

# Sony CSL

Sony Computer Science Laboratory Paris  
<http://www.csl.sony.fr>

**President :**  
Dr. Mario Tokoro

**Director :**  
Prof. Dr. Luc Steels

**Date et adresse :**  
6 octobre,  
9h30-16h00,  
Sony CSL Paris,  
6, rue Amyot,  
75005 Paris

**Site web :**  
<http://www.csl.sony.fr>

**Contact :** Sophie Boucher  
[sophie.boucher@csl.sony.fr](mailto:sophie.boucher@csl.sony.fr)

## Journée portes ouvertes 2004 Sony CSL Paris

### Avancées en robotique, sciences cognitives, musique, et neurosciences computationnelles

#### Sony CSL Paris : un laboratoire de recherche fondamentale

Sony CSL Paris a été créé en 1996 sur le modèle de Sony CSL Tokyo, inauguré 10 ans auparavant. L'objectif de ce laboratoire est de contribuer largement aux progrès de la science et de faire de la recherche fondamentale et originale, sans contraintes liées à des applications commerciales. Notre politique est de rendre nos résultats publics et de collaborer étroitement avec les institutions académiques. Les chercheurs de CSL Paris viennent de pays très divers. Leur travail est interdisciplinaire, relevant des domaines de la robotique, des sciences cognitives, de la musique et des neurosciences computationnelles.

#### Comprendre la communication humaine, construire des robots sociaux



**The Playground  
Experiment (2004)**  
<http://playground.csl.sony.fr>

Un des thèmes fondateurs de Sony CSL Paris est la communication. Il s'agit, en tant que défi scientifique, d'essayer de comprendre les fondations de la communication humaine et, en tant que défi technologique, de construire des robots avec lesquels nous pouvons avoir des interactions sociales complexes. Ces deux aspects sont intrinsèquement liés de notre point de vue, car les robots que nous construisons nous servent comme modèles des mécanismes qui sont en œuvre dans l'apprentissage du langage par les humains. Cette méthodologie synthétique nous permet de mieux comprendre le développement du langage humain, et cette compréhension nous fournit en retour des intuitions pour construire des robots parlants.

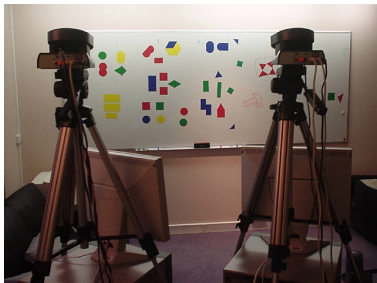


**Le projet Talking AIBO  
(2002)**

L'expérience des **Têtes Parlantes (1999)** a été notre première avancée dans ce domaine. Ce système robotique a montré comment un lexique et un système de catégories partagées pouvaient s'auto-organiser dans une société de robots au cours de leurs interactions culturelles. Cette expérience a posé les premières briques d'une nouvelle théorie des origines du langage : nous avons montré que le langage pouvait être considéré comme un système adaptatif complexe, dans lequel les interactions sociales locales pouvaient mener à la formation de structures culturelles globales dans une dynamique d'auto-organisation. Les résultats de cette expérience ont donné lieu à

de nombreux articles et à deux ouvrages (1). Le système robotique a été installé dans plusieurs musées, en particulier à Paris, Londres et Cambridge.

Continuant dans cette direction, nous avons développé l'expérience de **Maïdo et Gurby (2001)**, montrant comment le même cadre théorique pouvait rendre compte de la formation culturelle des véhicules physiques du langage : les systèmes de sons partagés (2). Cela nous a aussi permis de fournir une explication de la diversité des sons qui composent les langues humaines, et en même temps d'expliquer pourquoi il existe un certain nombre de régularités statistiques dans leurs phonologies. Le système robotique a été installé pendant deux ans dans l'exposition "L'Homme Transformé", à la Cité des Sciences et de l'Industrie de Paris.



