'espèce socio<sup>41</sup>. Il s'ensuit nécessairement des approches très différentes de ce qu'est pour le linguiste un fait de langue (Cf. Labov 1975) et l'on retrouve l'opposition *datum/exemplum* que nous venons de voir. Ainsi, ce qui est en jeu n'est pas uniquement l'inconsistance du jugement de grammaticalité très souvent démontrée en situation de variation sociolinguistique (Cf. Labov 1996), mais la caractérisation même de ce que l'interrogation du linguistique constitue comme fait de langage. Après avoir souligné que l'approche fonctionnelle du langage comme outil de communication dans des groupes sociaux différents est portée à prendre en considération essentiellement les différences qui s'éparent les langues pour, à l'aide d'une méthode structurale, en proposer une description systématique, Berwick et Chomsky (2011)<sup>42</sup> soulignent ainsi que l'approche matérialiste non fonctionnaliste qui voit le langage comme une propriété de l'esprit humain est au

---

<sup>39</sup>"It has been conventional to regard language as a system whose function is communication. This is indeed the widespread view invoked in most selectionist accounts of language, which almost invariably start from this interpretation. However, to the extent that the characterization has any meaning, this appears to be incorrect, for a variety of reasons to which we turn below." Berwick et Chomsky (2011, 25 également 35, 36).

« Accordingly, any approach to evolution of language that focuses on communication [...] may well be seriously misguided" Chomsky (2011, 61). Cf. également Hauser, Chomsky et Fitch (2002, 1569)

<sup>40</sup> "The word "language" has highly divergent meanings in different contexts and disciplines. In informal usage, a language is understood as a culturally specific communication system (English, Navajo, etc.). In the varieties of modern linguistics that concern us here, the term "language" is used quite differently to refer to an internal component of the mind/brain (sometimes called "internal language" or "I-language"). We assume that this is the primary object of interest for the study of the evolution and function of the language faculty. However, this biologically and individually grounded usage still leaves much open to interpretation (and misunderstanding)". Hauser, Chomsky et Fitch (2002, 1569)

<sup>41</sup> «What is language? The idealist conception is that language is a property of the individual, a species-specific and genetically inherited capacity to form rules of a particular type, relatively isolated from other activities of the human intelligence. The materialistic conception is that language is a property of the speech community, an instrument of social communication that evolves gradually and continuously throughout human history, in response to a variety of human needs and activities. » Labov (1987, x)

<sup>42</sup> Il est assez rare de trouver sous la plume de Chomsky une caractérisation, même critique, de la linguistique structurale et des linguistes qui l'on précédé. La référence à Sapir, Boas et Bloomfield, comme au structuralisme européen via Troubetzkoy, et la mise en exergue de Harris, tous considérés comme des linguistes de la diversité linguistique alors que lui-même défend l'unicité linguistique, mérite d'être soulignée.

contraire portée à réduire les différences de surface ou d'actualisation pour, par induction formelle, dégager l'unicité et la singularité du fonctionnement cognitif impliqué<sup>43</sup>

On voit ainsi que l'opposition entre linguistique structurale et linguistique générative n'est pas fondée sur des arguments techniques ou méthodologiques concernant la puissance des modèles explicatifs des faits grammaticaux, ou sur la capacité descriptive, ou encore prédictive, des analyses proposées, mais sur une différence radicale d'approche de ce qu'on doit considérer comme un sujet parlant, un fait de langage, le rôle de la communication où l'existence même de communautés de locuteurs. Avec le tournant cartésien des années soixante, la linguistique chomskienne a ainsi porté au premier plan une approche mentaliste, idéaliste et abstraite qui, comme au 17<sup>ème</sup> siècle, s'appuyait nécessairement sur une linguistique de *l'exemplum* excluant *ex hypothesis* toute variation et toute hétérogénéité structurales.

### 3.1 Un ressourcement empirique

Dès la fin des années 1990 et au début des années 2000 pourtant, le champ international de la linguistique devait considérablement évoluer et les linguistiques du *datum* retrouver la place qui était la leur cinquante ans auparavant. Dans tous les domaines de l'analyse linguistique, les travaux descriptifs de grande ampleur se multiplient et la base empirique des analyses s'en trouve singulièrement accrue. La notion même de corpus de données linguistiques, tant décriée dans le cadre génératif, s'en trouve réaffirmée.

Il n'est pas de domaine qui échappe à ce ressourcement empirique, lequel conduit à une critique directe des limites et des impasses des analyses construites sur la base d'un nombre très limité d'exemples repris d'un auteur l'autre<sup>44</sup>. Pour n'en prendre qu'un seul exemple, emblématique entre tous, les hypothèses innéistes de la Grammaire Générative fondées sur la thèse de la pauvreté du stimulus incapable par la même de conduire l'acquisition du langage (Berwick et Chomsky 2008, Chomsky 1968), peuvent depuis les années 2000 utilement être réévaluées à la lumière des corpus d'observation CHILDES qui comptaient en 2004 44 millions de mots dans 32 langues et pesait 750 gigabits<sup>45</sup>. A ce jour, plus de 3200 publications s'y réfèrent.

---

<sup>43</sup> Cette unicité et cette singularité résident dans le principe de récursivité dont la fonction d'unification FUSION (MERGE) est une instance. J'y reviens ci-après.

<sup>44</sup> Ainsi l'analyse principe de la liaison et du e muet en français dans un cadre génératif (Schane 1965) propose 41 règles ordonnées fondées sur 73 exemples (Cf. Laks 2011a). A contrario, le programme « Phonologie du Français Contemporain » offre, pour les mêmes phénomènes, une base de 190 000 et 47 500 sites phonologiques pertinents qui permet d'en ressourcer complètement la phonémologie. Il est alors possible de proposer pour ces questions classiques de la phonologie du français des analyses radicalement nouvelles (Cf. Durand, Laks, Calderone et Tchobanov 2011)

<sup>45</sup> Pour une présentation générale Cf. MacWhinney (2000, 2007). Pour une vue synthétique Cf. Gleason et Thompson (2002). Il faut ajouter à ces chiffres le corpus Talkbank qui pèse 450 gigabits pour 55 millions de mots en 18 langues, Cf. <http://childepsy.cmu.edu/>

Plus généralement, le développement considérable des capacités de stockage et de traitement informatiques, l'apparition d'outils technologiques nouveaux et le développement de réseaux coopératifs de partage de données permis par internet induisent à partir des années 2000 une réorientation empirique très sensible de la linguistique qui, pour reprendre la dichotomie construite par Berwick et Chomsky (2011) réoriente les sciences du langage vers la documentation de la diversité des phénoménologies à chaque niveau d'analyse comme condition préliminaire à toute construction d'hypothèses abstraites concernant l'unicité du phénomène cognitif. Il s'agit certes pour les tenants du cadre génératif d'une réorientation méthodologique sans grandes conséquences parce qu'elle affecte tous les niveaux (phonologie, morphologie, syntaxe, sémantique pragmatique etc.) tenus par eux pour périphériques, alors que leur modèle innéiste réductionniste ne porte que sur le noyau syntaxique central universel (*Cf. infra*), mais ce ressourcement empirique généralisé affecte durablement le paysage théorique de la linguistique. Des figures majeures de la pensée linguistique, longtemps décriées ou marginalisées comme Harris reviennent au centre du débat inter théorique avec une acuité insoupçonnée (Nevin 2002). La réitération des caractérisations anciennes qui considéraient avec quelque dédain le cadre théorique de la grammaire transformationnelle (Harris 1951) comme offrant une simple approche méthodologique<sup>46</sup>, ne suffisent plus à contrer ce qu'une relecture de ces propositions apporte pour la linguistique du 21<sup>ème</sup> siècle (Goldsmith 2005b). Comme je l'ai montré ailleurs (Laks 2008), et contrairement à ce qui continue d'être affirmé dans le cadre génératif, l'approche méthodologique de la diversité linguistique et l'analyse descriptive construite sur corpus de grande taille n'interdisent aucunement la construction d'hypothèses générales et abstraites et pour tout dire la construction d'un modèle théorique explicatif. Tout au contraire, se réaffirme au tournant du siècle, la conception classique pour les historiens de la science, qui voit dans la description systématique, le classement raisonné des faits, et la construction progressive d'hypothèses systémiques ascendantes, le mouvement même de la pensée scientifique moderne vers la construction de théories explicatives de plus en plus compréhensives.

Ce tournant vers ce que Goldsmith (à paraître) définit comme un néoempiricisme se marque donc par l'apparition de très nombreux corpus de grande taille et par la construction d'outils méthodologiques et théoriques de traitement particulièrement puissants. Comme il le montre également, les développements récents en statistique, probabilistique, et plus généralement en stochastique mathématique renouvelent les approches formelles et modélisatrices. La linguistique

---

<sup>46</sup> « The publication that was the foundation of American structural linguistics in the 1950s, Zellig Harris's *Methods in Structural Linguistics* (1951), was called "methods" because there seemed to be little to say about language beyond the methods for reducing the data from limitlessly varying languages to organized form. European structuralism was much the same. Nikolai Troubetzkoy's classic introduction to phonological analysis was similar in conception. More generally, structuralist inquiries focused almost entirely on phonology and morphology, the areas in which languages do appear to differ widely and in complex ways, a matter of broader interest, to which we will return Berwick et Chomsky (2011, 2)

de corpus qui constitue la base empirique et descriptive nécessaire aux développements de ces analyses s'impose ainsi dans toutes les dimensions de la recherche sur les langues et le langage, y compris les dimensions de grammaires formelles ou d'apprentissage psycholinguistique<sup>47</sup>.

Dans le cadre de ce néo-empirisme, les questions de la variation et de l'hétérogénéité des données, et partant de la variabilité et de la plasticité des modèles capables d'en rendre compte, qui furent longtemps forcées par la linguistique de *l'exemplum* reviennent au centre de la scène. La linguistique variationniste labovienne ne peut plus dès lors être tenue comme le faisait Chomsky (1977, 74-75) pour une dialectologie périphérique sans intérêt théorique. L'analyse quantitative de grands corpus et l'approche systémique de la variation interne des langues et des grammaires qui constituent le cœur de son programme scientifique retrouvent ainsi une actualité de première importance (labov 2004)<sup>48</sup>. Cette réorientation néo-empiriciste, que Labov appelle matérialiste, ne renonce pas pour autant à l'ambition théorique telle qu'elle est proclamée par la linguistique cartésienne. Tout au contraire, en reprenant les concepts fonctionnalistes de communication interpersonnelle en situation, de communauté linguistique et d'interaction, les années 2000 voient émerger de nouvelles approches théoriques et analytiques du phénomène langagier.

### 3.2 Des linguistiques de corpus aux linguistiques de l'usage

Ces nouvelles s'approchent empiricistes, tout en s'appuyant sur un certain nombre de thèses structuralistes (approches fonctionnalistes, rôle de la communication interpersonnelle socialisée, circonscription des communautés linguistiques) trouvent leur origine dans un clivage ancien au sein de la Grammaire Générationnelle alors naissante. Comme le montrent Goldsmith et Huck (1995), la polémique de la « sémantique générative », bien loin de constituer un épiphénomène historique, simplement anecdotique ou de pure surface, actualise au milieu des années 1960 une rupture fondamentale et extrêmement profonde dans le jeune courant générativiste et transformationnaliste. Même si à la fin des années 1970 le courant de la sémantique générative semble définitivement défait et si la polémique se tarit progressivement, les principaux porteurs de la contestation du modèle générativiste orthodoxe restent très actifs et la plupart se retrouvent au cœur des recompositions théoriques qui occupent le devant de la scène cognitive dans les années 2000. Ces

---

<sup>47</sup> Cf. par exemple Goldsmith et Aris (2009) pour une analyse quantitative et formelle de l'apprenabilité *tabula rasa* des catégories phonologiques.

<sup>48</sup> C'est sans doute sur le terrain de la phonologie, spécialement optimaliste, que la relation entre modélisations phonologiques et variationnisme a été récemment la plus productive. Cf. par exemple Antilla (2007), Antilla et Cho (1998), Boersma (1998). Pour une approche de la relation entre phonétique et phonologie sous ce rapport Cf. Hayes et Czirák Londe (2006), Hayes, Kirchner et Steriade (2004). Cf. également les actes de « Workshop on Variation, Gradience and Frequency in Phonology (Stanford, CA, 2007) : <http://www.stanford.edu/dept/linguistics/linginst/nsf-workshop/workshop-july-2007.html>

modèles étant centraux dans le débat actuel, il est utile de revenir rapidement sur ce moment crucial de rupture où nombre d'auteurs ont rompu avec le modèle chomskyan standard<sup>49</sup>.

De fait, lorsqu'on examine les principaux thèmes promus par les sémanticiens générativistes et les principales critiques qu'ils portent<sup>50</sup>, on est frappé par la contiguïté avec nombre de questions contemporaines. Pour Lakoff (1973a) par exemple il est absolument nécessaire de prendre en compte le contexte social, culturel et interactionnel dans lequel se déploie le langage. Sa critique du tournant cartésien réaffirme le caractère fonctionnel de la langue et il estime nécessaire de placer la communication interpersonnelle au cœur du dispositif linguistique. Il s'ensuit, dès cette période, une appréciation de la nécessité d'une base empirique et descriptive solide. Contre l'innéisme cartésien, il défend une conception sociale et culturelle des faits langagiers qui conduit à donner le primat au contenu sémantique et pragmatique des occurrences linguistiques pour en analyser la forme et les caractéristiques tant spécifiques qu'universelles. Ainsi, un certain nombre des lignes de force qui structurent le débat le plus contemporain étaient déjà actives dans la polémique de la sémantique générative des années 1960 et 1970. La plus grande partie de ses acteurs se retrouveront à partir des années quatre-vingt-dix dans la mouvance de la linguistique cognitive.

