

# Expériences sur l'origine et l'évolution du langage.

Luc Steels (1,2) and Frédéric Kaplan (1)

(1) Sony Computer Science Laboratory

6, Rue Amyot - Paris 75005

(2) AI Laboratory, Vrije Universiteit Brussel

## 1. Est-ce-que l'on peut mener des expériences en sciences cognitives ?

Les sciences naturelles ont une longue tradition dans la conduite d'expériences. Une expérience recrée un phénomène naturel dans des conditions contrôlées. En modifiant ces conditions et en observant l'évolution du phénomène étudié, on peut identifier des relations de cause à effet. L'expérimentation est le complément naturel de la modélisation. La modélisation consiste à développer une description formelle d'un système, par exemple sous la forme d'un ensemble d'équations, et de tester l'adéquation des données obtenues à partir de ce modèle avec celles observées dans la réalité. Bien que ce type de modélisation soit évidemment très important pour l'avancée de la connaissance scientifique, il ne permet pas d'obtenir le même niveau de validation qu'une expérience parce que le modèle procède nécessairement d'une abstraction importante par rapport à la réalité.

Dans quelle mesure une démarche expérimentale pourrait s'appliquer à l'étude du comportement cognitif humain? L'expérimentation avec des sujets humains ou des animaux a des limites fondamentales. La première d'entre elles relève de l'éthique. Des expériences du type de celles conduites en physique ou en chimie, sont tout simplement inacceptables moralement pour l'étude du comportement cognitif. Les psychologues conduisent des expériences avec des êtres humains pour étudier des phénomènes comme la perception des couleurs, mais ces expériences sont nécessairement limitées à des activités sans danger pour les sujets et elles se bornent à des instrumentations non intrusives. Même au sein de ces expériences, la grande complexité de la cognition humaine rend difficile l'isolement et l'identification du phénomène étudié. Par exemple, les expériences qui utilisent des mesures de type stimulus-réponse ou des scanners de l'activité cérébrale ne donnent que des idées vagues sur les activités du cerveau. De plus, ces expériences ne portent en général que sur un nombre très réduit de sujets, alors que beaucoup de phénomènes cognitifs ont une dimension culturelle qui, pour être étudiée, nécessiterait la mise en place d'expériences sur des populations importantes.

On peut aussi envisager de construire des systèmes cognitifs artificiels (d'ordinaire appelé agents) constitués de senseurs, d'actuateurs, d'un corps et d'une architecture cognitive particulière pour mener à bien des expériences en sciences cognitives. Ces agents peuvent être amenés à effectuer des tâches dans des environnements réels et à interagir avec d'autres agents ou avec des êtres humains. Grâce aux avancées récentes de la robotique et des technologies de réseaux, il est désormais possible de construire ces agents, mais ceci encore au prix de grandes difficultés techniques. Avec ce type d'agents artificiels, il est possible de changer tous les aspects de leur structure cognitive, de l'environnement dans lequel ils évoluent et donc d'étudier les relations de cause à effet comme dans les expériences menées en physique, en chimie ou en biologie.

## 2. Les origines du langage

L'expérience des "Têtes Parlantes" (Talking Heads), qui est installée au Palais de la Découverte, est un exemple de ce nouveau type d'expérimentation. L'expérience étudie une des questions les plus fascinantes des sciences cognitives, mais qui pour l'essentiel reste à ce jour non résolue : les origines et l'évolution du langage. Cette question concerne à la fois l'acquisition du langage et du sens chez l'enfant (ontogénèse), l'émergence d'un langage complexe et de la culture qui lui est associée dans l'espèce humaine (glossogénèse). Le langage et le sens sont considérés comme étroitement liés dans la mesure où il n'est pas possible d'étudier la question des origines du langage sans étudier la question de la construction du sens et vice-versa.

L'expérience est associée à l'exposition "Evolution de la communication" qui se tient au Palais de la Découverte à partir de Mai 2000. L'exposition montre différents types de communications animales, des plus simples (signaux chimiques, phéromones) au plus complexes (cris d'alarmes des singes vervet). Si le langage humain est aujourd'hui considéré comme la forme de communication la plus complexe, une question reste ouverte: quelle évolution a conduit à un tel saut qualitatif ?

Par certains aspects, le problème des origines du langage et du sens est similaire au problème des origines de la vie et de l'évolution des espèces. Il n'y a pas de données sur les développements des premiers langages humains dont on estime aujourd'hui l'apparition entre 30 000 et 100 000 ans avant notre ère. De même, il est difficile d'observer l'acquisition du langage chez l'enfant compte tenu du temps nécessaire à un tel apprentissage et de l'impossibilité de suivre le développement des structures sémantiques dans la tête de l'enfant.

L'expérience des "Têtes Parlantes" est la première expérience à grande échelle mise en place pour étudier certains aspects du problème des origines du langage. L'expérience étudie plus particulièrement l'acquisition de catégories ancrées visuellement (gauche/droite, rouge/vert, petit/grand, etc) et des mots utilisés pour nommer ces catégories. Il s'agit de savoir quelles capacités cognitives sont suffisantes pour qu'une population d'agents s'accordent sur un lexique donné. Une telle expérience permettra peut-être ainsi de mieux délimiter la spécificité du langage humain par rapport à certaines formes de communication animale, en particulier chez les primates.

