Du protolangage au langage : modèle d'une transition

Jean-Louis Dessalles

▶ **To cite this version:**

Jean-Louis Dessalles. Du protolangage au langage : modèle d'une transition. Marges Linguistiques, 2006, 11, pp.142-152. hal-00614810

HAL Id: hal-00614810
https://hal.science/hal-00614810

Submitted on 27 Aug 2011

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

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

Du protolangage au langage : modèle d'une transition

Jean-Louis Dessalles

Ecole Nationale Supérieure des Télécommunications
46 rue Barrault - 75013 Paris
dessalles@enst.fr

Résumé

L'existence des capacités syntaxiques qui permettent aux êtres humains de manier des langues complexes reste mystérieuse. Pour certains auteurs, ces capacités seraient apparues totalement par hasard au cours de l'évolution et leur application à la communication serait fortuite. Nous essayons ici de montrer comment la modélisation de l'interface syntaxe-sémantique permet d'envisager un tout autre scénario. L'aptitude à manier des structures syntaxiques serait apparue en deux temps et serait liée à une nouvelle capacité sémantique, la formation des prédicats. La récursivité serait apparue lors de la deuxième étape, comme un moyen de lier les prédicats entre eux pour permettre la détermination de leurs arguments.

mots clés : sémantique, syntaxe, évolution, cognition

Abstract

The existence of syntactic abilities allowing human beings to process complex languages remains mysterious. According to some authors, these abilities appeared by mere chance at some point in evolution, and their use in communication is, in some way, fortuitous. We try here to show how a simple model of the syntax-semantic interface allows us to consider a quite different scenario. The ability to process syntactic structures would have appeared in a two-step evolution and would be the consequence of a new semantic ability, the ability to form predicates. Recursion is claimed to have appeared in the second step, as a way to link predicates for their arguments to be determined.

Keywords : sémantique, syntaxe, évolution, cognition

1. Introduction : l'anatomie d'une transition[1]

Le comportement langagier humain repose sur plusieurs capacités cognitives propres à notre espèce. Les animaux n'utilisent pas de système de communication qui soit directement comparable au nôtre. Les tentatives pour leur enseigner différents aspects du langage humain (Premack & Premack 1983 ; Savage-Rumbaugh & Lewin 1994 ; Pepperberg 1999) révèlent que certaines des capacités cognitives qui nous permettent de manier le langage semblent ne pas être présentes chez les animaux testés. L'une de ces capacités, maintes fois soulignée, concerne l'usage d'une syntaxe. Certes, les oiseaux, par exemple les rossignols, sont connus

[1] Ce travail a été effectué dans le cadre de l'initiative OHLL (Origine de l'Homme, du Langage et des Langues) du CNRS.

pour produire des chants présentant plusieurs niveaux de combinatoire (Hauser 1996 :286). Cependant, rien n'indique que les différentes structures ainsi produites servent à exprimer des significations différenciées. Les primates entraînés produisent parfois spontanément des groupements de deux mots sémantiquement reliés au contenu concret exprimé, avec une tendance à observer un ordre systématique (Savage-Rumbaugh & Lewin 1994 :161). Ce type de performance, qui reste éloignée de ce que les êtres humains produisent avec le langage, amène certains auteurs à voir là l'expression d'une capacité protolangagière (Bickerton 1990 ; 1995). Le protolangage a été défini, par extrapolation à partir de l'observation des pidgins, comme une forme d'expression consistant à assembler quelques mots dans un énoncé court, sans support grammatical : absence de mots grammaticaux, absence de dépendance à distance dans la phrase, absence de marques de flexion, absence d'ordre défini (Bickerton 1990). Le protolangage, vers lequel nous régressons spontanément dès que nous sommes en difficulté linguistique, est ainsi présenté par Bickerton comme un précurseur du langage, une sorte de capacité intermédiaire entre la communication spontanée des primates et le langage proprement dit qui est universellement pratiqué dans notre espèce. Une expression moderne du protolangage peut s'observer sur les moteurs de recherches utilisés sur la toile mondiale. Bien que ces moteurs de recherche permettent de s'exprimer à l'aide d'expressions booléennes et reconnaissent des groupes nominaux, les utilisateurs ont tendance à s'exprimer par des suites de mots sans liens grammaticaux entre eux (Guichard 2002).

Quelle que soit l'attitude que l'on adopte vis-à-vis de la réalité du protolangage, la question de la transition menant, au cours de l'évolution, à la forme de langage que nous connaissons, ne peut être ignorée. Certaines de nos aptitudes cognitives nous permettent de manier des constructions langagières signifiantes et de réaliser des opérations non triviales sur ces structures, comme l'interrogation, la mise au passif, le changement modal ou temporel, etc. Sans prendre ici parti sur la question de savoir si ces dispositions cognitives sont spécifiques au langage ou correspondent à des fonctions plus générales (Piattelli-Palmarini 1979), nous cherchons dans la structure du langage des indices susceptibles de nous renseigner sur leur ordre d'apparition. Pour tenter de reconstituer l'anatomie de la transition vers le langage, nous nous intéressons particulièrement à l'interface entre la syntaxe et la sémantique, en adoptant le point de vue de la modélisation. En considérant la complexité relative des opérations de synchronisation entre la structure syntaxique et le sens, nous en viendrons à proposer un ordre dans l'apparition de certaines capacités qui sous-tendent la maîtrise du langage.