La linguistique cognitive constitue aujourd'hui un pôle de regroupement de nombreux courants contemporains, de la sémantique et la pragmatique, aux grammaires de constructions et au variationisme sociolinguistique, en passant par le connexionnisme. Ces courants, tout en gardant chacun leurs orientations spécifiques coopèrent et convergent sur un certain nombre de grandes orientations. Langacker (1987) a proposé de regrouper ces courants sous le vocable commun de « modèles linguistiques fondés sur l'usage ». Cette appellation synthétique commode désigne un large secteur de la recherche linguistique internationale qui partage des positions communes, sans être unifié autour d'une théorie standard rigide. A l'exact contraire de la Grammaire Générative en effet, ce courant contemporain défend la très grande richesse des données d'usage mais prône une certaine austérité de l'appareil conceptuel et du cadre théorique et formel<sup>51</sup>. Dans ce cadre, l'usage

<sup>49</sup> Lakoff, cite Ross et MacCawley comme constituant le noyau initial auquel s'adjoignent plus tard Fillmore, Talmy et Langacker, puis dans le développement de la linguistique cognitive, Fauconnier, Rosch, Kay, MacDaniel. Dans les années 1980 et 1990, ils convergent avec les défenseurs du connexionnisme, Rumelhart, Feldman etc. Lakoff (1973a) établit ainsi un lien entre des linguistes qui sur une quarantaine d'années furent actifs successivement dans différent cercles : la sémantique générative, la pragmatique, les grammaires casuelles, les grammaires de construction, la théorie de la métaphore, le connexionnisme etc.

<sup>50</sup> Une analyse des fondements théoriques de la polémique n'entre pas dans mon propos. On se reportera à Harris (1993) pour une analyse historique et à Goldsmith et Huck (1995) pour une analyse plus épistémologique. Les textes de synthèse des sémanticiens générativistes ne portant pas sur un point technique ou un argumentaire précis, sont assez rares. Je renvoie néanmoins à Lakoff (1973a, b). Le second constitue une réponse à l'analyse, elle-même critique, du débat entre les sémanticiens générativistes et les tenants du modèle standard par Searle (Searle 1972).

<sup>51</sup> "I have argued, both on methodological and on empirical grounds, that the principle of generality have received in linguistics a commonly accepted interpretation that is in fact not appropriate to its subject matter. Current doctrine favors a minimalist account of linguistic knowledge, described in accordance with complex array of theoretical apparatus featuring specialized devices for the various 'components' of the linguistic system. By contrast, cognitive grammar pursues a maximalist account of linguistic knowledge, and tends toward austerity in adoption of theoretical constructs; it seeks a unified treatment of the various facets of linguistic structure, attributing their differences to the content of the domain in question rather than the basic constructs invoked to handle them". Langacker (1988, 160)

sédimente en effet un savoir linguistique pratique considérable, actif dans tous les domaines de la production et de l'interprétation des occurrences, tandis que les principes cognitifs abstraits et généraux, réduits au minimum, ne sont pas propres au domaine du langage mais correspondent à des fonctions cognitives majeures omnibus.

Dans ces approches, les pratiques linguistiques ont le statut d'observables situés, la richesse du contexte social et culturel d'arrière-plan déterminant pour une grande part leur interprétation partagée (Barlow et Kemmer (2000, XXVI). Ainsi, dans un cadre structuré par les relations interpersonnelles et la fonction de communication, les pratiques langagières et leur interprétation sémantique et pragmatique sont premières<sup>52</sup>. La grammaire n'est plus une condition de la production d'évènements langagiers elle en est au contraire le sous-produit routinisé et appauvri (Barlow et Kemmer (2000) XI). En accordant une place centrale au processus de grammaticalisation, vu à la fois comme cognitif et diachronique, les linguistiques de l'usage renversent totalement la dichotomie classique chomskyenne compétence/performance, généralisant la performance comme moteur pratique et fonctionnel de la communication et ramenant la compétence à une sous partie sédimentée interne à cette dernière (Langacker 2000, 6-9). Faisant écho à Benveniste et aux courants discursifs qui en sont issus, Langacker fait clairement du discours en interaction la source de la langue et de sa routinisation grammaticale et adopte ainsi la perspective neo-empiriciste évoquée ci-dessus<sup>53</sup> Comme je l'ai montré ailleurs (Laks 2011b), telle était déjà la position de Saussure qui défendait le primat de la linguistique de la parole comme condition *sine qua non* d'une grammaire de la langue.

Thématisant ainsi les pratiques langagières dans leur contexte écologique, social et culturel, les linguistiques de l'usage redonnent un statut central à la description linguistique, aux analyses distributionnelles, statistiques et fréquentielles, et se situent clairement dans la mouvance des linguistiques de corpus<sup>54</sup>. Ce sont des linguistiques du *datum* et corrélativement, la variation et l'hétérogénéité internes y sont reconnues pour ce qu'elles sont. Mais surtout, dans les modèles

<sup>52</sup>“A usage-based model is one in which the speaker's linguistic system is fundamentally grounded in 'usage events': instances of a speaker's producing and understanding language [...] In this view, it does not make sense to draw a sharp distinction between what is traditionally called 'competence' and 'performance,' since performance is itself part of a speaker's competence. Instead of viewing language processing as something external to the system, which happens only to the outputs of competence, processing is rather to be seen as an intrinsic part of the linguistic knowledge system, which cannot be treated separately from it? Kemmer et Barlow (2000 VIII- IX)

<sup>53</sup> Pour fonder ce primat du discours, on s'est souvent autorisé de Benveniste (1966, 131) sans toujours voir que la sentence qu'il forge à ce dessein « *nihil est in lingua quod non prius fuerit in oratione.* » n'est que le calque de la sentence scolaistique parfaitement aristotélicienne « *nihil est in intellectu quod non prius fuerit in sensu* » qui fait le fond de l'empirisme sensualiste de Hume ou de Locke

<sup>54</sup> “A usage-based theory, whether its object of study is the internal or external linguistic system, takes seriously the notion that the primary object of study is the language people actually produce and understand. Language in use is the best evidence we have for determining the nature and specific organization of linguistic systems. Thus, an ideal usage-based analysis is one that emerges from observation of such bodies of usage data, called corpora. But even if not based primarily on such data, at a minimum, analyses must ultimately be at least consistent with production data [...] “The importance of frequency: Because the system is largely an experience-driven one, frequency of instances is a prime factor in its structure and operation. Since frequency of a particular usage pattern is both a result and a shaping force of the system, frequency has an indispensable role in any explanatory account of language” Kemmer et Barlow (2000, XV, IX)

basés sur l'usage, cette variation sociolinguistique, tant synchronique que diachronique, retrouve un statut systémique et fonctionnel<sup>55</sup>. Enfin, tant du point de vue de l'instanciation cognitive des régularités linguistiques que du point de vue de l'apprentissage situé de la langue et de la grammaire, ces approches sont parallèles aux approches connexionistes et neuromimétiques et plus généralement ressortissent aux modélisations dynamiques du langage. Avec toutes leurs variantes et approches connexes, grammaires de construction, grammaires exemplaristes et occurentialistes, grammaires discursives, neurales et cognitives, grammaires stochastiques et probabilistiques etc.<sup>56</sup>, les modèles basés sur l'usage et les linguistiques du *datum* ont profondément modifié le champ linguistique international et, marginalisant la linguistique cartésienne, constituent comme on va le voir le paradigme dominant de ce début de 21<sup>ème</sup> siècle.

### 3.3 La grammaire et l'usage, un bilan critique

Dans son adresse présidentielle au congrès de la *Linguistic Society of America*, Newmeyer (2003), dont on sait qu'il est pourtant un défenseur fervent de la linguistique chomskienne (cf. Newmeyer 1988), confirme ce changement et dresse une analyse assez désabusée du paysage international de la recherche linguistique. Il le confirme, le paradigme de la Grammaire Générative est devenu internationalement minoritaire tandis que domine aujourd'hui celui des modèles d'usage. Pour Newmeyer, la sémantique générative qui s'est ressourcée en linguistique cognitive a fini, avec les linguistiques de l'usage, par imposer le rejet de toute distinction claire entre savoir linguistique et usage de la langue, distinction qui est pourtant au cœur du paradigme chomskien (*op. cit.* 683). L'autonomie et la stabilité cognitives de la grammaire ont été battues en brèche au profit de dynamiques transitoires, de structurations partielles, d'organisation stochastiques. Toutes ces approches stipulent un stockage important de formes et d'occurrences concrètes et minimisent la portée et l'effet des contraintes grammaticales formelles (*op. cit.* 683-684)<sup>57</sup>. Ces approches se sont très largement diffusées, jusque dans modèles optimalistes, pourtant initialement proches du générativisme. Le connexionnisme et les systèmes neuromimétiques, tant attaqués dans les années

---

<sup>55</sup> "[There is an] intimate relation between usage, synchronic variation, and diachronic change : Patterns in usage data are in general patterns of variation along different dimensions of various kinds, from formal to social. In a cognitive usage-based model, variant linguistic forms can be thought of as alternate possibilities licensed by the linguistic network. The selection of a given entrenched variant for activation is governed by a complex set of motivating factors, including system-internal as well as contextual, situational factors. As observed in the seminal work of Labov, variation is highly structured, not only in the individual's system, but across groups of speakers. The effects of usage on the linguistic system [...] lead us to expect that speakers' language will be influenced by the productions they hear in particular speech communities of which they are members. [...] "the more speakers talk to each other the more they will talk alike, and so linguistic variation will pattern along lines of social contact and interaction." Kemmer et Barlow (2000, XVII)

<sup>56</sup> Goldberg (2006), Goldberg (1995), Bybee (2001), Bybee (2006), Feldman (2006), Lakoff et Johnson (1999), Chater et Manning (2006), Goldsmith et Aris (2009), Manning (2003)

<sup>57</sup> Cf. Langacker 1998 cité *supra* note 52

1980 (Pinker et Mehler 1989) se sont développés et fournissent un appui computationnel à ces approches. Tel est le constat du Président de la LSA en 2003.

Aucun domaine de recherche n'échappe à ce changement du paradigme de référence. Le traitement automatique des langues et la recherche en grammaire formelle sont affectés, comme la phonologie, la syntaxe ou la sémantique.<sup>58</sup> Mais plus, la distinction chomskienne compétence/performance elle-même est devenue obsolète<sup>59</sup> et la psycholinguistique strictement générative s'est réduite comme peau de chagrin<sup>60</sup>. Finalement, l'option fonctionnaliste et communicative domine à nouveau le paysage linguistique et force est de le constater, des arguments de poids ont été produits à l'appui des linguistiques de l'usage<sup>61</sup>. Avec ce nouveau paradigme, la variation, l'hétérogénéité et la sociolinguistique, pourtant tant décriées comme épiphénoménales (*Cf. supra*), reviennent sur le devant de la scène théorique. Si de plus, la notion même de phrase est contestée, et si la disparité complète entre ce que prédit la grammaire et ce dont atteste l'usage devient admise par tous, que reste-il de la linguistique cartésienne et de la Grammaire Générative?<sup>62</sup>

Comme très souvent lorsque le paradigme chomskien est majoritairement mis en question, plus tôt qu'à un débat avec les contradicteurs, à une évolution ou une adaptation prenant en compte les arguments opposés, on a assisté à partir des années 2000 à une réorganisation complète du cadre de référence de la Grammaire Générative<sup>63</sup>. Avec l'apparition de la Biolinguistique cette

<sup>58</sup> "I am quite sure that Christopher Manning is right when he writes that '[during] the last 15 years, there has been a sea change in natural language processing (NLP), with the majority of the field turning to the use of machine learning methods, particularly probabilistic models learned from richly annotated training data, rather than relying on hand-crafted grammar models' (Manning 2002b:441)" Newmeyer (2003, 682).

<sup>59</sup> "I believe that the great majority of psycholinguists around the world consider the competence-performance dichotomy to be fundamentally wrongheaded". (Newmeyer 2003, 682)

<sup>60</sup> La constatation avait déjà été faite par Tomasello (1995, 135). "The list [of innate aspects of language] contains things that no nonlinguist would ever recognize: such things as the projection principle, the empty category principle, the subjacency constraint, and the coordinate structure constraint. All of these universals are described in linguistically specific terms such that it is very difficult to relate them to cognition in other psychological domains".

<sup>61</sup> "First and most importantly, there is the evidence that has mounted in the past quarter-century that significant aspects of grammars are motivated by considerations of use. Functional linguists and generative linguists with a functional bent have provided (to my mind) incontrovertible evidence that grammars are shaped in part by performance considerations" (Newmeyer 2003, 683)

<sup>62</sup> "Reinforcing skepticism about classical generative models is the disparity between sentences generated by these grammars and actual utterances produced by language users. This disparity has led some linguists to conclude that grammar itself bears no relation to the proposition-like structures posited by formal linguists; structures specified by formal rules that take the sentence to be the basic unit of grammar, where sentences are in a rough mapping with propositions, verbs with predicates, and noun phrases with logical arguments. The priority of the sentence is dismissed by some critics of the generative program as a carryover from the Western logical tradition, reinforced by the conventions of written language" (Newmeyer 2003, 683)

<sup>63</sup> Comme cela a été remarqué par ailleurs (Encrevé 2000), le courant chomskien est souvent enclin à une réécriture hagiographique de sa propre histoire. Ainsi, parce que Chomsky (2007, 9) fait remonter la création du terme biolinguistique au compte-rendu que Piatelli-Palmarini fit en 1974 du débat qu'il eut à Royaumont avec Piaget, tous les tenants du nouveau cadre datent sa naissance de cette période (par exemple Di Sciullo et Boeckx (2011), Berwick (2011), Boeckx et Grohmann (2007)). Les historiens qui auraient noté l'absence du terme dans le titre de la version française ou anglaise du débat (Piatelli-Palmarini 1980, Piatelli-Palmarini et Noizet 1979) et auraient souligné une référence plus constante à la question de l'apprentissage dans un cadre innéiste vs constructiviste auraient donc proposé une lecture historique fautive. Chomsky (2007, 9) va plus loin et fait remonter le cadre biolinguistique à ses premiers travaux des années cinquante. Van Riemsdijk est un peu dubitatif : "I sort of said jokingly that in retrospect, if you tell Chomsky that he only came up with that stuff later, he would probably deny it and say that it had been clear to him right from the start — and, you know, who am I to say that he would be lying? All I'm saying is there was no real evidence in the writing that that was the main goal he was pursuing. It would actually be interesting at some point to ask him this question." Grohmann (2007, 138)

réorganisation de la théorie générative tente d'imposer un changement radical des termes du débat. La question se situerait désormais sur le terrain de l'évolution des espèces où la Bolinguistique chomskienne propose de démontrer la spécificité de la faculté de langage propre à l'homme.

