

## Termes en \*k<sup>w</sup>-, récursivité et origine du langage

### Introduction

Les recherches sur l'origine du langage connaissent depuis une quinzaine d'années un essor assez surprenant, qui tranche avec la période antérieure où les réflexions sur ce sujet étaient considérées comme des spéculations intellectuelles d'ordre philosophique, auxquelles les linguistes en tant que tels n'avaient rien à apporter<sup>1</sup>. A vrai dire, même aujourd'hui, un grand nombre de linguistes restent sceptiques sur l'intérêt de participer à cet élan pluridisciplinaire qui leur paraît relever d'un phénomène de mode encouragé à tort par certaines institutions scientifiques<sup>2</sup>. D'une certaine manière, l'esprit du célèbre interdit de la Société linguistique de Paris, même s'il est obsolète depuis plus d'un siècle<sup>3</sup>, semble encore rôder dans la communauté des linguistes...

Le problème est double : on peut se demander d'abord si l'apparition du langage au cours de l'homínisation est susceptible aujourd'hui d'être appréhendée par des méthodes scientifiques (toutes disciplines confondues), et ensuite si la linguistique elle-même a quelque chose de spécifique à apporter dans ce domaine.

Nous allons essayer de montrer ici que l'on doit répondre oui à ces deux questions, en nous centrant essentiellement sur la seconde. Notamment, nous nous appuyerons sur les travaux que Pierre Le Goffic (2001, 2002) a consacrés aux termes indo-européens en \*k<sup>w</sup>-, parce qu'ils nous semblent exemplaires d'une démarche proprement linguistique pouvant contribuer à nous éclairer sur les processus d'émergence de certaines propriétés caractéristiques du langage, en l'occurrence la récursivité telle qu'elle est mise en œuvre dans le mécanisme de subordination.

### 1. L'origine du langage : un problème scientifique

#### 1.1 Une nouvelle donne

S'il est légitime aujourd'hui de considérer l'origine du langage comme un problème scientifique, c'est en raison des progrès considérables de toute une série de disciplines, archéologie, paléanthropologie, éthologie, neurophysiologie comparée, génétique des populations, paléogénétique, etc., qui permettent de reconstituer les différentes étapes de l'évolution des homínidés de manière relativement assurée. Bien entendu, il ne s'agit pas de certitudes, et de

<sup>1</sup> Sur l'histoire des recherches sur l'origine du langage, voir Bergougnieux (2002, 2005) et Auroux (2006).

<sup>2</sup> Le CNRS a lancé en 2000 un programme interdisciplinaire pluriannuel de grande envergure intitulé *Origine de l'homme, du langage et des langues*. Ce programme a été ensuite repris et élargi à toute l'Europe par l'ESF (*European Science Foundation*) sous l'intitulé *Origin of Man, Language and Languages*.

<sup>3</sup> L'article 2 des statuts de la Société linguistique de Paris (SLP) promulgués en 1866, stipulait : « la société n'admet aucune communication concernant soit l'origine du langage soit la création d'une langue universelle ». Cet article disparaît dix ans plus tard lors d'une refonte des statuts, avec cette fois un article 1 qui formule une interdiction tout aussi sévère, mais plus générale et du coup moins claire en ce qui concerne l'origine du langage : « la société a pour objet l'étude des langues et l'histoire du langage. Tout autre sujet est rigoureusement interdit ». On peut en effet se demander dans quelle mesure le problème de l'origine fait partie ou non de l'histoire du langage...

nouvelles découvertes peuvent modifier sur tel ou tel point les hypothèses qui font actuellement consensus, mais il semble peu probable que le canevas général soit complètement bouleversé.

Ainsi la question de l'émergence du langage se pose dans un cadre de plus en plus précis, au fur et à mesure que l'on sait mieux où, quand et comment sont apparus et se sont développés les différents traits qui caractérisent notre espèce, comme la bipédie, la fabrication d'outils, la maîtrise du feu, la division sociale du travail, etc. Toutes ces connaissances sont autant de contraintes qui doivent être prises en compte par tout scénario d'émergence du langage. Autrement dit, même si ces scénarios restent spéculatifs puisque l'on n'a pas de traces directes de l'activité de langage chez nos lointains ancêtres, on peut les discuter et les confronter de manière rationnelle, en évaluant dans quelle mesure ils satisfont ces contraintes. Cette activité est donc de plein droit une activité scientifique, au même titre que les recherches sur les autres grands problèmes d'origine, comme l'origine de la vie ou l'origine de l'univers, et avec le même statut épistémologique : il s'agit dans tous les cas d'élaborer les scénarios les plus plausibles dans l'état actuel de nos connaissances, sans masquer le caractère conjectural de ces travaux.

L'intérêt de ces recherches est double. D'abord, l'élaboration de scénarios rationnels plausibles suffit à démontrer que l'on n'est pas forcé de recourir à des explications surnaturelles (intervention divine ou autre) que des prédicateurs de tout poil ne manqueraient pas de présenter comme les seules possibles si la science leur abandonnait ce terrain. Encore de nos jours, on en conviendra, cela reste un enjeu de société dont on ne doit pas sous-estimer l'importance. Ensuite, et de manière sans doute plus positive, l'établissement de conjectures de ce type présente l'intérêt de guider les recherches futures en fournissant des objectifs de validation ou de réfutation des hypothèses proposées. De ce point de vue, il faut souhaiter que le plus grand nombre possible de scénarios soient mis en concurrence, de manière à élargir au maximum l'éventail de recherches qui, au-delà de cette première motivation, peuvent s'avérer fécondes pour la compréhension de bien d'autres questions scientifiques. Ainsi, il semble tout à fait légitime que des organismes de recherche financent des programmes sur ce sujet, même si ces recherches ne présentent pas d'applications prévisibles dans l'immédiat<sup>4</sup>.

