

# Amorçage d'une sémantique lexicale dans une population d'agents autonomes, ancrés et situés

---

## Résumé

Cet article décrit l'amorçage d'une ontologie et d'un lexique partagé dans une population d'agents robotiques dotés de capacités visuelles. Cette évolution a lieu alors que les agents jouent un jeu de langage, appelé "guessing game". Nous montrons comment la synonymie et l'ambiguïté du système sémantique, qui émergent dans un premier temps, sont progressivement réduites au fur et à mesure que l'environnement physique se complexifie. Nous discutons également, comment des interactions mixtes entre agents humains et agents artificiels sont possibles dans ce cadre expérimental ainsi que les perspectives applicatives de tels systèmes.

---

Mot-clés: sémantique lexicale, dialogue homme-machine, évolution du langage.

## 1. Introduction

Durant les dernières décennies, la linguistique computationnelle a atteint un haut degré de sophistication dans les méthodes qu'elle met en oeuvre, migrant peu à peu vers des applications industrielles. Néanmoins certains enjeux de base dans le traitement du langage naturel restent à ce jour des problèmes ouverts, dans la mesure où très peu d'outils théoriques ou pratiques existent pour les traiter. Deux de ces problèmes sont étudiés dans cet article:

- La plupart des systèmes du traitement automatique du langage naturel sont des systèmes formels symboliques non ancrés dans la réalité par des senseurs et des actuateurs. Ceci limite nécessairement leur degré de compréhension du langage naturel et donc leurs possibles applications. Par exemple, beaucoup de concepts ancrés perceptuellement sont différents d'une langue à l'autre. Ceci rend la traduction très difficile quand le sens ne peut être relié à une expérience sensorielle. De la même manière, beaucoup de phrases produites dans un certain contexte ne peuvent être vraiment désambiguïsées que si le contexte peut être perçu, interprété et intégré dans le processus de compréhension du langage.
- La plupart des systèmes de traitement automatique du langage naturel suppose que le langage est fixe. Le lexique et la grammaire sont précodés (souvent à la main) et ne sont pas supposés changer au cours des conversations. Pourtant, les conventions linguistiques sont en flux constant. Locuteur et interlocuteur transforment leur langage en fonction du succès qu'ils obtiennent lorsqu'ils communiquent ou des informations extra-linguistiques complémentaires quand leurs systèmes linguistiques sont incompatibles ou incomplets.

Ceci suggère que les locuteurs humains ont des stratégies de réparation de leur connaissance linguistique en cas d'échec en communication et des stratégies de construction pour entendre l'étendue langage quand cela est nécessaire.

Nous avons examiné ces deux problématiques en utilisant de la modélisation formelle, des simulations informatiques et des expériences avec des agents robotiques. Nous étudions les dialogues en langage naturel sous la forme de *jeux de langage adaptatifs* au cours desquels locuteur et interlocuteur conversent au sujet de scènes du monde réel perçues au travers de leur appareil sensoriel. Contrairement à un apprentissage dirigé par un intervenant humain, c'est seulement à l'occasion d'un échec en communication que les agents réparent leur système linguistique. Dans nos expériences, des groupes d'agents autonomes commencent sans connaissance linguistique et doivent amorcer un langage et un système de concepts pour décrire le monde qu'ils perçoivent, sans intervention humaine. Jusqu'à présent, nous avons étudié principalement la sémantique lexicale, mais de premiers résultats dans le domaine de la syntaxe (Steels, 1998) ont également été obtenus.

Cet article décrit nos expériences sur la formation d'un lexique. Il présente succinctement les "Talking Heads", la plateforme expérimentale avec laquelle nous travaillons, et les stratégies utilisées par les agents pour construire leur langage. Nous montrons que des dynamiques sémiotiques très riches apparaissent avec l'élimination de la synonymie pure et des phénomènes de désambiguïsation. Nous étudions également les interactions possibles entre des agents humains et des agents artificiels. Nos expériences sont construites pour que des humains puissent jouer le rôle d'agents artificiels. Nous montrons que des humains peuvent ainsi influencer le langage des agents de manière à le rendre plus proche du nôtre et donc que le cadre des jeux de langage adaptatifs est une approche possible pour construire des systèmes automatiques du traitement du langage naturel capables d'apprendre de façon continue.