S'appuyant sur la distinction des deux extensions du concept de faculté de langage déjà posée dans le programme minimaliste (Chomsky 1995) et reformulée dans le nouveau cadre évolutionniste biogénétique (Hauser, Chomsky et Fitch 2002), Newmeyer propose une défense et illustration de la Grammaire Générative en défendant une distinction radicale entre grammaire et usage. Dans une argumentation qui concède tout aux modèles fondés sur l'usage, il reprend ainsi la thématique chomskienne selon laquelle la Grammaire Générative et la Biolinguistique ne sont pas des linguistiques en ceci qu'elles ne prennent pas pour objet la langue au sens circulant du terme. Elles s'intéressent uniquement au dispositif cognitif qui est à son principe. Ainsi, par une politique de la terre brûlée, il concède que les linguistiques du *datum* et les modèles fondés sur l'usage couvrent avec satisfaction tout l'empan du champ linguistique, mais il reste, dit-il, la grammaire, et en son principe elle doit être radicalement distingué de l'usage. Ce reste cognitif et grammatical est en propre l'objet de la grammaire Générative et de la Biolinguistique qui n'est concernée que par la faculté de langage au sens étroit et aucunement par le sens large.

En effet, la faculté de langage au sens large du terme (Faculty of Language in the Broad sense – FLB) n'a rien de spécifique à l'homme. Elle couvre tous les aspects communicationnels, interactionnels, sociaux et culturels. Comme l'avait déjà noté Chomsky (1995) à la suite de Fodor (1983a), elle s'applique à tous les systèmes périphériques et couvre la phonétique, la phonologie, la morphologie, la sémantique et la pragmatique des langues. En bref, tout ce qui n'est pas la syntaxe. Et encore, la faculté de langage au sens étroit du terme (Faculty of Language in the Narrow sense – FLN) qui lui est opposée ne couvre pas toute la syntaxe, mais porte exclusivement sur la syntaxe noyau. Plus précisément, sur l'un de ses principes computationnels : la récurrence telle qu'elle s'exprime au travers du principe de FUSION (MERGE). Dans sa tentative de sauvegarde du cœur même de la grammaire, Newmeyer concède ainsi toute la FLB aux linguistiques de l'usage pour, avec Hauser, Chomsky et Fitch (2002), ne centrer la Biolinguistique que sur l'analyse de la FLN<sup>64</sup>.

Dans le débat entre linguistique du *datum* et linguistique de *l'exemplum*, entre modèles fondés sur l'usage et linguistique cartésienne, on en revient donc *in fine*, à la question cognitive.

---

<sup>64</sup> On ne peut suivre Newmeyer lorsque pour ce faire il s'autorise de Saussure. Comme je l'ai montré ailleurs avec beaucoup d'autres (Laks 2011b), et contrairement à la vulgate du *Cours*, Saussure distingue nettement la grammaire du linguiste et celle du locuteur. La seconde étant par principe impénétrable parce qu'elle constitue une formation de compromis de la variété des usages. Pour Saussure, définitivement, la Science du Langage est du côté de l'analyse des usages (ce qu'il appelle la linguistique de la parole) dans leurs dimensions sociales et culturelles (*i. e.* sémiologique).

## 4 Hétérogénéité et variation : la question cognitive

Comme on vient de le voir, avec le paradigme de la Biolinguistique, la linguistique chomskienne se donne un programme qui n'est nouveau qu'en apparence. Depuis le Programme minimaliste, la distinction entre FLB et FLN était en place ainsi que le centrage sur les capacités cognitives individuelles et les états représentationnels internes via la distinction entre langage-E et langage-I<sup>65</sup>. Ce qui est nouveau, c'est le centrage exclusif de la linguistique cartésienne sur le la FLN et le langage-I. Toutes les autres dimensions sont abandonnées aux modèles basés sur l'usage et aux linguistiques du *datum*. Le sociologue de la science ne pourra manquer d'y voir un effet des rapports de forces inter théoriques tels que Newmeyer en a fait le constat *supra*.

### 4.1 Innéisme, essentialisme et rationalisme

Avec ce recadrage de son programme, la Biolinguistique n'a donc plus à s'interroger que sur un objet extrêmement restreint et, en droit, n'a plus ni à prendre en compte ni à répondre, à d'autres arguments qu'à ceux portant sur celui-ci : la faculté de langage au sens étroit<sup>66</sup>. Mais le réductionnisme chomskien est encore plus radical : la FLN ne contient que les propriétés strictement spécifiques aux langues humaines et parmi celles-ci uniquement les propriétés qui ne peuvent être acquises par une exposition aux données externes. Ces propriétés spécifiques se réduisent en fait à une seule, le principe computationnel de récursion, et encore, est-il saisi sous la forme très spécifique du principe syntaxique de FUSION (MERGE)<sup>67</sup>. Tout le reste ressortit à la FLB, est donc extérieur à l'investigation biolinguistique, et est en conséquence abandonné aux linguistiques de l'usage. Ces dernières peuvent parfaitement être pensées dans un cadre darwinien et évolutionniste où l'apparition d'une forme supérieure et complexe de communication procure un avantage sélectif décisif à l'espèce. FUSION ne le peut pas. Pour Chomsky, cette opération en quoi se résume toute la syntaxe noyau ne procède pas d'une adaptation graduelle mais d'une rupture abrupte. Contre le *motto* darwinien classique, Berwick (2011) le souligne avec force : *natura [syntax] facit saltum* et

---

<sup>65</sup> "We are concerned, then with states of language faculty, which we understand to be some array of cognitive traits and capacities, a particular component of the human mind/brain. The language faculty has an initial state, genetically determined; in the normal course of development it passes through a series of states in early childhood, reaching a relatively stable steady state that undergoes little subsequent change (...) To a good first approximation, the initial state appears to be uniform for the species. (...) we call the theory of the state attained its grammar and the theory of the initial state universal grammar. (...) The initial state is in crucial respects a special characteristic of humans, with properties that appear to be unusual in the biological world. (...) When we say that Jones has the language L, we now mean that Jones's language faculty is in the state L (...) To distinguish this concept of language from others, let us refer to it as I-language, where I is to suggest 'internal', 'individual', and 'intentional'" Chomsky (1995 18-19).

<sup>66</sup> « We hypothesize that FLN only includes recursion and is the only uniquely human component of the faculty of language. We further argue that FLN may have evolved for reasons other than language, hence comparative studies might look for evidence of such computations outside of the domain of communication (for example, number, navigation, and social relations)». Hauser, Chomsky et Fitch (2002, 1569)

<sup>67</sup> "In this sense, there is no possibility of an 'intermediate' language between a non-combinatorial syntax and full natural language syntax — one either has Merge in all its generative glory, or one has effectively no combinatorial syntax at all, [...], in a sense there is only a single grammatical operation: Merge. Once Merge arose, the stage for natural language was set. There was no turning back." Berwick (2011, 99). Cf. également Berwick et Chomsky (2011), Chomsky (2004), Hauser, Chomsky et Fitch (2002) Chomsky (2011)

l'apparition de la syntaxe noyau est pure de toute détermination évolutionniste<sup>68</sup>. En cohérence avec le rationalisme cartésien, la Biolinguistique chomskienne pose ainsi FUSION comme *l'ultima ratio* en quoi réside la nature (restrictivement) humaine de l'homme.

En effet, si comme l'admettent Berwick et Chomsky (2011) le principe computationnel de récursion se retrouve dans différents domaines de la cognition animale et humaine (numération, organisation de scènes etc.)<sup>69</sup>, FUSION n'en découle pas directement mais en constitue une exaptation. Le terme n'est pas utilisé ici dans son sens originel (Gould et Vrba 1982) où la réutilisation pour de nouveaux objectifs d'une fonction cognitive existante était conduite par la pression adaptative et donc par la dynamique sélective, mais dans un sens purement discontinuiste et catastrophique : Chez l'homme, la fonction de récursion, active dans différents domaines cognitifs, se trouve soudain, et sans aucune raison, appliquée aux systèmes de communication existants. Cette rupture, non motivée, qui a lieu dans le cerveau d'un seul individu est décisive : la syntaxe et avec elle la grammaire, le langage humain et avec lui l'espèce homo, sont nés. L'avantage sélectif qui en résulte est considérable, mais ce n'est pas lui qui motive le processus<sup>70</sup>.

Ce point est décisif. En effet, si la syntaxe résulte seulement d'une évolution adaptative, s'il n'y a pas de rupture décisive, et si cette adaptation a lieu au sein du groupe pour en accroître l'organisation grégaire, sociale, culturelle et symbolique, alors c'est l'évolution du groupe qui en constitue le moteur. Pour le dire autrement, il s'agit alors d'une approche fonctionnaliste qui place la fonction de communication interpersonnelle au centre du dispositif et fait de l'apparition du langage humain un moment de l'évolution, conduit et motivé par l'avantage sélectif que procure une meilleure organisation sociale et une meilleure régulation symbolique du groupe. Dans cette hypothèse, FLB couvre l'entièreté de la scène et avec elle les linguistiques de l'usage. Il n'y a plus ni nécessité de FLN, ni paradigme rationaliste en tant que tel. De plus comme nous l'avons vu *supra*, si nous sommes du côté des usages et du *datum*, hétérogénéité et variation se retrouvent au cœur du dispositif, et même mieux, c'est l'existence de l'hétérogénéité et de la variation sociales qui impose un mode de gestion communicationnel de ces dernières. *Ergo*, comme le disait Saussure après Whitney, la langue est bien alors une institution sociale.

---

<sup>68</sup> “However, unlike Linnaeus' and Darwin's slogan shunning the possibility of discontinuous leaps in species and evolution generally—*natura non facit saltum* — we advocate a revised motto that turns the original on its head: *syntax facit saltum* — syntax makes leaps — in this case, because human language's syntactic phenotype follows from interactions amongst its deeper components, giving it a special character all its own, apparently unique in the biological world” Berwick (2011, 6)

<sup>69</sup> “Only FLN is uniquely human. [...] we hypothesize that most, if not all, of FLB is based on mechanisms shared with nonhuman animals. In contrast, we suggest that FLN—the computational mechanism of recursion—is recently evolved and unique to our species”. Hauser, Chomsky et Fitch (2002, 1572)

<sup>70</sup> “Unbounded Merge (hence displacement) must have arisen from some rewiring of the brain, hence in an individual, not a group. The individual so endowed would have had many advantages: capacities for complex thought, planning, interpretation, and so on. The capacity would be partially transmitted to offspring, and because of the selective advantages it confers, it might come to dominate a small breeding group, though as with all such novel mutations, there is an issue about how an initially small number of copies of such an allele might survive, despite a large selective advantage” Berwick et Chomsky (2011, 13).

Une telle conclusion doit à tout prix être réfutée. C'est pourquoi Hauser, Chomsky et Fitch (2002) qui ont parfaitement saisi cet enjeu s'emploient à récuser les arguments en faveur de l'hypothèse fonctionnaliste communicationnelle en contestant que les besoins croissants d'organisation, de coopération et d'échange au sein du groupe puissent motiver l'exaptation de la récursivité en FUSION. Mais les arguments en faveur d'une rupture de continuité ayant conduit à l'apparition catastrophique de la syntaxe noyau restent tous extrêmement discutables et même les espoirs mis dans la découverte du « gène du langage »(FOXP2) ont été déçus<sup>71</sup>. En définitive, comme le souligne Fodor (2001), l'argument chomskien est strictement de nature épistémologique. Critiquant ceux qui avaient cru pouvoir ouvrir une voie darwinienne à la Grammaire Générative en en s'appuyant précisément sur une lecture fonctionnaliste de la communication humaine et de ses avantages adaptatifs (Cf. par exemple Pinker 1997, Pinker et Bloom 1989), Fodor réaffirme que dans un cadre rationaliste bien compris, ceci ne peut être défendu et que c'est en définitive avec Platon que Chomsky argumente en défendant le caractère inné de la faculté de langage<sup>72</sup>. Le langage n'a définitivement rien à voir avec la communication interpersonnelle et tout avec la théorie nativiste de la connaissance, au moins pour ce qui concerne son mécanisme central.

Dans l'approche biolinguistique, l'universalité de FUSION tout comme l'impossibilité de la dériver de la fonction de communication par une évolution darwinienne graduelle constituent des stipulations axiomatiques. Comme on vient de le voir, si les principes de la syntaxe noyau pouvaient être dérivés de la communication interpersonnelle, toute la construction biolinguistique en serait définitivement ruinée. Il en est de même de son caractère universel. On notera néanmoins, que si le caractère d'axiomes posés a priori ne peut être démontré pour ces deux thèses, chacune peut se trouver empiriquement contestée. De fait, de nombreuses approches ont proposé de dériver les principes cognitifs de la syntaxe noyau à partir de fonctions cognitives non spécifiques au domaine linguistique, actives dans d'autres domaines de l'intelligence humaine ou des relations

---

<sup>71</sup> « If so, then the entire FOXP2 story, and motor externalization generally, is even further removed from the picture of core syntax/semantics evolution. [...] Summarizing, FOXP2 does not speak to the question of the core faculty of human language because it really has nothing to do with the core language phenotype, Merge and syntax. [...] To be sure, FOXP2 remains a possibly necessary component of the 'language system' [...] But it is not human language *tout court*. If all this is so, then the explanation 'for' the core language phenotype may be even more indirect and difficult than Richard Lewontin (1998) has sketched" Berwick et Chomsky (2011, 12-13)

<sup>72</sup> "Chomsky's ideas about innateness would have been intelligible to Plato; and they would have been intelligible in much the terms of the present debate. This is because Chomsky's nativism is primarily a thesis about knowledge and belief; it aligns problems in the theory of language with those in the theory of knowledge. Indeed, as often as not, the vocabulary in which Chomsky frames linguistic issues is explicitly epistemological. [...] much of the knowledge that linguistic competence depends on is available to the child a priori (i.e., prior to learning). [...] it is, to repeat, primarily epistemological nativism that Chomsky shares with the rationalists. When Plato asks what the slave boy knows about geometry, and where on earth he could have learned it, it really is much the same question that Chomsky asks about what speaker/hearers know about their language and where on earth they could have learned that »Fodor (2001, 10)

interpersonnelles de communication<sup>73</sup>. Contre ces hypothèses fonctionnalistes et communicationnelles, on vient de rappeler le radicalisme de la réponse chomskienne.