**L'expérience des Têtes Parlantes (1999)**

En plus des avancées purement scientifiques, ces deux expérimentations nous ont permis de faire des réalisations technologiques innovantes. Par exemple, nous avons utilisé les connaissances acquises dans l'expérience des Têtes Parlantes pour construire un système et un cadre interactionnel dans lequel un humain peut enseigner à un robot comment nommer les objets qui sont dans son environnement. Le projet **Talking AIBO** illustre cette approche avec un premier prototype robotique (3). Dans le cadre de l'expérience "Maïdo et Gurby", nous avons aussi développé des technologies de synthèse vocale émotionnelle et de reconnaissance des émotions dans la voix humaine (4). Cette technologie est actuellement utilisée dans le système de synthèse vocale du robot **humanoïde Qrio**, développé par Sony (5).

Nos recherches se développent maintenant dans deux directions. Ces nouveaux résultats seront présentés lors de la journée portes ouvertes 2004. Un des enjeux est de montrer que les dynamiques collectives identifiées pour l'auto-organisation lexicale peuvent être étendues pour rendre compte de l'émergence de formes grammaticales. L'objectif est de montrer comment une syntaxe complexe et des structures sémantiques non-triviales peuvent émerger et évoluer culturellement. Pour cela, un nouveau cadre formel a été développé, appelé les "**grammaires constructionnelles fluides**" (6). Il repousse les limites imposées par les systèmes déjà existants.



**Maïdo et Gurby (2001)**

L'autre direction consiste à étudier les prérequis nécessaires à l'émergence de la communication. Il s'agit de montrer comment un certain nombre de mécanismes qui étaient pré-programmés dans l'expérience des Têtes Parlantes et dans celle de Maïdo et Gurby peuvent être développées de manière autonome par les robots.

Nous menons ainsi des travaux sur **l'attention partagée**. Il s'agit de montrer comment des robots peuvent développer ensemble des protocoles sociaux pour interpréter et manipuler

les comportements attentionnels. Nous montrerons de telles techniques lors de la journée portes ouvertes.

Un autre exemple est notre travail sur **les origines auto-organisées des systèmes de vocalisations**. Nous présentons un modèle des origines de la parole, dans lequel les agents ne possèdent aucune capacité pré-programmée d'interaction sociale coordonnée. Ce travail (7), réalisé dans le contexte d'une thèse de doctorat, est à la base d'une avancée conceptuelle dans le domaine académique des origines du langage. Il sera publié l'année prochaine dans un livre édité par Oxford University Press. Il a aussi reçu le "Prix Le Monde de la recherche universitaire", qui élit chaque année les 15 meilleures thèses soutenues en France et dans un domaine de sciences exactes.

Enfin, nous essayons de comprendre comment des agents peuvent développer des interactions coordonnées et être intéressés par la communication sans que cela soit pré-programmé. Il s'agit de comprendre les dynamiques qui sont à la base du développement du bébé durant sa première année. Quels sont les mécanismes qui le guident vers des activités de complexité croissante, depuis les boucles sensori-motrices les plus simples jusqu'aux interactions linguistiques ? Pour répondre à ces questions nous travaillons sur des modèles informatiques de la **curiosité**. Nous équipons les robots de motivations internes qui les poussent à apprendre de manière active. Pendant la journée portes ouvertes, nous présenterons l'expérience du tapis d'éveil ("**The Playground Experiment**") (8), dans laquelle un robot muni de ce système de motivation joue sur un tapis d'éveil de bébé. Nous montrons comment ce robot se développe et s'engage de manière autonome dans des activités de plus en plus complexes.

Ces nouveaux projets font partie d'un **projet européen** plus large, appelé **ECagents**, lancé en janvier 2004 et auquel nous participons activement. Dans le cadre de ce projet, des concepts, des outils et des modèles pour l'analyse des sociétés d'agents naturels et artificiels sont élaborés, ainsi que des algorithmes pour mettre au point des sociétés d'agents robotiques communicants et en perpétuelle évolution. Le consortium réunit des scientifiques de renom spécialistes de la cognition et des systèmes complexes, qui viennent d'Italie, de France, de Suisse, de Belgique, de Hongrie, de Suède, d'Allemagne, d'Espagne et du Japon. Pour plus d'information, un site web existe: <http://www.ecagents.org>.