## 1.2 Les principales étapes de l'hominisation

Quelles sont donc ces contraintes que doit prendre en compte toute réflexion sur l'origine du langage ? Autrement dit, que sait-on de l'histoire de l'hominisation qui constitue le cadre dans lequel doivent être élaborés les scénarios d'émergence ? C'est ce que nous allons maintenant résumer très brièvement<sup>5</sup>.

La séparation de la branche des hominidés avec celle qui conduit aux chimpanzés, les singes anthropoïdes qui nous sont les plus proches, date de six à huit millions d'années. Les premières données fossiles d'importance datent d'environ quatre millions d'années : plusieurs espèces d'australopithèques apparaissent à cette période en Afrique et ils vont occuper le terrain, si l'on peut dire, pendant deux millions d'années. Il s'agit d'animaux qui sont déjà bipèdes (comme la célèbre Lucy) et qui vivent dans la savane. On a toutes les raisons de penser (taille du cerveau, mode de vie,

<sup>4</sup> : A vrai dire, on doit même s'en réjouir, puisque c'est le signe que ces organismes restent convaincus de l'importance de la recherche fondamentale « gratuite », à une époque où la tendance est plutôt de privilégier la rentabilité à court terme voire à courte vue.

<sup>5</sup> Il existe de nombreux ouvrages sur le sujet. On pourra consulter par exemple Donald (1991), Tattersall (1998) et Klein (1999).

etc.) que leurs capacités cognitives étaient tout juste comparables à celle des singes anthropoïdes actuels. Il en est vraisemblablement de même pour le premier représentant du genre *Homo*, à savoir *Homo habilis*, qui apparaît il y a deux millions et demi d'années. Celui-ci doit l'honneur de l'appartenance à notre genre au fait qu'on lui a attribué les premiers outils lithiques. Mais d'une part, il semble qu'il ait été précédé sur ce point par certains australopithèques, et, d'autre part, on sait aujourd'hui que les chimpanzés, comme bien d'autres espèces d'ailleurs<sup>6</sup>, utilisent de manière systématique (et parfois fabriquent) des outils, ce qui banalise quelque peu ce trait que l'on a cru longtemps l'apanage exclusif de l'homme.

Les choses ne changent vraiment qu'avec l'apparition, il y a moins de deux millions d'années d'une nouvelle espèce : *Homo erectus*<sup>7</sup>, qui va connaître un succès évolutif tout à fait spectaculaire. En effet, il est le premier hominidé à sortir d'Afrique pour se répandre dans tout l'Ancien Monde, de l'Europe à l'Extrême Orient. Il est aussi le premier à posséder une capacité crânienne nettement plus grande que celle des chimpanzés et des gorilles, capacité qui va croître de manière continue au cours du temps. Son outillage va aussi se perfectionner progressivement, de même que ses capacités à maîtriser son environnement<sup>8</sup>. Cette évolution continue va aboutir aux différents *Homo sapiens* archaïques, qui apparaissent dans différentes régions du globe il y a quelques centaines de millions d'années, le plus célèbre étant l'Homme de Neandertal, qui a vécu en Europe entre -350 000 et -30 000 ans environ.

Cependant, cette évolution continue ne se poursuit pas jusqu'à nos jours. En effet, entre -100 000 et -80 000 ans, un petit groupe de ces *Homo sapiens*, vivant quelque part en Afrique de l'Est, va prospérer de manière prodigieuse, puisqu'il va en quelques dizaines de milliers d'années conquérir à son tour tout l'Ancien Monde<sup>9</sup>, jusqu'en Australie. En revanche, toutes les autres branches d'*Homo sapiens* archaïques, et notamment les Néandertaliens, disparaissent sans laisser de descendance. Tous les humains actuels, les *Homo sapiens* modernes, sont donc les descendants de ce petit groupe d'*Homo sapiens* africains, dont la réussite tranche avec l'effondrement, encore aujourd'hui mal compris, de la totalité de ses congénères.

### 1.3 L'hypothèse du protolangage

Les scénarios de l'émergence du langage doivent donc s'inscrire dans cette histoire de l'homínisation qui met en évidence deux étapes décisives : l'apparition d'*Homo erectus* il y a près de deux millions d'années et l'apparition de notre espèce proprement dite, l'*Homo sapiens* moderne, il y a quelque cent mille ans.

Il semble raisonnable de faire l'hypothèse qu'avant *Homo erectus*, les hominidés avaient un système de communication du même type que celui des grands singes, et que les premiers changements importants sont corrélés avec le succès évolutif d'*Homo erectus*. C'est en tout cas l'hypothèse que Dereck Bickerton (1990) a formulée, et elle a été assez largement adoptée par la suite. Il a appelé *protolangage* ce nouveau système de communication, et il l'a caractérisé comme un langage dépourvu de syntaxe et de mots grammaticaux, mais doté d'un lexique. Comme on le sait,

---

<sup>6</sup> Cf. sur ce point Lestel (2001).

<sup>7</sup> Pour être précis, les premiers représentants d'*Homo erectus* sont appelés *Homo ergaster*, qui constitue donc une sous-espèce d'*Homo erectus* (encore que, pour certains, il s'agisse d'une espèce à part entière).

<sup>8</sup> On lui attribue notamment la domestication du feu (il y a quelque quatre cent mille ans).

