

# **Modélisation de l'émergence du langage**

DESSALLES Jean-Louis

# Action “ Origine de l’Homme, du Langage et des Langues ”

## A. FICHE ADMINISTRATIVE

### Titre du projet :

**M.E.L.: Modélisation de l’émergence du langage**

### Mots-clés :

modélisation, émergence, évolution, sciences cognitive, épistémologie, systèmes multi-agents, théorie linguistique, syntaxe, sémantique, protolangage, systèmes adaptatifs, éthologie.

### Résumé du projet :

Les données disponibles concernant l’émergence de la faculté de langage et des capacités cognitives qui lui sont associées sont indirectes. Contrairement aux transformations morphologiques que l’on peut reconstituer grâce aux fossiles, l’apparition d’un comportement que l’on observe dans une seule espèce pose un problème difficile. La modélisation permet de progresser dans la compréhension des phénomènes qui ont permis l’émergence du langage : (1) en imposant des contraintes sur les scénarios possibles ; (2) en permettant d’identifier les paramètres critiques ; (3) en démontrant la cohérence interne des différents modèles proposés.

Le projet MEL regroupe plusieurs initiatives, toutes concernées par le souci de modéliser certains aspects de l’émergence de la faculté langagière. Citons notamment :

- tenter une reconstitution acceptable de l’émergence du langage en plusieurs phases, avec notamment la possibilité d’une forme intermédiaire dite ‘protolangage’ ; chaque phase devant apparaître comme fonctionnelle sur le plan social et cognitif.
- montrer que certains aspects de la syntaxe et de la sémantique peuvent émerger dans des conditions écologiques et cognitives particulières.

Au-delà de l’utilisation d’une même méthodologie : la modélisation, les participants de ce projet ont le souci de confronter leurs approches et de les soumettre à une réflexion critique d’ordre épistémologique, permettant d’estimer la portée des résultats obtenus.

## 1. Responsable scientifique du projet

Nom : **DESSALLES**

Prénom : **JEAN-LOUIS**

Discipline du responsable scientifique : **Intelligence Artificielle**

Établissement de rattachement : **Ecole Nationale Supérieure des Télécommunications - URA 820  
Dép. InfRes**

Adresse professionnelle : **46 rue Barrault**

Code postal **75634** Commune **Paris Cedex 13**

Tél : **01 45 81 75 29**

Fax : **01 45 81 31 19**

E-Mail : **dessalles@enst.fr**

## 2. Laboratoire ou organisme de rattachement de l’équipe de recherche

Intitulé : **Laboratoire Traitement et Communication de l’Information (LTCI)**

- Unités CNRS :  unité propre du CNRS

unité associée ou mixte du CNRS

Préciser le code unité **URA820**  
Préciser la délégation régionale : .....

- Unités hors CNRS :  unité universitaire (*Préciser l'université*)  
.....  
 Autre (*Préciser : EPST, Association, équipe étrangère...*)

Nom du directeur de l'organisme : Marc Peyrade (ENST) - Jean-Pierre Tubach (URA 820)

Adresse professionnelle : **Dep. Informatique et Réseaux - ENST**  
**46 rue Barrault**

Code postal **75634** Commune **Paris Cedex 13**

Tél : **01 45 81 75 29**

Fax : **01 45 81 31 19**

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : LATTICE- UMR CNRS  
Ecole Normale Supérieure

Discipline(s) couverte(s) par l'équipe : Philosophie de la fonction symbolique, modélisation du langage

Nom du directeur de l'équipe : Catherine FUCHS

Membre(s) partenaire(s) :

Nom : **LASSÈGUE**

Grade : CR1

Nom : **VICTORRI**

Grade : DR2

Nom : **VISETTI**

Grade : CR1

Prénom : **Jean**

E-Mail : Jean.Lassegue@ens.fr

Prénom : **Bernard**

E-Mail : victorri@ens.fr

Prénom : **Yves-Marie**

E-Mail : yves-marie.visetti@ens.fr

Adresse : 1 rue Maurice Arnoux

Code postal : **92120**

Commune : Montrouge

Tél : 01 42 53 75 97

Fax : 01 42 53 75 86

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Ecole Normale Supérieure

Discipline(s) couverte(s) par l'équipe : Ethologie et épistémologie

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **LESTEL**

Grade :

Adresse : 46 rue d'ULM

Code postal : **75230**

Tél :

Prénom : **Dominique**

E-Mail : lestel@ens.fr

Commune : Paris Cedex 05

Fax :

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Laboratoire de Physique Statistique de l'Ecole Normale Supérieure  
Ecole Normale Supérieure

