



CONACYT-Mexique  
CNRS-INRIA-Université Joseph Fourier Grenoble I –France

**Fiche-projet**  
**2004**

Laboratoire Franco-Mexicain d'Informatique  
LAFMI

Nom coordinateur/coordinatrice du projet en France : **Peter Ford Dominey**

Nombre del Coordinador de Proyecto en México : **Alfredo Weitzenfeld**

**1. Titre du projet** : Une Interface en Langage Naturel pour Commande et Contrôle Robotique

Mots-clés (4 maximum) : **human language technology ; autonomous robots ;**

**2. Établissements :**

**En France :**

Etablissement du coordinateur ou de la coordinatrice : **Institut des Sciences Cognitives**

Laboratoire (ou équipe) : **Sequential Cognition and Language Group**

Statut de l'unité : **CNRS UMR 5015**

**Au Mexique :**

Etablissement du coordinateur ou de la coordinatrice : **ITAM**

Laboratoire (ou équipe) : **Div. Académica de Ingeniería**

**3. Coordinateurs du projet :****En France :**Nom et prénom : **Dominey, Peter Ford**Grade : **CNRS CR-1, HDR**Adresse administrative<sup>1,2</sup> :**Institut des Sciences Cognitives, CNRS UMR 5015, 67 Bd Pinel, 69675 Bron Cedex**Téléphone : **0437911266**Télécopie : **0437911210**Mél : **dominey@isc.cnrs.fr****Au Mexique :**Nom et prénom : **Weitzenfeld, Alfredo**Grade : **Professor**Adresse administrative<sup>1,2</sup> :**ITAM, Río Hondo #1, San Angel Tizapán, México DF, CP 0100**Téléphone : **55-56284060**Télécopie : **55-56284065**Mél : **alfredo@itam.mx**

<sup>1</sup> Auquel appartient le responsable scientifique du projet.

<sup>2</sup> Indiquer le type et le numéro de l'Unité dans l'organisme Université, INRIA, CNRS, autre .... (exemple : **UMR 691 CNRS**).

## 5. Description détaillée du projet:

### 5.1 Identification du sujet et état des connaissances:

#### **Sujet : interaction homme-robot en temps réel dirigé vers un but**

Les robots deviennent de plus en plus capables de fonctions complexes, sensorielles et motrices, avec la capacité d'exécuter des tâches de plus en plus complexes. Ainsi, la possibilité d'interagir avec eux d'une manière ergonomique, adaptative et en temps réel devient un objectif de plus en plus important. Le projet actuel cherche à combiner le contrôle de robot de pointe autonome avec les systèmes cognitifs de pointe en vue d'obtenir une l'interaction parlée, c'est à dire une Interface Homme-Robot (**IHR**) parlée robuste.

Idéalement, la recherche en interaction Homme-Robot permettra la communication et la coopération naturelle, ergonomique et optimale entre les humains et les systèmes robotisés. Afin d'avancer dans cette direction, nous avons identifié deux pré-requis majeures : Premièrement, nous devons étudier un environnement de robotique vrai dans lequel les ingénieurs et les chercheurs ont déjà développé des expériences et besoins par rapport à la IHR. Deuxièmement, nous devons étudier un système pour le traitement automatique des langues (TAL) validé au niveau psychologique, et qui peut être appliqué aux problèmes robotiques arbitraires. En réponse à la première condition en ce qui concerne le contexte robotisé, nous étudierons la Commande et le Contrôle du Robot dans le contexte international de « foot Robot » ou « RoboCup », dans lequel le groupe Mexicain concourt au niveau international. Dans le contexte d'un modèle de langue psychologiquement valide, nous étudierons un modèle de correspondance entre la langue et le sens, développé par le laboratoire français. Le modèle est basé sur les aspects neurologiques et comportementaux de la langue humaine, et a été déployé dans les contextes robotisés.

#### **Etat de l'art :**

##### A. Commande et Contrôle de Robots dans le Domaine de Robocup de AIBO

**RoboCup** [3] (l'Initiative de Coupe Du Monde de Robot) est une recherche internationale et une initiative d'éducation fournissant un domaine de problèmes unifiés de football pour la robotique intelligente. Dans un tel environnement dynamique, des robots rapides multiples doivent collaborer pour atteindre leur objectif de but-marquant avec des stratégies de groupe différentes. RoboCup présente des ligues multiples, chacune avec ses propres défis tels que la vision globale vs. locale, la roue vs. les jambes, le « hardware » vs. le soft, simulation vs. temps réel, etc. Le dénominateur commun parmi ces ligues est l'autonomie du robot.

**La Ligue de AIBO de Sony** [1] est d'un intérêt particulier par rapport à notre projet puisque c'est la seule ligue où aucun développement de « hardware » robotique n'est exigée. La ligue de AIBO utilise une plate-forme de logiciel standardisée, Open-R [2], où les équipes doivent développer leurs propres algorithmes de logiciel basés sur une architecture de matériel unique. Chaque équipe consiste en quatre robots à quatre pattes, où chaque robot incorpore son propre processeur, ses articulateurs de jambe et tête, la camera et les détecteurs vidéo qui lui permettent de réagir réciproquement avec le monde extérieur. Il y a plusieurs défis à résoudre pour permettre aux AIBOs de jouer au football. Premièrement, le robot doit pouvoir traiter les images en temps réel afin d'identifier les objets dans le champ de jeu (les joueurs, le but, la balle, etc.) En plus, le robot doit pouvoir connaître son emplacement dans le champ, doit pouvoir marcher avec ses quatre jambes et exécuter des actions spécialisées telles que « donne un coup de pied à la balle vers le but » et « bloque des coups reçus, » dans le cas d'un gardien de but. Le robot doit pouvoir prendre des décisions sur le type et le moment des actions (les comportements), en partageant cette information avec les autres robots en vue de développer une stratégie d'équipe. La ligue de AIBO présente plusieurs défis intéressants pour des robots autonomes comme mes suivants :

- **Perception.** La reconnaissance d'objets visuels dépend de facteurs internes et externes au robot. En particulier, les conditions externes telles que la luminosité et le calibrage de la couleur affectent d'une manière critique l'exécution du robot. Par exemple, quand l'illumination est faible et que les couleurs peuvent être confondues les unes avec les autres, le robot peut percevoir l'information de manière erronée, affectant ainsi la localisation.