## 3. L'expérience

L'expérience présentée au Palais de la Découverte se présente sous la forme d'une plate-forme robotique comprenant deux "Têtes Parlantes". Chaque "Tête Parlante" est constituée d'une caméra qui bouge verticalement et horizontalement, d'un ordinateur qui implémente les fonctions cognitives (perception, catégorisation, recherche dans le lexique, etc.), d'un écran où sont visualisés les états internes de l'agent chargé dans le corps de la "Tête Parlante", d'un moniteur qui montre la scène que voit la caméra et des entrées et sorties audio. Ce système est placé en face d'un tableau blanc sur lequel sont exposées des formes géométriques simples : des triangles, des cercles et des rectangles de différentes couleurs.



Figure 1: Dans chaque site, deux caméras mobiles sont placées devant un tableau blanc (Site d'Anvers - Première phase de l'expérience)

L'installation du Palais de la Découverte fait partie d'un réseau de plates-formes robotiques situées dans d'autres expositions (Exposition N01SE à Londres et à Cambridge) et dans des laboratoires de recherche (Sony CSL Paris, VUB AI Lab Bruxelles). Ces plates-formes sont reliées les unes aux autres par Internet. Un agent, chargé dans le corps "physique" d'une "Tête Parlante", peut se téléporter dans un autre système de "Têtes Parlantes". Ainsi il peut découvrir de nouveaux environnements.

Les agents présents sur chaque plate-forme peuvent interagir les uns avec les autres, en participant à un jeu de langage nommé le "guessing game". Un agent joue le rôle de locuteur et un autre, celui de l'interlocuteur. Les agents jouent en alternance l'un ou l'autre rôle. Le locuteur doit désigner un objet à l'interlocuteur en utilisant une forme verbale. L'interlocuteur interprète cette forme verbale et pointe vers l'objet supposé être désigné par le locuteur. Le jeu est un succès si le locuteur considère que l'interlocuteur a deviné juste. Si c'est un échec, le locuteur indique, de façon non verbale, le sujet qu'il voulait désigner à l'interlocuteur, et les deux agents adaptent leurs structures internes pour être plus efficaces dans les futurs jeux.

Par exemple, si le contexte contient [1] un carré rouge, [2] un triangle bleu, [3] un cercle vert, le locuteur peut alors dire quelque chose comme "le rouge" pour identifier [1]. Si le contexte contient aussi un triangle rouge, il doit être plus précis et dire quelque chose comme "le carré rouge". Bien sûr, si les agents n'interagissent qu'entre eux, ils ne diront pas "le carré rouge" mais

utiliseront leur propre langue et concepts qui, a priori, ne ressembleront pas au Français. Par exemple, un agent pourra dire "Malewina" pour signifier [EN-HAUT A-GAUCHE NIVEAU-ELEVE-DE-ROUGE].

Nous jouons souvent aux "guessing games" lorsque nous parlons. Lors d'un déjeuner, si je demande à un ami de me passer le "vin", je m'attends à ce qu'il me tende la bouteille située de l'autre côté de la table. S'il me passe le "pain" ou s'il me regarde avec un visage interrogatif (peut-être ne comprend-il pas le français ?), je lui montrerai sans doute, par un geste de la main, l'objet désiré. Les échecs en communication ne sont pas rares lors de "guessing games". En effet les deux interlocuteurs ne voient pas la scène du même point de vue, ne catégorisent peut être pas les différents objets qui la composent de la même manière et les indications gestuelles, sensées lever les éventuelles ambiguïtés, sont elles-mêmes vagues et imprécises.

Dans l'expérience des "Têtes parlantes", lorsqu'un nouvel agent est créé il ne possède ni lexique, ni système de catégories. Ce n'est qu'en interagissant avec d'autres agents qu'il va pouvoir construire des catégories pour décrire les scènes qu'il voit et associer à ces catégories des mots particuliers partagés par les autres agents.

On peut suivre l'évolution de cette expérience mondiale sur une interface web (<http://talking-heads.csl.sony.fr>) où divers outils d'analyse sont proposés. Les utilisateurs peuvent explorer le lexique collectif qui se construit en observant ce que recouvrent les mots utilisés. Chaque utilisateur peut également créer son propre agent et définir la succession des sites qu'il doit visiter. L'agent revient sur le serveur central lorsqu'il a achevé son parcours, et son créateur peut alors inspecter l'état de son lexique, consulter les photographies des interactions auxquelles il a participé. Il peut aussi influencer le vocabulaire de l'agent en suggérant des mots pour les situations présentées. Ainsi le lexique global émergent est un mélange de mots inventés par les agents et de mots suggérés par des humains.