Cette réponse est encore plus ferme, s'il se peut, lorsque l'universalité de la récursivité des langues est mise en cause. La polémique à propos de la non récursivité du Pirahã en est un bon exemple. Everett (2005) avait mis en cause l'universalité linguistique de la récursion à partir de l'analyse de cette langue amérindienne dont il est l'un des spécialistes. Il en tirait des conséquences générales sur les relations entre langue et culture qui invalidaient la position biolinguistique<sup>74</sup>. Le caractère dirimant de cette critique n'avait pas échappé aux commentateurs.<sup>75</sup> Nevins, Pesetsky et Rodrigues (2009) se sont donc employés à démontrer que l'analyse syntaxique des faits exhibés par Everett ne remettait nullement en cause l'analyse de Hauser, Chomsky et Fitch (2002) concernant les rapports entre langue et culture, pour autant que l'on pose une caractérisation suffisamment abstraite et générale de FUSION<sup>76</sup>. Dans sa réponse détaillée, Everett (2009, 439) souligne à juste titre que ceci revient à restreindre à nouveau le caractère falsifiable des principes proposés pour la grammaire universelle car « this [more abstract] version of MERGE can neither be supported nor criticized by facts because it is definitional and therefore not falsifiable ». On en conclut que pour la Biolinguistique et le rationalisme chomskien, le caractère irréductiblement spécifique au langage des principes de la syntaxe noyau, leur absence de corrélation à toute dimension sociale, culturelle ou informationnelle et leur nativisme constituent les prémisses non contestables d'une approche fondamentalement essentialiste pour ce qui concerne la nature de l'homme, approche qui suppose une discontinuité évolutive et est, de ce point de vue, proche du créationnisme métaphysique.

---

<sup>73</sup> Sur la question de l'origine des langues Cf. Laks, Cleuziou, Demoule et Encrev  (2007). Cf. également pour des approches différentes Bickerton (1998), Bickerton et Calvin (2000), Dessalles (2000), Dessalles, Picq et Victorri (2006), Pinker (1997), Pinker (1999).

<sup>74</sup> “[My] article also offers more detailed argumentation for the hypothesis that culture can exert an architectonic effect on grammar. It concludes that Pirah  falsifies the single prediction made by Hauser, Chomsky, and Fitch (2002) that recursion is the essential property of human language”. Everett (2009, 455)

<sup>75</sup> Dans leur défense de la biolinguistique chomskienne, Nevins, Pesetsky et Rodrigues (2009, 671), écrivent ainsi : « The New Scientist (March 18, 2006) suggested, for example, that Pirah  might constitute ‘the final nail in the coffin for Noam Chomsky’s hugely influential theory of universal grammar’; and the Chicago Tribune (June 10, 2007), under the headline ‘Shaking language to the core’, reported that Everett had ‘fired a volley straight at the theory when he reported that the Brazilian tribe he was studying didn’t use recursives [sic]’. More recently, the Times of London (October 24, 2008) has characterized Everett’s claim that ‘Pirah  lack the grammatical principle of recursion’ as an ‘astonishing find’. If the conclusions in NP&R are correct, of course, Pirah  presents us with no nail, no coffin, no volley, and no astonishing find.”

<sup>76</sup> “As NP&R (n. 11) pointed out, no construction in a given language (be it English or Pirah ) constitutes a demonstration of recursion or its absence independent of the analysis that this construction receives in the context of a particular theory. Hauser, Chomsky, and Fitch (2002) presupposed, rightly or wrongly, an approach to syntactic structure in which all phrase structure—not just clausal embedding or possessor recursion—serves as a demonstration of recursion. We had this in mind when we noted in NP&R that if Pirah  really were a language whose fundamental rule is a non recursive variant of Merge, no sentence in Pirah  could contain more than two words” Nevins, Pesetsky et Rodrigues (2009, 679). Je souligne.

## 4.2 Du non apprentissage des langues

Il n'a pas été suffisamment remarqué que depuis le tournant cartésien des années 1965, la linguistique chomskyenne se présentait comme une théorie du non apprentissage des langues. Cette thèse est encore renforcée avec l'approche Principes et Paramètres et le Programme Minimaliste (Chomsky 1995). L'apprentissage s'y trouve réduit à un réglage paramétrique des principes de la grammaire universelle (GU). Ce réglage est assuré par un dispositif d'acquisition du langage (DAL) qui confronte GU aux données obvies afin de produire une grammaire particulière (GP). L'apprentissage est donc modélisé par la formule bien connue :

$$(GU \times DAL) (data) = GP$$

Dans cette formule, deux des trois facteurs sont considérés comme innés et sont donc donnés au départ. Il s'ensuit que l'apprentissage linguistique proprement dit ne correspond à rien de plus qu'à une spécification locale de mécanismes génétiquement inscrits dans l'esprit du locuteur. Cette thèse, typique du rationalisme cartésien qui voit la logique et les principes de la grammaire universelle comme immanents (Arnauld et Lancelot 1660), était déjà celle de Platon lorsqu'il démontrait à Ménon que son esclave, connaissait déjà parfaitement, et de manière innée, la géométrie du carré<sup>77</sup>.

Dans le corpus de la Grammaire Générative, cette thèse nativiste n'est pourtant jamais présentée comme une prise de position philosophique mais comme la conclusion d'un raisonnement empirique. Ce raisonnement, exactement comme celui de Platon d'ailleurs, est fondé sur la supposée pauvreté du stimulus accessible à l'enfant au cours de son apprentissage<sup>78</sup>. Compte tenu de son importance centrale pour l'équilibre de la théorie chomskyenne, on s'attend à ce qu'elle ait suscité de nombreuses études fondées sur des observations empiriques de longue durée et qu'elle ait été solidement démontrée. Or, dans le cadre chomskyen, force est de constater qu'il n'en a rien été. La thèse de la pauvreté du stimulus, constitue l'un des items les plus récurrents de la littérature générative et pourtant l'un des moins empiriquement établi dans ce cadre.

Tel n'est pas le cas dans le cadre des modèles fondés sur l'usage. En linguistique du *datum*, de très nombreux corpus, particulièrement volumineux, ont été spécialement construits pour tester l'hypothèse de la pauvreté du stimulus. Les analyses quantitatives et qualitatives avec suivi longitudinal régulier d'enfants en apprentissage sont nombreuses. Pour de très nombreuses langues

<sup>77</sup> Ménon interroge Socrate : «Mais qu'est-ce qui te fait dire que nous n'apprenons pas et que ce que nous appelons le savoir est une réminiscence ? ». Après sa leçon de géométrie au jeune esclave, Socrate conclut : « C'est donc que ces opinions se trouvaient déjà en lui. N'est-ce pas vrai? [...] S'il ne les a pas acquises dans la vie présente, il faut bien qu'il les ait eues dans un autre temps et qu'il s'en trouvât pourvu d'avance. » Platon (1849 Trad. Victor Cousin, 80d-86c)

<sup>78</sup>« Likewise, the central problem of language acquisition arises from the poverty of the "primary linguistic data" from which the child effects this construction; and the proposed solution of the problem is that much of the knowledge that linguistic competence depends on is available to the child a priori (i.e., prior to learning) Fodor (2001, 10).

différentes, les productions d'enfants du babillage précoce à jusqu'à une compétence stabilisée, y sont très bien documentées<sup>79</sup> (Cf. Supra le programme coopératif CHILDES). Or ces analyses conduisent précisément à contester radicalement l'hypothèse chomskyenne. Newmeyer s'en faisait justement l'écho *supra*. Le domaine extrêmement dynamique de la psycholinguistique de l'acquisition propose ainsi de très nombreuses réfutations empiriques de l'impossibilité d'acquérir telle ou telle fonction syntaxique sur la base seulement des données disponibles dans l'environnement<sup>80</sup>.

Les données linguistiques auxquelles un locuteur est quotidiennement confronté sont nombreuses et riches de tout un contexte social, culturel et interactionnel qui vient les renforcer et les sanctionner<sup>81</sup>. Mehl, Vazire, Ramirez-Esparza, Statcher et Pennebaker (2007) ont ainsi estimé à 16 000 mots environ la production journalière d'un locuteur, avec bien entendu une large dispersion autour de cette moyenne. Quant à l'apprentissage linguistique natif, Morgan (1989, 352) a estimé qu'un enfant acquérait sa langue après avoir été confronté à 4 280 000 phrases environ. De tels ordres de grandeurs suggèrent immédiatement de tester des modèles statistiques et probabilistiques pour rendre compte de l'acquisition. De fait, de nombreuses analyses utilisant des techniques bayesiennes, neurocomputationnelles ou d'autres outils stochastiques sont venues contester, phénomène syntaxique par phénomène syntaxique, la thèse de la pauvreté du stimulus et récuser de façon convaincante l'approche nativiste<sup>82</sup>.

L'approche nativiste de la Grammaire Générative ne se fonde pas uniquement sur l'argument de la pauvreté du stimulus dont je viens de monter qu'il est particulièrement sujet à caution. Un second argument, formel cette fois, est cité à l'appui de la nécessité pour l'enfant apprenant une langue de disposer d'un riche répertoire de connaissances linguistiques *a priori*. Il s'agit du théorème de Gold (1967) que j'ai déjà évoqué. Comme c'était déjà le cas avec la thèse de la pauvreté du stimulus, le théorème de Gold est souvent cité comme une preuve définitive dans la littérature générative mais très rarement commenté ou analysé dans sa portée pratique. Comme le souligne Johnson (2004), la préoccupation de Gold était fort éloignée de tout propos linguistique et a fortiori

<sup>79</sup> Comme l'écrivait le jeune Chomsky dans sa critique de Skinner « The manner in which factors operate and interact in language acquisition is completely unknown. It is clear that what is necessary in such a case is research, not dogmatic and perfectly arbitrary claims, based on analogies to that small part of the experimental literature in which one happens to be interested. » Chomsky (1959, 43). Pour une revue des arguments empiriques contre la thèse de la pauvreté du stimulus Cf. par exemple Pullum et Scholz (2002), Sampson (2002). Pour une défense de cette thèse Cf. Berwick, Pietroski, Yankama et Chomsky (2011)

<sup>80</sup> De très nombreuses fonctions syntaxiques, coordination, subordination, inversion du sujet etc. sont analysées dans cette perspective. Cf. pour une présentation Parisse (2005)

<sup>81</sup> Ce que Mufewene (2001) a justement appelé l'écologie du langage.

<sup>82</sup> Cf par exemple Elman et Lewis (2001) pour une approche neurocomputationnelle, Perfors, Tenenbaum et Regier (2006) pour une analyse bayésienne du placement de l'auxiliaire dans les interrogatives, Foraker, Regier, Khetarpal, Perfors et Tenenbaum (2009) pour une analyse bayésienne des anaphores, Reali et Christiansen (2005) pour une analyse statistique du placement des auxiliaires dans les questions polaires.

de toute hypothèse sur l'apprentissage humain. Le théorème de Gold est une démonstration strictement mathématique dans le cadre de la théorie générale des langages formels. Gold ne se pose aucun problème d'acquisition ou de sélection de grammaire par un locuteur. Comme le titre l'indique, son théorème porte exclusivement sur l'identification d'une grammaire formelle parmi toutes celles appartenant à une classe générant un même ensemble de suites symboles. Après avoir très précisément critiqué chacune des interprétations du théorème de Gold et montré leur caractère largement interprétatif, Johnson conclut à son absence de pertinence dans le débat cognitif<sup>83</sup>. Il s'ensuit que le théorème de Gold n'apporte aucun argument à la théorie de la pauvreté du stimulus et laisse la thèse innéiste pour ce qu'elle est : une construction de nature uniquement épistémologique tout à fait liée à l'orientation rationaliste de la linguistique de *l'exemplum*.

Mais, si le théorème de Gold ne permet pas de fonder la non apprenabilité des grammaires des langues humaines, il reste qu'étant donné un corpus d'occurrences linguistiques, il n'existe pas de solution unique pour en construire un modèle formel. Dans le cadre du néo empirisme qu'il défend, Goldsmith (2010) s'est appuyé sur les travaux de Marcken (1996) concernant l'apprentissage automatique non supervisé. En distinguant nettement le terrain de l'argumentation cognitive et celui des modèles formels de l'apprentissage linguistique, il a démontré qu'une application du principe statistique de la « Longueur Minimale de Description (Minimum Description Length -MDL, Rissanen 2007) permettait de faire converger et d'optimaliser un dispositif d'apprentissage automatique de la morphologie (Goldsmith 2005a, 2011), sans aucune connaissances a priori. Précisément parce qu'il rejette l'interprétation métaphorique que nous avons vu à l'œuvre dans l'exploitation du théorème de Gold, Goldsmith souligne à juste titre que l'interprétation cognitive de MDL et des résultats obtenus en apprentissage automatique non supervisé reste une question théorique ouverte.