<sup>9</sup> et aussi d'ailleurs le continent Américain, mais plus tard (il y a environ vingt mille ans).

on peut très bien communiquer de l'information à l'aide d'énoncés courts, constitués de quelques mots lexicaux, pour peu que la situation d'énonciation s'y prête. C'est notamment le cas quand les interlocuteurs sont engagés dans une activité commune, et que les objectifs de la communication sont clairement établis, bref quand suffisamment d'éléments extralinguistiques de type pragmatique peuvent suppléer le manque de marques grammaticales et de structures syntaxiques<sup>10</sup>.

Prenant en compte l'évolution progressive des *Homo erectus*, qui s'est poursuivie pendant plus d'un million d'années, Bickerton suggère que le lexique du protolangage, ou plutôt des protolangages (les systèmes de communication des différents groupes d'*Homo erectus* ayant inévitablement divergé après leur dispersion géographique) soit devenu de plus en plus riche. Une augmentation des compétences cognitives de catégorisation auraient accompagné l'évolution de ces hominidés, expliquant en partie leurs plus grandes capacités d'adaptation et d'apprentissage, et cela se serait traduit par une augmentation du stock de mots lexicaux permettant la communication d'informations de plus en plus précises sur leur environnement.

Le passage du protolangage au langage aurait constitué une deuxième étape, qu'il est bien sûr tentant d'associer à l'émergence de notre propre espèce, *Homo sapiens* moderne. L'acquisition de cette nouvelle faculté aurait doté nos ancêtres de capacités symboliques sans aucune commune mesure avec celles des autres *Homo sapiens* archaïques et cela pourrait expliquer leur différence de destinée. Il est alors intéressant de se demander en quoi les propriétés du langage ont pu constituer un avantage sélectif suffisant pour rendre compte du développement spectaculaire des uns et de l'extinction des autres. Cependant ce n'est pas cette question qui va nous occuper ici<sup>11</sup>. Nous allons plutôt nous pencher sur un problème *a priori* plus spécifiquement linguistique : comment doit-on imaginer ce passage du protolangage au langage ? Comme un processus graduel d'enrichissements successifs ou comme un saut brutal ? Autrement dit, peut-on concevoir des états intermédiaires viables entre un protolangage réduit à son lexique et le langage tel que nous le connaissons, ou bien faut-il le voir comme un phénomène indécomposable, l'acquisition d'une première propriété essentielle entraînant forcément la mise en place de tout le système ?

Mais avant de nous attaquer à cette question précise, il nous faut d'abord discuter de la pertinence pour les linguistes de s'occuper de ce genre de sujet.

## 2 L'origine du langage : un problème linguistique

### 2.1 Retour sur un interdit

Que faut-il penser de l'interdit formulé par la Société linguistique de Paris ? Pour reprendre les termes de Sylvain Auroux (2006), qui montre que le débat à ce sujet n'a jamais cessé, l'interdit doit-il être conçu comme conjoncturel, destiné à être levé dès que les conditions s'y prêteraient, ou bien comme indépassable, car motivé par des raisons scientifiques de fond ? On trouve encore aujourd'hui des partisans des deux thèses. Pour Gabriel Bergougnoux, il s'agit d'une question de politique scientifique, donc entièrement conjoncturelle :

---

<sup>10</sup> Voir Victorri (2005) pour une présentation plus détaillée de l'hypothèse de Bickerton.

<sup>11</sup> Nous renvoyons le lecteur intéressé à cette question à Victorri (2002), Dessalles *et al.* (2006) et Dessalles (2007).

*C'est au compromis difficilement négocié entre les fondateurs catholiques, les jeunes universitaires et les hauts fonctionnaires du Ministère de l'Instruction Publique que la SLP dut, jusqu'à son abandon en 1876, l'article 2 de ses statuts qui, moins qu'un dogmatisme de définition du champ scientifique, visait à écarter la concurrence des adhérents de la SAP<sup>12</sup> tout en neutralisant les catholiques venus de la Société d'Ethnographie. (Bergougnieux, 2002, p. 11).*

En revanche, François Rastier, commentant le même article des statuts, écrit :  
*Cet article reflète ainsi une réflexion épistémologique assurée : la linguistique est une science descriptive et historique, qui ne se mêle pas d'imaginer des langues, que ce soient celles de premiers ou des derniers hommes. Une science définit son objet en récusant les faux problèmes dont elle se prive, et cette privation fondatrice la distingue décisivement de la métaphysique. (Rastier, 2006, p. 302).*

Et il ajoute plus loin :  
*La question de l'origine du langage ne se pose pas si l'on admet que le langage est une création culturelle : son histoire n'est autre que celle des langues, et se confond avec celle des sociétés humaines. (Rastier, 2006, p. 305).*

Quant à Auroux, après avoir notamment critiqué les travaux de Meritt Ruhlen, sur lequel nous allons revenir, il conclut :

*A un moment, il faut bien en revenir à la valeur cognitive des discours « scientifiques » : les conséquences théoriques de l'interdit de la Société de linguistique sont infiniment plus solides que ne le laissent supposer quelques affirmations tonitruantes de ces trente dernières années. (Auroux, 2006, p. 85).*

Le fait que l'interdit ait été levé dix ans plus tard plaide néanmoins en faveur de raisons conjoncturelles. Au motif évoqué par Bergougnieux, on peut en ajouter un autre, encore plus pragmatique : il s'agissait sans doute aussi d'écarter les thèses farfelues qui fleurissaient à l'époque sur le sujet, produites par ceux que Marina Yaguello (1984) a appelés « les fous du langage » : des amateurs « éclairés », dopés par les succès de la linguistique historique, dont les élucubrations auraient inutilement surchargé de travail la jeune Société. Les mathématiciens avaient été aux prises avec les mêmes difficultés à propos de la quadrature du cercle, et l'Académie Royale des Sciences de Paris avait eu, elle aussi, recours à un interdit à ce propos en 1775 (cf. Jacob 2006)<sup>13</sup>.