Dans la suite, nous commençons par poser le problème en reprenant certaines caractéristiques formelles de la syntaxe qui sont généralement invoquées pour démontrer l'originalité de notre mode de communication, notamment la récursivité et les systèmes de flexion. Nous tentons de dégager un modèle simplifié du rôle de ces caractéristiques dans l'expression du sens. Nous tirerons argument de la structure de ce modèle simplifié pour proposer une succession plausible dans l'établissement des capacités cognitives qui sous-tendent ces aspects du langage. Nous conclurons en replaçant ces propositions dans le cadre des recherches actuelles sur l'origine de la capacité langagière.

2. Quelques aspects notables de la capacité syntaxique

La variété des langues est grande, mais elle n'est pas infinie. En particulier, toutes les langues spontanées du monde, à l'exception des pidgins récents, possèdent une grammaire qui impose la flexion de certains mots et leur agencement dans la phrase. Aucun peuple au monde ne s'exprime spontanément en groupant, sans ordre particulier, des mots invariables[1]. Certes, les règles syntaxiques diffèrent largement d'une langue à l'autre. Certaines langues comme le

---

[1] Les pidgins de création récente constituent clairement une exception. Toutefois, ils sont utilisés comme langue principale seulement dans des conditions sociales anormales qui imposent un brassage d'individus adultes de communautés linguistiques différentes (Bickerton 1990). L'histoire de la formation des créoles et l'exemple des enfants sourds du Nicaragua (Kegl, Senghas & Coppola 1999) indiquent que les pidgins sont une forme de communication instable qui disparaît en tant que moyen de communication principal dès que la structure sociale permet aux enfants de moins de sept ans de former des communautés.

latin ou le dyirbal semblent très permissives en ce qui concerne l'ordre des mots dans les phrases simples (Dixon 1972), alors que d'autres comme le français ou l'anglais sont nettement moins tolérantes sur cet aspect de la grammaire. Certaines langues, comme le lakhota, mettent en œuvre un système morphologique complexe affectant principalement le verbe, alors que des langues comme l'anglais ou le chinois sont très économes en flexions (Van Valin & LaPolla 1997). Malgré ces différences flagrantes, il est possible de faire quelques observations générales sur la structure des langues qui suffiront pour la suite de notre propos.

La première observation porte sur l'existence de dépendances entre les mots. Dans un énoncé protolangagier du type *maison-voisin-feu*, chacun des trois mots entretient un rapport sémantique direct avec la scène perçue décrite. La communication peut fonctionner sans que les mots d'un tel énoncé ne dépendent les uns des autres. Les langues humaines introduisent plusieurs formes de dépendance, qui sont clairement indiquées par la grammaire. Dans la phrase *la maison du voisin est en feu*, les mots entretiennent des relations formelles, indépendamment de leur signification. Ces relations, par exemple, rapprochent la préposition *en* et le nom *feu*, ou les noms *maison* et *voisin*, alors qu'elles laissent *voisin* et *feu* sans lien direct. Ces dépendances priment souvent sur le sens : la grammaire d'une expression comme *la souris mange le chat* impose une relation sémantique qui va à l'encontre de ce que suggère la simple association entre les sens évoqués par les mots *souris*, *chat*, *manger*.

La deuxième observation concerne le fait que les dépendances syntaxiques sont asymétriques. Dans une expression latine comme *domus domini*, le génitif du mot *dominus* indique sans ambiguïté que ce mot est un satellite du mot *domus*. En français, les deux expressions *la maison du maître* et *le maître de la maison* se distinguent par la position relative des mots par rapport à la préposition *de*. Cette dépendance asymétrique s'interprète de diverses manières, comme résultant d'une opération de fusion syntaxique (Chomsky 1995) ou d'une dépendance sémantique (Van Valin & LaPolla 1997). Du fait de la présence de dépendances asymétriques, les énoncés langagiers comportent des syntagmes, dont la tête domine, au sens structurel, l'élément dépendant.

La troisième observation concerne la structure "moléculaire" des énoncés langagiers. Comme les relations de dépendance peuvent s'enchaîner, les syntagmes se retrouvent emboîtés les uns dans les autres comme des poupées russes. Dans une phrase simple comme *J'ai moyennement aimé l'acoustique du concert auquel j'ai assisté hier*, on trouve maints emboîtements : le syntagme verbal centré sur le verbe *assister* dépend (indirectement) du syntagme nominal centré sur le mot *concert*, qui fait partie intégrante du syntagme nominal centré sur *acoustique*, lui-même inclus dans le syntagme verbal dont la tête est le verbe *aimer*. Les différentes théories syntaxiques peuvent différer par le choix précis des types de syntagmes, mais toutes admettent le principe de l'imbrication. Il en résulte une structure qui peut être qualifiée d'arborescente (si l'on représente le graphe des dépendances), de moléculaire (par analogie avec la chimie des macro-molécules, qui sont constituées d'autres molécules), de fractale (par référence aux structures vivantes ou physiques invariantes par changement d'échelle) ou de récursive. Ce dernier qualificatif fait référence à une propriété de la procédure qui permet de lire ou d'engendrer la structure syntaxique. Les procédures récursives ont la propriété d'appeler une copie d'elles-mêmes au cours de leur exécution. C'est donc la même procédure qui permet d'analyser le syntagme nominal *acoustique du concert auquel j'ai assisté hier* et le syntagme nominal inclus *concert auquel j'ai assisté hier*. D'un point de vue technique, le fait remarquable est que la récursivité à l'œuvre dans le langage humain est une récursivité centrale, c'est-à-dire une récursivité qu'il est impossible de ramener à une simple itération répétitive.