Discipline(s) couverte(s) par l'équipe : Sciences Physique, Modélisation en sciences cognitives,  
en biologie et en sociologie

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **NADAL**

Grade :

Nom : **WEISBUCH**

Grade :

Adresse : 46 rue d'ULM

Code postal : 75230

Adresse : 24 rue Lhomond

Code postal : 75005

Tél : 33.1 - 44 32 32 75

Prénom : **Jean-Pierre**

E-Mail : nadal@lps.ens.fr

Prénom : **Gérard**

E-Mail : weisbuch@lps.ens.fr

Commune : Paris Cedex 05

Commune : Paris

Fax : 33.1 - 44 32 34 33

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Laboratoire Dynamique Du Langage - UMR CNRS 5596  
Université Lumière Lyon 2

Discipline(s) couverte(s) par l'équipe : Modélisation en sciences du langage

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **COUPÉ**

Grade :

Nom : **MARSICO**

Grade :

Nom : **PELLEGRINO**

Grade :

Nom : **PERRONE**

Grade :

Adresse : 14, rue Berthelot,

Code postal : 69363

Tél : 04 72 72 64 63

Prénom : **Christophe**

E-Mail : ccoupe@ens\_lyon.fr

Prénom : **Egidio**

E-Mail : Egidio.Marsico@ish\_lyon.cnrs.fr

Prénom : **François**

E-Mail : Francois.Pellegrino@univ\_lyon2.fr

Prénom : **Emmanuelle**

E-Mail : Emmanuelle.Perrone@univ\_lyon2.fr

Commune : Lyon Cedex 07

Fax : 04 72 72 65 90

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : CREA  
Ecole Polytechnique

Discipline(s) couverte(s) par l'équipe :

Nom du directeur de l'équipe : Jean Petitot

Membre(s) partenaire(s) :

Nom : **BOURGINE**

Grade :

Nom : **PETTITOT**

Grade :

Nom : **SCUBLA**

Grade :

Adresse : 1 rue Descartes

Code postal : 75005

Tél :

Prénom : **Paul**

E-Mail : bourgine@poly.polytechnique.fr

Prénom : **Jean**

E-Mail : petitot@poly.polytechnique.fr

Prénom : **Lucien**

E-Mail : scubla@poly.polytechnique.fr

Commune : Paris

Fax :

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire :  
Sony Computer Science Laboratory

Discipline(s) couverte(s) par l'équipe : Modélisation en sciences du langage

Nom du directeur de l'équipe : Luc Steels

Membre(s) partenaire(s) :

Nom : **OUDEYER**  
Grade :  
Nom : **STEELS**  
Grade :  
Adresse : 6 rue Amyot  
Code postal : 75005  
Tél : 01 44 08 05 02

Prénom : **Pierre\_Yves**  
E-Mail : py@csl.sony.fr  
Prénom : **Luc**  
E-Mail : steels@csl.sony.fr  
Commune : Paris  
Fax : 01 45 87 87 50

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : COSTECH, Département Technologie et Sciences de l'Homme  
Université de Technologie de Compiègne

Discipline(s) couverte(s) par l'équipe : Modélisation en biologie

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **STEWART**

Prénom : **John**

Grade :

E-Mail : John.Stewart@utc.fr

Adresse : -

Code postal : 60206

Commune : Compiègne Cedex

Tél : 03 44 23 43 64

Fax : (0)3 44 23 52 12

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Centre de linguistique française - EA 188  
Université Paris IV

Discipline(s) couverte(s) par l'équipe : Sémiologie

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **RASTIER**

Prénom : **François**

Grade :

E-Mail : lpe2@idf.ext.jussieu.fr

Adresse : 1 rue Victor Cousin

Code postal : 75230

Commune : Paris Cedex 05

Tél : 01 40 46 32 71

Fax : 01 40 46 32 54

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Laboratoire d'étude des mécanismes cognitifs

Discipline(s) couverte(s) par l'équipe : neuro-psychologie cognitive

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **BEDOIN**

Prénom : **Nathalie**

Grade :

E-Mail :

Adresse :

Code postal :

Commune : Bron

Tél :

Fax :

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Institut Charles Darwin International

Discipline(s) couverte(s) par l'équipe : Epistémologie de l'évolution

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **TORT**

Prénom : **Patrick**

Grade :

E-Mail :

Adresse : 23 rue de la République

Code postal : 93230

Tél : 01 48 43 76 08

Commune : Romainville

Fax : 01 48 43 76 08

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Autres laboratoires

Discipline(s) couverte(s) par l'équipe : Paléo-anthropologie

Membre(s) partenaire(s) :