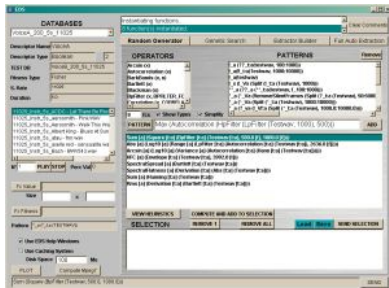
**Un enfant jouant avec le Continuator (2003)**

### **Enrichir l'expérience musicale: accès, écoute et pratique**

Le but principal de la recherche dans le domaine musical à Sony CSL Paris est d'inventer des technologies qui enrichissent l'expérience musicale du point de vue de l'utilisateur. Pour atteindre ce but, nous développons des prototypes variés qui atténuent la distinction traditionnelle entre l'écoute et la pratique

instrumentale. Ces travaux se répartissent en deux grands domaines: les systèmes d'interaction musicale, où nous essayons de développer de nouvelles manières de faire de la musique, et les systèmes d'accès à la musique, où le but est d'arriver à maîtriser la complexité des grands catalogues de morceaux.

Les projets concernant l'interaction musicale ont générés des systèmes comme **MusicSpace** (10), un système de contrôle de la spatialisation de la musique dans lequel les utilisateurs peuvent modifier en temps réel la position des sources sonores comme celles des instruments. Une représentation du savoir-faire des ingénieurs en acoustique, sous forme de mécanismes à bases de contraintes, s'assure que toutes les actions de l'utilisateur sont corrigées pour que le résultat produisent un mixage satisfaisant.



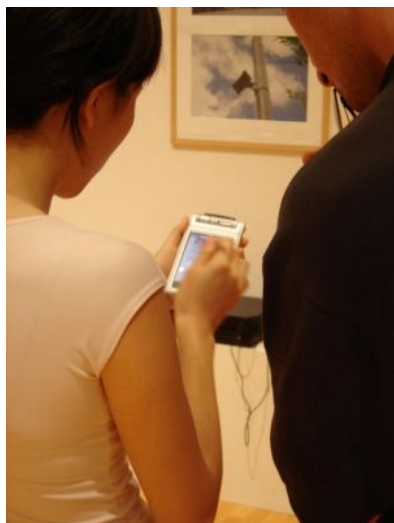
**Le système EDS (2004)**

Le projet du **Continuator** (11) a produit un système qui permet une interaction en temps réel avec un instrument qui peut apprendre un style musical. Avec le Continuator, l'utilisateur peut jouer de la musique quand il le veut, et le système s'engage automatiquement dans un dialogue et produit des phrases musicales qui sont la continuation de celles de l'utilisateur. Progressivement, le dialogue devient de plus en plus intéressant car le système apprend en continu et à partir de toutes les interactions précédentes. Nous étudions justement l'intérêt qui est produit chez l'utilisateur et essayons de relier ce travail à la théorie psychologique du "Flow" (12).

Les projets concernant l'accès à la musique se focalisent sur la gestion des grandes bases de données en utilisant des techniques orientées sur le contenu (13). Nous travaillons sur des représentations des titres musicaux (et des catalogues eux-mêmes) qui permettent à des utilisateurs de trouver des morceaux qui les intéressent, en particulier quand ils ne les connaissent pas déjà. Cela implique deux axes de recherche. Le premier est celui de l'extraction automatique d'information pertinente à partir des titres et du signal acoustique, comme la structure rythmique ou le timbre (14). Le **projet EDS** (15) est une tentative de généralisation de ce processus d'extraction de propriétés complexes qui met en oeuvre une approche novatrice: c'est le système lui-même qui va inventer et améliorer des mesures pertinentes pour un problème audio particulier et à partir des signaux acoustiques. Nous avons montré que EDS a de meilleures performances que les approches classiques sur des problèmes standards d'acoustique (e.g. détection de la voix chantée).

Le second axe de recherche concerne les interfaces homme-machine, qui permettent aux utilisateurs de trouver de la musique à partir de recherches basées sur des mesures de similarités. Nous avons développé plusieurs browsers musicaux,

comme par exemple le browser musical Cuidado, et plus récemment des systèmes d'accès et de contrôle de la musique basés sur les gestes.