## 2. L'expérience des "Talking Heads"

La plateforme robotique utilisée dans les expériences décrites dans cet article consiste en un ensemble de "Talking Heads" connectées par l'Internet. Chaque Talking Head est constituée d'une caméra SONY EVI-D31 pouvant bouger verticalement et horizontalement (figure 1), d'un ordinateur qui implémente les fonctions cognitives (perception, catégorisation, recherche dans le lexique, etc.), d'un écran présentant les états internes de l'agent chargé dans le corps de la "Talking Head", d'un moniteur montrant la scène telle que la voit la caméra et des entrées et sorties audio. Un agent, chargé dans le corps "physique" d'une Talking Head, peut se téléporter dans une autre Talking Head connectée par l'Internet. Deux agents ne peuvent interagir ensemble que dans le cas où ils sont instanciés physiquement dans deux corps partageant un même environnement. La plateforme contient également un commentateur artificiel qui décrit et commente l'évolution des dialogues ainsi qu'un ensemble de moniteurs permettant des statistiques sur les jeux et des mesures sur les ontologies et les langages des agents (succès moyen en communication, cohérence lexicale, taille moyenne des lexiques et des ontologies, etc.)

Dans les expériences présentées dans cet article, l'environnement partagé consiste en un tableau magnétique sur lequel différentes formes géométriques sont placées: des triangles, des cercles et des rectangles de différentes couleurs. La simplicité de cet environnement permet une étude plus facile des dynamiques complexes dans la population d'agents.



FIG. 1 – Deux talking heads et leur moniteur associé montrant ce que chaque caméra perçoit.

### Le "guessing game"

Les agents interagissent dans le cadre d'un jeu de langage nommé "guessing game". Un "guessing game" peut être joué par deux agents dotés de capacités visuelles. Un agent joue le rôle de *locuteur* et l'autre, celui d'*interlocuteur*. Les agents alternent dans ces deux rôles et tous développent la capacité d'être locuteur ou interlocuteur. Les agents sont capables de ségmenter en "objets" l'image perçue par la caméra et de collecter un ensemble d'informations sensorielles concernant chacun de ces objets: couleurs (canaux R G B), niveau moyen de gris (canal GRAY), position (canaux HPOS et VPOS). L'ensemble des objets et les données les concernant constitue le *contexte*. Le locuteur choisit un objet dans ce contexte : c'est le *sujet* de l'interaction.

Le locuteur donne un indice linguistique à l'interlocuteur afin qu'il identifie le sujet par rapport aux autres objets du contexte. 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].

A partir de cet indice linguistique, l'interlocuteur essaie de deviner quel est le sujet choisi par le locuteur. Il lui indique son choix en pointant vers un des objets. Cette étape est réalisée par la transmission de la direction dans laquelle la Talking Head regarde. Le jeu est un succès quand l'interlocuteur a deviné juste. C'est un échec si l'interlocuteur pointe vers un autre objet du contexte, ou si un des deux agents n'a pas été capable de réaliser une des étapes précédentes du jeu. Dans le cas d'un échec, le locuteur indique à l'interlocuteur, de façon extra-linguistique, le sujet qu'il voulait désigner et les deux agents adaptent leur structures internes afin d'être efficaces dans les jeux futurs.

L'architecture des agents est constituée par deux composants: [1] un module de conceptualisation qui permet la catégorisation du monde perçu et la recherche d'un référent dans l'image et [2] un module de verbalisation permettant d'associer une forme verbale à un concept et

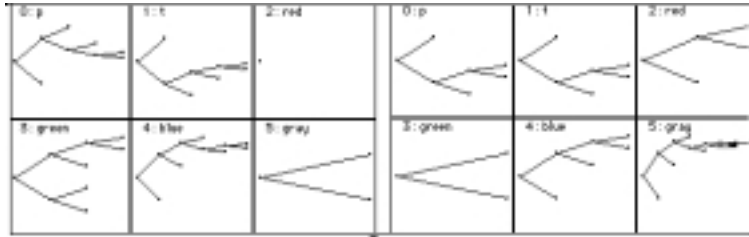


FIG. 2 – *Les arbres de discrimination de deux agents.*

d'interpréter une forme pour reconstruire son sens. Les agents commencent sans lexique ni ontologie préconstruite. Une ontologie partagée et un lexique doivent donc émerger au cours d'un processus d'auto-organisation. Les agents construisent, étendent et adaptent leur ontologie et leur lexique, au fur et à mesure qu'ils participent à des jeux de langage.