Nom : **PTCQ**

Grade :

Adresse : 10 rue de l'ancien Parc

Code postal : 95820

Tél : 01 39 37 66 41 - 63 28

Prénom : **Pascal**

E-Mail :

Commune : Bruyeres / Oise

Fax :

### 3. Autre(s) laboratoires ou organisme(s) partenaires

Intitulé laboratoire : Autres laboratoires

Discipline(s) couverte(s) par l'équipe : Socio-écologie animale et anthropologie

Nom du directeur de l'équipe :

Membre(s) partenaire(s) :

Nom : **GUILLE-ESCURET**

Grade :

Adresse : 76 rue du Cardinal Lemoine

Code postal : 75005

Tél :

Prénom : **Georges**

E-Mail :

Commune : Paris

Fax :

Date : 13.06.2000

Signature du Responsable Scientifique :

**Jean-Louis Dessalles**

## B. PROJET SCIENTIFIQUE

### M.E.L.: Modélisation de l'émergence du langage

#### Objectifs

Ce projet vise à comprendre l'émergence du langage humain, en étudiant les conditions biologiques et sociales qui ont pu présider à son apparition et les mécanismes susceptibles d'expliquer ce processus d'émergence. Une part importante du travail consistera à étayer la réflexion par la confection de modèles mathématiques et informatiques permettant d'évaluer la cohérence et la portée des hypothèses sur lesquelles nous travaillerons.

#### Problématique

Le langage apparaît comme l'un des traits distinctifs de l'espèce humaine. Son émergence au cours de notre phylogénèse, bien qu'ayant pu être déclenchée par des circonstances fortuites, a obéi aux contraintes de l'évolution par sélection naturelle. *Notre projet est de progresser, grâce aux outils de la modélisation, dans la compréhension des conditions de cette émergence.* En particulier, il s'agit d'essayer d'étudier, pour différents aspects du langage, la possibilité qu'ils soient le fruit d'une émergence phylogénétique ou qu'il soient le résultat d'une émergence culturelle.

Selon une périodisation bien reçue actuellement, due au linguiste Derek Bickerton (1990, 1995), l'apparition du langage proprement dit, tel qu'on peut l'appréhender à travers la diversité des langues actuelles, se serait effectuée en deux temps : apparition d'un protolangage avec l'Homo erectus africain (appelé aussi Homo ergaster), il y a plus d'1,5 million d'années, expliquant en partie le succès évolutif de cette espèce qui a colonisé tout l'Ancien Monde ; apparition beaucoup plus récente du langage proprement dit, au cours de la période qui va de -150 000 ans (date approximative présumée d'apparition de notre espèce en Afrique, qui aurait à son tour colonisé la planète) à -40 000 ans (début de "l'explosion symbolique"). Sur le fond de cette périodisation très simple, s'affrontent diverses théories concernant les

conditions d'émergence du langage ("grooming and gossip" (Dunbar 1996), théorie du prestige au sein des coalitions (Dessalles 2000), mimésis et ritualisation (Donald 1991), capacité narrative et mythique (Victorri 1999), etc.).

La possibilité d'une émergence par paliers est conforme à la vision moderne de l'évolution par sélection naturelle qui prévoit une succession d'équilibres dans la filiation des espèces (Gould & Eldredge 1977). L'existence de paliers est due au fait que, contrairement à une idée reçue, l'évolution (la micro-évolution) par sélection naturelle est un processus rapide qui conduit une espèce, en quelques dizaines ou centaines de générations, à occuper un optimum local. La transition entre paliers (macro-évolution) est un phénomène rare et imprévisible qui correspond au passage d'un équilibre à un autre (Dessalles 1996).

L'existence probable de paliers dans la phylogenèse du langage, bien qu'elle diminue l'amplitude du "saut" qui sépare le langage humain de la communication animale, demande que les phénomènes évolutifs ayant conduit à l'apparition des différents aspects du protolangage et du langage (phonologie, syntaxe, sémantique et pragmatique) soient logiquement ordonnés dans le temps et correctement corrélés. En l'absence de données directes sur ces événements passés, la modélisation constitue un outil incontournable lorsqu'il s'agit de cerner les conditions ayant rendu l'apparition du langage possible, voire probable : non seulement parce qu'elle permet de construire une image dynamique effective (causale si l'on veut), mais aussi parce qu'elle permet de mettre en évidence des phénomènes que la seule réflexion n'aurait pas permis de prévoir.