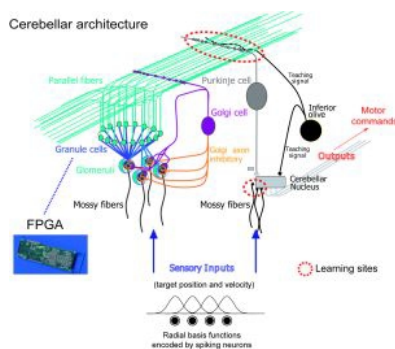
**Les communautés créatives (2004)**

## Les communautés créatives et le partage de contenu

Les technologies digitales et l'internet ont accéléré les modes de partage et de production des contenus qui reposent sur la collaboration et ne nécessitent pas de contrôle centralisé. Nos recherches sur les communautés créatives ont pour but de mieux comprendre ces modes de production. Nous essayons de développer des technologies qui permettent de les améliorer.

On peut prendre l'exemple du partage des données dans les systèmes peer-to-peer. Les participants de ces systèmes organisent souvent leurs propres données d'une manière qui est particulière à chacun, ce qui complique la formulation des recherches automatiques. Une solution à ce problème est d'imposer un standard d'organisation des informations pour tous les utilisateurs. Cependant, l'expérience des Têtes Parlantes (1999) a montré comment des agents pouvaient auto-organiser un lexique partagé comme effet co-latéral de leurs interactions. La journée portes ouvertes nous donnera l'occasion de montrer comment l'utilisation de ces techniques de sémantique émergente permet l'échange d'informations dans de tels systèmes distribués.

Une seconde démonstration se concentrera sur la participation du public pour la production d'un contenu. Il s'agit du système que nous avons expérimenté à l'exposition "Social Capital: Formes et Interaction", organisée par le Whitney Independent Study Program de New-York. Dans cette installation, nous avons créé un espace virtuel et partagé dans lequel les visiteurs pouvaient mettre des images et les annoter avec du texte.



**Modèle du cervelet (2004)**

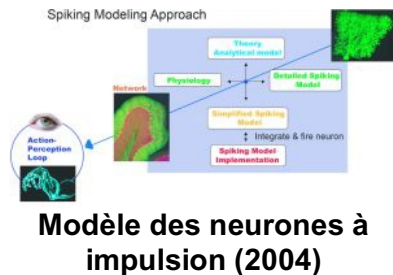
## Comprendre les mécanisme cérébraux

Bien que des avancées récentes aient été faites dans les domaines de l'apprentissage automatique des systèmes artificiels, des problèmes cruciaux demeurent sans solution. Il s'agit de la compréhension de capacités comme l'adaptation, la généralisation, l'apprentissage en continu, ou la conceptualisation. Est-ce que les cellules nerveuses du cerveau sont un support computationnel dont les caractéristiques et les représentations permettraient à des machines de résoudre ces problèmes ? Est-ce que les structures neurales du cerveau utilisent une méthode de computation unique et différente de celle qui est utilisée par les circuits intégrés et nos ordinateurs ? Le groupe de neuroscience de Sony CSL Paris essaie de répondre à ces questions en étudiant les mécanismes en oeuvre dans le cerveau et son substrat neuronal

dans le cerveau et son substrat neuronal.

Les neurones dans le cerveau traitent l'information de manière analogique, mais ils communiquent entre eux sous la forme d'impulsions électriques qui sont essentiellement digitales (0 ou 1) et asynchrones. Nous essayons de comprendre comment des neurones à impulsion permettent de prendre des décisions et d'apprendre rapidement, dans le but de construire des systèmes adaptatifs temps-réel qui apprennent en continu.

Le groupe Neurosciences de Sony CSL Paris et ses partenaires du projet européen SpikeFORCE, ont présenté une technologie innovante à la conférence internationale « Neural Information Processing System » (NIPS) qui s'est tenue à Vancouver en décembre 2003. Le système était un hybride software-hardware pour le contrôle moteur adaptatif. Ce système est l'un des premiers qui réalise la boucle d'action-perception complètement à base de neurones à impulsions.