D'autre part, des tons différents de couleurs similaires peuvent confondre aussi les systèmes de perception de robots. Alors que des algorithmes de reconnaissance complexes pourraient résoudre ces problèmes, une telle approche affectera directement la vitesse de traitement.

- **Mouvement.** La marche à quatre pattes des robots exige un mouvement commun coordonné. Les mouvements rapides ont un effet direct sur l'exécution du jeu. D'autre part, la perception et le mouvement doivent être synchronisés pour optimiser les résultats.

- **Coordination.** Jouer en équipe a un effet important sur les résultats du jeu. Alors que les robots ont besoin d'être efficaces individuellement, leur coordination est la clé indispensable à une exécution de jeu réussie contre les équipes adversaires. La coordination sous forme de stratégies varie selon les capacités individuelles des robots, le nombre de joueurs à un moment particulier du jeu, etc. Il est crucial de pouvoir expérimenter des stratégies différentes et évaluer leur exécution contre des équipes et des conditions de jeu différentes.

- **Modèle de Monde.** Grâce à une communication sans fil entre robots, il est possible de construire un modèle partagé, unique du champ, de l'emplacement de balle et du robot. Un tel modèle réduit l'incertitude de la localisation du robot et enrichit les stratégies de coordination entre les robots. En percevant la balle, un robot éloigné peut informer ses équipiers leur permettant ainsi de se réorienter et de se diriger vers la balle.

- **Reconnaissance d'Action.** Il existe beaucoup de possibilités d'amélioration des stratégies de jeu de AIBO actuelles. Par exemple, la plupart des équipes ne font pas l'attention à la reconnaissance d'action, par opposition à la reconnaissance d'objets plus traditionnelle. En effet, la reconnaissance d'action, les robots doivent prévoir des actions d'une équipe adverse en reconnaissant des situations de jeu, telles que « donner un coup de pied à la balle » ou « le blocage d'un équipier ».

**Comportement :** Les comportements suivants sont les stratégies de AIBO les plus communément utilisées par les équipes : (i) le comportement individuel sans collaboration, (ii) la collaboration avec rôles statiques, et (iii) la collaboration avec rôles dynamiques.

**Comportement individuel :** Une stratégie de comportement individuel classique utilisée par plusieurs équipes [6] nécessitera deux types de joueurs, un gardien de but et un attaquant. La Figure 1 montre les transitions de l'état comportemental pour un gardien de but.

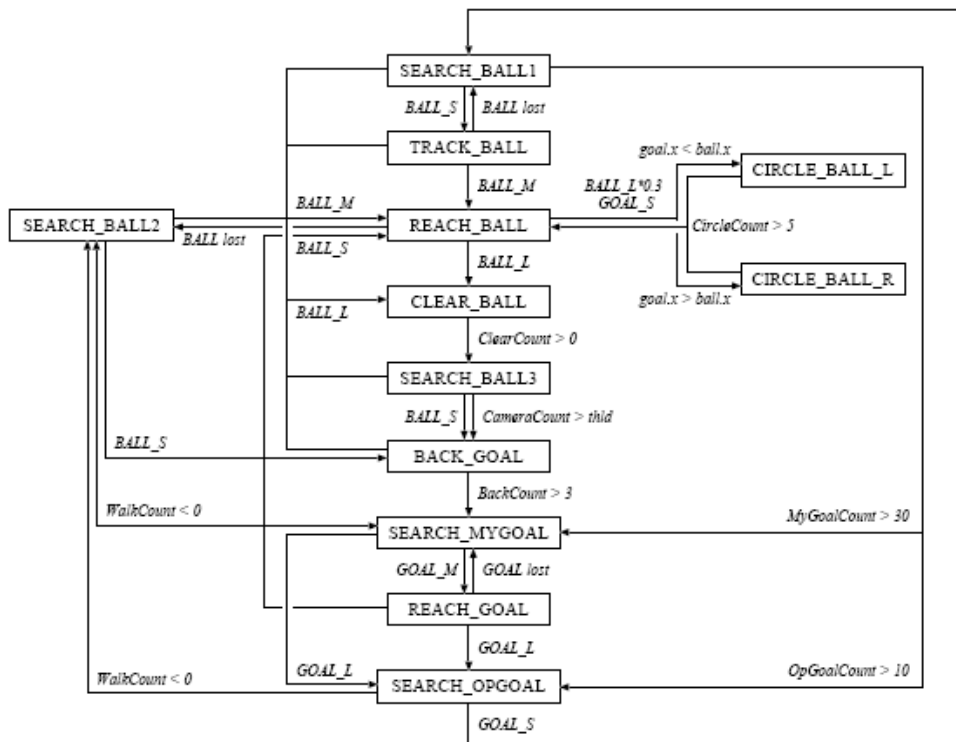


Figure 1. Transition des états pour le gardien de but

La stratégie de l'attaquant (attacker) implique habituellement de chercher la balle en encerclant le champ. Une fois que le robot trouve la balle, il s'approche et s'arrête devant elle, regarde autour afin de vérifier son orientation pour tirer la balle dans la bonne direction. L'auto-localisation est importante en vue d'une orientation correcte, et d'éviter la règle de « défenseur illégal » où aucun joueur outre le gardien de but ne doit être présent dans le secteur du but défendant.

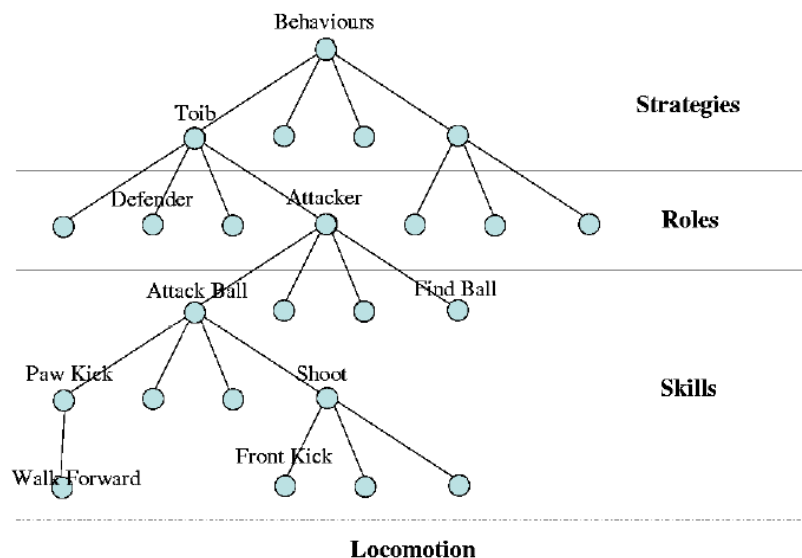
### La collaboration avec les Rôles Statiques :

Un autre genre de stratégie devra à assigner des rôles statiques aux joueurs. Par exemple, dans [7] les robots ont trois rôles distincts : l'attaquant primaire, le supporter offensif, et le supporter défensif. Dans ce modèle, chaque robot doit négocier avec les autres agents d'une manière coordonnée. Pour atteindre une telle collaboration, un modèle du monde global est construit, où les robots partagent l'information. L'information partagée inclut la position actuelle du robot et une estimation d'incertitude de cette position ; le jugement de la position de la balle et l'incertitude associée.