La modélisation intervient également lorsqu'il s'agit de faire la part entre l'émergence biologique de nouvelles capacités et l'apparition, au niveau collectif, de nouveaux modes d'organisation de la communication. Par exemple, certaines caractéristiques d'une langue donnée, comme son système vocalique, émergent de la nécessité d'intercompréhension, en réalisant un compromis entre différentes contraintes (nombre de formes, discrimination, facilité de production, etc.) (Lindblom 1986,1998). La modélisation permet de tester des modèles prévoyant que certaines caractéristiques du langage émergent de l'interaction entre les agents communicants.

## Les contraintes révélées par la modélisation

Pendant longtemps, le problème de l'origine du langage est resté un problème sous-contraint : la compréhension trop partielle des mécanismes évolutifs et le manque d'outils pour appréhender les phénomènes collectifs laissaient le champ libre à toutes les spéculations concernant l'apparition du langage dans la lignée humaine. Le rapport s'est maintenant inversé, en partie grâce aux modélisations de la biologie théorique et aux simulations des phénomènes collectifs. De nos jours, les modèles concernant l'origine du langage doivent passer un certain nombre de tests, ce qui exclut tout danger de spéculation gratuite. Citons quelques exemples de contraintes fournies par la modélisation :

### 1. modélisation biologique

- contrainte d'optimalité biologique locale : La théorie des équilibres ponctués (Gould & Eldredge 1977) et la théorie du parallélisme implicite (Holland 1975) montrent que la sélection naturelle conduit les espèces à occuper des optima locaux. La conséquence importante est que toute caractéristique naturelle doit être localement optimale pour une fonction (cette observation ne doit pas être confondue avec le "panglossisme" dénoncé par Gould et Lewontin (1979)). En particulier, la structure de la faculté de langage, dans tous ses aspects biologiques, doit être localement optimale pour sa fonction. Par exemple, si l'on identifie une fonction narrative, il faut montrer comment le langage est localement optimal pour relater des événements passés. Cette contrainte, imposée par la modélisation biologique, exclut de fait tous les modèles qui présenteraient notre espèce, ou celles qui l'ont précédée, comme en déséquilibre évolutif, arguant de la lenteur de l'évolution pour justifier le fait que telle ou telle faculté syntaxique (par ex. la récursivité), phonologique (par ex. la structure de la syllabe), sémantique (par exemple la capacité à gérer des relations thématiques) ou physiologique (par exemple la position basse du larynx) se soit développée lentement pour n'atteindre sa pleine efficacité qu'avec l'espèce humaine.
- contrainte d'optimalité culturelle locale : La même contrainte d'optimalité locale s'applique dans le cas de l'évolution culturelle, en ce qui concerne ses aspects adaptatifs. Par exemple, le système vocalique d'une langue donnée doit présenter des aspects localement optimaux (pour un ensemble de critères souvent contradictoires, par ex. le nombre des formes sonores et la qualité de discrimination).
- contrainte de fonctionnalité : Il est habituel de diviser la faculté de langage en niveaux bien individualisés : phonétique, morphologie, syntaxe, sémantique, pragmatique. Tout modèle, même partiel, de l'évolution de ces niveaux, doit tenir compte du caractère fonctionnel de l'ensemble. En particulier, si l'on veut défendre la thèse du protolangage, il faut accorder à notre espèce parente une protosémantique, une protophonologie et une protopragmatique qui forment un ensemble fonctionnel. A l'inverse, imaginer qu'une espèce comme *erectus* ait pu disposer d'une capacité à former des représentations sémantiques similaire à la nôtre, sans disposer des moyens syntaxiques pour exploiter cette capacité, est incohérent.
- contrainte de fonction biologique : La fonction biologique du langage, perçue longtemps comme allant de soi, se révèle extrêmement problématique. Le fait que l'échange d'informations par le langage puisse, dans certaines conditions, bénéficier au groupe qui le pratique, n'a aucune influence sur le fait que la sélection naturelle favorisera la propension à fournir des informations aux autres. Au contraire ! Fournir des informations utiles à ses concurrents génétiques semble, à première vue contraire à la stratégie darwinienne la plus élémentaire. Tout scénario de l'émergence de la faculté de langage se doit de résoudre ce paradoxe.