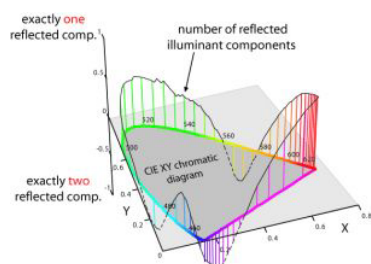
Ce système imite la réponse des neurones dans le cervelet en calculant leurs impulsions ; ce sont ces impulsions que les neurones utilisent pour transmettre des informations. Le cervelet est une partie du cerveau qui est très impliquée dans l'apprentissage et la coordination des activités motrices. Le système hybride combine les réponses de différents types de neurones qui sont générées sur des puces électroniques et dans des logiciels qui tournent sur des ordinateurs connectés en réseau. Plus de détails sont disponibles sur le site web : [www.spikeforce.org/nipsdemo.html](http://www.spikeforce.org/nipsdemo.html).

Le consortium SpikeFORCE réunit des chercheurs en neurosciences, des physiciens et des ingénieurs de quatre institutions européennes. Son but est de produire un modèle du cervelet, basé sur les derniers résultats physiologiques et d'analyses computationnelles, qui peut être implémenté sur un ordinateur et/ou sur puces électroniques et tourner en temps réel. Le consortium s'intéresse aussi aux avantages des neurones à impulsion et de leur implémentation physique, avec pour objectif de construire des algorithmes d'apprentissage qui permettraient de construire des systèmes adaptatifs en temps réel.

Nous avons élaboré un logiciel qui permet de simuler des réseaux de neurones impulsifs en temps réel. A partir de celui-ci, nous avons construit un modèle du cervelet composé de milliers de neurones et qui apprend à suivre visuellement des cibles en mouvement suivant différentes trajectoires. Durant la journée portes ouvertes, nous présenterons ce modèle du cervelet tournant et apprenant aussi en temps-réel.

## Comprendre les lois sensori-motrices

Perception et contrôle moteur ont longtemps été considérés comme des problèmes distincts. Il est aujourd'hui clair que ce n'est pas le cas, et l'absence d'un cadre théorique général permettant de les rapprocher se fait de plus en plus ressentir. Nous travaillons au développement d'un tel cadre, permettant de formaliser les lois sensori-motrices et de comprendre leurs propriétés fondamentales.



### Les lois sensori-motrices de la perception des couleurs (2004)

Ce cadre a été utilisé récemment pour caractériser les lois sensori-motrices auxquelles sont confrontés les sujets trichromates lors de l'exploration visuelle de surfaces colorées. Il a permis d'étudier les propriétés de reflectance des surfaces sous un jour nouveau, et de montrer que les surfaces rouges, vertes, bleues d'une part, et jaunes, cyan, violettes d'autre part, correspondent à certaines singularités des propriétés de reflectance au regard des photopigments humains.

Les psychologues expérimentaux ont accumulé de nombreux arguments soutenant l'idée que toutes les couleurs ne sont pas équivalentes d'un point de vue perceptif, certaines couleurs telles que le rouge ou le vert semblant perceptuellement irréductibles, par opposition à d'autres couleurs comme le orange ou le jaune verdatre.

La question de l'origine de cette asymétrie perceptuelle est débattue depuis des décennies entre neurophysiologistes, linguistes, cognitivistes, et philosophes. Cette découverte apporte des arguments inédits à ce débat, puisqu'aucune singularité physique justifiant la singularité de certaines couleurs n'avait jusqu'à présent été clairement identifiée.

Notre travail sera présenté à la journée portes ouvertes de Sony CSL Paris, et est effectué en collaboration avec le CNRS et l'Ecole Normale Supérieure de Paris.

Une autre avancée concernant la compréhension de notre perception des couleurs nous a permis d'expliquer les interactions entre la catégorisation des couleurs et la manière dont on les nomme dans une langue donnée. Ce travail a été accepté comme article principal dans la prestigieuse revue Behavioral and Brain Sciences (17). Dans une série d'expériences systématiques, en collaboration avec des chercheurs du laboratoire d'intelligence artificielle de l'université de Bruxelles, nous avons montré comment une société d'agents pouvait construire culturellement à la fois un système de catégories de couleurs et un système de noms pour ces couleurs qui soient partagés par tous les membres de la communauté. Ces expériences, basées sur les mécanismes d'interactions et les structures cognitives développées dans les Têtes Parlantes (1999), a permis de montrer comment des individus pouvaient partager les mêmes catégories sans que cela soit inné.