Les joueurs ont un rôle assigné d'une manière fixe comme : l'attaquant primaire, le supporter offensif, et le supporter défensif. L'attaquant primaire est spécifié en premier, suivi par le supporter défensif, et finalement le supporter choquant. Cet ordre est conçu pour rendre le système plus robuste. Si un ou deux des robots échouent, les membres restants de l'équipe peuvent continuer de jouer.

### La collaboration avec les Rôles Dynamiques

Pour collaborer en ayant des rôles dynamiques, une approche courante est l'usage d'un arbre de décision combinant l'information sur les données visuelles et le modèle du monde existant afin de définir la prochaine série d'actions à exécuter. L'arbre est généralement fractionné en trois niveaux, correspondant aux stratégies, aux rôles et aux compétences [5,6], comme indiqué dans la figure 2. Le sommet définit les stratégies qui peuvent être utilisées. Par exemple, une stratégie pourrait être la dispersion de robots en passant la balle. Une autre stratégie pourrait être de faire jouer tous de robots comme les attaquants, et une autre pourrait être de faire jouer tous les robots de manière défensivement. Ces stratégies différentes peuvent être développées et peuvent être essayées les unes contre les autres. Pendant un jeu, les robots peuvent être programmés pour changer de stratégies en fonction du score, ou en fonction de combien de temps restant dans le jeu.



**Figure 2.** L'arbre de décision pour la collaboration des robots avec les rôles dynamiques. Chaque stratégie incorpore un nombre de rôles que les robots peuvent exécuter. Les exemples incluent des attaquants, des supporteurs et des défenseurs. Ces stratégies sont en général dynamiquement changées pendant un jeu, sauf le rôle de gardien de but donné à un robot particulier. A un niveau inférieur chaque robot contient un certain nombre de compétences.

### Etat de l'Art B. Interaction Homme-Robot:

Principalement basée sur les développements récents dans l'analyse du discours, de la robotique, et de la puissance du calcul informatique, l'Interaction Homme-Robot est devenue un domaine de recherche important et en voie de développement rapide. Ceci est révélé par la création en 1997 de la Conférence Internationale sur les Interfaces Utilisateur Intelligent (Intelligent User Interfaces), et l'inclusion d'Interaction Humaine dans le IEEE-RAS/RSJ Conférence Internationale sur les Robots Humanoïdes, inaugurée en 2000. La recherche dans ce domaine a révélé que quatre aspects cruciaux doivent être considérés pour une interaction homme-robot réussie. Premièrement, l'environnement comportemental doit être bien défini, et assez riche pour profiter d'une telle interaction. L'environnement de RoboCup de AIBO décrit ci-dessous satisfait clairement à ces conditions. Deuxièmement, l'utilisation de supports perceptifs extralinguistiques, y compris la vision, doit être clairement caractérisée. Troisièmement, le système doit utiliser un modèle de langage approprié à l'extraction des sens des phrases, et finalement, cet ensemble doit être intégré dans un environnement de logiciel robuste pour construire un système technologique du langage humain fiable.

**Le Traitement visuel d'Événement :** Un de ses problèmes principaux de la vision informatique a été de développer des algorithmes robustes pour l'interprétation de scènes et d'événements, une fois l'objet identifié et suivi. Exploitant l'idée que des perceptions primitives (perceptual primitives) peuvent servir de base à cette robustesse, Siskind [15] a démontré des perceptions primitives de « dynamique de force » de contact, de soutien, d'attachement peuvent être extraites de séquences d'événement vidéo et utilisés pour reconnaître des événements tels que prendre, poser, empiler, sur la base de leur caractérisation dans une logique d'événement (event logic). Les résultats similaires ont été atteints par Steels et Baillie [16]. L'utilisation de ces représentations intermédiaires rend le système robuste (résistant) à la variabilité dans les paramètres de mouvement et vision. Le plus important, cette recherche a démontré que la sémantique lexicale d'un certain nombre de verbes pourrait être établie par le traitement automatique d'image. Nous avons récemment exploité récemment cette approche, en utilisant la perception primitive de « contact » pour classer des événements physiques, comme illustré dans la Figure 3 [19].

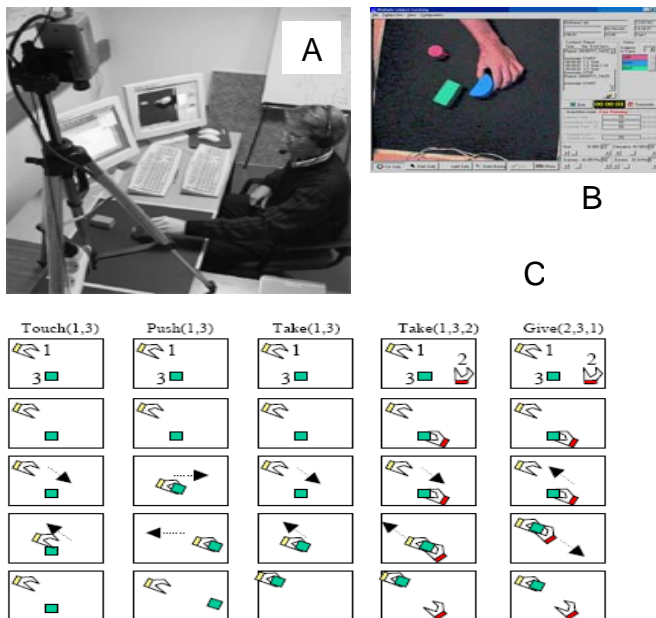


Figure 3. A. Aperçu général du Système d'Interaction d'homme-robot. L'opérateur humain exécute des actions avec les blocs, et les décrit au robot, pour que celui-la apprenne, et pose ensuite des questions au robot sur les événements pendant la démonstration d'apprentissage.