## 2. modélisation cognitive

- parcimonie cognitive : Le défi de la modélisation cognitive consiste à montrer que la complexité des productions langagières n'est qu'apparente, qu'elle résulte de mécanismes simples appliqués de manière itérée ou récursive.
  - Par exemple, la complexité apparente de la structure des phrases du langage peut résulter de l'application récursive d'une procédure qui connecte les syntagmes par leur spécifieur et leurs compléments. Dans une perspective évolutive, il s'agit d'étudier si cette capacité peut émerger à partir de systèmes plus simples, non récursifs.
  - Au niveau sémantique, de nombreux auteurs ont suggéré que les relations argumentales obéissent à un schéma fixe, de nature spatiale et topologique (Langacker 1986, Jackendoff 1983, Gruber 1965, Porzig 1950 (cité dans Lorenz 1973)). Ce schéma identifie un thème, un repère, une relation topologique entre le thème et le repère, et un acteur éventuel qui exerce une force sur le thème (Talmy 1988). Ce seraient toujours ces mêmes relations thématiques qui seraient appliquées, de manière systématique, aux relations prédicatives contenues dans toute phrase pour en déterminer le sens. L'enjeu d'une perspective évolutive en modélisation cognitive est de montrer comment l'apparition de cette nouvelle capacité de segmentation thématique a profondément modifié le système de communication humain.
  - Au niveau pragmatique, il s'agit de montrer comment la structure des dialogues et celle du discours argumentatif résultent de l'application d'une stratégie récursive simple, qui consiste chaque fois à détecter ou à résoudre un "conflit cognitif". Dans une perspective évolutive, il s'agit de montrer le rôle central que le conflit cognitif a pu jouer, et ses conséquences sur la motivation biologique à la prise de parole.

## 3. modélisation des phénomènes collectifs

- émergence des représentations : Le langage repose sur un certain nombre de structures émergentes, propres à une communauté d'individus communicants. C'est le cas par exemple de certaines structures phonologiques (en français du Sud, une voyelle moyenne est ouverte si et seulement si elle est non-terminale au sein de la syllabe (Durand 1990)), de certaines structures syntaxiques (en anglais, les adjectifs possessifs s'accordent avec le possesseur), de certaines structures sémantiques (par ex. structuration métaphorique du temps inversée chez les Aymaras (Núñez, communication personnelle)), ou de certaines représentations conceptuelles du niveau pragmatique (Sperber 1996). La modélisation des phénomènes collectifs peut permettre d'étudier les paramètres critiques qui permettent l'émergence de telles structures.
- émergence de communautés de locuteurs : L'utilisation du langage suppose l'existence de groupes partageant, au moins en partie, le même code et les mêmes représentations. La dynamique de la formation de ces groupes pose un problème de "poule et d'œuf" : le code n'existe que parce que la communication le met en œuvre, et la communication n'est possible que parce que le code est déjà là. La modélisation permet d'étudier ce problème de double émergence.
- émergence des conditions sociales d'usage du langage : Certains modèles de l'apparition du langage font référence à l'existence de coalitions au sein desquelles le langage joue un rôle structurant (Dunbar 1996). Le langage a été décrit soit comme une forme de coopération à l'intérieur de la coalition, soit comme un moyen d'établir et de maintenir des liens entre partenaires de coalition. Ces scénarios peuvent être modélisés et mis à l'épreuve de la simulation. On peut en particulier étudier les conditions de stabilité de la coopération, ou déterminer les conditions dans lesquelles des coalitions sont susceptibles d'apparaître et de se maintenir. Ces questions se révèlent fondamentales si l'on veut comprendre les raisons pour lesquelles le langage n'est apparu que dans notre seule lignée (Dessalles 2000).

## L'approche proposée

### 1. Modélisation évolutive

La théorie contemporaine de l'évolution prévoit l'occurrence de transitions brusques dans la filiation entre espèces. La direction de ces changements majeurs et le moment exact où ils se produisent sont essentiellement dus au hasard. En revanche, *a posteriori*, de tels changements respectent plusieurs contraintes : ils sont adaptatifs et leur probabilité d'apparition augmente dans certaines zones, dites critiques, de l'espace des paramètres. Par exemple, la probabilité d'apparition du langage pourrait avoir considérablement augmenté dans une structure sociale où la taille des coalitions dépassait un certain seuil.

Par la modélisation, nous pouvons explorer certains espaces de paramètres de manière à déterminer des ensembles de conditions dans lesquelles le langage a pu émerger. Nous pouvons également montrer les faiblesses de telle ou telle approche théorique (par exemple, les difficultés rencontrées par des modèles trop simplement coopératifs).

### 2. Modélisation cognitive

Notre problématique concerne non seulement les conditions biologiques et sociales d'utilisation du langage, mais aussi le langage lui-même dans ses formes et ses valeurs spécifiques. Il est possible de refléter dans les modèles certaines différences de structure formelle entre les langues et divers types de communication animale (exemples : aspects discrets, récursivité des structures, etc.). De même, il apparaît possible et nécessaire de proposer une modélisation du niveau sémantique, dans le prolongement des travaux de la linguistique cognitive. Il s'agit là de dimensions du langage qu'aucune théorie n'a encore réussi à intégrer de façon pleinement satisfaisante : or, en parlant, les sujets construisent et partagent des représentations (évoquant *in absentia* de scènes, fictions, inférences) dont certains traits fonctionnels ont dû être déterminants pour le problème d'émergence qui nous occupe.