### **Le module de conceptualisation**

Les "sens" sont des catégories qui permettent de distinguer le sujet des autres objets du contexte. Les catégories sont organisées en arbres de discrimination (figure 2) dans lesquels chaque noeud contient un discriminateur capable de filtrer l'ensemble des objets en un sous-ensemble qui satisfait une certaine contrainte. Par exemple, il peut exister un discriminateur basé sur la position horizontale (HPOS) du centre d'un objet (normalisée entre 0.0 et 1.0) qui classe les objets du contexte en deux sous-ensembles: l'ensemble "gauche" quand  $HPOS < 0.5$  (que nous désignerons par [HPOS-0.0,0.5]) et l'ensemble droit quand  $HPOS > 0.5$  ([HPOS-0.5,1.0]). Des sous-catégories supplémentaires sont créées en sub-divisant la région de chaque catégorie. Par exemple, la catégorie "très à gauche" (ou [HPOS-0.0,0.25]) s'applique aux objets dont la valeur HPOS est dans la région [0.0, 0.25].

Un ensemble de catégories distinctives est trouvé en filtrant les objets du contexte en partant du sommet de chaque arbre de discrimination de façon à isoler uniquement le sujet. Il est, par exemple, possible que l'ensemble ([HPOS-0.5,1.0] et [VPOS-0.0, 0.25]) identifie le sujet sans ambiguïté car aucun autre objet du contexte n'a de valeur dans le domaine ainsi défini. Souvent, il y a plus d'une solution possible mais toutes les solutions sont transmises au module de verbalisation.

Les arbres de discrimination de chaque agent sont formés en utilisant des dynamiques de croissance et d'élagage couplées avec l'environnement. Les arbres de discrimination croissent de façon aléatoire par l'addition de nouveaux noeuds subdivisant les régions de catégories existantes. Les différents noeuds des différents arbres sont en compétition les uns avec les autres. L'utilisation et le succès de chaque noeud sont mesurés et les noeuds non pertinents pour les environnements rencontrés par les agents sont élagués. De plus amples détails au sujet des jeux de discrimination peuvent être trouvés dans (Steels, 1997).

### **Le module de verbalisation**

Le lexique de chaque agent consiste en une mémoire associative où sont rangées des associations entre des formes (ici des mots) et des sens (ici des catégories simples). Chaque association a un score. Les mots sont des combinaisons aléatoires de syllabes. Quand un locuteur doit exprimer une catégorie, il sélectionne tous les mots associés à cette catégorie, les ordonne et

choisit celui avec le score le plus élevé pour le transmettre à l'interlocuteur. Quand l'interlocuteur doit interpréter un mot, il sélectionne tous les sens possibles de ce mot et teste ceux qui sont utilisables dans le présent contexte, i.e. ceux qui identifient un référent unique. Parmi ceux-ci, l'interlocuteur choisit celui avec le score le plus élevé.

En fonction du résultat du guessing game, le locuteur et l'interlocuteur modifient le score de certaines associations. Quand le jeu est un succès, ils augmentent chacun le score de leur association gagnante et diminuent celui des associations en compétition, implémentant ainsi un processus d'inhibition latérale. Quand le jeu est un échec, chacun diminue le score de l'association qu'il a utilisée. De nouvelles associations peuvent être mémorisées. Un locuteur crée un nouveau mot quand il n'a pas encore de mot pour le sens qu'il veut exprimer. Un interlocuteur peut rencontrer un nouveau mot qu'il n'a jamais entendu auparavant et dans ce cas, créer une nouvelle association entre ce mot et son sens probable. Pour déterminer ce sens, l'interlocuteur utilise d'abord l'information extra-linguistique donnée par le locuteur pour identifier le référent et catégorise ce dernier en utilisant ces propres arbres de discrimination. Ces mécanismes d'amorçage du lexique ont été décrits et validés dans des papiers plus anciens (Steels & Kaplan, 1998) et sont similaires à ceux étudiés par Oliphant (Oliphant, 1996).