B. Image de données visuelles envoyées au robot.

C. Décours temporel des événements de contact qui permettent au système de vision reconnaître les actions suivantes : *toucher, pousser, prendre et donner.*

### Technologie du Langage Humain (HLT) et Modélisation de langage :

La Technologie de Langage Humain (HLT) se réfère aux systèmes de reconnaissance de la parole et sa conversion au texte, à la synthèse de discours et au contrôle de flux de dialogue (dialog control). Ces questions font l'objet d'un important subvention de la DARPA (la Defense Advanced Research Projects Agency), ce qui a eu pour résultat la mise au point d'outils de HLT disponible dans le domaine publiques comme le Communicateur de CMU [17], et la Trousse à outils de CSLU [18]. Nous avons utilisé le « Rapid Application Development » tool-kit de CSLU (RAD) comme décrit au-dessous.

Une question cruciale dans les interfaces homme-robot est la manière dont le sens est échangé via la langue naturelle. Dans ce contexte, différentes méthodes classiques sont utilisables y compris les modèles de grammaires context-free et n-gramme, et ces grammaires peuvent être directement utilisées dans les logiciels existants comme la Trousse à outils de CSLU, et la Suite de CPK-NLP. Dans ce contexte, nous avons développé un modèle de langage basé sur la théorie de grammaire de construction qui a l'avantage par rapport à nombreux autres d'approches de permettre l'apprentissage de nouvelles structures grammaticales par l'entraînement. Nous avons démontré l'utilité de l'approche par la grammaire de construction dans une plate-forme d'interaction d'homme-robot (Dominey et al. 2004) [19].

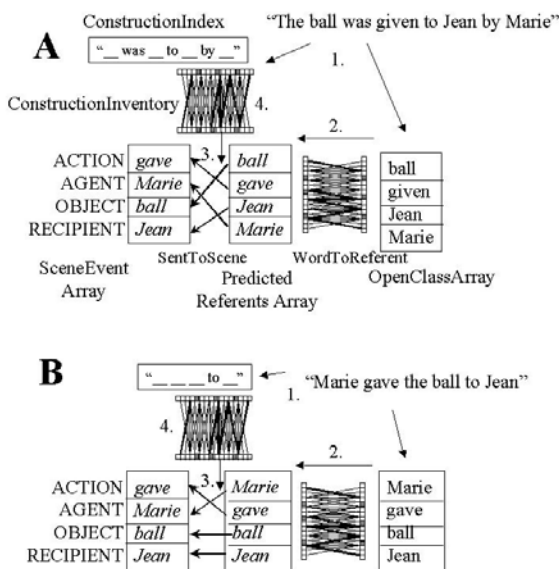


Figure 4. Architecture des constructions grammaticales. Traitement des phrases actives et passives, respectivement en A et en B. En entrée, les mots de classe ouverte sont représentés dans le tableau « Open Class Array » (OCA) et les mots de classe fermée sont représentés dans le « Construction index ». L'analyse de scène visuelle fournit les informations au tableau « Scene Event Array » (SEA) avec le sens extrait en tant qu'éléments de scène. Les mots dans « OCA » sont traduits dans « Predicted Referents » via le mapping « WordToReferent » afin d'être représentés dans le tableau « Predicted Referents Array » (PRA). Les éléments du PRA sont mis en correspondance avec leurs rôles dans le tableau « Scene Event Array » (SEA) par le mapping « SentenceToScene » spécifique à chaque type de phrases. Ce mapping est récupéré dans « Construction Inventory », via « ConstructionIndex » qui encode les mots de classe fermée qui caractérisent chaque type de phrase.

La figure 3 illustre le montage de notre plate-forme d'interaction homme-robot. Cette plate-forme a été initialement utilisée pour produire des paires <Phrase, Sens> pour entraîner le modèle de construction grammatical illustré dans la Figure 4. Les phrases sont extraites du discours parlé en utilisant des logiciels commerciales, et le sens est extrait d'image vidéo basée sur la séquence temporelle des contacts, décrits dans la Figure 3C. Une fois que le modèle a été entraîné, il peut utiliser les constructions grammaticales apprises pour produire la phrase qui correspond à un sens donné. Ceci a été utilisé pour la description de scène automatique, et un système de question-réponse (Dominey et al. 2004) [19].

## 5.2 Objectifs spécifiques :

Dans ce contexte, l'objectif principal du projet est de développer une interface homme-robot basé sur le langage naturel robuste et fiable. L'interface sera située entre l'AIBO (Robot) dans le contexte de RoboCup, et l'entraîneur/coach (Homme) qui dirige l'activité du robot. L'environnement de RoboCup a été choisi parce que c'est un environnement robotique standardisé et bien documenté dont l'équipe de ITAM a une grande expérience vaste avec, et qui permet la quantification et l'évaluation des performances de notre système. Cet objectif est conforme aux axes principaux de l'appelle d'offre LAFMI concernant les systèmes des agentes distribués, le traitement d'images, la robotique et la communication.

Bien que la communication homme-robot doive être relativement libre, elle devra toute de même être organisé autour de la structure de la tâche, c.-à-d. le domaine de RoboCup. En effet, un sous-objectif sera de démontrer que la représentation en *prédicat-argument* des sens de déclarations de langue naturelle peut se généraliser aux domaines différents, y compris les commandes dans la tâche de RoboCup. Ainsi, dans cette tâche, les interactions homme-robot peuvent se classer dans quatre catégories :

1. Commandes de jeu: les instructions spécifiques quoi faire - comprenant : tirer, passer le pas la balle à X, défendre le but, etc.

2. Interrogations sur l'état : les questions incluent - Que faites-vous ?, Où êtes-vous ?, A quelle distance de la balle êtes vous?, etc.
3. Interrogations de justification : Concerne principalement le fait de savoir pourquoi le robot a exécuté une action donnée.
4. Transfert de connaissance stratégique: Transmission de la connaissance de stratégie, par ex. "Si vous êtes bloqué devant, passez la balle à un de nos joueurs derrière vous."