### 3. Modélisation des phénomènes collectifs

Depuis plusieurs années, des travaux s'inscrivant dans le paradigme de la modélisation multi\_agents traitent de phénomènes linguistiques en rapport avec l'origine et l'évolution du langage. Qu'ils s'intéressent au lexique, à la syntaxe ou à la phonologie, une des principales difficultés inhérentes à ces modélisations est de les confronter au système fonctionnel que constitue le langage humain. En particulier, si les notions d'auto\_organisation permettent de retrouver des tendances universelles rencontrées dans les systèmes réels (De Boer, 2000), l'implémentation d'un système multi\_agents permettant de faire émerger la communication parlée et le langage doublement articulé par auto\_organisation reste utopique. L'hypothèse sur laquelle nous travaillerons considère que le langage " complexe " que nous utilisons aujourd'hui s'est développé à partir d'un stade de langage " simple " défini par l'existence d'un lien direct entre séquence de sons et référents ainsi que par l'absence de double articulation.

Les simulations multi\_agents appliquées à la linguistique mettent rarement l'accent sur les aspects écologiques du langage (à l'exception des travaux de Cangelosi, 2000). Dans le cadre de ce projet, nous souhaitons tenir compte du rôle fonctionnel du langage, en modélisant les interactions entre le monde et les agents, en particulier dans les rapports ressources / prédation, la survie des individus dépendant en grande partie de leur capacité d'adaptation. L'objectif est de fournir un modèle de type multi\_agents permettant d'évaluer si un scénario d'évolution d'un langage simple vers un langage complexe peut émerger dans le cadre d'interactions tenant compte d'aspects cognitifs, sociologiques et écologiques.

D'un point de vue cognitif, les agents modélisés sont soumis à des contraintes, en particulier au niveau de leurs mémoires, tant à court terme qu' " encyclopédique " auxquelles s'ajoutent les limites du processus de traitement des informations. Ils peuvent également être pourvus d'un modèle de représentation interne des connaissances.

Nous nous proposons d'étudier par ailleurs plusieurs aspects sociaux du langage qui modifient la manière dont les interactions entre agents se déroulent (existence de classes sociales et de registres, influence des tailles de populations, etc.). A cela s'ajoute la possibilité de définir des stratégies de communications individuelles différenciées, en intégrant par exemple des facteurs stochastiques de communicabilité.

Enfin, la prise en compte des processus de transmission du savoir et des concepts entre agents d'une même génération et/ou de générations différentes est sans conteste un aspect fondamental du processus étudié.

## Actions envisagées

Les actions que nous proposons dans le cadre de ce projet sont d'ordre théorique, expérimental (simulations) et épistémologique. L'objectif est (1) d'obtenir des résultats nouveaux et (2) de progresser dans l'intégration des résultats dans un ensemble conceptuel cohérent, de manière à offrir une image plausible de l'origine du langage humain.

### 1. Aspects théoriques

Il est possible de formaliser plusieurs phénomènes qui ont pu être à l'œuvre lors des transitions majeures qui ont conduit au langage. Citons :

- la modélisation de la coopération symétrique et des conditions dans lesquelles elle est évolutivement stable. Evaluation de la pertinence de ces conditions dans le cas du langage.
- la modélisation d'un niveau intermédiaire entre l'individu et le groupe : le système de coalitions. Le groupe est l'unité au sein duquel s'effectue la reproduction. Les coalitions sont caractérisées par le fait que les individus se choisissent pour s'allier. Le langage peut être un élément essentiel de ce choix. Il s'agit de modéliser les conditions dans lesquelles le langage peut émerger, en tant que critère d'alliance, dans un système favorisant l'existence de coalitions.
- la modélisation des structures de signification et de l'avantage qu'elles procurent aux individus capables de les mettre en œuvre. On distinguera notamment deux aspects de la construction du sens : (1) la capacité à former des scènes imagées par combinaison d'éléments ; (2) le repérage des relations thématiques. Il s'agit d'étudier, dans une perspective évolutive, si l'un de ces deux systèmes sémantiques peut, à lui seul, constituer un système fonctionnel (un système d'attribution de sens est fonctionnel s'il permet de produire des inférences adéquates). Si tel est le cas, on peut légitimement faire l'hypothèse que ce système a précédé l'autre.
- l'adéquation des aspects fonctionnels des différents aspects du langage : l'approche évolutive nous oblige à poser la question, assez inhabituelle en sciences du langage, de l'adéquation des structures aux fonctions :