#### 4. Résultats.

Durant la première phase de l'expérience qui s'est déroulée de Juin à Novembre 1999, cinq plateformes ont été actives (Paris, Bruxelles, Anvers, Tokyo, Lausanne). Un demi million d'interactions ancrées ont été jouées, un millier d'agents ont été créés, plus de 10 000 mots inventés et 300 catégories construites. Le coeur du lexique qui a émergé est constitué par une dizaine de mots qui désignent des catégories simples pour décrire les positions et les couleurs. Ce lexique de base mêle des mots inventés par les agents ("Kazozo" pour les objets verts, "Sesubipu" pour les objets à droite, etc.) et des mots de différentes langues introduits par des humains ("down" pour les objets en bas, "rouge" pour les objets rouges, etc.). Ce lexique permet d'identifier les formes géométriques simples dans les environnements utilisés durant la première phase de l'expérience.

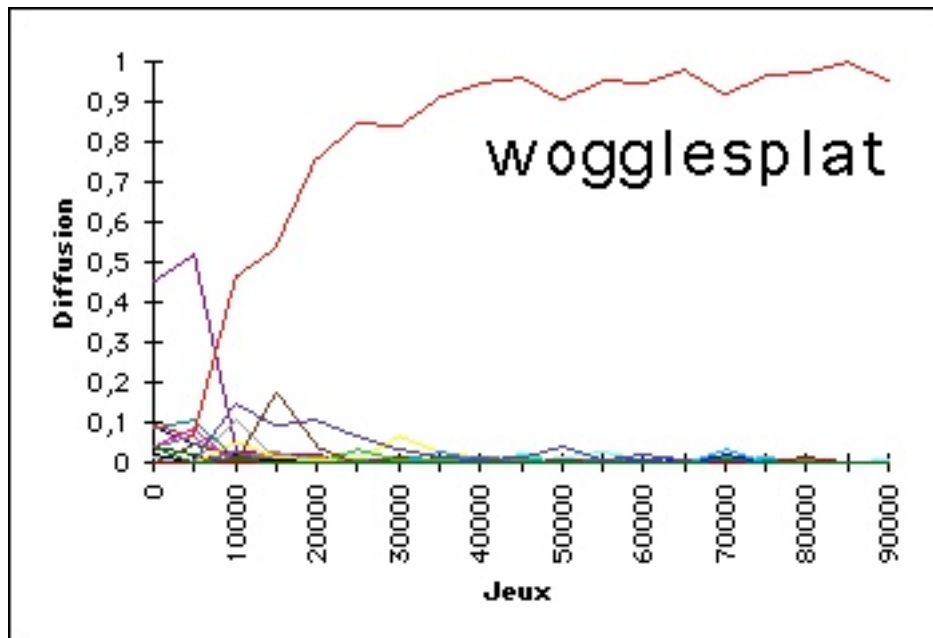


Figure 2: Compétition entre différents mots pour nommer le sens "à gauche" ([HPOS 0 0.5]), Après une première période de compétition intense, le mot "Wogglesplat" s'impose comme le mot utilisé par la majorité des agents.

Une seconde phase de l'expérience a débuté en Janvier 2000 avec les sites de Paris, Bruxelles, Cambridge et Londres. Dès Avril 2000, le site de Cambridge sera remplacé par celui du Palais de la Découverte. Une nouvelle population d'agents tente, comme lors de la première phase, de s'accorder sur le sens de mots inventés ou appris pour désigner des éléments des environnements qui leur sont présentés.

L'expérience permet d'identifier un ensemble de conditions suffisantes qui permet l'émergence d'un lexique commun. Ces conditions portent sur la structure cognitive des agents, le type d'interactions auxquelles ils participent et le milieu dans lequel ils évoluent. Par contraste, une telle expérience montre également que certaines hypothèses ne sont pas nécessaires pour obtenir le même phénomène.

## 5. Conclusions.

L'expérience des "Têtes Parlantes" présentée dès Avril 2000 au Palais de la Découverte montre comment des agents artificiels dotés de capacités cognitives spécifiques peuvent développer collectivement un lexique pour désigner des objets de monde réel. Il s'agit d'une expérience d'un type nouveau. Dans des domaines où l'expérimentation n'est plus possible, le débat d'idées menace de s'enliser faute d'évidence empirique. Construire des expériences comme les "Têtes Parlantes" permet de tester et de comparer diverses hypothèses explicatives. Le débat se réorganise autour de nouveaux "résultats expérimentaux" et les méthodes scientifiques classiques peuvent être dès lors, appliquées. Nous espérons que cette démarche, qui est aujourd'hui mise en œuvre par un nombre croissant de laboratoires de par le monde, constituera dans un proche avenir un langage commun et un support de débats dans ce type de recherches hautement interdisciplinaires.

## 6. Remerciements

Les "Têtes Parlantes" sont le fruit de l'effort collectif d'une équipe de chercheurs aux laboratoires Sony CSL Paris et VUB AI Lab. Nous remercions en particulier Angus McIntyre et Joris van Looveren. Le site web [talking-heads.csl.sony.fr](http://talking-heads.csl.sony.fr) a été développé en collaboration avec Silvère Tajan et Alexis Agahi.