La dernière observation que nous faisons ici à propos des dépendances syntaxiques est qu'elles s'exercent à distance. Dans *l'acoustique du concert auquel j'ai eu la chance d'assister hier*, on comprend que le nom *concert* sert de complément à la fois au nom *acoustique* et au verbe *assister*. Pourtant, *concert* et *assister* sont séparés de sept mots. La distance de ce type de dépendance peut être arbitrairement grande (*le concert auquel la fille du copain que j'ai connu quand j'étais à la fac a eu la chance d'assister avait une mauvaise acoustique*). L'interprétation de ces dépendances à distance varie d'une théorie à l'autre, selon que l'on parle de mouvement (le complément de *assister* migrant en tête de clause) ou de duplication silencieuse.

Pour plusieurs auteurs, certaines de ces propriétés syntaxiques sont indissociables. L'existence de dépendances asymétriques, l'emboîtement des syntagmes et les différentes fo

rmes de dépendance à distance iraient de pair, elles constitueraient une sorte de paquet cadeau que la Nature nous aurait offert en nous dotant de la capacité syntaxique. L'ensemble de ces propriétés résulterait de la seule opération de fusion (Berwick 1998) ou, de manière équivalente, de l'apparition fortuite dans notre espèce d'un accès à la récursivité (Hauser, Chomsky & Fitch 2002). De fait, il est difficile d'imaginer des intermédiaires dans l'accès à la récursivité. Dans ces conditions, la transition du protolangage (ou du non-langage) vers le langage ne peut être que soudaine. Certains ont pu invoquer l'intervention d'une improbable macro-mutation (Bickerton 1990). Sa cause aurait été purement fortuite, et ne devrait rien à un quelconque accroissement du pouvoir communicationnel (Chomsky 1975 :75 ; Piattelli-Palmarini 1989 ; Lightfoot 2000). La nouvelle capacité récursive, peut-être à la faveur d'un décloisonnement fonctionnel (Hauser, Chomsky & Fitch 2002), aurait entraîné l'ensemble des propriétés de la syntaxe qu'il nous est donné d'observer dans les différentes langues, avec parmi elles les propriétés de dépendance que nous avons soulignées ici.

La question de la transition vers la capacité syntaxique, s'il s'est agi d'un changement abrupt et fortuit, perd beaucoup de son intérêt scientifique, car elle n'a pu produire qu'un dispositif peu adapté aux utilisations que les êtres humains en font.

> Language design as such appears to be in many respects "dysfunctional," yielding properties that are not well adapted to the function language is called upon to perform. (Chomsky, 1995:162).
>
> […] full-blown evolutionary novelty can also suddenly arise, so to speak, for no reason, because novelty caused by sheer proximity between genes is not governed by function and it, therefore, eludes strict adaptationism. » (Piattelli-Palmarini 1989:8)

Le point de vue selon lequel l'évolution pourrait ainsi produire des structures ou des compétences complexes peu fonctionnelles apparaît comme non fondé sur le plan biologique (Dessalles 1996), et ne saurait s'appliquer au langage (Pinker & Bloom 1990 ; Pinker & Jackendoff 2005). La théorie de la sélection naturelle et l'étude cognitive du langage suggèrent toutes deux que la syntaxe remplit une fonction précise, qui est celle de l'expression du sens sur le canal de l'expression vocal, et que sa structure est la conséquence évolutive de cette fonction. Il est donc légitime d'étudier la fonction de la syntaxe pour tenter de proposer une reconstitution des grandes étapes par lesquelles est passée l'évolution du langage.

Dans la suite, nous allons nous attacher à remettre en question ces deux propriétés supposées de la transition du protolangage au langage que seraient son caractère abrupt et son caractère fortuit. En observant, du point de vue de la modélisation, l'interface entre la syntaxe et la sémantique, nous allons suggérer une transition en deux étapes au moins, dont chacune serait motivée, sur le plan biologique, par ses conséquences sur le type de communication qu'elle rend possible.

3. À l'interface entre la grammaire et le sens

La manière dont le sens d'un énoncé est construit à partir de sa structure grammaticale a fait l'objet de nombreux travaux et continue à être l'enjeu de nombreuses recherches. Les paradigmes divergent largement, allant d'une conception mettant en jeu deux constructions isomorphes (Montague 1974 ; Jackendoff 1990 ; Carpenter 1997) à diverses conceptions dans lesquelles la structure grammaticale ne fait qu'imposer des contraintes à une construction sémantique qui se déroule en parallèle (Johnson-Laird 1977 ; Kamp & Reyle 1993 ; Talmy 2000). Pour notre propos, une théorie globale de la construction du sens n'est pas nécessaire. Il nous suffit de considérer certains aspects de la synchronisation entre le code syntaxique et le niveau sémantique.

Le premier aspect de cette synchronisation concerne l'identification de la relation prédicat-arguments. Nous nous plaçons dans l'hypothèse où la représentation sémantique d'un énoncé

met en jeu un prédicat logique[1]. Cette hypothèse est compatible avec la plupart des modèles de la sémantique[2]. Il est important de noter que l'hypothèse n'implique pas l'existence d'un langage de la pensée, dans la mesure où le prédicat en question peut être une représentation transitoire, construite en contexte à partir d'éléments non symboliques (Ghadakpour 2003).