- modélisation du caractère adapté de la capacité à manier des structures phonologiques compte tenu d'un certain débit de parole et d'une taille de lexique
  - modélisation du caractère adapté de la capacité à utiliser un système de syntagmes et un système de marquage morphologique pour l'identification des relations thématiques.
  - modélisation du caractère adapté des mécanismes de construction du sens pour les capacités inférentielles
  - modélisation du caractère adapté des mécanismes de génération d'argument dans le dialogue.
- la modélisation des processus communicationnel émergents : détermination de la validité théorique des résultats expérimentaux.

## 2. Aspects expérimentaux

Nous disposons de plusieurs plates-formes de simulation pour tester les théories évoquées ci-dessus. En particulier :

- une plate-forme de simulation génétique, développée à l'ENST, permettant de tester la viabilité évolutive des différentes stratégies de communication.
- des plates-formes multi\_agents, l'une développée au laboratoire Dynamique Du Langage, l'autre au laboratoire d'informatique de Sony.
  - Dans la première, les agents possèdent un ancrage phonologique et les interactions tiennent compte de contraintes articulatoires et perceptives (Perrone, 2000). L'architecture du système permet de prendre en compte des contraintes de différents types, et, à l'heure actuelle, les interactions entre agents intègrent déjà plusieurs aspects sociologiques (notions de familles, de rangs sociaux et de différences de prestige entre populations). Les aspects évoqués dans le paragraphe précédent seront implémentés en relation avec les modules considérés, en s'appuyant sur la structure " objet " du système.
  - La plate-forme développée au Sony-CSL permet de tester l'ancrage des symboles émergents dans un contexte sensori-moteur, par l'emploi de robots dotés de caméras et de moteurs leur permettant d'orienter le regard. Cet ancrage permet de s'affranchir en partie de l'idéalisation liée aux simulations informatiques.
- une plate-forme de traitement du langage naturel, en cours de développement à l'ENST. L'originalité de cette plate-forme est d'offrir un niveau sémantique permettant de tester la séparation entre deux composantes sémantiques fonctionnellement distinctes.

Dans le cadre de ce projet, nous comptons développer et compléter ces plates-formes et les utiliser pour explorer les espaces de paramètres analysés dans l'approche théorique.

## 3. Aspects épistémologiques

La modélisation comporte le danger de donner une image par trop caricaturale des phénomènes observés ou imaginés. Il est d'autant plus important de souligner que notre groupe rassemble des compétences variées qui permettront de mieux rapporter les modèles à la multiplicité des données éthologiques, anthropologiques, archéologiques, linguistiques et cognitives.

Notre souci est de conserver un regard critique sur la validité des modèles que nous développons. La modélisation permet de ramener sur le terrain de l'expérience des phénomènes dont on pouvait penser qu'ils étaient non répétables, et donc inaccessibles à l'investigation scientifique. L'analyse des conditions de validité de l'approche modélisatrice est partie intégrante de notre projet.

Notre ambition est, à terme, d'intégrer à notre cadre de réflexion des questions réputées difficiles, notamment la relation entre le langage et les autres activités et traces symboliques, les mécanismes d'évolution des représentations culturelles véhiculées par le langage et l'apparition de nouvelles conditions de sélection (interaction gènes-culture).

## Composition du groupe

Notre groupe comporte un fort contingent de modélisateurs, spécialistes de l'utilisation de la théorie des systèmes dynamiques dans divers domaines allant de la biologie aux sciences cognitives et aux sciences socio-économiques, en passant bien sûr par les sciences du langage, fortement représentées. Il comporte aussi des spécialistes de plusieurs disciplines directement impliquées dans l'étude de l'origine de l'homme et du langage (paléo-anthropologie, éthologie des primates, socio-écologie animale, anthropologie sociale, philosophie). Il comporte enfin des spécialistes de la modélisation multi-agents des phénomènes de langage.