#### 4.3 Les modèles basés sur l'usage et les grammaires de construction : un modèle cognitif alternatif

Comme je l'ai déjà dit, avec les grammaires basées sur l'usage, la question cognitive peut être pensée dans un nouveau cadre. Les questions concernant la variation et l'hétérogénéité que nous avons posées y trouvent toute leur place. Depuis une trentaine d'années, le champ de l'anthropologie cognitive a connu des développements spectaculaires. La collaboration de

<sup>83</sup> “In fact, as long as the notion of identifiability in the limit from any environment has no obvious psychological interpretation, there is little of psychological interest to be concluded from Gold's Theorem [...] Despite its simplicity, many authors have taken Gold's Theorem to threaten some fundamental views about the mind, and they have responded with various criticisms. However many of these attacks are misguided for largely formal reasons. But a look at the details shows that Gold's Theorem is still of questionable direct relevance to cognitive science.” Johnson (2004 587)

neuropsychologues, sociologues, anthropologues, ethnologues, linguistes et psychologues a permis des avancées décisives dans la compréhension de la genèse et de l'évolution des systèmes symboliques, culturels et sociaux<sup>84</sup>. L'analyse comparative de ces systèmes, tels qu'ils existent dans le règne animal et chez l'homme, tout comme l'analyse de leur développement phylogénétique et de leur maturation dans les groupes humains, a permis de dégager des conclusions solides dans un champ qui est majoritairement dominé par ce que Changeux (1983) a appelé un néo-darwinisme neuronal. Les questions posées par le fonctionnement des systèmes linguistiques, culturels et sociaux et par leur reproductibilité intergénérationnelle, avaient déjà été abordées dans une perspective évolutionniste et néo darwinienne par Dawkins (1976) qui proposait de les traiter métaphoriquement comme des allèles d'un gène spécifique nouveau apparu avec l'homme moderne.

Ces débats et ces propositions sont bien entendu de première importance pour une perspective de recherche comme celle de la Biolinguistique, mais ils sont en très grande part restés extérieurs à la réflexion des linguistes générativistes. Le domaine de l'anthropologie cognitive est effet très largement dominé par un paradigme avec lequel nous venons de voir qu'elle est incompatible. Dans sa dimension fonctionnaliste ce paradigme recherche dans l'analyse des fonctions remplies par un dispositif les motivations de son existence. Dans sa dimension interactionniste et communicationnelle il accorde à la communication et aux relations interpersonnelles une place centrale. Or, comme je l'ai déjà signalé ci-dessus, Hauser, Chomsky et Fitch (2002), Berwick et Chomsky (2011), Di Sciullo et Boeckx (2011), et avec eux linguistes inscrits dans le courant biolinguistique, mettent beaucoup de soin à réfuter toute approche fonctionnelle en matière de langue et à nier que la communication soit en rien impliquée dans la genèse et le développement de la faculté de langage spécifique à l'homme comme je l'ai déjà dit. Chomsky (2007) est très clair sur ce point : la fonction de récursion propre à FLN est d'abord totalement endogène, créant un langage interne de la pensée qui ne s'externalise comme support de la communication interpersonnelle que de façon totalement secondaire et accessoire<sup>85</sup>. Les approches fonctionnelles et communicationnelles sont donc récusées *ex definitio*. Au total, si, comme c'est généralement le cas en anthropologie cognitive, on attribue à la fonction d'organisation grégaire et à la fonction de communication régulatrice du groupe un statut de motivation première, on quitte *de facto* le cadre

---

<sup>84</sup> On trouvera dans le volume issu d'un symposium interdisciplinaire coordonné par la Fondation Fyssen pour l'anthropologie un état des lieux récent : Levinsonet Jaisson (2006)

<sup>85</sup> "Emergence of unbounded Merge in human evolutionary history provides what has been called a "language of thought," an internal generative system that constructs thoughts of arbitrary richness and complexity, exploiting conceptual resources that are already available or may develop with the availability of structured expressions. If the relation to the interfaces is asymmetric, as seems to be the case, then unbounded Merge provides only a language of thought, and the basis for ancillary processes of externalization. [...] The capacity would be transmitted to offspring, coming to dominate a small breeding group. At that stage, there would be an advantage to externalization, so the capacity would be linked as a secondary process to the sensorimotor system for externalization and interaction, including communication". Chomsky (2007,22, 23)

rationaliste pour rejoindre celui de la linguistique descriptive du *datum*. La variété et l'hétérogénéité des systèmes se retrouvent alors sur le devant de la scène et une systématique analytique est dès lors convoquée pour rendre compte de leur extension et de leur variété<sup>86</sup>. On explique ainsi pourquoi les linguistiques fondées sur l'usage et les psychologues reprenant ce cadre, se sont trouvés particulièrement impliqués dans le développement récent du champ de l'anthropologie cognitive, et aussi pourquoi ceci n'a pas été le cas des linguistes générativistes.

C'est dans ce cadre d'une anthropologie cognitive généralisée que Tomasello élabore depuis une quinzaine d'années un modèle linguistique et cognitif s'inscrivant dans ligne générale de l'évolution telle qu'elle est vue par les néo darwiniens en s'appuyant sur les résultats de la linguistique des usages<sup>87</sup>. Il s'inscrit ainsi explicitement dans une linguistique du *datum*<sup>88</sup>

Comme de très nombreux anthropologues cognitivistes l'on souligné, la communication verbale interpersonnelle et l'élaboration de formes très sophistiquées de comportements culturellement et socialement régulés procurent à l'espèce humaine un avantage sélectif décisif. Mais cette aptitude communicationnelle n'est pas propre à l'espèce humaine. On en trouve de nombreux prodromes chez tous les animaux sociaux, depuis les plus éloignés comme les insectes jusqu'aux plus proches de l'homme comme les primates et les grands singes. L'altruisme réciproque (Trivers 1971, 2002) tel qu'il s'illustre par exemple dans les stratégies complexes d'épouillage ou le partage de ressources a été proposé comme l'un des ressorts des régulations sociales dont la complexité va croissant avec l'évolution. Avec ce qu'il suppose de calcul anticipatif, d'ajustement des comportements et de stratégies individuelles au sein des groupes, un certain nombre de sociobiologistes y ont vu l'origine de comportements sociaux complexes propres à l'homme<sup>89</sup>. Pour

---

<sup>86</sup> Positions que Chomsky (2007, 1) attribue explicitement aux behavioristes skinneriens et plus spécialement aux structuralistes (américains et européens) soucieux avant tout de corpus et de méthodes de recueil de donnée, au premier rang desquels il cite Harris et son ouvrage *Methods in structural linguistics*. Langacker ne recuse pas ce constat lorsqu'il note : "language has two basic and closely related functions: a semiological function, allowing thoughts to be symbolized by means of sounds, gestures, or writing, as well as an interactive function, embracing communication, expressiveness, manipulation and social communion. A pivotal issue in linguistic theory is whether the functions language serves should be taken as foundational or merely subsidiary to the problem of describing its form. The recognition of their foundational status is the primary feature distinguishing functionalist approaches to language from the formalist tradition (notably generative grammar)". Langacker (1998, 1)

<sup>87</sup> Dans les deux volumes qu'il a réuni sous le titre suffisamment explicite « The new psychology of language : cognitive and functional approaches to linguistic structure » Tomasello (1998), Tomasello (2008a) présente ainsi les contributions de linguistes travaillant dans le cadre des modèles basés sur l'usage dont beaucoup participèrent aux débats de la sémantique générative et des linguistiques qui en sont issues : Langacker, Talmy, Fillmore, Fauconnier, Givon, mais aussi Croft, Hopper Bybee, Goldberg, Haspelmath, Van Valin etc., ce qui identifie assez bien le courant linguistique sur lequel il s'appuie.

<sup>88</sup> « In diametric opposition to [generative] methodological assumptions, cognitive-functional linguists take as their object of study all aspects of natural language understanding and use, including unruly idioms, metaphors, and irregularities. They [...] take as an important part of their data not disembodied sentences derived from introspection, but rather utterances or other longer sequences from naturally occurring discourse ». Tomasello (2008a, XII)

<sup>89</sup> De la question des « tricheurs » naît ainsi une interrogation sur la dissymétrie entre bénéfices et obligations. On trouvera des indications sur les développements de l'altruisme réciproque en sociologie morale et politique, en théorie des jeux et jusqu'en modélisation mathématique des comportement et des marchés par exemple dans Clavien (2010)

de nombreux anthropologues cependant, cette dynamique n'est pas suffisante pour expliquer la rupture de continuité dans laquelle s'inscrit l'homme.

Tout en accordant au développement de l'altruisme réciproque un rôle très important dans l'évolution vers l'homme, Tomasello construit son analyse sur la rupture de continuité fondamentale qu'il introduit dans les systèmes sociaux et les modes de communication régulatrice impliqués, déjà très sophistiqués, l'apparition chez l'homme d'une attribution d'intentionnalité<sup>90</sup>. Cette attribution d'intentionnalité modifie radicalement les fonctionnalités de l'altruisme réciproque. Elle est spécifique à l'espèce humaine et est motivée par le renforcement de la fonction grégaire<sup>91</sup>. Du point de vue ontogénétique, Tomasello en voit le substrat dans les activités de pointage, de pantomime et de mimique imitative du jeune enfant, comme dans les activités associées de deixis partagée qui règlent la dénomination commune des objets. La « lecture des intentions », fonctionnalité nouvelle qui émerge entre 9 et 12 mois chez l'enfant lui permet de se construire une véritable théorie de l'esprit, base d'une pensée abstraite (Tomasello 2003, 3). C'est sur cette base, que la communication est analysée comme une fonction qui permet à une personne de manipuler symboliquement les états intentionnels et mentaux des personnes avec lesquelles elle interagit. Le langage est alors conçu comme « un inventaire structuré de symboles » (Langacker 1998, 1) assumant cette fonction de manipulation des états mentaux. La double orientation, communicative et fonctionnelle, de ces modalités est donc très claire. Leur ancrage dans la structuration sociale et l'organisation culturelle et symbolique des groupes humains est, de ce point de vue, fondamental<sup>92</sup>

L'attribution d'intentionnalité et la lecture des actions, des comportements et des événements sous ce rapport est en elle-même génératrice d'une forme de pensée abstraite parce qu'elle se focalise d'emblée sur leur signification pratique dans la relation interpersonnelle. Confronté à la très grande variabilité des actes et des objets, l'esprit humain développe ainsi une habileté particulière pour la reconnaissance de similarités partielles. La recherche de patrons comparables, tant au plan perceptif que conceptuel, débouche sur un développement très

---

<sup>90</sup> « Specifically, human cooperation is structured by what some modern philosophers of action call shared intentionality or "we" intentionality. In general, shared intentionality is that is necessary for engaging in uniquely human forms of collaborative activity in which a plural subject "we" is involved: joint intentions, mutual knowledge, shared beliefs—all in the context of various cooperative motives. Tomasello (2008b, 7)

<sup>91</sup> "specifically, human beings cooperate with one another in species-unique ways involving processes of shared intentionality [... ]This fundamentally cooperative process makes human communication utterly different from the communicative activities of all other species on the planet". Tomasello (2008b, 72, 99)

<sup>92</sup> "At some point in human evolution, Homo Sapiens evolved the ability to communicate with another symbolically. [...] These transformations of linguistic structure occur as a result of social-interactive processes" Tomasello (2008b), "And what about language? The current hypothesis is that it is only within the context of collaborative activities in which participants share intentions and attention, coordinated by natural forms of gestural communication, that arbitrary linguistic conventions could have come to existence evolutionarily [...] this perspective on human communication and language thus basically turns the Chomskian proposal on its head, as the most fundamental aspects of human communication are seen as biological adaptations for cooperation and social interaction in general, whereas the more purely linguistic, including grammatical, dimensions of language are culturally constructed and passed along by individual linguistic communities." Tomasello (2008b, 9, 11, 163)

spectaculaire de la capacité à catégoriser, c'est-à-dire à schématiser, à extraire des formes récurrentes abstraites, et à rassembler en classes d'équivalence des objets ou des actes partiellement dissemblables, mais équivalents du point de vue fonctionnel. Cette dynamique ascendante d'abstraction et de conceptualisation va de pair avec une très grande sensibilité statistique et probabilistique à la récurrence du même, à la reconnaissance de patrons distributionnels complexes, au-delà des similitudes et des différences, et ceci sur de très larges séquences temporelles. On voit ainsi que très loin d'être anomale, dans ce schéma général, l'existence d'une hétérogénéité structurée et d'une variation intrinsèque des objets et des actions est en fait ce qui conduit et motive le travail ascendant d'abstraction, de schématisation et de catégorisation. En effet, si il faut interpréter des occurrences variables et hétérogènes pour reconstruire l'intentionalité pratique qui leur est sous-jacente, il devient nécessaire d'analyser les similitudes et les différences partielles afin de les classer comme équivalentes en construisant un système de catégories plus ou moins abstraites. On peut souligner qu'il s'agit très précisément de la thèse taxinomique de la systématique que j'ai évoquée ci-dessus en présentant les sciences du *datum*.

Dans cette approche, les fonctions cognitives impliquées, contrairement à la thèse de Fodor (1983b) et de Chomsky (1986), n'ont donc rien de spécifique. Sensibilité aux similitudes partielles, extraction de schémas, catégorisation, routinisation, sensibilité statistique correspondent à des fonctions très générales de la cognition humaine qui trouvent à s'appliquer dans le domaine des interactions verbales mais ne leur sont pas particulières<sup>93</sup>. Hors l'attribution d'intentionalité et les dimensions sociales et culturelles qu'elle introduit, le domaine du langage n'est donc en rien spécifique. S'y appliquent aux relations interpersonnelles les mêmes fonctions cognitives supérieures qu'ailleurs. La discontinuité dans l'évolution est donc ici vue comme ayant sa source dans la lecture des états mentaux des alter ego et dans le changement complet que cela introduit dans les relations interpersonnelles, leur gestion pratique, leur organisation routinisée et la catégorisation conceptuelle abstraite que leur variabilité de surface conduit à mettre en œuvre pour les traiter du point de vue cognitif.

Comme on le voit, il s'agit d'une approche parfaitement neoempiriciste pour laquelle la pratique communicative est première par rapport aux codes qui finissent par la réguler et l'organiser. Ceci est cohérent avec les sciences du *datum* et les linguistiques de l'usage qui voient le système de la langue comme second par rapport aux pratiques langagières. Parce que la langue est abordée comme une modalité sociale et culturelle routinisée, la grammaire qui en constitue la systématique

---

<sup>93</sup> On aura reconnu ici la thèse défendue par les linguistiques cognitives et les linguistiques des usages (Langacker 1987, 1991), Fauconnier (1997), Lakoff et Johnson (1999). Bien évidemment ce n'est pas un hasard.

est vue comme un sous-produit de l'activité communicationnelle. La grammaire n'est alors que le produit social et culturel d'une dynamique de grammaticalisation progressive des actes de langage. A l'inverse des approches générativistes où elle est conçue comme une condition *ab origine* de la langue, dans cette approche, comme tous les phénomènes culturels et sociaux, la grammaire est vue comme le produit historique d'une activité sociale<sup>94</sup>.