## **2.2 Quelle légitimité pour les linguistes ?**

Cependant, même si l'on admet que l'interdit ait pu être conjoncturel, cela n'invalide pas en soi la position des auteurs qui pensent que les linguistes n'ont aucune légitimité à s'occuper de l'origine du langage.

La critique de Auroux vise avant tout les tenants de « la méthode multilatérale » et de « l'étymologie globale », comme Joseph Greenberg et Meritt Ruhlen, qui ont cherché à dépasser les limites de la méthode comparatiste pour établir l'existence d'une douzaine de « macrofamilles »

---

<sup>12</sup> Société d'Anthropologie de Paris.

<sup>13</sup> Contrairement à ce que laisse entendre Auroux (2006, p. 59), l'interdit de la quadrature du cercle n'a pas pu être édicté en vertu d'une « impossibilité scientifique » en 1775 puisque cette impossibilité ne sera rigoureusement démontrée que plus d'un siècle plus tard (en 1882, quand Lindemann établit la transcendance du nombre  $\pi$ ).

regroupant toutes les langues du monde, et, en ce qui concerne Ruhlen, de remonter même à la langue originelle dont il reconstruit une vingtaine de racines (Ruhlen 1997). On ne peut que donner raison à Auroux sur les faiblesses des méthodes utilisées par ces auteurs. Comme il a été démontré à plusieurs reprises<sup>14</sup>, les affirmations de Ruhlen, en tout cas en ce qui concerne la langue originelle<sup>15</sup>, ne valent guère mieux que celles des fous du langage décrits par Yaguello.

La linguistique historique semble donc effectivement impuissante à contribuer aux recherches sur l'origine du langage. Mais d'autres branches de la linguistique peuvent être mises à contribution. Sur ce point, les déclarations péremptoires de Rastier, affirmant que la question de l'origine du langage ne se pose pas, ou qu'il s'agit d'un faux problème, ne sont absolument pas justifiées. Le simple fait que l'on soit passé d'un état où le langage n'existait pas à un état où il existe (ce que l'on est bien obligé d'admettre, à moins d'être fixiste) suffit pour que la question se pose. Quand Rastier affirme que le langage est une création culturelle dont l'histoire se confond avec celle des sociétés humaines, il défend une position tout à fait respectable, mais on ne voit pas en quoi cela invaliderait tout questionnement sur ce processus de création. D'ailleurs, Rastier lui-même, dans l'article d'où sont issues ces citations, après avoir ironisé sur tous ceux qui ont proposé des scénarios d'origine du langage ces dernières années, y va de sa propre hypothèse, très stimulante par ailleurs<sup>16</sup>. De manière assez surprenante pour un chercheur aussi prompt à dénoncer les erreurs méthodologiques et épistémologiques de ses collègues, il se met alors à utiliser exactement les mêmes méthodes que celles dont il a sévèrement condamné l'utilisation par les autres, ce qui, on en conviendra, affaiblit singulièrement son argumentation<sup>17</sup>.

### 2.3. Quelle place pour les linguistes ?

La légitimité des linguistes dans les entreprises pluridisciplinaires qui se sont mises en place ces dernières années se justifie d'autant plus que la plupart des débats mettent en jeu de manière cruciale la conception du langage que l'on adopte, en ce qui concerne ses propriétés aussi bien structurelles que fonctionnelles. Qui serait mieux placé que les linguistes pour apporter les clarifications théoriques nécessaires à ce propos ? De plus, sur un certain nombre de points, ce sont des réflexions sur les langues elles-mêmes, dans des domaines spécifiquement linguistiques (phonologie, morphologie, syntaxe, sémantique), qui sont à l'origine des hypothèses avancées : le rôle du linguiste

---

<sup>14</sup> Cf., entre autres, Victorri 2000 et Boë *et al.* 2003.

<sup>15</sup> Reste posée la question de savoir si l'on peut remonter plus loin dans le temps que l'ont décrété les comparatistes (pour qui l'horizon temporel du proto-indo-européen, de l'ordre de six à huit mille ans) est indépassable. Depuis que d'autres disciplines, et notamment la génétique des populations, semblent en mesure de reconstituer l'histoire des grandes migrations humaines, on peut en effet imaginer que la linguistique historique pourrait mettre au point d'autres méthodes s'appuyant plus systématiquement sur les connaissances acquises par les autres disciplines. Mais de toute façon, même si l'on double, voire triple la profondeur temporelle ainsi atteinte, il est impensable que l'on puisse aller jusqu'à l'origine des langues humaines, qui rappelons-le, réclamerait de remonter à plus de cinquante mille ans...

<sup>16</sup> Il est vrai que plutôt que de traiter de l'émergence du langage, il préfère parler de « l'émergence du symbolique », mais celui-ci englobe celui-là, comme il le dit explicitement : « Plus que l'origine du langage, il nous paraît plus utile d'explorer les conditions de l'émergence du symbolique et de la constitution propre de l'entour humain, où le langage occupe évidemment une place éminente mais non exclusive. » (Rastier 2006, p. 307).