Pour chacune de ces interactions nous établirons la correspondance entre le langage et le sens (dans le contexte du contrôle du robot), avec ces correspondances représentées comme des constructions communicantes qui peuvent être utilisées de manière généralisée. Ainsi, les ordres/commandes de jeu impliqueront des constructions communicantes qui transforment des ordres tels que "Passe la balle à John" dans la syntaxe d'ordre de robot de la forme " passer (la balle, John) ". Les interrogations sur l'état et interrogations de justification seront transformées en questions qui interrogent les structures de contrôle de système. La réponse à ces questions sera alors transformée en une phrase compréhensible par l'homme qui sera prononcé par le synthétiseur de parole. Les interactions de transfert de connaissance stratégique modifieront une base de données de règles qui définissent le comportement tel que "si posséder (balle) et le but (bloqué) alors passer(balle)".

### 5.3 Méthodologie:

#### 5.3.1 Définition des Constructions Communicantes pour différents types d'interaction

Ceci impliquera d'établir pour chaque type d'ordre décrit ci-dessus, une spécification de la correspondance entre la structure grammaticale de l'ordre et la structure *prédicat-argument* de commande AIBO.

Exemples des instructions de l'entraîneur aux attaquants :

#### a. A un attaquant :

1. Tirer. Quand un joueur a la balle, l'entraîneur peut ordonner à ce joueur de donner un coup de pied à la balle. Cette action peut être utilisée pour envoyer la balle vers le but d'équipe adverse ou de la dégager de son propre but.

2. Passer la balle. Quand un attaquant différent de celui qui est près de la balle a une meilleure position pour tirer, l'entraîneur peut ordonner à l'attaquant près de la balle de passer la balle à l'autre attaquant.

3. Défendre un coup de pied libre (free kick). Actuellement, le jeu ne s'arrête pas pour un coup de pied libre, cependant cette règle peut changer à l'avenir. Dans ce cas, l'entraîneur peut commander au robot d'aller défendre les buts afin d'éviter un coup direct au but d'un joueur adverse.

#### b. Aux attaquants multiples:

*Attaquants défendent.* Quand un attaquant perd la balle l'équipe peut être plus vulnérable à une contre-attaque de l'équipe adverse. L'entraîneur peut demander aux attaquants de revenir vers le but et le défendre.

Exemples d'instructions de l'entraîneur au gardien de but :

1. Avance de gardien de but. Dans quelques cas le gardien de but ne sortira pas attraper la balle, car celui-ci est hors jeu. Dans quelques cas si le contraire serait désiré, par exemple, pour éviter un coup d'un attaquant adverse, l'entraîneur peut commander au gardien de but de sortir et attraper la balle.

Exemple d'instructions de l'entraîneur au défenseur :

1. Retenir la balle. Il y a des cas où nous voulons qu'un joueur retienne la balle. Cette action peut être utilisée quand les autres joueurs se sont rentrés du terrain. L'entraîneur peut commander à un défenseur de retenir la balle.

2. Passer la balle. Similaire à l'ordre à l'attaquant de passer la balle.

Exemple des instructions d'entraîneur à n'importe quel joueur :

1. Arrêter. Arrêter toutes actions afin d'éviter une pénalité et pour éviter de bloquer une tire de sa propre équipe.
2. Localiser. Quand l'entraîneur voit qu'un joueur est perdu dans le champ, il peut demander au joueur de se localiser dans le champ.

Exemple d'instructions de l'entraîneur à tous les joueurs :

1. Défendre. Défendre avec tous les joueurs. Tout le monde prend une position défensive.
2. Attaquer. Attaquer avec tous joueurs (sauf le gardien de but). Tout le monde prend une position d'attaque.

Exemple de questions d'entraîneur à n'importe quel joueur :

1. Votre action ? Le joueur donne l'action qu'il est en train d'accomplir.
2. Votre localisation ? Le joueur donne sa position dans le champ.
3. Votre distance à la balle ? Le joueur donne la distance à la balle.
4. Les objets que vous pouvez voir ? Le joueur nomme tous les objets qu'il voit (les bornes, les joueurs, le but et la balle).
5. Pourquoi avez-vous fait cette action ? Le joueur donne les raisons de l'action qu'il a réalisée. (Par exemple, le joueur était près de la balle et a vu le but, donc le joueur a tiré vers le but.)
6. Votre comportement actuel ? Le joueur donne son comportement actuel (attaquant, défenseur, etc.)

Pour chaque type d'interaction décrit ci-dessus a traité ci-dessus, nous définirons la construction communicante qui identifie la correspondance structurale entre les phrases et grammaticales les ordres dans le protocole contrôle de robot.

### 5.3.2 Définition le modèle de dialogue et le protocole d'interaction

Une fois l'interface du niveau d'ordre défini, elle doit être insérée dans la structure d'échange "aller-retour" du dialogue Homme-Robot. Concrètement, ci-dessous sont présentés deux exemples de dialogues entre l'entraîneur et l'attaquant, et l'entraîneur et le gardien de but.

Exemple 1. Entraîneur instruisant un attaquant.	Exemple 2. Entraîneur instruisant le gardien de but.
L'entraîneur : voyez-vous la balle ? AIBO : Non. L'entraîneur : La balle est derrière vous. Tournez 180 degrés. AIBO : bien. L'entraîneur : Quels objets voyez-vous ? AIBO : je vois seulement la balle. L'entraîneur : quelle est votre distance à la balle ? AIBO : 30 centimètres. L'entraîneur : Allez vers la balle AIBO: bien. L'entraîneur : Maintenant passez la balle à AIBO 2. AIBO : Quelle est la position de AIBO 2 ? L'entraîneur : La position du AIBO 2 est x, y. AIBO : bien. L'entraîneur : quelle est votre action actuelle ? AIBO : je tourne à droite 40 degrés. AIBO : Maintenant je passe la balle à AIBO 2. L'entraîneur : Bien, maintenant rentre à votre but. AIBO : bien.	L'entraîneur : voyez-vous la balle ? AIBO : Oui. L'entraîneur : quelle est la distance à la balle ? AIBO : plus de 60 centimètres. L'entraîneur : Faites attention. L'équipe adverse a la balle. AIBO : bien. L'entraîneur : Si vous voyez la balle à une distance inférieure à 40 centimètres, sortez pour attraper la balle. AIBO : bien. L'entraîneur : quelle est votre action actuelle ? AIBO : je sors pour attraper la balle. L'entraîneur : Pourquoi avez-vous fait cette action ? AIBO : j'ai vu la balle à 30 centimètres de ma position, donc je suis votre ordre. L'entraîneur : bien.