Prenons l'exemple d'un prédicat à deux places *P(x,y)*. Ce sera par exemple un prédicat associé à l'énoncé *Pierre frappe Paul*. L'une des fonctions fondamentales de la syntaxe consiste à offrir les moyens de repérer *x*, *y* et *P* dans l'énoncé. En particulier, il s'agit de permettre de distinguer les énoncés associés à *P(x,y)* et à *P(y,x)*. Les langues utilisent divers moyens pour parvenir à ces identifications des actants. Les plus évidents sont la position et le marquage flexionnel. Ainsi, dans une langue comme le français, la position des mots est suffisamment stricte pour permettre le repérage du prédicat et de ses arguments, et pour distinguer par leur sens l'énoncé *Pierre frappe Paul* et l'énoncé *Paul frappe Pierre*. En latin, la présence des cas nominatif et accusatif rend le même service dans un énoncé comme *Paulus Petrum ferit*. La grammaire d'une langue donnée permet également de repérer arguments et prédicat dans le cas de prédicats à trois places (*Paul vend sa voiture à Jacques*), à une place (*Paul est malade*) ou à zéro place (*il pleut*).

Dans la précédente section, nous présentions l'existence de dépendances asymétriques au sein de la phrase comme une caractéristique fondamentale du langage. L'un des effets fondamentaux de ces dépendances est précisément de permettre l'identification de la relation entre le prédicat et ses arguments. En d'autres termes, le syntagme constitue le reflet visible de la prédication. Appelons fonction argumentale cette propriété de la syntaxe.

Le second aspect de la synchronisation syntaxe-sémantique auquel nous allons nous intéresser est la détermination. Pour être utilisable au niveau pragmatique, la représentation sémantique doit être suffisamment instanciée. Dans certains contextes, ce n'est pas le fait qu'un quelconque être vivant ait mangé quelque chose qui importe, mais bien que le chat de la maison a mangé le rôti destiné aux invités. Dans un autre contexte, par exemple celui d'un malade qui recouvre l'appétit, l'identité du mangeur est pertinente, alors que l'entité mangée peut rester non précisée. Voyons comment la syntaxe se met au service de la détermination des arguments des prédicats.

Les dispositifs syntaxiques permettant la détermination des arguments logiques sont variés. Intéressons-nous à deux d'entre eux, la modification et la relativisation. Dans l'énoncé :

Va me chercher le petit livre que Paul a offert à Jacques,

la modification par l'adjectif *petit* et la clause relative *que Paul a offert à Jacques* aident l'interlocuteur à mieux comprendre de quel livre il s'agit. Si l'on représente le sens correspondant à l'aide de prédicats, on obtient une représentation du type *Petit(x) & Livre(x) & Offert(Paul,x,Jacques)*, en supposant que l'auditeur de l'énoncé soit capable de former, dans le contexte courant, les prédicats notés *Petit*, *Livre* et *Offert*. Dans ce type de représentation (qui rappelle le langage informatique Prolog), il est convenu que deux variables de même nom, au sein de la même clause logique, désignent le même objet. On pourrait imaginer que les êtres humains s'expriment de cette manière, à l'aide de propositions simples et de variables :

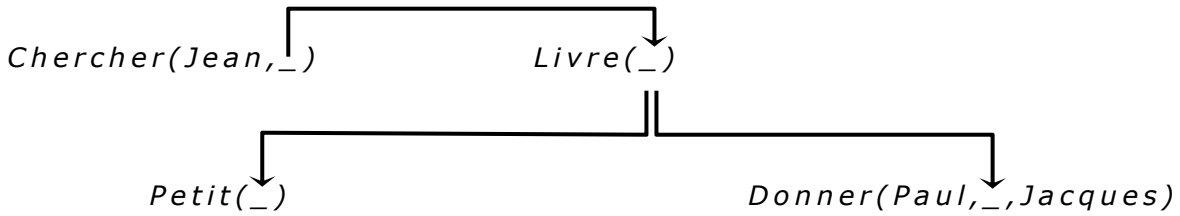
Va me chercher quelque chose ;
ce quelque chose est un livre ;
ce quelque chose est petit ;
Paul a offert ce quelque chose à Jacques.

---

1 Les prédicats logiques sont le moyen privilégié par lequel le sens des énoncés est interfacé avec les aspects pragmatiques, notamment l'argumentation (Dessalles 2000).

2 En particulier, il est possible de convertir en prédicats logiques les constructions produites dans les modèles qui utilisent des formalismes à base de traits, d'éléments logiques, d'éléments fonctionnels, de graphes ou d'éléments schématiques. Les modèles qui ne font intervenir que des représentations analogiques du sens, par exemple des images, pourraient faire exception.

Ce mode d'expression par partage de variables entre propositions simples semble lourd à la plupart d'entre nous (bien qu'une minorité d'informaticiens utilisateurs du langage Prolog l'affectionnent pour exprimer des connaissances précises). La syntaxe des langues naturelles permet le plus souvent de faire l'économie de variables explicites, conformément au schéma suivant.



Dans ce schéma, les flèches représentent la liaison sémantique qui existe entre les syntagmes. Cette liaison est obtenue grâce à un certain nombre de dispositifs syntaxiques, parmi lesquels le principe de liaison sémantique qui stipule que deux syntagmes en relation syntaxique (complémentation ou adjonction) doivent partager une variable au niveau sémantique[1] (Dessalles 2000). Ce principe correspond aux liaisons entre fragments de discours qui apparaissent dans un formalisme comme celui de la DRT (Kamp & Reyle 1993). Le principe de liaison sémantique n'est pas un moyen infaillible d'assurer la liaison entre arguments logiques. En français, le syntagme *le dépassement de la voiture* est sémantiquement ambigu, car on ne sait pas si la voiture en question dépasse ou est dépassée. Ce phénomène s'explique par le fait que le principe de liaison sémantique sous-spécifie, dans ce cas, l'argument connecté dans *Dépasser(_,_) & Voiture(_)*, où l'un quelconque des deux arguments de *Dépasser* peut être lié à l'argument de *Voiture*.