<sup>17</sup> C'est ainsi qu'après avoir fustigé l'emploi d'arguments s'appuyant sur le développement de l'enfant (« Selon Haeckel, l'ontogenèse récapitulait la phylogenèse. Cette thèse simpliste a été abandonnée depuis un siècle par les sciences de la vie », p. 301), il y a lui-même recours, par exemple pour justifier son hypothèse –intéressante– d'une origine rythmique de la textualité (« Dans l'ontogenèse, non seulement les vocalisations rythmiques précèdent l'apprentissage de la parole, mais elles le conditionnent et l'enfant apprend à parler par répétition de formes mélodiques ». p. 319).

n'est pas alors simplement de contribuer au débat, mais bien d'être l'élément moteur autour duquel doit s'organiser la réflexion.

C'est le cas notamment pour la question que nous avons soulevée ici : doit-on imaginer le passage du protolangage au langage comme un processus graduel d'enrichissements successifs, passant par des états intermédiaires stables ou, au contraire, comme un saut brutal, les caractéristiques essentielles du langage étant trop interdépendantes pour permettre de tels stades intermédiaires ?

Pour Bickerton lui-même (cf. Bickerton 1998), ce passage du protolangage au langage s'est produit brutalement (il parle d'évolution « catastrophique »). Il considère que l'élément décisif a été l'acquisition d'un mécanisme cognitif capable de distinguer les rôles actanciels (qu'il appelle *theta analysis component*) pour que puisse se mettre en place l'ensemble des structures qui manquaient au protolangage. Ce mécanisme aurait d'ailleurs préexisté dans le système cognitif des hominidés, où il se serait forgé au cours du développement de leur intelligence sociale. Le saut aurait consisté en la mise en rapport de ce module préexistant avec le sous-système cognitif responsable de la communication.

Noam Chomsky est aussi à ranger dans la catégorie des partisans du saut brutal. En effet, après avoir longtemps écarté le problème de l'origine du langage en insistant sur le caractère exceptionnel de la faculté du langage en tant qu'organe biologique (Chomsky 1995), il a défendu récemment<sup>18</sup> l'idée que la faculté de langage « au sens étroit » (*FLN : faculty of language in the narrow sense*) pourrait être réduite à une seule capacité computationnelle, la récursivité :

*We propose in this hypothesis that FLN comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mapping to the interfaces.* (Hauser et al. 2002, p. 1573).

Comme chez Bickerton, cette capacité aurait pu préexister à son utilisation dans l'activité de langage :

*It is possible, as we discuss below, that key computational capacities evolved for reasons other than communication, but after they proved to have utility in communication, were altered because of constraints imposed at both the periphery ... and more central levels.* (Hauser et al. 2002, p. 1569-70).

A l'opposé, d'autres auteurs<sup>19</sup> ont défendu très tôt l'idée d'un processus plus progressif, les différentes propriétés spécifiques du langage ayant été acquises les unes après les autres, chacune d'elles constituant un avantage adaptatif dans la mesure où elle représentait une amélioration du système de communication à elle seule. Mais c'est sans aucun doute la contribution de Ray Jackendoff (1999) qui est la plus aboutie. Celui-ci propose en effet une évolution comportant une dizaine d'étapes, allant d'un système de communication du type de ceux des grands singes au langage humain, l'un des stades intermédiaires étant identifié au protolangage de Bickerton. Chaque étape, correspondant à l'acquisition d'un sous-système qui confère au système de communication des propriétés nouvelles, est soigneusement argumentée. Nous allons donc commencer par présenter rapidement ce cadre général, avant de montrer comment le travail de Pierre Le Goffic permet de l'enrichir sur un point précis, à peine effleuré par Jackendoff : l'émergence de la récursivité. Comme

---

<sup>18</sup> dans un article écrit avec deux autres chercheurs, Hauser et Fitch (Hauser et al. 2002).

<sup>19</sup>Cf. notamment Pinker et Bloom 1990 et Pinker 1994.

on le verra, il ne s'agit pas de la récursivité sous la forme générale, globalisante et quelque peu magique mise en avant par Chomsky, mais de la caractéristique des langues dans laquelle la récursivité intervient le plus fortement : l'enchâssement de propositions subordonnées.

### 3 Protosyntaxe, récursivité et subordination

#### 3.1 Les propositions de Jackendoff

La première étape de Jackendoff concerne la faculté de symbolisation. Le système de communication des hominidés se serait d'abord différencié de ceux des grands singes par l'utilisation d'énoncés constitués d'un seul mot-symbole dont la signification dépend du contexte, contrairement aux cris d'alarme des autres primates. Ainsi l'énoncé *sanglier* aurait pu vouloir dire, en fonction de la situation<sup>20</sup>, aussi bien *Tiens, on dirait un sanglier qui s'approche* que *Je reprendrais bien un morceau de sanglier*. Jackendoff fait remarquer que pourraient aussi avoir fait partie de ce système des énoncés comme *oui, non, chut, aïe, zut, salut, etc.*, qui seraient donc autant de « fossiles » de ce premier stade encore présents dans le langage humain.

Deux voies indépendantes de complexification auraient été ensuite empruntées, sans que l'on puisse spécifier dans quel ordre :

- d'une part, le développement d'un système combinatoire phonologique (peut-être d'ailleurs syllabique avant d'être phonémique), à la suite de l'utilisation d'un lexique de plus en plus abondant ;
- d'autre part, le développement d'un mécanisme de concaténation de symboles pour produire des énoncés plus complexes, la position des symboles pouvant, dans un deuxième temps, exprimer certaines relations sémantiques entre eux.