Le module de conceptualisation propose plusieurs solutions au module de verbalisation qui choisit de préférence celles qui ont déjà été lexicalisées. Les agents mesurent le succès de chaque catégorie dans le jeu et l'utilisent ce score pour diriger les dynamiques de croissance et d'élagage. Le langage, dans ces conditions, influence de façon forte la forme des ontologies créées par les agents. Les deux modules sont couplés structurellement et se coordonnent en l'absence d'un coordinateur central.

## Exemples

Nous commencerons par un jeu de langage tout simple. Le locuteur, **a1**, a choisi comme sujet un triangle placé en bas de la scène. Il n'y a qu'un autre objet dans la scène, un rectangle, en haut. Dans ces conditions, la catégorie  $[VPOS-0.0,0.5]_{a1}$ , qui est valide quand la position verticale  $VPOS < 0.5$ , est applicable pour le triangle mais pas pour le rectangle. Si l'on suppose que **a1** a une association dans son lexique entre  $[VPOS-0.0,0.5]_{a1}$  et le mot "lu", alors **a1** sélectionne cette association et transmet le mot "lu" à l'interlocuteur qui est l'agent **a2**.

Si l'on suppose maintenant, que **a2** a rangé dans son lexique une association entre "lu" et  $[R-0.0,0.5]_{a2}$ , il fait l'hypothèse que le sens de "lu" est  $[R-0.0,0.5]_{a2}$ . Quand il applique cette catégorie à la scène, autrement dit quand il filtre les objets de la scène dont le niveau dans le canal rouge (R) n'est pas dans la région  $[0.0, 0.5]$ , il sélectionne un objet unique, le triangle. Ainsi **a2** conclut qu'il s'agit du sujet de l'interaction et pointe vers lui. Le locuteur reconnaît que l'interlocuteur a désigné le bon objet et le jeu est un succès.

Ce jeu illustre une situation dans laquelle le locuteur et l'interlocuteur identifient le même référent avec deux sens différents. Le locuteur utilise la position verticale et l'interlocuteur utilise le niveau du canal rouge dans l'espace RGB.

Dans ce deuxième exemple, le locuteur est à nouveau **a1** et utilise la même catégorie et le même mot pour désigner le triangle. Mais l'interlocuteur, **a3**, interprète "lu" dans des termes de position horizontale  $[HPOS-0.0,0.5]_{a3}$  (partie gauche de la scène). Comme plusieurs objets dans la scène satisfont à cette catégorie, l'interlocuteur est incapable d'identifier un objet unique. Le locuteur pointe alors vers le sujet et l'interlocuteur crée une nouvelle association entre "lu" et

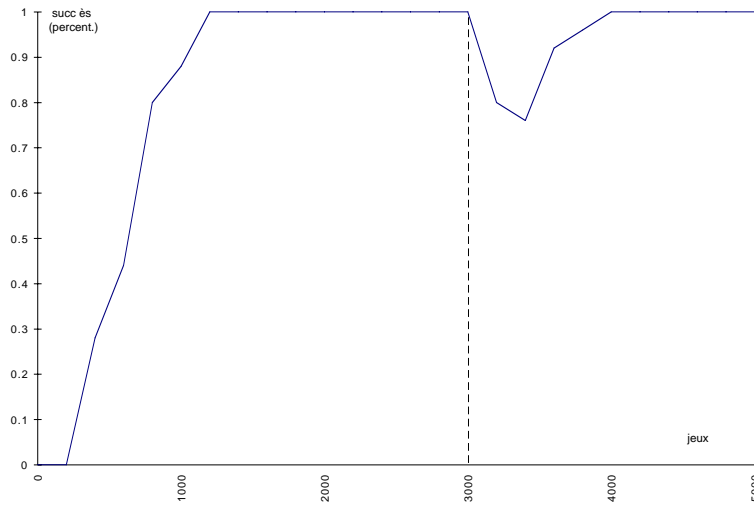


FIG. 3 – Evolution du succès moyen tous les 200 jeux sur une série de 5000 jeux avec 20 agents. Les agents atteignent un haut niveau de succès en environ 1000 jeux. Le changement dans l’environnement introduit au jeu 3000 provoque une chute temporaire du succès moyen.