La fonction essentielle du principe de liaison sémantique et des différents dispositifs syntaxiques qui le servent est de permettre la détermination des objets du discours. Si l'on conçoit chacun des prédicats comme une contrainte posée par le locuteur sur le monde perçu, chaque liaison avec un nouveau prédicat limite le champ des possibilités, et donc facilite la compréhension pour l'interlocuteur. Identifier l'objet du discours consiste à résoudre une équation comme *Petit(x) & Livre(x) & Offert(Paul,x,Jacques)*, où *x* est l'inconnue. Un locuteur pertinent s'attachera à donner suffisamment de contraintes prédicatives pour que la solution apparaisse comme unique à son interlocuteur (l'utilisation d'un déterminant défini annonce cette unicité pour l'interlocuteur). Dans un contexte déjà contraint, exprimer simplement *Livre(x)* pourra suffire. Le principe de liaison sémantique permet de mieux comprendre la raison d'être de la récursivité syntaxique. Chaque nouveau prédicat introduit pour spécifier l'argument d'un autre prédicat peut se voir, à son tour, adjoindre un nouveau prédicat qui précisera ses propres arguments. Il en résulte une structure sémantique arborescente qui lie les prédicats les uns aux autres.

Le principe de liaison sémantique agit localement sur des syntagmes proches, compléments ou adjoints. Paradoxalement, ce même principe de liaison sémantique permet de comprendre la raison d'être de certaines dépendances à distance. Notamment, il confère un rôle au mouvement (ou à la duplication silencieuse) des syntagmes nominaux, lorsqu'un rapprochement doit être obtenu en dépit d'une trop grande distance structurelle. Dans le syntagme *l'acoustique du concert auquel j'ai eu la chance d'assister hier*, le problème consiste à lier les arguments de *Acoustique(_)*, de *Concert(_)* et de *Assister(_,_)*. Or, le mot *concert* ne peut être complément à la fois du nom *acoustique* et du verbe *assister*. Le pronom *auquel*, qui renvoie au complément silencieux de *assister*, permet d'assurer cette double liaison. Le principe de liaison sémantique s'applique entre *acoustique* et son complément

---

[1] Ce principe doit s'appliquer avec discernement. Par exemple, il demande de distinguer un emploi prédicatif de la proposition *sur*, dans *elle voit le livre sur la table*, d'un emploi non prédicatif comme dans *elle compte sur Paul*. Dans le deuxième cas, il n'y a qu'un seul prédicat, et donc aucune liaison sémantique.

concert, et entre *concert* et le pronom *auquel*[1]. La liaison sémantique est ensuite assurée par le lien direct entre le pronom *auquel* et le complément silencieux de *assister*. Si l'on analyse de manière classique notre exemple comme résultant d'un mouvement en début de clause du complément de *assister*, ce déplacement a pour rôle d'amener le mot *concert* en position d'établir un lien sémantique avec le mot *acoustique.*[2]

Nous venons d'évoquer deux grandes fonctions assurées par la syntaxe, la fonction argumentale et la détermination. Dans un scénario où la capacité à gérer les relations syntaxiques serait un pur produit du hasard, parachutée, pour ainsi dire, dans le paysage cognitif humain, l'existence de ces deux fonctions, et le fait que la syntaxe les assure plutôt bien, apparaissent comme miraculeux. Nous proposons d'inverser la perspective, et de voir l'aptitude à manier la syntaxe, non comme un événement miraculeux dépourvu de toute cause, mais comme le résultat d'une évolution en plusieurs étapes.

4. Du modèle à la reconstitution évolutive

Du point de vue de la modélisation, la syntaxe apparaît comme un appareil bien agencé, capable notamment de remplir deux fonctions essentielles, l'identification des arguments des prédicats et la détermination des objets du discours par liaison sémantique entre prédicats. Il est naturel de chercher à aller au-delà de ce constat, en s'interrogeant sur l'origine phylogénétique de cet appareil. Si l'on renonce à considérer l'aptitude syntaxique comme une capacité cognitive monobloc fournie clé en main par un hasard de l'évolution, il devient pertinent de tenter une reconstitution de l'ordre dans lequel les composantes de cette capacité sont apparues au cours de la phylogenèse.

La démarche que l'on peut adopter dans ce type de recherche est identique à celle des anatomistes évolutionnistes, qui distinguent les caractères ancestraux des caractères dérivés. Lorsqu'un caractère morphologique ne peut exister qu'en présence d'un autre caractère morphologique, il est nécessairement dérivé. Les ocelles qui apparaissent sur les ailes des papillons sont des caractères dérivés, apparus après les ailes au cours de la phylogenèse. Dans le cas des capacités à manier les relations syntaxiques, il est légitime de se demander laquelle, des deux fonctions que nous avons considérées, a été assurée en premier.

Posée sous cette forme, la question semble devoir recevoir une réponse dépourvue d'ambiguïté. L'identification des arguments des prédicats est sans conteste un préalable à la possibilité de lier les prédicats entre eux. En d'autres termes, la maîtrise des outils syntaxiques qui nous permettent d'identifier les arguments des prédicats *a dû précéder* l'utilisation du principe de liaison sémantique à des fins de détermination.