On obtient alors l'équivalent du protolangage de Bickerton. La principale différence avec Bickerton consiste donc en l'existence dès ce niveau d'une protosyntaxe, puisque l'ordre des mots serait déjà en partie contraint. Jackendoff propose notamment un principe général de rapprochement (*Grouping*) des mots plus fortement liés sémantiquement, et deux principes d'ordonnement, qui se seraient d'ailleurs eux aussi fossilisés, sous forme d'universaux, dans nos langues : une règle actancielle stipulant que l'agent serait préférentiellement en tête d'énoncé (*Agent First*) et une règle discursive, suivant laquelle le rhème se situerait plutôt en fin de phrase (*Focus Last*).

En ce qui concerne le passage du protolangage au langage, Jackendoff propose une première étape, capitale, de mise en place d'une structure syntagmatique hiérarchique, les syntagmes, plus ou moins complexes, remplaçant désormais le mot dans l'organisation linéaire de l'énoncé.

Trois types d'innovations sont alors postulées par Jackendoff pour aboutir au langage proprement dit, sans que l'on puisse préciser là encore d'ordre entre elles, étant donnée leur indépendance relative :

- l'utilisation d'un vocabulaire relationnel, composé des unités qu'on appelle mots grammaticaux ou mots outils (prépositions, conjonctions, etc.), qui servent à exprimer explicitement des relations sémantiques abstraites ;

---

<sup>20</sup> et de la prosodie (mais Jackendoff ne mentionne pas explicitement cet aspect).

- la mise en place d'une morphologie flexionnelle (déclinaisons, conjugaisons, etc.), qui contribue notamment (souvent d'ailleurs de manière partiellement redondante avec d'autres mécanismes) à l'expression des relations sémantiques entre constituants ;
- le « reste » de la syntaxe (*further aspects of syntax*) qui complète l'organisation syntaxique de l'énoncé, notamment l'ordonnancement des syntagmes, qui, rappelons-le, n'était qu'ébauché sous la forme de quelques principes de base s'appliquant à l'ordre des mots dans la protosyntaxe.

### 3.2 La récursivité

On notera que Jackendoff, contrairement à Chomsky, n'accorde pas à la récursivité une place centrale, qui en ferait l'unique agent de la spécificité et de la complexité des langues. Au contraire, même s'il ne le mentionne pas explicitement, on peut penser que Jackendoff distribue la récursivité dans au moins deux acquisitions successives, d'abord au moment de la mise en place d'une structure syntagmatique hiérarchique, et ensuite dans la dernière étape où aurait été acquis le reste de la syntaxe. Notamment, on peut penser que c'est au cours de cette dernière étape qu'il situe l'acquisition du mécanisme de subordination.

Un tel découpage semble tout à fait justifié. En effet, on peut soutenir que la récursivité dans les langues est loin d'être une règle générale qui s'appliquerait de manière uniforme aux différents types de constituants.

D'abord un certain nombre de phénomènes syntaxiques qui sont traités comme des phénomènes récursifs par les grammaires génératives ne le sont pas vraiment. Prenons comme exemple l'adjonction d'adjectifs à une tête nominale en français. Le nombre d'adjectifs épithètes que l'on peut ainsi adjoindre est généralement très limité, et leur ordre très contraint. Ainsi, on ne peut pas changer l'ordre des adjectifs dans une expression comme *la troisième jolie petite poupée brune exposée dans la vitrine*. Traiter ces expressions par des règles récursives du type  $N' \rightarrow A N'$  et  $N' \rightarrow N' A$  n'est pas justifié, dans la mesure où pour éviter la surgénération produite par ces règles, on est de toute façon forcé de les restreindre par d'autres règles (sur l'ordre et la position des différents adjectifs en fonction de leur valeur sémantique) qui rendent inutiles les règles récursives en question.

Là où la récursivité se justifie plus parce que le nombre de constituants à engendrer est moins limité, on a souvent affaire à ce que l'on appelle en informatique de la *récursivité terminale* (*tail recursion*), c'est-à-dire une récursivité qui construit un arbre qui branche toujours à droite (ou toujours à gauche), comme dans l'expression suivante : *une poupée exposée dans la vitrine du magasin du quartier*. Cette récursivité est beaucoup plus facile à traiter cognitivement parce qu'elle n'impose qu'une charge mémorielle supplémentaire minimale<sup>21</sup>. On peut penser que ce type de structure hiérarchique est apparu plus tôt que les formes plus complexes d'enchâssement (*center embedding*).

Bien entendu, les syntagmes prépositionnels à l'intérieur d'un syntagme nominal peuvent aussi être enchâssés au milieu les uns des autres, comme dans *la poupée de cire dans la vitrine de gauche du magasin de jouet du quartier*, mais ces enchâssements sont de profondeur très limitée.

<sup>21</sup> D'ailleurs, en informatique, la récursivité terminale est généralement implémentée comme de la simple itération.

Il n'y a vraiment qu'avec la proposition subordonnée que la récursivité au sens fort devient un mécanisme pleinement utilisé. Non pas qu'elle n'ait pas aussi des limites, très vite atteintes dans certaines conditions, comme le montre avec humour le petit article satirique suivant<sup>22</sup> :

*New speech disorder linguists contracted discovered!*

*An apparently new speech disorder a linguistics department our correspondent visited was affected by has appeared. Those affected our correspondent a local grad student called could hardly understand apparently still speak fluently. The cause experts the LSA sent investigate remains elusive. Frighteningly, linguists linguists linguists sent examined are highly contagious. Physicians neurologists psychologists other linguists called for help called for help called for help didn't help either. The disorder experts reporters SpecGram sent consulted investigated apparently is a case of pathological center embedding.*

Néanmoins, le mécanisme de subordination permet de multiplier les niveaux d'enchâssement, et surtout d'introduire à chaque niveau une complexité sans aucune commune mesure avec les seuls emboîtements de syntagmes nominaux et de syntagmes prépositionnels, comme le montre une phrase telle que