Il en est de même des catégories et des fonctions grammaticales qui, loin de préexister à leurs instantiation, comme dans les approches cartésiennes, ne sont que des construction cognitives progressives, des taxons et des super taxons, permettant de regrouper en un ensemble unique des occurrences variées et variables. Ainsi, comme le souligne par exemple Langacker (1991), la catégorie « verbe » ne préexiste pas à ses actualisations. Ce qui constitue un verbe dans une langue donnée, n'est autre que l'ensemble des éléments qui se comportent régulièrement comme tels, et qui sont donc progressivement cognitivement et mémoriellement rassemblés sous une catégorie abstraite définie et construite à partir des fonctionnalités concrètes, hétérogènes et variables, qui se manifestent en discours. Comme je l'ai souligné, c'est la perception de similitudes fonctionnelles en même temps que de différences de formes, la perception de l'hétérogénéité et de la variation donc, qui motive ce travail d'abstraction taxinomique de construction de taxons et de super taxons, ce que les grammairiens appellent des catégories<sup>95</sup>. On comprend dès lors mieux pourquoi Tomasello (2003) construit son approche ontogénétique et phylogénétique en s'appuyant directement sur la linguistique des usages et l'on voit que l'opposition entre linguistique du *datum* et de *l'exemplum*, linguistique cartésienne et linguistique des usages, en définitive entre rationalisme et empiricisme s'applique à tous les niveaux de l'analyse linguistique et cognitive. La variabilité et l'hétérogénéité des faits de langue, niée ou affirmée comme fondamentale, occupent dans ces constructions théoriques une place centrale.

On voit également pourquoi j'ai pu dire (Laks 1996) que ces approches relèvent d'un neo structuralisme ou la taxinomie et le dégagement progressif d'une systématique explicative s'érigent sur une analyse des données observables. Le corpus, pour peu qu'il soit suffisamment étendu et représentatif, est riche. La taxinomie qui en émerge directement et l'analyse statistique de son

---

<sup>94</sup> « If grammatical structure do not come directly from the human genome, as above-reported data suggest they do not, and if children do not invent them de novo, as they clearly cannot, then it is legitimated to ask, Where do grammatical structures come from? The answer is that, in the first instance they come from processes of grammaticalization in language history. [...] "Even so, grammaticalization by itself is not enough because, it does not account for the abstractness of linguistic structures. [...] children make this contribution in more extended developmental processes in which they apply their general cognitive, social-cognitive, and vocal-auditory processing skills to the historical products of grammaticalization." Tomasello (2008b, 163)

<sup>95</sup> On retrouve ici un débat qui traverse toute l'histoire de la grammaire et de la philosophie : les catégories de l'entendement (e.g. les catégories grammaticales) sont-elles données a priori ou construites au travers de l'expérience cognitive. Ce débat a récemment été réactivé sur le terrain des mathématiques : tandis que Connes défend une position platonicienne sur le caractère a priori des éléments et des lois mathématiques que le chercheur ne fait que (re)découvrir, Changeux défend qu'il s'agit toujours de constructions mentales qui naissent du rapport entre les fonctions cognitives supérieures et les données de l'expérience et du contexte. Cf. Changeux et Connes (1989)

organisation informe la théorie que l'on peut en construire. Dans cette perspective, la linguistique est significativement une méthode de traitement des données telle qu' Harris (1951) par exemple l'a théorisée et systématisée. La relation entre analyse linguistique et analyse cognitive n'y est pas, comme en grammaire générative, posée un comme un *a priori* théétique<sup>96</sup>, mais résulte de la convergence entre ce que la psycholinguistique cognitive construit indépendamment comme schéma fonctionnel et acquisitionnel d'une part et ce que la structurale propose comme systématique du corpus de l'autre. Comme je l'ai rappelé Laks (2011b) contre toutes les vulgates du Cours de Linguistique Générale, telle était exactement la position de Saussure<sup>97</sup>

## 5 Conclusion : Variatio omnibus

La variation occupe un place centrale dans la pensée de Darwin qui ouvre ainsi son ouvrage de référence Darwin (1859) sur un premier chapitre entièrement consacré à la variation, son origine, ses effets et ses mécanismes. Certes, comme le rappelle Hoquet (2009), Darwin ne disposait pas encore à son époque d'un modèle génétique totalement articulé qui puisse lui permettre d'analyser en détails le mécanisme de génération aléatoire de diversité dont il décrit les effets. Cette avancée décisive sera accomplie quelques années plus tard par Wallace. Mais Mayr (1982) rappelle dans son histoire de la pensée biologique en quoi et comment les trois moments postulés par Darwin, variation, sélection, et héritage restent centraux pour la génétique et la théorie moderne de l'évolution. On sait aujourd'hui avec une grande précision comment fonctionne ce générateur aléatoire de diversité : la reproduction des organismes vivants est imparfaite, la recopie et la recombinaison du génome introduit à chaque génération des différences aléatoires. Avec la reproduction sexuée, ces imperfections sont d'autant plus importantes que deux génomes sont recombinés. L'évolution s'appuie directement sur cette variabilité en sélectionnant certains des caractères aléatoirement produits par le générateur. Sont ainsi sélectionnés, et reproduits, ceux qui procurent à l'individu et à sa parentèle un avantage adaptatif décisif. On comprend alors qu'au cours de l'évolution la diversité et la variabilité aillent croissants, l'espèce humaine étant elle-même marquée par une variabilité de très grande ampleur, sous tous les rapports possibles.

<sup>96</sup> « La grammaire est un objet systématiquement ambigu qui désigne à la fois l'objet mental construit par l'enfant qui apprend sa langue maternelle et l'objet abstrait construit par le linguiste pour rendre compte de cet apprentissage : "Nous employons le terme "grammaire" pour désigner à la fois le système de règles représenté dans le cerveau du locuteur-auditeur , système acquis normalement dans la petite enfance et utilisé dans la production et l'interprétation des énoncés, et la théorie que le linguiste construit en tant qu'hypothèse concernant la grammaire intériorisée réelle du locuteur-auditeur ». (Chomsky Halle 1968, 26).

<sup>97</sup> « La langue ne peut pas procéder comme le grammairien, elle a un autre point de vue et les mêmes éléments ne lui sont pas donnés, elle fait ce qui par le grammairien est considéré comme des erreurs mais qui n'en sont pas, car il n'y a de sanctionné par la langue que ce qui est immédiatement reconnu par elle. [...] Entre l'analyse subjective des sujets parlants eux-mêmes (qui seule importe!) et l'analyse objective des grammairiens, il n'y a donc aucune correspondance, quoiqu'elles soient fondées toutes deux en définitive sur la même méthode (confrontation de séries) Saussure (1916 édition Engler 1968, 2759)

Si j'ai tenu à rappeler la place centrale qu'occupe la variabilité dans la théorie moderne de l'évolution c'est pour souligner que contrairement à la façon dont l'envisage le rationalisme chomskyen, il ne s'agit en rien d'une modalité anomale ou disruptive. Tout au contraire, la variation apparaît dans le vivant comme une modalité fondamentalement dynamique et structurante. En matière de langage et de cognition il n'en va pas autrement. Récemment, Labov (2001, 3-33) a rappelé que concernant le changement linguistique, sur les 16 critères sur-lesquels Darwin lui-même avait construit son parallèle entre évolution des espèces et évolution des langues, 15 étaient appropriés du point de vue des acquis de la recherche variationniste contemporaine. Le 16ème portant sur l'avantage adaptatif et l'amélioration du rendement communicationnel que le changement linguistique devrait procurer constitue ce que Labov appelle le paradoxe darwinien. En matière de changement linguistique en effet, on ne peut mettre en avant une dynamique sélective qui sélectionnerait une forme, une structure ou une fonction linguistique sur la base de l'amélioration qu'elle produirait. Il reste, que le concept central du darwinisme, celui de générateur aléatoire de diversité ainsi que la tension entre unicité de l'espèce, mesurée notamment par la capacité à combiner deux génotypes dans une descendance non stérile (.i.e la barrière d'espèce), et hétérogénéité des phénotypes trouvent à s'appliquer aux systèmes culturels et sociaux que sont les langues : extrême variabilité des formes et des occurrences limitée pourtant par les nécessités de l'intercompréhension et le marquage de l'appartenance à une même communauté linguistique.

Mais l'importance du concept de variation et d'hétérogénéité structurée ne se limite pas aux phénomènes de changement linguistique. J'ai rappelé ici que son impact était bien plus large et affectait l'ensemble de la linguistique, pour ne pas dire des sciences cognitives. Si la linguistique de *l'exemplum* n'a de cesse d'éliminer toute hétérogénéité et toute variation, les linguistiques du *datum*, les modèles fondés sur l'usage et la théorie de la cognition culturelle s'appuient au contraire fondamentalement sur ses potentialités structurantes et sur les dynamiques qu'elle introduit. C'est nous l'avons vu, la variabilité des formes dans l'usage qui constitue le moteur et la motivation de leur organisation taxinomique de plus en plus en abstraite et la cristallisation progressive de catégories fonctionnelles de traitement. Cette variabilité n'affecte pas seulement les différences interindividuelles. Parce qu'elle est une des dimensions fondamentale de tout usage, et parce que la grammaire s'érige sur l'usage, cette variabilité et cette hétérogénéité interne affecte également le dispositif langagier, cognitif et pratique, de tout locuteur, en synchronie comme en diachronie. Dans cette approche, la compétence linguistique d'un locuteur situé n'est ni stable, ni homogène. Pour autant qu'il s'agit bien d'une compétence pratique, socialement constituée et socialement exercée, c'est bien au plan individuel comme au plan social un produit historique et culturel. Ce qui l'unifie relativement et la constraint dans des limites de variabilité données, c'est précisément ce qui fait

l'unité historique et sociale des communautés humaines, le partage inégal mais contraignant de normes, de règles et de routines, en un mot le partage d'une même culture.

Au passage, cela règle un problème jamais réellement argumenté dans l'approche cartésienne : l'identité stricte supposée des grammaires de tous les locuteurs d'une langue qui conduit au fameux problème de la convergence des apprentissages et indirectement motive l'utilisation, métaphorique on l'a vu, du théorème de Gold. Dans une linguistique des usages, rien n'impose une telle convergence des apprentissages et une telle identité stricte des grammaires mentales. La communication interpersonnelle dans une communauté réelle, tissée de variation et d'hétérogénéité structurales, n'impose pas l'identité stricte des compétences. Tout au contraire, comme on le voit pour tous les dispositifs culturels et sociaux, c'est le partage des mêmes normes et des mêmes modalités d'évaluation qui, quand bien même les sujets sociaux en auraient un usage inégalitaire et différencié, assure la cohérence du tissu social et contraint son hétérogénéité interne en même temps qu'elle limite la variation des pratiques et des usages.

S'il ne suppose pas une convergence complète et une identité parfaite des états stabilisés à l'âge adulte, l'apprentissage *in situ* des compétences communicationnelles, linguistiques culturelles et sociales, n'impose pas non plus une homogénéité des données dont il se nourrit. On sait que les modélisations statistiques utilisées dans les différents systèmes d'apprentissage automatiques, symboliques, sub-symboliques ou connexionnistes, ne peuvent converger si les données sont trop régulières et trop homogènes. Une certaine quantité de bruit ou d'incertitude est toujours nécessaire aux systèmes et parfois il faut en garantir la présence en les introduisant explicitement sous forme de biais. Comme je l'ai souligné à la suite de Tomasello (2008b), la cognition humaine est extrêmement sensible aux régularités et aux différences, à leur récurrence et à leur organisation dans le temps, en bref l'intelligence humaine des phénomènes et des actes est en très grande partie de type statistico-probabiliste. C'est pourquoi, bien loin de constituer un frein ou une entrave, l'existence d'une variabilité structurée limitée et contrainte correspond au contraire à un avantage formel et à une facilitation cognitive décisive.

Réfléchissant aux arguments permettant de fonder l'idéalisat<sup>ion</sup> du locuteur-auditeur et l'homogénéisation *a priori* des communautés linguistiques, Chomsky (1980, 27-28), faute d'aucun argument empirique ou d'aucune observation factuelle écrit : « Nous restons ainsi avec ce qui doit être la question de fond : notre idéalisat<sup>ion</sup> déforme-t-elle à ce point le monde réel qu'elle ne peut produire aucun aperçu véritable de la faculté de langage, ou bien, au contraire, nous donne-t-elle la possibilité de découvrir des propriétés fondamentales de cette faculté ? Bref, est-elle légitime ? Supposons que l'on réponde par la négative. Cela condamne à admettre l'une ou l'autre des deux thèses suivantes : 1. les hommes sont ainsi faits qu'ils seraient incapables d'apprendre le langage

dans une communauté linguistique homogène ; la variabilité ou l'incohérence des données accessibles sont une condition nécessaire de l'apprentissage ; 2. les hommes pourraient apprendre le langage dans une communauté linguistique homogène, mais les propriétés de l'esprit qui le leur permettraient ne relèvent pas de l'acquisition normale dans le monde réel, fait de diversités, de conflits dialectaux, etc. Je ne puis croire que quiconque aura réfléchi à la question aille adhérer à l'une ou l'autre de ces affirmations, qui apparaissent bien aussi désespérément improbables l'une que l'autre. Rejetons-les donc ; ce faisant, nous admettons que l'être humain possède une propriété de l'esprit qui lui permettrait d'apprendre la langue d'une communauté linguistique homogène ».

Le panorama de la recherche linguistique contemporaine que j'ai esquissé ici, appuyé sur la distinction épistémologique entre linguistique de *l'exemplum* et linguistique du *datum* conduit à contester point par point cette *reductio ad absurdum*. Tout au contraire, dès que l'on prend au sérieux les données massives de l'usage, le fonctionnement des communautés réelles, la variation structurée apparaît comme une motivation et un moteur des dynamiques de l'apprentissage et comme organisatrice et régulatrice des communications interpersonnelles.