La conséquence la plus importante de ce constat est que l'aptitude à manier les relations syntaxiques a pu apparaître en deux temps, *la récursivité n'intervenant que dans le second temps*. Le scénario est le suivant. Premier temps, apparaissent les clauses simples n'exprimant qu'une seule prédication. Cette prédication peut être qualifiée de principale : c'est celle qui est utilisée au niveau pragmatique pour l'argumentation (Dessalles 2000). Ainsi, l'énoncé *Pierre a frappé Paul* est directement exploitable par l'auditeur qui peut adopter une attitude conversationnelle, voire s'engager dans une action. À ce stade, la prédication trouve une trace explicite dans le langage, alors qu'elle était absente du protolangage (dans une expression protolangagière du type *combat Pierre Paul*, où les trois mots ne sont ensemble que parce qu'ils se rapportent à la même scène, la constitution d'une éventuelle prédication est entièrement à la charge de l'auditeur, qui doit de plus deviner qui est le frappeur et qui est le frappé).

---

1 Le statut du lien entre ce type de pronom et son antécédent varie d'un modèle à l'autre. Il nous suffit ici de considérer que ce lien transmet la liaison sémantique.

2 En français, la préposition "migre" avec le complément, alors qu'en anglais, elle peut rester sur place, apparaissant alors comme une postposition.

Dans un deuxième temps, la capacité d'exprimer des prédicats acquiert une fonction dérivée. Les énoncés continuent d'inclure une prédication principale, celle qui est utilisée par le niveau argumentatif, mais ils comportent maintenant des prédications auxiliaires servant à la détermination. Ces prédications auxiliaires apparaissent notamment sous la forme de modifications par adjonction ou complémentation. Grâce au principe de liaison sémantique, les prédications auxiliaires peuvent jouer leur rôle de spécification des objets du discours[1]. Le phénomène remarquable est que la spécification des arguments concerne non seulement la prédication principale, mais également les prédications secondaires, dont les arguments demandent eux aussi à être spécifiés. Il en résulte un enchaînement récursif des prédications auxiliaires qui, grâce au principe de liaison sémantique, peuvent s'enchaîner les unes aux autres. C'est donc à ce stade que la capacité de gérer les enchaînements récursifs de syntagmes trouve sa raison d'être.

5. Discussion

Le scénario que nous venons d'esquisser a pour principal mérite de rendre inutile l'hypothèse improbable selon laquelle la capacité syntaxique serait arrivée en bloc par hasard, pour trouver ensuite une utilité dans la communication. La prise en compte de la synchronisation syntaxe-sémantique permet de distinguer deux aptitudes fondamentales, apparues successivement, dans la capacité syntaxique, chacune étant motivée par un rôle fonctionnel. La première a permis d'exprimer la prédication principale, en mettant en œuvre un énoncé simple isolé pour chaque prédication. Il n'est pas besoin d'invoquer un quelconque événement miraculeux ici. Plusieurs études de simulation évolutive ont montré que des formes de syntaxe pouvaient émerger dans un contexte où il est utile d'exprimer des prédications (Batali 1994, 1998 ; Kirby 2002). Le rôle de la syntaxe, dans ce cas, est de donner des indices visibles des dépendances asymétriques entre le prédicat et ses arguments. De ce point de vue, les différents systèmes flexionnels (systèmes de cas, de genres, de classes, etc.) apparaissent comme concurrents des systèmes de position (Jackendoff 1999 ; 2002). S'il existe deux dispositifs en compétition pour la même fonction, la satisfaction de cette fonction ne tient vraisemblablement pas du miracle, contrairement à ce que supposent les auteurs cités précédemment (cf. section 2).

Ce qui motive certains à penser que la maîtrise de la syntaxe est le fruit d'une innovation unique et improbable est lié au caractère moléculaire des énoncés langagiers. La capacité d'emboîter les syntagmes les uns dans les autres, de manière récursive, ne semble pas pouvoir être acquise de manière progressive. Dans notre scénario, cette capacité apparaît dans la seconde étape, lorsque les prédications auxiliaires sont utilisées comme contraintes de détermination. Nous ne savons pas si l'émergence d'une syntaxe récursive est spécialement improbable lorsqu'il s'agit d'exprimer des prédications auxiliaires sémantiquement liées. Nous pouvons cependant mentionner deux indices qui limitent cette improbabilité. Premièrement, certains chercheurs ont pu faire émerger des syntaxes récursives dans d'autres contextes, pour répondre à un problème d'économie de formes (Nowak, Plotkin & Jansen 2000) ou de généralisation (Kirby 2002). Deuxièmement, la syntaxe récursive n'est pas le seul moyen d'exprimer l'enchaînement des prédications auxiliaires. On peut parfaitement imaginer un mode d'expression à la Prolog fonctionnant par partage de variables (cf. section 3). Si un tel système n'a pas émergé pour servir l'expression humaine spontanée, c'est sans doute que le principe de liaison sémantique est apparu plus rapidement, dotant les énoncés d'une structure moléculaire et de mécanismes de dépendance à distance.