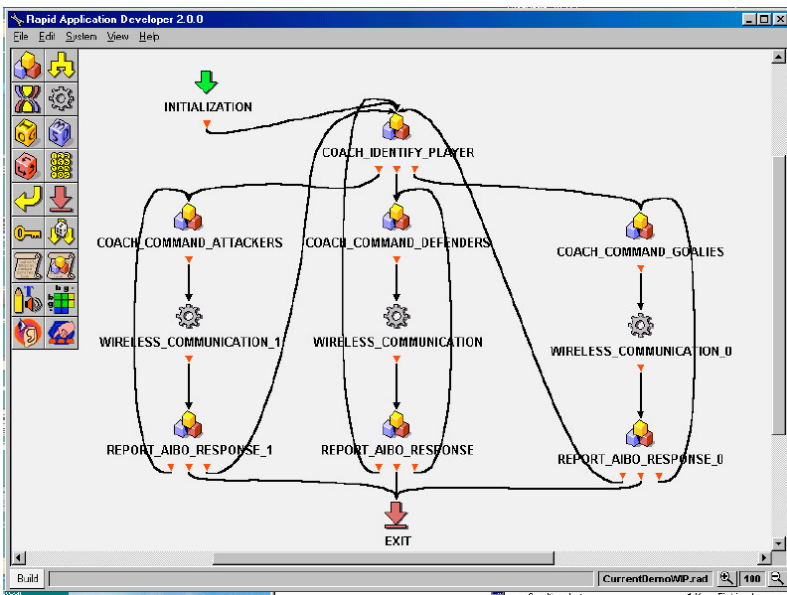


Figure 5. Modèle de dialogue. Afin de contraindre le traitement de discours, l'entraîneur identifiera d'abord le joueur (les joueurs) il s'adresse, et alors entrera dans un domaine une interaction d'ORDRE spécifique. Les ordres seront transmis au Aibo par la communication sans fil, le Aibo agira, répondra et sa réponse sera transmise oralement à l'entraîneur par un synthétiseur. L'entraîneur alors peut choisir de continuer à interagir avec le même joueur (comme dans l'exemple 1), ou changer de joueur (comme dans la transition de l'exemple 1 à l'exemple 2).

### 5.3.3 Attribution de ces fonctions au logiciel et au matériel «hardware»:

Architecture du système. Le système de jeu de football AIBO inclut la perception spécialisée et les algorithmes de contrôle avec lien au système d'exploitation Open-R. Open-R offre une série d'interfaces modulaires permettant accéder aux différents composants de matériel (hardware) dans le AIBO. Les équipes sont responsables de la programmation du niveau d'application, y compris la conception d'une architecture de système contrôlant la perception et le mouvement. Dans la Figure 6 nous montrons une architecture typique de système de AIBO. L'architecture inclut les modules suivants :

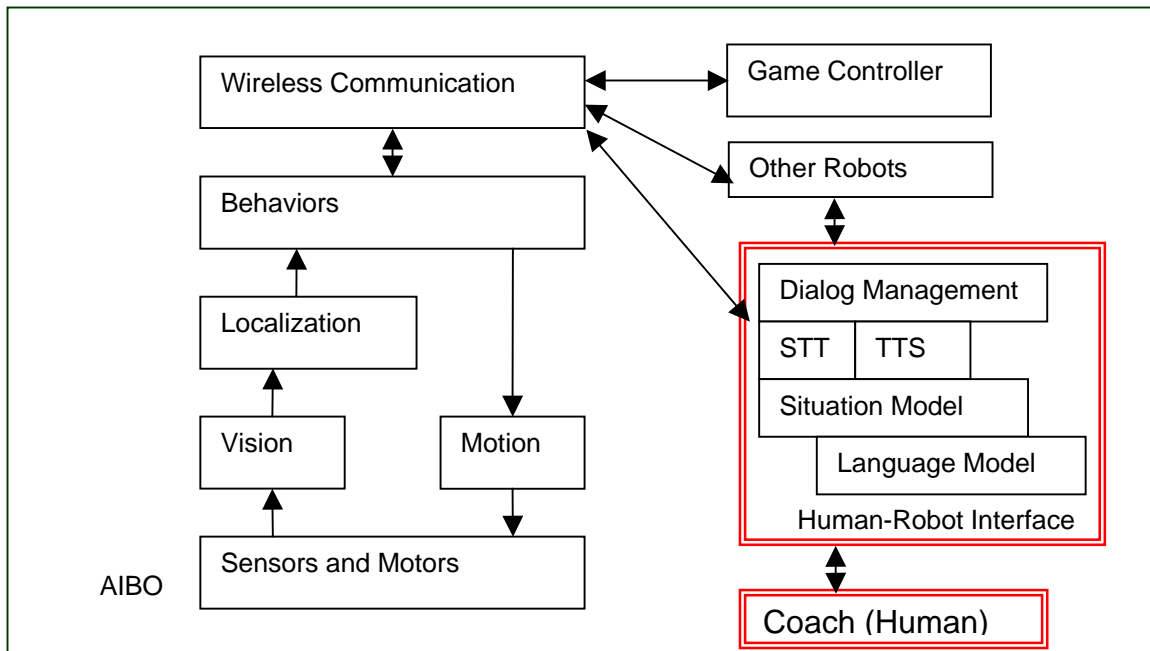


Figure 6. L'architecture de système de robot de AIBO.

Les modules sont développés par chaque équipe avec accès au matériel via les appels de système Open-R (Open-R system calls). Les sous-systèmes "Entraîneur" et Interface de "Homme-Robot" correspondent aux nouveaux composants pour l'interaction homme-robot. Ceci inclut le Directeur de Dialogue (développé en CSLU RAD), le Discours au Texte et le Texte au Discours (RAD), le modèle de situation, et le modèle de langage.

- **Capteurs et Moteurs.** Ce module interagit avec les appels de système Open-R pour contrôler les moteurs physiques et obtenir l'information de la camera.
- **Vision.** Reçoit en temps réel une séquence d'images de la camera de AIBO. Exécute le traitement nécessaire à identifier des objets dans le champ. Le niveau le plus simple de reconnaissance implique l'identification des objets par la couleur avec « addition filtering » pour éviter des inconsistances et dans un but d'optimisation.
- **Localisation.** Déterminer la position du robot en tenant compte des buts, des bords de champ et des bornes aux coins du champ. Différents algorithmes sont utilisés pour augmenter le degré de fiabilité de la position de chaque robot. Les robots partagent cette information pour obtenir un modèle de monde.
- **Communication sans fil.** Transfère les informations entre les robots pour développer un modèle de monde ou une stratégie coordonnée. Reçoit les informations du Contrôleur de Jeu, un ordinateur satellite envoyant des informations sur l'état du jeu (but, pénalité, début et fin de jeu) contrôlé par un arbitre humain. Fournit les bases pour l'Interaction Homme-Robot.
- **Comportements.** Contrôle les mouvements du robot à partir de comportements programmés en réponse à l'information venant des autres modules, comme ceux de la vision, de la localisation et de la communication sans fil. Les comportements sont affectés par la stratégie de jeu, les rôles spécifiques qui prennent joueurs, tel que l'attaquant ou le gardien de but, et par l'interaction humaine.
- **Motion.** Reçoit les ordres du module de comportements correspondant aux actions du robot telles que marcher, courir, tirer, tourner à droite, etc. Ces actions contrôlent des moteurs dans les jambes et la tête.