Dans une perspective fonctionnaliste, nous en revenons ainsi 50 ans après à ce qui fût l'intuition géniale de Weinreich : « The solution, we will argue, lies in the direction of breaking down the identification of structuredness with homogeneity. The key to a rational conception of language change — indeed, of language itself — is the possibility of describing orderly differentiation in a language serving a community. We will argue that nativelike command of heterogeneous structures is not a matter of multidialectalism or “mere” performance, but is part of unilingual linguistic competence. One of the corollaries of our approach is that in a language serving a complex (i.e., real) community, it is absence of structured heterogeneity that would be dysfunctional ». Weinreich, Labov et Herzog (1968, 96)

## Références

- Aarts, Bas (2000): Corpus linguistics, Chomsky and fuzzy tree fragments, *in* C. Mair et M. Hundt (dirs.), *Corpus Linguistics and Linguistic Theory.*, Amsterdam/Atlanta: Rodopi, 5-13
- Antilla, Arto (2007): Variation and optionality. , *in* P. de Lacy (dir.) *The Cambridge Handbook of Phonology*, Cambridge: Cambridge University Press, 519-536.
- Antilla, Arto et Cho, Young-Mee Yu (1998): Variation and change in Optimality Theory, *Lingua* 104 31-56
- Arnauld, Antoine et Lancelot, Claude (1660): *Grammaire générale et raisonnée, contenant les fondements de l'art de parler*, Paris: Républications Paulet 1969
- Bachelard, Gaston (1938): La formation de l'esprit scientifique; contribution à une psychanalyse de la connaissance objective, Paris: J. Vrin
- Baratin, Marc, Desbordes, Françoise, Hoffman, Philippe et Pierrot, Alain (1981): *L'analyse linguistique dans l'Antiquité classique. I, Les théories*, Paris: Éditions Klincksieck
- Barlow, Michael et Kemmer, Suzanne (dirs.) (2000): *Usage based models of language*; Stanford Cal.: CSLI,
- Béguelin, Marie-José (1990): Conscience du sujet parlant et savoir du linguiste *in* R. Liver, I. Werlen et P. Wunderli (dirs.), *Sprachtheorie und Theorie der Sprachwissenschaft. Festschrift für Rudolf Engler*, Tübingen: Gunter Narr, 208-220
- Benveniste, Émile (1966): *Problèmes de linguistique générale*, Paris: Gallimard
- Berwick, Robert C. et Chomsky, Noam (2008): Poverty of the Stimulus' revisited: recent challenges reconsidered. *Proceedings of the 30th Annual Conference of the Cognitive Science Society* Austin, TX: Cognitive Science Society, 383
- Berwick, Robert C., Pietroski, Paul, Yankama, Beracah et Chomsky, Noam (2011): Poverty of the Stimulus Revisited, *Cognitive Science* 35/7, 1207-1242,
- Berwick, Robert C; (2011): Syntax Facit Saltum Redux:Biolinguistics and the Leap to Syntax, *in* A. M. Di Sciullo et C. Boeckx (dirs.), *The Biolinguistic Enterprise. New Perspectives on the Evolution and Nature of Human Language Faculty*, Oxford: Oxford University Press, 1-18
- Berwick, Robert C; et Chomsky, Noam (2011): The Biolinguistic Programm : The Current State of its Developpment, *in* A. M. Di Sciullo et C. Boeckx (dirs.), *The Biolinguistic Enterprise*, Oxford: Oxford University Press, 18-42
- Bickerton, Derek (1998): Catastrophic evolution: the case for a single step from protolanguage to full human language, *in* J. R. Hurford, M. Studdert-Kennedy et C. Knight (dirs.), *Approaches to the evolution of language*, Cambridge: Cambridge University Press, 341-358
- Bickerton, Derek et Calvin, William H. (2000): *Lingua ex Machina:Darwin and Chomsky Reconciled*, Cambridge Mass.: MIT Press
- Bloomfield, Leonard (1926): A Set of Postulates for the Science of Language, *Language* 2, 153-164
- Bloomfield, Leonard (1933): *Language*, New York: H. Holt and Company
- Boeckx, Cedric et Grohmann, Kleanthes K. (2007): *The BIOLINGUISTICS Manifesto Biolinguistics 1*, Open Journal Systems,
- Boersma, Paul (1998): Functional Phonology : Formalizing the interactions between articulatory and perceptual drives, Amsterdam: LOT
- Bouquet, Simon (1997): *Introduction à la lecture de Saussure*, Paris: Payot
- Bourdieu, P., Rey, A., Milner, J. C., Delesalle, S., Encrevé, P. et Fauconnier, G. (1977): Table ronde « linguistique et sociologie du langage », *Langue Française* 34, 35-51
- Bourdieu, Pierre (1980): *Le sens pratique*, Paris: Éditions de Minuit
- Bourdieu, Pierre (1982): Ce que parler veut dire : L'économie des échanges linguistiques, Paris: Fayard
- Bourdieu, Pierre (1994): *Raisons pratiques : sur la théorie de l'action*, Paris: Editions du Seuil
- Bourdieu, Pierre (1997): *Méditations pascaliennes*, Paris: Editions du Seuil
- Bourdieu, Pierre et alii, et. (1993): *La Misère du monde*, Paris: Editions du Seuil

- Buffon, Georges-Louis Leclerc de (1749): *Histoire naturelle, générale et particulière, avec la description du Cabinet du Roy.* , Réédition Paris: Éditions Honoré Champion, 2007-2009 volumes 1, 2, 3
- Bybee, Joan (2001): *Phonology and language use*, Cambridge: Cambridge University Press
- Bybee, Joan (2006): *Frequency of use and the organization of language.*, Oxford: Oxford University Press
- Chambers, J. K. et Trudgill, Peter (1980): *Dialectology*, Cambridge; New York: Cambridge University Press
- Changeux, J.-P. (1983): *L'homme neuronal*, Paris: Fayard
- Changeux, J.-P. et Connes, A. (1989): *Matière à penser*, Paris: Odile Jacob
- Chater, Nick et Manning, Christopher D. (2006): Probabilistic models of language processing and acquisition, *Trends in cognitive sciences*. 10 7, 335
- Chevalier, Jean-Claude (2007): Les exemples et la norme dans les grammaires : étude historique, in G. Siouffi et A. Steuckardt (dirs.), *Les linguistes et la norme*, Berlin: Peter Lang, 151-163
- Chomsky, Noam (1957): *Syntactic structures*, La Haye: Mouton
- Chomsky, Noam (1959): A review of B.F. Skinner's *Verbal Behavior*, *Language* 35 1, 26-58
- Chomsky, Noam (1965): *Aspects of the theory of syntax*, Cambridge, Mass.: M.I.T. Press Trad. Fr. Paris 1969 Editions du Seuil
- Chomsky, Noam (1966): *Cartesian linguistics: a chapter in the history of rationalist thought*, New York,: Harper & Row
- Chomsky, Noam (1968): *Language and mind*, New York: Harcourt Brace & World Trad. française Louis-Jean Calvet, Paris Payot, 1976
- Chomsky, Noam (1977): Langue, linguistique, politique : dialogues avec Mitsou Ronat, Paris: Flammarion
- Chomsky, Noam (1980): *Rules and representations*, New York: Columbia University Press Trad. française A. Kihm, Paris, Flammarion, 1985
- Chomsky, Noam (1986): *Knowledge of language : its nature, origin, and use*, New York: Praeger
- Chomsky, Noam (1995): *The minimalist program*, Cambridge, Mass.: The MIT Press
- Chomsky, Noam (2004): Biolinguistics and the Human Capacity (Talk delivered at MTA, Budapest, May 17, 2004).
- Chomsky, Noam (2007): *Of Minds and Language*; Biolinguistics 1, Open Journal Systems,
- Chomsky, Noam (2011): Some simple evo-devo theses: how true might they be for language?, in R. K. Larson, V. Déprez et H. Yamakido (dirs.), *The Evolution of Human Language, Biolinguistic Perspectives*, Cambridge: Cambridge University Press, 43-63
- Chomsky, Noam et Halle, Morris (1968): *Sound Pattern of English*, New York: Harper and Row
- Clavien, Christine (2010): Je t'aide, moi non plus : biologique, comportemental ou psychologique, l'altruisme dans tous ses états, Paris: Vuibert
- Currie, Haver C. (1952): A Projection of Sociolinguistics: The Relationship of Speech to Social Status, *Southern Speech Journal* 17 28-37
- Currie, Haver C. (1981): Sociolinguistics and American linguistic theory, *International Journal of the Sociology of Language* 31, 29-41
- Darwin, Charles (1859): L'origine des espèces au moyen de la sélection naturelle ou la préservation des races favorisées dans la lutte pour la vie., Traduction Edmond Barbier, Nouvelle édition Daniel Becquemont, 1992, Paris: Flammarion
- Dawkins, Richard (1976): *The selfish gene*, New York ; Oxford: Oxford University Press
- Dessalles, Jean-Louis (2000): Aux origines du langage : une histoire naturelle de la parole, Paris: Hermès
- Dessalles, Jean-Louis, Picq, Pascal et Victorri, Bernard (2006): *Les origines du langage*, Paris: Le Pommier
- Di Sciullo, Anna Maria et Boeckx, Cedric (2011): The Biolinguistic Enterprise. New Perspectives on the Evolution and Nature of Human Language Faculty, Oxford: Oxford University Press

- Durand, Jacques, Laks, Bernard, Calderone, Basilio et Tchobanov, Atanas (2011): Que savons nous de la liaison aujourd'hui?, *Langue Française* 169, 103-135
- Durkheim, Émile (1927): *Les règles de la méthode sociologique*, Paris: F. Alcan
- Elman, J.L. et Lewis, J.D. (2001): Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. *in* B. Skarabela, S. Fish et A. H.-J. Do (dir.) *26th annual Boston University conference on language development*: Cascadilla Press, 359-370
- Encrev , Pierre (1976): *Pr sentation In Labov William Sociolinguistique*, Paris: Editions de Minuit
- Encrev , Pierre (1982): A propos du march  linguistique, *in* N. Dittmar et B. Schlieben-Lange (dir.), *Die Soziolinguistik in romanischsprachigen L ndern*, T bingen: Narr, 97-104
- Encrev , Pierre (1986): Variation et structure,  tudes de phonologie et de pragmatique sociolinguistiques, Th se d'Etat, Paris VIII
- Encrev , Pierre (2000): The old and the new : Some remarks on phonology and its history., *Folia Linguistica* XXXIV Special issue published by John Goldsmith and Bernard Laks, 56-84
- Everett, Daniel (2005): Cultural constraints on grammar and cognition in Pirah : Another look at the design features of human language, *Current Anthropology* 46, 621-646
- Everett, Daniel (2009): Pirah  culture and grammar : A response to some criticism, *Language* 85, 405-42,
- Feldman, Jerome A. (2006): From Molecule to Metaphor. A Neural Theory of Language Cambridge Mass.: MIT Press
- Fodor, Jerry.(1983a): The modularity of mind : an essay on faculty psychology, Cambridge: MIT press
- Fodor, Jerry (1983b): *The Modularity of Mind*, Cambridge MA: MIT Press
- Fodor, Jerry (2001): The mind doesn't work that way : the scope and limits of computational psychology, Cambridge Mass.: MIT Press
- Foraker, Stephani , Regier, Terry , Khetarpal, Naveen , Perfors, Amy et Tenenbaum, Joshua (2009): Indirect Evidence and the Poverty of the Stimulus: The Case of Anaphoric One, *Cognitive Science* 33, 287-300
- Foucault, Michel (1967): La Grammaire g n rale de Port-Royal, *Langages* 7, 7-15.
- Fukuyama, Francis (1992): *La fin de l'histoire et le dernier homme*, Paris: Flammarion
- Glatigny, Michel (1982): La notion de r gle dans la « grammaire » de Meigret, *Histoire Epist mologie Langage* 4-2, 93-106
- Gleason, Jean Berko et Thompson, R. Bruce (2002): Out of the Baby Book and Into the Computer: Child Language Research Comes of Age, *Apa Review of books* 47-4, 390-394
- Gold, Mark E. (1967): Language identification in the limit, *Information and Control* 16, 447-474,
- Goldberg, Adele E. (2006): *Constructions at Work : The nature of generalization in language*, Oxford: Oxford University Press
- Goldberg, Ad le E. (1995): *Constructions: A construction grammar approach to argument structure*, Chicago: University of Chicago Press
- Goldsmith, J (2005a): An Algorithm for the Unsupervised Learning of Morphology, *Natural Language Engineering* 1 1,
- Goldsmith, J (2011): *The Linguistica Project*; Chicago: University of Chicago
- Goldsmith, J (  para tre): Towards a new empiricism:  
<http://hum.uchicago.edu/~jagoldsm//Papers/empiricism1.pdf>
- Goldsmith, John (2005b): Review : The Legacy of Zellig Harris: Language and information into the 21st century, *Language* 81-3, 719-736
- Goldsmith, John (2010): Towards a new empiricism for linguistics, *A para tre*,
- Goldsmith, John et Aris, Xanthos (2009): Learning phonological categories *Language* 85- 1, 4-39
- Goldsmith, John et Huck, Geoffrey (1995): *Ideology and linguistic theory : Noam Chomsky and the deep structure debates.*, New York: Routledge
- Gould, Stephen J. et Vrba, Elizabeth (1982): Exaptation:a missing term in the science of form *Paleobiology* 8- 4-15
- Groupe Ars Grammatica (2005): Pr sentation du De Adverbio de Priscien, *Histoire Epist mologie Langage* 27-2, 7-28