## Quelques références bibliographiques des participants

- DESSALLES, J-L. (1998). "Altruism, status, and the origin of relevance". In J. R. Hurford, M. Studdert-Kennedy & C. Knight (ed), *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press, 130-147.
- DESSALLES, J-L. (1999). "Coalition factor in the evolution of non-kin altruism". *Advances in complex systems*, 2(2), 143-172.
- DESSALLES, J-L. (2000). *Aux origines du langage - Une histoire naturelle de la parole*. Paris : Hermès.
- DESSALLES, J-L. (2000). "Language and hominid politics". In C. Knight, M. Studdert-Kennedy & James Hurford (eds), *The evolutionary emergence of language: social function and the origins of linguistic form*. Cambridge : Cambridge University Press.
- DESSALLES, J-L. (2000). Proceedings of the Third Conference on the Evolution of Language. Paris : E.N.S.T. 2000-S-002.
- DESSALLES, J-L. (2000). "Two stages in the evolution of language use". *Proceedings of the Third Conference on the Evolution of Language*. Paris : E.N.S.T. 2000-S-002.
- LAKS B., VICTORRI B. (éditeurs) : Origine des langues et du langage, numéro de la revue *Langages*, en préparation.
- LASSEGUE, J. séminaire "Théories du symbole" (textes disponibles sur : <http://www.lattice.ens.fr/Lassegue.html/semsymb1.html>)
- MARSICO E., COUPÉ C & PELLEGRINO P. (2000), "Evaluating the influence of language contact on lexical changes", in Proc. of the 3<sup>rd</sup> Conference on The Evolution of Language, Paris, 3\_6 April 2000
- PERRONE E. (2000), "The emergence of phonology in a population of artificial agents in a phonetic naming game", in Proc. of the 3<sup>rd</sup> Conference on The Evolution of Language, Paris, 3\_6 April 2000
- PETITOT J. (1992). *Physique du Sens*. Editions du CNRS.
- VICTORRI B. (1997). "Débat sur la langue mère", Pour la Science, dossier hors série "*Les langues du Monde*", octobre 1997.
- VICTORRI B. (1999). "La place de la fonction narrative dans l'émergence du langage et la structure des langues". *T.L.E. (Théorie, Littérature, Enseignement)*, 17, 23-38.
- VICTORRI B. (2000). "The role of narration in the emergence of human language". *Third International Conference on the Evolution of Language*, Paris.
- VISSETTI, Y.-M. (1999). "L'expression symbolique et le problème de l'origine du langage". *Actes du séminaire "Les signes et les techniques"*. Université Technologique de Compiègne, pp. 137-150.

## Autres références bibliographiques

- BICKERTON, D. (1990). *Language and Species*. University of Chicago Press.
- BICKERTON, D. (1995). *Language and human behavior*. London : UCL Press, ed. 1996.
- CANGELOSI A. (2000), "Evolution of symbolisation in chimpanzees and neural nets", in *Proc. of the 3<sup>rd</sup> Conference on The Evolution of Language*, Paris, 3\_6 April 2000
- DE BOER B. (2000), "Emergence of sound systems through self-organisation", in In C. Knight, M. Studdert-Kennedy & James Hurford (eds), *The evolutionary emergence of language: social function and the origins of linguistic form*. Cambridge : Cambridge University Press.
- DESSALLES, J-L. (1996). *L'ordinateur génétique*. Paris : Hermès.
- DONALD, M. (1991). *Origins of the modern mind*. Cambridge : Harvard University Press.
- DUNBAR, R. I. M. (1996). *Grooming, gossip, and the evolution of language*. Cambridge : Harvard University Press.
- DURAND, J. (1990). *Generative and non-linear phonology*. London : Longman.
- GOULD, S. J. & Eldredge, N. (1977). "Punctuated equilibria : the tempo and mode of evolution reconsidered". *Paleobiology*, 3, 115-151.

- GOULD, S. J. & Lewontin, R. C. (1979). "The spandrels of San Marco and the Panglossian program: a critique of the adaptationist program." *Proceedings of the Royal Society of London*, 205, 281-288.
- GRUBER, J. S. (1965). *Lexical structures in syntax and semantics*. Amsterdam : North Holland, ed. 1976.
- HOLLAND, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor : The University of Michigan Press.
- JACKENDOFF, R. (1983). *Semantics and cognition*. Cambridge : The MIT Press, ed. 1995.
- LANGACKER, R. (1986). *Foundations of cognitive grammar*. Stanford : Stanford University Press.
- LINDBLOM, B. (1986). "Phonetic universals in vowel systems". *Experimental Phonology*, 13-44.
- LINDBLOM, B. (1998). "Systemic constraints and adaptive change in the formation of sound structure". In J. R. Hurford, M. Studdert-Kennedy & C. Knight (ed), *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press.
- LORENZ, K. (1973). *L'envers du miroir - Une histoire naturelle de la connaissance*. Paris : Flammarion, ed. 1975.
- SPERBER, D. (1996). *La contagion des idées*. Paris : Odile Jacob.
- TALMY, L. (1988). "Force dynamics in language and thought". *Cognitive Science*, 12, 49-100.