## References:

### (1) Sur l'expérience des têtes parlantes

#### Articles

Steels, L. and Kaplan, F. [Bootstrapping grounded word semantics](#). In Briscoe, T., editor, Linguistic evolution through language acquisition: formal and computational models, pages 53-73, Cambridge University Press. Cambridge, 2002.

Steels, L., Kaplan, F., McIntyre, A. and Van Looveren, J. [Crucial factors in the origins of word-meaning](#). In Wray, A., editor, The Transition to Language, pages 252-271, Oxford University Press. Oxford, UK, 2002.

#### Livres

Steels, L. The Talking Head Experiment, Limited edition for the Laboratorium exhibition.

Kaplan, F. La naissance d'une langue chez les robots, Hermès Science, 2001.

The Talking Heads Experiment web site : <http://talking-heads.csl.sony.fr/>

### (2) Sur l'expérience Maïdo et Gurby

Oudeyer, P-Y. Origins and Learnability of Syllable Systems, a Cultural Evolutionary Model, in Artificial Evolution, pp. 143-155, eds. P. Collet, C. Fonlupt, J.K. Hao, E. Lutton, M. Schonenaus, LNCS 2310, Springer Verlag.

Videos available on: <http://www.csl.sony.fr/~py/videos.html>

(4) Oudeyer P-Y. (2002) The production and recognition of emotions in speech: features and algorithms, International Journal in Human-Computer Studies, vol. 59/1-2, pp. 157-183, special issue on Affective Computing.

(3) Steels, L. and Kaplan, F. [AIBO's first words: The social learning of language and meaning](#). Evolution of Communication, 4(1):3-32 2001.

The Talking AIBO Project web site: <http://www.fkaplan.com>

(5) Qrio web site: <http://www.sony.net/SonyInfo/QRIO/>

(6) Steels, L. [Constructivist Development of Grounded Construction Grammars](#). In Daelemans, W., editor, Proceedings Annual Meeting of Association for Computational Linguistics, 2004.

(7) Oudeyer P-Y. (2005) From Holistic to Discrete Vocalizations, to appear in the volume "Language Origins: Perspectives on Evolution", ed. Tallerman M., Oxford University Press.

Oudeyer, P-Y. (to appear as a book) The Self-Organization of Speech Sounds, Studies in the Evolution of Language, Oxford University Press.

The Origins of Speech Sounds Project web site: [www.csl.sony.fr/~py](http://www.csl.sony.fr/~py)

(8) Oudeyer, P-Y. and Kaplan, F. [Intelligent adaptive curiosity: a source of self-development](#). In Luc Berthouze and Hideki Kozima and Christopher G. Prince and Giulio Sandini and Georgi Stojanov and G. Metta and C. Balkenius, editor, Proceedings of the 4th International Workshop on Epigenetic Robotics, vol. 117, pages 127-130, 2004. Lund University Cognitive Studies.

The Playground Experiment web site: <http://playground.csl.sony.fr>

(10) Pachet, F. and Delerue, O. (2000) [On-The-Fly Multi-Track Mixing](#). Proceedings of AES 109th Convention, Los Angeles, 2000. AES.

(11) Pachet, F. (2004) [Enhancing Individual Creativity with Interactive Musical Reflective Systems](#). Psychology Press.

(12) Pachet, F., Addressi, Anna-Rita (2004) [When Children Reflect on Their Playing Style: The Continuator](#). ACM Computers in Entertainment, 1(2)

(13) Pachet, F. (2003) [Content Management for Electronic Music Distribution: The Real Issues](#). Communications of the ACM, April

(14) Aucouturier, J.-J., Pachet, F. and Sandler, M. (2004) [The Way It Sounds : Timbre Models For Analysis and Retrieval of Polyphonic Music Signals](#). IEEE Transactions of Multimedia.

(15) Zils, A. and Pachet, F. (2004) [Automatic Extraction of Music Descriptors from Acoustic Signals using EDS](#). Proceedings of the 116th AES Convention, May

(17) Steels, L. and T. Belpaeme (2005) Coordinating Perceptually Grounded Categories through Language. A Case Study for Colour. Target article Behavioral and Brain Sciences.