- Grohmann, Kleanthes K. (2007): *An Interview with Henk van Riemsdijk* Biolinguistics 1 Open Journal Systems,
- Gumperz, John J. (1982): *Discourse strategies*, Cambridge: Cambridge University Press
- Harris, Randall (1993): *Linguistic Wars*, Oxford: Oxford University Press
- Harris, Zellig (1951): *Methods in Structural Linguistics*, Chicago: University of Chicago Press
- Hauser, Marc D., Chomsky, Noam et Fitch, Walter Tecumseh (2002): The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?, *Science* 29822 November, 1569-1580
- Hayes, Bruce et Czirák Londe, Zsuzsa (2006): Stochastic Phonological Knowledge: The Case of Hungarian Vowel Harmony, *Phonology* 23, 59-104
- Hayes, Bruce, Kirchner, Robert et Steriade, Donca (dirs.) (2004): *Phonetically-Based Phonology* Cambridge: Cambridge University Press.,
- Hockett, Charles F. (1942): A System of Descriptive Phonology, *Language* 18, 3-21
- Holtz, Louis (1981): *Donat et la tradition de l'enseignement grammatical : étude sur l'Ars Donati et sa diffusion (IVe-IXe siècle) et édition critique*, Paris: Centre national de la recherche scientifique
- Hoquet, T. (2009): *Darwin contre Darwin: comment lire "L'origine des espèces" ?*, Paris : Éditions du Seuil
- Horace (457/1944): *L'Art poétique ou Épître aux Pisons*, Paris, Garnier, 1944:
- Hymes, Dell H. (1972): Models of the interaction of language and social life., in J. J. Gumperz et D. H. Hymes (dirs.), *Directions in sociolinguistics*, New York: Holt, Rinehart and Winston, 35-71
- Johnson, Kent (2004): Gold's Theorem and Cognitive Science, *Philosophy of Science* 71, 571-592
- Keller, Madeleine (2009): Exemples et citations chez Priscien : examen de deux passages du livre xv « De aduerbio » des *Institutiones grammaticae* in B. Bortolussi, M. Keller, S. Minon et L. Sznadjer (dirs.), *Traduire, transposer, transmettre : dans l'Antiquité gréco-romaine*, Paris: Picard,
- Koerner, Konrad (1991): Toward a History of Modern Sociolinguistics, *American Speech* 66-1, 57-70
- Labov, William (1966): *The Social Stratification of English in New York City*, Washington: Center for Applied Linguistics
- Labov, William (1972): *Sociolinguistic patterns*, Philadelphia: University of Pennsylvania Press
- Labov, William (1975): *What is a linguistic fact?* , Lisse: Peter de Ridder Press
- Labov, William (1976): *Sociolinguistique*, Paris: Editions de Minuit
- Labov, William (1979): *Le parler ordinaire*, Paris: Editions de Minuit
- Labov, William (1981): Resolving the neogrammarian controversy, *Language* 57-2, 267-309
- Labov, William (1987): *Some Observations on the Foundation of Linguistics*;  
<http://www.ling.upenn.edu/~wlabov/Papers/Foundations.html>
- Labov, William (1994): *Principles of Linguistic Change : Internal factors*, Oxford: Blackwell 1
- Labov, William (1996): When Intuitions Fail in L. McNair, K. Singer, L. Dolbin et M. Aucon (dirs.) *Papers from the Parasession on Theory and Data in Linguistics* Chicago Linguistic Society 77-106
- Labov, William (2001): *Principles of Linguistic Change : Social factors*, Oxford: Blackwell
- Labov, William (2004): Quantitative Analysis of Linguistic Variation, in U. Ammon, N. Dittmar, K. J. Mattheier et P. Trudgill (dirs.), *Sociolinguistics. An International Handbook of the Science of Language and Society*, La Haye, Berlin: 3-1, 6-22
- Lakoff, George (1973a): Deep Language in New-York Review of Books February 8,
- Lakoff, George (1973b): Interview with Herman Parret, in H. Parret (dir.) *Discussing language*, The Hague: Mouton, 151-178
- Lakoff, Georges et Johnson, Marc H. (1999): *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, New-York: Basic Books
- Laks, Bernard (1996): *Langage et cognition : l'approche connexionniste*, Paris: Hermès
- Laks, Bernard (2008): Pour une phonologie de corpus, *Journal of French Language Studies* 18-1, 3-32
- Laks, Bernard (2011a): La phonologie du français et les corpus. , *Langue Française* 169, 3-17
- Laks, Bernard (2011b): La phonotactique saussurienne : Système et loi de la valeur s *Langages* 178
- Laks, Bernard, Cleuziou, Serge, Demoule, Jean-Paul et Encrevé, Pierre (dirs.) (2007): *Origin and Evolution of Languages : Approaches, Models, Paradigms*; Londes: Equinox,

- Lamarck, Jean Baptiste Pierre Antoine de Monet de (1809): *Philosophie zoologique, ou, Exposition des considérations relative à l'histoire naturelle des animaux*, Paris: Dentu Paris , Flammarion, coll. GF, 1994
- Langacker, R (1998): Conceptualization, Symbolisation and Grammar, in M. Tomasello (dir.) *The New Psychology of Language : Cognitive and functional approaches to linguistic structure*, Mahwah, NJ: Lawrence Erlbaum, 1-39
- Langacker, Ronald (1987): *Foundations of cognitive grammar I: Theoretical prerequisite*, Stanford: Stanford University Press
- Langacker, Ronald (1991): *Foundations of cognitive grammar II : Descriptive applications*, Stanford: Stanford University Press
- Langacker, Ronald (2000): A Dynamic Usage-Based Model, in M. Barlow et S. Kemmer (dirs.), *Usage Based Models of Language*, Standford: CSLI, Standford University, 1-65
- Langacker, Ronald W. (1988): A usage-based model in B. Rudzka-Ostyn (dir.) *Topics in Cognitive Linguistics*, 127-165
- Lehmann, Winfred P. (1962): *Historical Linquistics : An Introduction*, New York: Holt, Rinehart and Winston
- Levinson, Stephen C., Jaisson, Pierre et Fyssen Foundation (2006): *Evolution and culture : a Fyssen Foundation symposium*, Cambridge, Mass.: MIT Press
- Linné, Carl von (1735): *Systema naturae, per regna tria naturae : secundum classes, ordines, genera, species cum characteribus, differentiis, synonymis, locis*. Leyde: Paris, Bibliothèque Nationale de France
- MacWhinney, Brian (2000): *The CHILDES Project: Tools for Analyzing Talk.*, Mahwah, NJ: Erlbaum
- MacWhinney, Brian (2007): The TalkBank Project. , in J. Beal, K. Corrigan et L. Moisl (dirs.), *Creating and Digitizing Language Corpora: Synchronic Databases*, Palgrave-Macmillan, 1
- Maniglier, Patrice (2008): Processing culture : new trends in artificial intelligence and linguistics in the light of structuralism, in S. Franchi et F. Bianchini (dirs.), *Toward an archeology of Artificial Intelligence*, Berlin: Springer
- Manning, Christopher D. (2003): Probabilistic syntax, in B. Rens, J. Hay et S. Jannedy (dirs.), *Probabilistic linguistics*, Cambridge, MA: MIT Press, 289-341
- Marcken, Carl de (1996): Unsupervised Language Acquisition. Phd, MIT
- Martinet, André (1962): *A functional view of language*, Oxford: Clarendon Press
- Mayr, Ernest (1982): *The growth of biological thought : diversity evolution, inheritance*, Cambridge MA: Belknap Press Trad. fr. , Paris, Fayard 1989
- Mehl, Matthias R., Vazire, Simine, Ramirez-Esparza, Nairán, Statcher, Richard B. et Pennebaker, James W. (2007): Are Women Really More Talkative Than Men?, *Science* 317-5834, 82-82
- Meigret, Louis (1542): *Traité touchant le commun usage de l'écriture françoise*, Lyon: Republications Slatkine Genève, 1972
- Meillet, Antoine (1921): *Linguistique historique et linguistique générale*, Paris: Champion
- Milner, Jean-Claude (1989): *Introduction à une science du langage*, Paris: Editions du Seuil
- Milroy, James et Milroy, Lesley (1985): *Authority in language : investigating language prescription and standardisation*, London u.a.: Routledge & Paul
- Morgan, J. L. (1989): Learnability considerations and the nature of trigger experiences in language acquisition, *Behavioral and Brain Sciences* 12, 352-353
- Mufwene, Salikoko S. (2001). *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Nevin, Bruce (dir.) (2002): *The legacy of Zellig Harris: Language and information into the 21st century : Philosophy of science, syntax and semantics.* ; Philadelphia:: John Benjamins,
- Nevins, Andrew, Pesetsky, David et Rodrigues, Cilene (2009): Pirahã exceptionality : a reassessment, *Language* 85, 355-404,
- Newmeyer, Frederick J. (1988): *Linguistics : the Cambridge survey*, Cambridge; New York; Melbourne: Cambridge university press
- Newmeyer, Frederick J. (2003): Grammar is grammar and usage is usage. : . *Language* 79 4, 682-797

- Parisse, Christophe (2005): New perspectives on language development and the innateness of grammatical knowledge., *Language Sciences* 27, 383-401
- Patrick, Peter L. (2002): The speech community, in P. Trudgill, J. K. Chambers et N. Schilling-Estes (dirs.), *The handbook of language variation and change*, Malden Mass.: Blackwell, 573-599
- Paul, Hermann (1909): *Prinzipien der sprachgeschichte*, Halle: M. Niemeyer
- Perfors, A., Tenenbaum, J. et Regier, T (2006): Poverty of the stimulus? A rational approach in R. Sun (dir.) *28th annual conference of the cognitive science society*: Erlbaum, 663-668
- Piatelli-Palmarini, Massimo (dir.) (1980): *Language and learning : the debate between Jean Piaget and Noam Chomsky*; Cambridge, Mass.: Harvard University Press,
- Piatelli-Palmarini, Massimo et Noizet, Yvonne (1979): *Théories du langage, théories de l'apprentissage : le débat entre Jean Piaget et Noam Chomsky*, Paris: Éditions du Seuil
- Pinker, S (1997): *How the Mind Works*, New York: Norton
- Pinker, Steven (1999): *L'instinct du langage*, Paris: Editions Odile Jacob
- Pinker, Steven et Bloom, Paul (1989): Natural language and natural selection, *Behavioral and Brain Sciences* 13, 707-784
- Pinker, Steven et Mehler, Jacques (dirs.) (1989): *Connections and symbols*; Cambridge, Mass.: MIT Press,
- Platon (1849): *Menon, de la vertu*, Paris: Rey : Traduction Victor Cousin
- Pullum, Geoffrey K et Scholz, Barbara C. (2002): Empirical assessment of stimulus poverty arguments, *The Linguistic Review*. 18/1-2, 9-50
- Quintilianus, Fabius *Institutio Oratoria*, Paris, 1842, Trad. M. Nisard:
- Ramus, P. (1562): *Gramère*, Paris: Bibliothèque Nationale de France
- Reali, F. et Christiansen, M (2005): Uncovering the richness of the stimulus: Structure dependence and indirect statistical evidence, *Cognitive Science* 29, 1007-1028
- Rissanen, Jorma (2007): *Information and Complexity in Statistical Modeling*, Berlin: Springer
- Sampson, Geoffrey (2002): Exploring the richness of the stimulus, *The Linguistic Review* 19/ 1-2, 73-104
- Sapir, Edward (1921): *Language, an introduction to the study of speech*, New York,: Harcourt Brace and Company
- Saussure, Ferdinand de (1916): *Cours de linguistique générale*, Paris: Payot 1972 Paris : Payot
- Saussure, Ferdinand de (2001): *Ecrits de linguistique générale*, Paris: Gallimard
- Schane, Sanford A. (1965): The phonological and morphological structure of French, Phd, MIT
- Schuchardt, Hugo (1909): Die Lingua Franca, *Zeitschrift für Romanische Philologie* XXXIII, 441-461
- Schuchardt, Hugo (1922): Hugo Schuchardt-Brevier . Ein vademekum der allgemeinen Sprachwissenschaft, Halle: Max Niemeyer
- Searle, John R. (1972): A Special Supplement: Chomsky's Revolution in Linguistics, *New-York Review of Books* June 29,
- Tomasello, Michael (1995): Language is Not an Instinct (Review of Pinker 1994), *Cognitive Development* 10, 131-156
- Tomasello, Michael (Dir.) (1998): *The new psychology of language : cognitive and functional approaches to language structure*, Mahwah, NJ : Erlbaum
- Tomasello, Michael (1999): *The cultural origins of human cognition*, Cambridge, Mass.: Harvard University Press
- Tomasello, Michael (2003): *Constructing a language : A usage based theory of language acquisition*, Cambridge Mass.: Harvard University Press
- Tomasello, Michael (Dir.) (2008a): *The new psychology of language* vol. 2, New York; London: Psychology Press
- Tomasello, Michael (2008b): *Origins of Human Cognition*, Cambridge, Mass.: MIT Press
- Tomasello, Michael (2008c): Some surprises for psychologists, in M. Tomasello (dir.) *The new psychology of language* 2, Mahwah, NJ: Lawrence Erlbaum 1-15
- Trivers, Robert (2002): *Natural selection and social theory selected papers of Robert Trivers*; Oxford; New York: Oxford University Press,

- Trivers, Robert L. (1971): The Evolution of Reciprocal Altruism, *The Quarterly Review of Biology* 46- 1, 35-57
- Troubetzkoy, Nicolas Sergueevitch (1939): *Grundzüge der Phonologie*, Trad. Fr J. Cantineau) : *Principe de Phonologie*, Paris, Klincksieck,
- Valéry, Paul (1941): *Tel quel, Edition Oeuvre II*, Paris, Gallimard, Bibliothèque de la Pleiade, 1960, 696
- Vaugelas, Favre de Claude (1647/1934): *Remarques svr la langue françoise, utiles à ceux qui veulent bien parler et bien escrire* Paris: Droz
- Weinreich, Uriel (1951): *Research Problems in Bilingualism, with Special Regard to Switzerland*. PhD. Columbia University, New York
- Weinreich, Uriel (1953): *Languages in contact : findings and problems*, La Haye: Mouton
- Weinreich, Uriel (1954): Is a Structural Dialectology Possible? . *Word* 10, 268-280
- Weinreich, Uriel, Labov, William et Herzog, Marvin (1968): Empirical Foundations for a Theory of Language Change, in W. Lehmann et Y. Malkiel (dirs.), *Directions for Historical Linguistics*, Austin: University of Texas Press, 95-188
- Whitney, William Dwight (1875): *La vie du langage*, Paris: Librairie Germer Baillièvre et compagnie Facsimile Didier Erudition