Certains travaux présentent la syntaxe comme une innovation ayant permis de répondre aux besoins d'une pression lexicale accrue. Devant l'augmentation du nombre des significations à communiquer, la combinaison syntaxique apparaît comme un moyen de limiter

---

1 De manière marginale, les prédications auxiliaires peuvent jouer un rôle argumentatif. Dans *J'ai vu la jolie Marie qui était très pressée*, les prédications introduites par l'adjectif *jolie* et par la relative *qui était très pressée* ne servent pas à spécifier les objets du discours, mais expriment des éléments de connaissance à des fins argumentatives. Dans notre scénario, cet usage des prédications auxiliaires est lui-même un usage dérivé par rapport à la détermination.

le nombre formes lexicales à mémoriser (Kirby 2000; Nowak, Plotkin & Jansen 2000). On comprend, d'après ce qui précède, que la pression lexicale justifie, au mieux, l'apparition du protolangage. Les énoncés du protolangage, en combinant les significations de mots juxtaposés, permettent d'évoquer des scènes pour lesquelles les locuteurs ne disposent pas de mots uniques adéquats (Dessalles 2000). La syntaxe réalise une performance bien différente. Elle permet d'exprimer la prédication. Selon ce scénario, le protolangage était limité à l'évocation de scènes concrètes. La transition du protolangage au langage a constitué une véritable innovation, par laquelle nos ancêtres, en accédant à la prédication, se sont engagés dans la voie de l'argumentation (Dessalles 2000).

La capacité sémantique de prédication a souvent été ramenée à l'activité de catégorisation, qui passe par la mise en relation d'un objet localisé avec un ensemble de propriétés (Hurford 2003). Présentée de cette manière, la prédication semble chose bien banale, et doit être l'apanage de nombreux animaux. Nous avons souligné ailleurs la différence fondamentale qui existe entre la capacité de catégorisation et la prédication argumentative exprimée par le langage[1] (Dessalles & Ghadakpour 2003). La prédication langagière serait, elle, une acquisition récente de notre espèce, dont l'avènement a entraîné, en deux phases, l'apparition de la capacité à maîtriser les relations syntaxiques.

En proposant ce scénario, nous sommes conscients de renvoyer une partie du mystère de l'évolution des capacités langagières dans le niveau sémantique. Si l'expression syntaxique de la prédication principale, puis des prédications auxiliaires, cessent de sembler miraculeuses, l'improbabilité de la transition du protolangage au langage repose maintenant entièrement sur les capacités cognitives permettant de penser de telles prédications. Nous assumons pleinement ce changement de perspective, que nous pouvons résumer de la manière suivante.

- Premier temps : apparition d'une nouvelle capacité sémantique, permettant aux individus de former des relations prédicatives à des fins argumentatives (Dessalles 2000). Une première forme de syntaxe apparaît alors pour exprimer ces prédications (un énoncé simple par prédication). Les relations entre le prédicat et ses arguments sont assurées par des systèmes comme la flexion (cas, genres, classes, …) ou la position.

- Deuxième temps : apparition d'une nouvelle utilisation sémantique de la prédication. Les prédicats peuvent être utilisés comme des contraintes au service de la détermination des objets du discours. De nouveaux outils syntaxiques apparaissent alors pour exprimer ces prédications auxiliaires. Le principe de liaison sémantique est utilisé pour lier localement les prédications entre elles. L'emboîtement récursif des syntagmes et les dépendances à distance émergent comme des moyens efficaces d'exprimer les enchaînements de prédications.

Compte tenu des résultats qui ont pu être obtenus dans les simulations d'évolution langagière, il est permis de penser que l'apparition des outils syntaxiques ne présente pas de caractère miraculeux, dès lors que l'on suppose données les capacités sémantiques correspondantes. En revanche, il est difficile, dans l'état actuel des connaissances, d'estimer les chances d'apparition de la capacité de prédication.

6. Conclusion

La modélisation de la synchronisation entre syntaxe et sémantique nous a permis de distinguer deux fonctions importantes de la syntaxe : la fonction argumentale et la détermination. Nous avons suggéré que ces deux fonctions étaient apparues successivement. La première, obtenue notamment par des mécanismes de flexion et de position, permet l'expression de la prédication principale, utilisée par le niveau argumentatif. La seconde, obtenue par l'emboîtement des syntagmes et les dépendances à distance, permet l'accumulation de prédications auxiliaires qui servent à contraindre les objets du discours pour l'interlocuteur.

---

[1] Une différence fondamentale est que la prédication langagière peut être systématiquement et explicitement niée en contexte : *cet objet n'est pas un livre parce qu'il n'est pas en papier*. La prédication langagière peut être formée en contexte par une opération de contraste entre représentations (Ghadakpour 2003).

Le scénario qui est proposé ici repose sur une modélisation simplificatrice de l'interface entre syntaxe et sémantique. La syntaxe comporte bien d'autres raffinements, que nous avons volontairement ignorés. Notre objectif était de proposer un cadre simple, un scénario en deux étapes, sur lequel il est ensuite possible de greffer d'autres hypothèses concernant l'apparition d'autres caractéristiques du langage. Nous avons voulu marquer la différence entre les outils requis pour l'expression de la prédication principale et celle des prédications annexes, en soulignant que la récursivité et les dépendances à distance ne sont utiles qu'à la seconde de ces fonctions.

Le moteur de la transition, selon notre scénario, n'est donc pas l'apparition d'une nouvelle capacité syntaxique, mais l'apparition d'une nouvelle capacité sémantique. En formant des prédicats, en les exprimant par des énoncés simples, puis par des énoncés moléculaires pour atteindre une détermination poussée des objets du discours, nos ancêtres ont pu affirmer des relations et surtout remettre en question les relations affirmées par les autres. Conformément à ce qui est généralement admis, nous avons tout lieu de penser que cette transition correspond à l'avènement d'*homo sapiens*.