#### 5.3.4 Définition des accords d'interface

A partir de la configuration d'architecture définie dessus, et le modèle de dialogue explicité dans la Figure 5, nous définirons les nouvelles interfaces entre l'entraîneur et le HRI, et le plus important, les interfaces techniques entre le HRI et le AIBO. Le composant WIRELESS\_COMMUNICATION du modèle d'interface de RAD exécutera deux fonctions d'interface : Il convertira les ordres de discours (défini dans 5.3.1) aux ordres de Aibo en Open-R. Il convertira aussi les réponses de Aibo en Open aux phrases qui seront présentées à l'entraîneur par le synthétiseur.

#### 5.3.5 Tester de modules de logiciel avec les interfaces simulées.

Valider (a) la capacité de l'AIBO à recevoir des ordres simulés et à répondre de manière appropriée, et (b) la capacité d'interface homme-robot à envoyer et recevoir respectivement des ordres et des réponses simulées.

#### 5.3.6 Dupliquer le plateforme AIBO à Lyon pour des testes initiales.

Le groupe de Lyon acquerra le robot AIBO avec les fonds de la ACI Neurosciences Intégrative et Computationnelle.

#### 5.3.7 Intégration et test à ITAM.

Au Laboratoire de Robotique de l'ITAM nous avons cinq robots AIBO et sommes actuellement en train d'acquérir cinq nouveaux.

#### 5.3.8 Evaluation de RoboCup.

Nous essaierons le système obtenu en jouant des matches amicaux contre les autres équipes de RoboCup. Au Mexique il y a plusieurs équipes participant à la ligue de AIBO, aussi qu'en Europe et aux Etats-Unis.

### **5.4 Résultats escomptés au terme de l'action :**

Tandis que ce travail sera fait dans l'interaction proche entre les deux groupes de recherche, Dominey sera responsable des aspects de langage naturels, et que Weitzenfeld sera responsable du contrôle du robot.

#### A. La Langue naturelle (Peter Dominey)

1. Spécification et le prototype de constructions communicantes.
2. Spécification et prototype de modèle de dialogue et de protocole d'interaction.
3. Spécification et prototype des accords d'interface.

#### B. Le Contrôle de robot (Alfredo Weitzenfeld)

1. Allocation de constructions communicantes, le modèle de dialogue, et le protocole d'interaction au logiciel et matériel « hardware ».
2. Essais de logiciels avec les interfaces simulées et AIBOs en temps réel.
3. RoboCup évaluation.

#### C. Présentations et Publications.

- 1 Présentations : Présentation des résultats aux réunions y compris :

International Conference on Intelligent User Interfaces 2006

IEEE-RAS/RSJ International Conference on Humanoid Robots

Competition at International RoboCup

- 2 Publications : dans les journaux appropriés y compris Cognitive Systems Research, IEEE Transactions on Systems Man and Cybernetics, Robotics and Automation.

### 5.5 Calendrier sur la durée du projet :

1. Rencontrant « kick-off » au Mexique –Description des Scénarios : Mois 1, Durée 1 Semaine, Objectif: Travailler sur les exemples de simulation des quatre catégories d'interactions. Identifier les interfaces majeures, et valider l'architecture de niveau 1. Le *livrable* sera le Document de Description de Scénarios qui inclura les descriptions détaillées en UML (Universal Modeling Language) d'interactions d'homme-robot « de bout en bout », indiquant toutes les interfaces, tous les appels de niveau de système, etc.

2. Préparation du Document de Description Fonctionnelle Version1 : Mois 2-6 ; Durée 4 Mois ; l'Objectif: Produire une description fonctionnelle détaillée de l'interface homme-machine, avec une description détaillée de la structure des ordres verbaux et comportement correspondant de robot. (5,4 A. 1, A. 2)

3. Développement de Système Version 1: Mois 3-6, Durée 3 Mois : l'Objectif –Développement d'une version préliminaire du système démontrant une communication fiable de bout en bout (de l'ordre d'utilisateur humain – à la réponse de système) pour un sous-ensemble limité du répertoire d'interaction. (5,4 B. 1)

4. Version 1 Intégration & Teste : Mois 7-8, Durée 2 mois ; Objectif: Intégration et test de version 1. (5,4 B. 2)

5. Version 2 Document de Description de Système et des Interfaces : Mois 7-10, Durée 4 Mois ; l'Objectif: Finaliser l'architecture et les interfaces y compris la correspondance des différents types d'interaction et l'ordre de robot de *prédicat-argument* correspondant. (5,4 A. 3)

6. Version 2 Développement de Système : Mois 9 – 12, Durée 3 Mois : l'Objectif Développement d'une version préliminaire du système démontrant une communication fiable de bout en bout (de l'ordre d'utilisateur humain – à la réponse de système) pour le répertoire complète d'interaction. (5,4 B. 1)

7. Version 2 Evaluation de Langage : Mois 11-12, Durée 2 Mois ; Objectif: l'Intégration et le test de composant langagier de la version finale (5,4 B. 2)

8. Version 2 Intégration & le Test : Mois 13-14, la Durée 2 Objectif de Mois : l'Intégration et le test de version finale (5,4 B. 2)

9. Essai de RoboCup : les Mois 15 – 18, la Durée 4 mois : l'Objectif - l'usage Intense du système. (5,4 B. 3)

10. Mise en valeur et Publication : Mois 19-24 : Ceci inclura la production d'un film documentaire scientifique court pouvant être utilisé par les agences françaises et mexicaines subventionnant comme démonstration d'un projet réalisé en collaboration franco-mexicane. (5,4 C)

## 5.6 Antécédents de coopération :

Alfredo Weitzenfeld est le concepteur principal de la « Neural Simulation Language » (NSL), un système pour de modélisation neuronale [11]. Le travail sur NSL a été commencé dans les 1990 à l'University of Southern California (USC), où un grand nombre de modèles neuronaux biologiques a été développé. Peter Dominey, au cours de son doctorat à USC a développé un modèle de système oculomoteur chez le singe en utilisant le système NSL. Son travail est inclus dans un chapitre de livre de NSL [12]. Un projet récemment subventionné par CONACYT au Mexique (C03-42440) dénommé "Visually-guided NeuroEthological Autonomous Robots: An Adaptive Middleware Approach to Distributed Embedded Mobile Systems" inclut une extension de le travail original de Peter Dominey dans le systèmes oculomoteur [13]. Dominey continue à utiliser NSL dans ses études de simulation.