*La poupée qui te plaisait tant parce qu'elle secouait la tête quand on cherchait à lui donner à manger en approchant une cuillère de sa bouche a eu visiblement beaucoup de succès si j'en juge par le fait que la boutique où on l'avait trouvée a été littéralement dévalisée entre le jour où on l'a vue, qui, si je ne m'abuse, ne remonte pas à plus d'une semaine, et hier où j'ai voulu en acheter une pour en faire cadeau à la fille du voisin qui nous a dépannés si gentiment quand on s'est retrouvés sans voiture l'autre jour.*

Autrement dit, si l'on considère que la récursivité est une propriété caractéristique du langage, c'est essentiellement la récursivité des structures de subordination que l'on doit prendre en compte puisque ce n'est pratiquement que dans ces structures qu'elle est incontournable en tant qu'opération capable d'engendrer une complexité dont il est impossible de rendre compte par des règles non récursives.

### **3.3 Les travaux de Pierre Le Goffic sur les termes en $*k^w$ -**

Si l'on veut savoir quand et comment les langues sont devenues « pleinement » récursives, il faut donc se demander quand et comment le mécanisme de subordination a émergé. Si l'on suit les propositions de Jackendoff, on doit situer cet avènement dans les toutes dernière innovations au cours du processus de constitution du langage. Mais cela ne nous dit pas comment cela a-t-il pu se produire.

C'est là qu'interviennent les travaux de Le Goffic (2001, 2002) sur les termes indo-européens en  $*k^w$ -. Celui-ci commence par montrer que ces termes (qui ont donné les termes en *qu-* en français, en *wh-* en anglais, etc.) sont à la fois des marqueurs de l'interrogation partielle, des indéfinis et des marqueurs syntaxiques de la subordination. Loin d'être spécifique de l'indo-européen, cette situation semble assez largement répandue dans les langues du monde. Dans la lignée de Culioli (1999), Le Goffic explique cette polyvalence en caractérisant  $*k^w$ - comme marqueur d'une opération cognitivo-linguistique fondamentale qu'il décrit comme un « parcours des valeurs d'instanciation possibles d'une variable ». Il introduit des variables en attente d'instanciation. Ainsi une proposition comme

<sup>22</sup> trouvé sur le web à l'adresse <http://specgram.com/CLI.2/03.bakery.disorder.html>

*Qui est venu* ne peut pas être un énoncé atomique assertif car *qui* reste non instancié. En revanche, ce peut être un énoncé interrogatif (*Qui est venu ?*) marquant que l'instanciation est à la charge de l'interlocuteur. L'assertion, elle, doit prendre la forme *Quelqu'un est venu*, où la forme *qu-* est complétée par un marqueur d'une opération d'extraction qui indique explicitement que la variable est instanciée par une valeur indéfinie. Les autres termes de la famille (*quoi, quel, quand, où*<sup>23</sup>, *comment, combien*) fonctionnent selon le même principe.

Dans les différents types de subordination, les termes en *\*k<sup>w</sup>-* servent de connecteurs liant la subordonnée et la principale par le fait que la variable doit admettre les mêmes valeurs d'instanciation dans les deux propositions. Le Goffic met d'abord en évidence cette propriété caractéristique pour un type particulier de subordination, où elle apparaît le plus clairement. Il s'agit du type qu'il appelle *subordination intégrative prototypique*, dont voici quelques exemples :

*Qui dort dîne.*

*Où tu iras j'irai*

*Quand on veut on peut.*

Il montre alors que de ce type de structure, qui est considéré comme le type par excellence de la phrase complexe indo-européenne<sup>24</sup>, on peut dériver tous les autres types de subordination, y compris les relatives et les complétives. Comme le fonctionnement du terme en *\*k<sup>w</sup>-* dans la subordonnée intégrative fait appel à exactement la même opération de parcours que celle qui est à l'œuvre dans les emplois interrogatifs et indéfinis, il en déduit que le mécanisme de subordination ne consiste en rien d'autre qu'en l'exploitation d'une potentialité de ces marqueurs qui est déjà présente dans leurs autres emplois :

*Replacé dans cette perspective, l'emploi subordonnant apparaît de même niveau que l'emploi interrogatif : c'est encore une sorte d'assertion provisoire, spéculative, à laquelle son énonciateur lui-même apporte viabilité par une seconde prédication en les articulant. A ce compte il n'est pas évident que la subordination intégrative nécessite des opérations qualitativement plus complexes que l'interrogation.* (Le Goffic, 2001, p. 51).

Il conclut alors :

*Les opérations de constitution de l'énoncé (comme la subordination intégrative) sont du même ordre que certaines opérations référentielles (produisant des expressions indéfinies) ou énonciatives (comme l'interrogation) ; la syntaxe ne doit pas être enfermée sur elle-même dans un domaine sui generis, et l'indo-européen, avec ses thèmes *\*k<sup>w</sup>-* (parcours), *\*to-* (déictique, résomptif), *\*yo-* (anaphorique, éventuellement substituable à *\*k<sup>w</sup>-* dans les emplois intégratifs), illustre peut-être une sorte d'équipement cognitivo-linguistique de base, présent sous des formes diverses depuis l'origine du langage, et qui serait à modéliser. On admettra sans doute que l'on ne saurait imaginer une forme de langage humain (une langue) qui ne comporterait pas ce type d'équipement ; alors on ne saurait imaginer de langage humain sans phrase complexe* (Le Goffic, 2001, p. 52-53).