Remerciements

Je remercie Laleh Ghadakpour qui a bien voulu commenter une version antérieure de cet article. Je voudrais exprimer ma gratitude aux personnes qui sont à l'origine de l'initiative "Origine de l'Homme, du Langage et des Langues" du CNRS, qui ont permis à la problématique des origines du langage de retrouver une juste place dans le débat scientifique français.

Références

Batali, J. (1994). "Artificial Evolution of Syntactic Aptitude". In A. Ram & K. Eiselt (Eds), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, N.J. : Lawrence Erlbaum Associates, 27-32.

Batali, J. (1998). "Computational simulations of the emergence of grammar". In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds), *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press, 405-426.

Berwick, R. C. (1998). "Language evolution and the minimalist program: the origin of syntax". In J. R. Hurford, M. Studdert-Kennedy & C. Knight (Eds), *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press, 320-340.

Bickerton, D. (1990). *Language and species*. Chicago : University of Chicago Press.

Bickerton, D. (1995). *Language and human behavior*. London : UCL Press, ed. 1996.

Carpenter, B. (1997). *Type-logical semantics*. Cambridge, MA : M.I.T. Press.

Chomsky, N. (1975). *Réflexions sur le langage*. Paris : Flammarion, ed. 1981.

Chomsky, N. (1995). *The minimalist program*. Cambridge (Ma) : MIT Press.

Dessalles, J-L. (1996). *L'ordinateur génétique*. Paris: Hermès.
http://www.infres.enst.fr/~jld/papiers/pap.evol/96061301.html

Dessalles, J-L. (2000). *Aux origines du langage : Une histoire naturelle de la parole*. Paris: Hermès-sciences.
http://www.enst.fr/~jld/papiers/pap.evol/99111703.html

Dessalles, J-L. & Ghadakpour L. (2003). "Object recognition is not predication - Commentary on James R. Hurford: 'The neural basis of predicate-argument structure'". *Behavioral and Brain Sciences* 26(3), 290-291.
http://www.ling.ed.ac.uk/~jim/BBSNEURO/dessalles.html

Dixon, R. M. W. (1972). *The Dyirbal language of North Queensland*. Cambridge : Cambridge University Press.

Ghadakpour, L. (2003). Le système conceptuel, à l'interface entre le langage, le raisonnement et l'espace qualitatif: vers un modèle de représentations éphémères. Paris: Thèse de doctorat, Ecole Polytechnique.
http://www.enst.fr/~jld/theses/laleh

Guichard, E. (2002). *L'internet : mesures des appropriations d'une technique intellectuelle*. Paris : Thèse de doctorat de L'École des hautes études en sciences sociales.

Hauser, M. D. (1996). *The evolution of communication*. Cambridge : The MIT Press.

Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002). "The faculty of language: what is it, wo has it, and how did it evolve ?". *Science*, *298*, 1569-1579.

Hurford, J. R. (2003). "The neural basis of predicate-argument structure". Behavioral and Brain Sciences 26(3).
http://www.ling.ed.ac.uk/~jim/newro.htm

Jackendoff, R. (1990). *Semantic structures*. Cambridge : The MIT Press.

Jackendoff, R. (1999). "Possible stages in the evolution of the language capacity". *Trends in cognitive sciences*, *3*(7), 272-279.

Jackendoff, R. (2002). *Foundation of language - brain, meaning, grammar, evolution*. Oxford: Oxford university press.

Johnson-Laird, P. N. (1977). "Procedural semantics". *Cognition*, *5*, 189-214.

Kamp, H. & Reyle, U. (1993). *From discourse to logic : Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht : Kluwer Academic Publisher.

Kirby, S. (2002). "Learning, bottlenecks and the evolution of recursive syntax". In T. Briscoe (Ed), *Linguistic evolution through language acquisition: Formal and computational models*. Cambridge, MA : Cambridge University Press, 173-203.

Montague, R. (1974). *Formal Philosophy*. London : Yale University Press.

Nowak, M. A., Plotkin, J. B. & Jansen, V. A. A. (2000). "The evolution of syntactic communication". *Nature*, *404*, 495-498.

Pepperberg, I. M. (1999). *The Alex studies - Cognitive and communicative abilities of grey parrots*. Cambridge, MA : Harvard University Press, ed. 2002.

Piattelli-Palmarini, M. (1979). *Théories du langage - Théories de l'apprentissage*. Paris : Seuil.

Piattelli-Palmarini, M. (1989). "Evolution, selection and cognition: From 'learning' to parameter setting in biology and in the study of language". *Cognition* 31, 1-44.

Pinker, S. & Bloom P. (1990). "Natural language and natural selection". *Behavioral and Brain Sciences* 13(4), 707-784.
http://www.bbsonline.org/documents/a/00/00/04/99/index.html

Pinker, S. & Jackendoff R. (2005). "The faculty of language: what's special about it?". *Cognition* 95, 201-236.

Premack, D. & Premack, A. J. (1983). *L'esprit de Sarah*. Paris : Fayard, ed. 1984.

Savage-Rumbaugh, E. S. & Lewin, R. (1994). *Kanzi: the ape at the brink of the human mind*. New York : John Wiley & Sons.

Talmy, L. (2000). *Toward a Cognitive Semantics: Concept Structuring Systems*. Cambridge, MA : MIT Press.

Van Valin, R. D. & LaPolla, R. J. (1997). *Syntax. Structure, meaning and function*. Cambridge University Press.