## 5.7 Publications :

- Dominey PF, Ramus F (2000) Neural network processing of natural language: I. Sensitivity to serial, temporal and abstract structure of language in the infant. *Lang. and Cognitive Processes*, 15(1) 87-127
- Dominey, P.F. (2003a) Learning Grammatical Constructions in a Miniature Language from Narrated Video Events, Proceedings of the 25<sup>th</sup> Annual Meeting of the Cognitive Science Society, Boston
- Dominey, P.F. (2003b) Learning Grammatical Constructions from Narrated Video Events for Human-Robot Interaction, Proc. IEEE Conf. On Humanoid Robotics, Karlsruhe.
- Dominey PF (2000) Conceptual Grounding in Simulation Studies of Language Acquisition, *Evolution of Communication*, 4(1), 57-85.
- Dominey PF, Hoen M, Lelekov T, Blanc JM (2003) Neurological basis of language in sequential cognition: Evidence from simulation, aphasia and ERP studies, (in press) *Brain and Language*
- Weitzenfeld, A., Arbib, M.A., Alexander, A., 2002, *The Neural Simulation Language: A System for Brain Modeling*, MIT Press.
- Weitzenfeld, A., Gutierrez-Nolasco, S., Venkatasubramanian, N., 2003, Controlling Mobile Robots with Distributed Neuro-Biological Systems, Proc. AINS 2003, 2<sup>nd</sup> Annual Symposium on Autonomous Intelligent Networks and Systems, June 30 - July 3, Menlo Park, California.
- Weitzenfeld A., 2003, Embedded Mobile Systems: From Brain Theory To Neural-based Robots, Proc MED '03, 11<sup>th</sup> Mediterranean Conference on Control and Automation, June 17-20, Rhodes, Greece.
- Weitzenfeld A., Gutierrez-Nolasco S., and Venkatasubramanian N., 2003, MIRO: An Embedded Distributed Architecture for Biologically inspired Mobile Robots, Proc ICAR-03, 11<sup>th</sup> International Conference on Advanced Robotics, June 30 - July 3, Coimbra, Portugal.
- Arkin, R.C., Ali, K., Weitzenfeld, A., and Cervantes-Perez, F., 2000, Behavioral Models of the Praying Mantis as a Basis for Robotic Behavior, *Journal of Robotics and Autonomous Systems*, 32 (1) pp. 39-60, Elsevier.

## Bibliography

- [1] AIBO official web site. Url: [www.AIBO.com](http://www.AIBO.com)
- [2] Open R SDK official web site: <http://openr.AIBO.com/openr/eng/index.php4>
- [3] RoboCup official web site: <http://www.robocup.org>
- [4] Chen, et al. "A Description of the rUNSWift 2003 Legged Robot Soccer Team". Australia, 2003. Url: <http://www.cse.unsw.edu.au/~robocup/teamDesc2003.pdf>
- [5] Chen, et al. "Rise of the AIBOs III -AIBO Revolutions". Australia, 2003. Url: <http://www.cse.unsw.edu.au/~robocup/report2003.pdf>
- [6] Mitsunaga, Toichi, Izumi and Asada. "Baby Tigers 2003: Osaka Legged Robot Team" Japan 2003. Url: <http://www.er.ams.eng.osaka-u.ac.jp/robocup/BabyTigers/BabyTigers-TechReport-2003.pdf>
- [7] Veloso, Winner, Lenser, Bruce, and Balch, "Vision-servoed localization and behavior-based planning for an autonomous quadruped legged robot", In Proceedings of AIPS-2000, Breckenridge, April 2000.
- [8] Rescue Robotics, IEEE Robotics & Automation Magazine, 9 (3), September 2002.
- [9] RoboCupRescue, Urban Search and Rescue Robot Competitions, 2004 (<http://www.isd.mel.nist.gov/projects/USAR/competitions.htm>).
- [10] Murphy, R., Rescue Robotics for Homeland Security, Special Issue on Emerging Technologies for Homeland Security, Communications of the ACM, 47 (3) 66 - 68, March 2004.
- [11] Weitzenfeld, A., Arbib, M.A., Alexander, A., 2002, The Neural Simulation Language: A System for Brain Modeling, MIT Press.
- [12] Dominey, P, Arbib, M.A., and Alexander, A., 2002, The Modular Design of the Oculomotor System in Monkeys, in The Neural Simulation Language: A System for Brain Modeling, Eds. Weitzenfeld, A., Arbib, M.A., Alexander, A., MIT Press.
- [13] Dominey P.F., and Arbib M.A., A Cortico-Subcortical Model for Generation of Spatially Accurate Sequential Saccades, Cerebral Cortex, 2:153-175, 1992.
- [14] Orfinger, B., Robot Responders at WTC Site Fit Into Tight Spaces, Disaster Relief, Oct 2001, (<http://www.disasterrelief.org/Disasters/011015robots/>).
- [15] Siskind JM (2001) Grounding the lexical semantics of verbs in visual perception using force dynamics and event logic. *Journal of AI Research* (15) 31-90
- [16] Steels, L. and Baillie, JC. (2002). Shared Grounding of Event Descriptions by Autonomous Robots. *Robotics and Autonomous Systems*, 43(2-3):163--173.
- [17] CMU Communicator: <http://www.speech.cs.cmu.edu/Communicator/>
- [18] CSLU Toolkit <http://www.cslu.ogi.edu/toolkit/index.html>
- [19] Dominey PF, Boucher J-D, Inui T (2004) Building an Adaptive Spoken Language Interface for Perceptually Grounded Human-Robot Interaction, submitted IEEE Conf. On Humanoid Robotics 2004, Los Angeles.