<sup>23</sup> où fait bien partie de la famille des termes indoeuropéens en *\*k<sup>w</sup>-*, même si l'initiale a été perdue dès le latin *ubi*.

<sup>24</sup> Plus précisément, il s'agit du type traditionnellement appelé *corrélatif* (cf. en latin *Quod habuit, id perdidit ; qualis pater, talis filius ; etc.*), mais Le Goffic montre que c'est le terme subordonnant initial qui caractérise la structure, et non le terme corrélatif (largement facultatif d'ailleurs) introduisant la principale.

Cette hypothèse vient donc compléter et enrichir les propositions de Jackendoff, sur le point précis de l'émergence de la subordination : elle serait concomitante de l'acquisition de cet équipement cognitivo-linguistique (vraisemblablement au cours de l'étape d'acquisition des unités grammaticales décrite par Jackendoff). La récursivité qui en résulte, même si elle est plus puissante que les autres formes de récursivité présentes dans le langage, ne serait pas pour autant à mettre au compte d'une capacité computationnelle spécifique, de nature essentiellement syntaxique, comme le postule Chomsky.

## Conclusion

Ainsi ces réflexions de Pierre Le Goffic, issues d'un travail purement linguistique, peuvent contribuer utilement aux recherches sur l'origine du langage, et, au-delà, aux débats théoriques sur la conception du langage, notamment à propos des rapports entre syntaxe et sémantique et des liens entre sémantique grammaticale et opérations cognitives.

Bien entendu, l'hypothèse de Le Goffic, comme il le souligne lui-même à plusieurs reprises, doit être confortée (ou infirmée) par des recherches plus poussées et plus étendues, notamment grâce à « une enquête empirique à mener sur les liens entre interrogatifs, indéfinis, et subordonnants à travers les langues du monde » (Le Goffic 2001, p. 51). Comme nous l'avons dit, c'est l'un des intérêts essentiels de conjectures de ce genre que de motiver des entreprises ambitieuses du type de cette enquête. Et c'est au fond pour cela avant tout que les linguistes ne doivent plus hésiter à s'emparer des débats sur l'origine du langage.

## Bibliographie

- Auroux S. (2006) Les embarras de l'origine des langues, *Marges linguistiques*, 11, p. 58-92.
- Bergounioux G. (2002) La sélection des langues : darwinisme et linguistique, *Langages*, 146, p. 7-18.
- Bergounioux G. (2005) L'origine du langage : mythes et théories, in J.-M. Hombert (éd.), *Aux origines des langues et du langage*, Fayard, p. 14-39.
- Bickerton D. (1990) *Language and Species*, University of Chicago Press.
- Bickerton D. (1998) Catastrophic evolution: the case for a single step from protolanguage to full human language, in Hurford J.R., Studdert-Kennedy M., Knight C. (eds.), *Approaches to the Evolution of Language*, Cambridge University Press, p. 341-358.
- Boë L.-J., Bessière P., Vallée N. (2003) When Ruhlen's "mother tongue" theory meets the null hypothesis, *Proceedings of the XV<sup>th</sup> International Congress of Phonetic Sciences*, Barcelone.
- Chomsky N. (1995) Language and Nature, *Mind*, 104, p. 1-61.
- Culioli A. (1999) *Pour une linguistique de l'énonciation*, tome 3, Ophrys.
- Dessalles J.-L. (2007) *Why We Talk - The evolutionary origins of language*, Oxford University Press.
- Dessalles J.L., Picq P., Victorri B. (2006) *L'émergence du langage*, Editions du Pommier.
- Donald M. (1991) *Origins of Modern Mind: Three Stages in the Evolution of culture and Cognition*, Cambridge, Harvard University Press.
- Hauser M.D., Chomsky N., Fitch W.T. (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298, p. 1569-79.

- Jackendoff R. (1999) Possible stages in the evolution of the language capacity, *Trends in Cognitive Sciences*, 3:7, p. 272-279.
- Jacob M. (2006) *La quadrature du cercle : un problème à la mesure des Lumières*, Fayard.
- Klein R.G. (1999) *The Human Career: Human Biological and Cultural Origins*, University of Chicago Press.
- Le Goffic P. (2001) Sur les sources et le développement de la subordination dans le langage : l'exemple de l'indo-européen, *Recherches en Linguistique Psychologie cognitive*, 16, Presses universitaires de Reims, p. 25-56.
- Le Goffic P. (2002) Marqueurs d'interrogation/indéfinition/subordination : essai de vue d'ensemble, *Verbum*, 24:4, p. 315-340.
- Lestel D. (2001) *Les origines animales de la culture*, Flammarion.
- Pinker S. (1994) *The Language Instinct*, New York, William Morrow and Co.
- Pinker S., Bloom P., (1990) Natural language and natural selection, *Behavioral and Brain Sciences*, 13, p. 707-784.
- Rastier F. (2006) De l'origine du langage à l'émergence du milieu sémiotique, *Marges linguistiques*, 11, p. 297-326.
- Ruhlen M. (1997) *L'origine des langues. Sur les traces de la langue mère*, Belin.
- Tattersall I. (1998) *Becoming Human, Evolution and Human Uniqueness*, New York, Harcourt Brace and Co.
- Victorri B. (2000) La langue originelle, *Science et Avenir*, hors-série 125, pp. 6-11.
- Victorri B. (2002) *Homo narrans : le rôle de la narration dans l'émergence du langage*, *Langages*, 146, p. 112-125.
- Victorri B. (2005) Les « mystères » de l'émergence du langage, in Hombert J.M. (éd.), *Aux origines des langues et du langage*, Fayard, p. 212-231.
- Yaguello M. (1984) *Les fous du langage*, Seuil.