[VPOS-0.0,0.5]<sub>a3</sub>. Cette association sera des lors en compétition avec celle qu’il a déjà.

Le tableau ci-dessous montre une partie du vocabulaire d’un agent après 3000 jeux de langage. Le score de chaque association est également indiqué.

Forme	Sens	Score	Forme	Sens	Score
wovota	[R-0.0,0.125]	1.0	sogavo	[G-0.5,1.0]	0.0
tu	[GRAY-0.25,0.5]	0.0	naxesi	[G-0.5,1.0]	0.0
gorepe	[VPOS-0.0,0.5]	0.3	ko	[G-0.5,1.0]	0.0
zuga	[VPOS-0.0,0.5]	0.1	ve	[G-0.5,1.0]	0.0
lora	[VPOS-0.25,0.5]	0.1	migine	[G-0.5,1.0]	0.0
wovota	[VPOS-0.25,0.5]	0.2	zota	[G-0.5,1.0]	0.9
di	[VPOS-0.25,0.5]	0.0	zafe	[G-0.5,1.0]	0.1
zafe	[VPOS-0.0,0.25]	0.2	zulebo	[HPOS-0.0,1.0]	0.0
wowore	[VPOS-0.0,0.25]	0.9	xi	[HPOS-0.0,1.0]	0.0
mifo	[HPOS-0.0,1.0]	1.0			

On voit dans cette table que pour certains sens (par exemple R-0.0,0.125]), une forme unique ”wovata” s’est clairement imposée. Pour d’autres sens, comme [GRAY-0.25,0.5], un mot était connu mais n’est plus désormais utilisé. Pour des sens comme [VPOS-0.0,0.5], deux mots sont encore en compétition: ”gorepe” et ”zuga”. De même, certains mots, comme ”zafe” peuvent avoir deux sens différents [VPOS-0.0,0.25] et [G-0.5,1.0].

### 3. Dynamiques de coévolution entre catégorisation et lexicalisation

Nous pouvons maintenant illustrer certaines dynamiques du guessing game, lorsqu’il est joué par des agents ancrés et situés qui interagissent dans un environnement physique partagé. La figure 3 montre une série de 5000 jeux pour un groupe de 20 agents. La première tendance est la suppression de la synonymie, par un processus de rétroaction positive qui décide quelle forme va être associée préférentiellement à un référent particulier. La figure 4 montre clairement

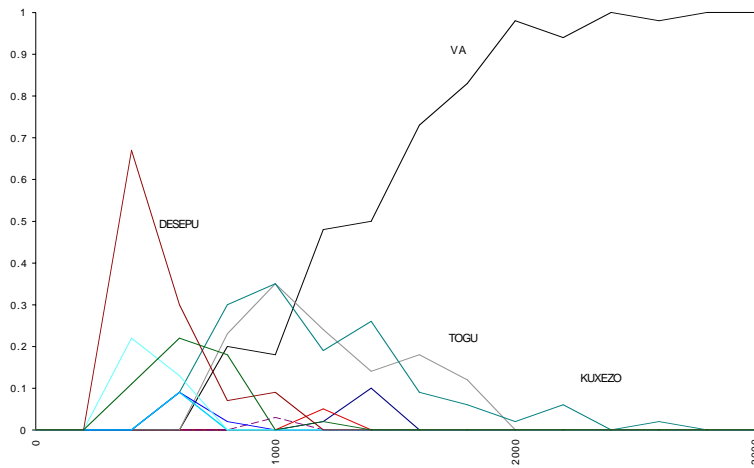


FIG. 4 – *Diagramme RF: évolution de la fréquence de toutes les formes utilisées pour le même référent en 3000 jeux de langage*

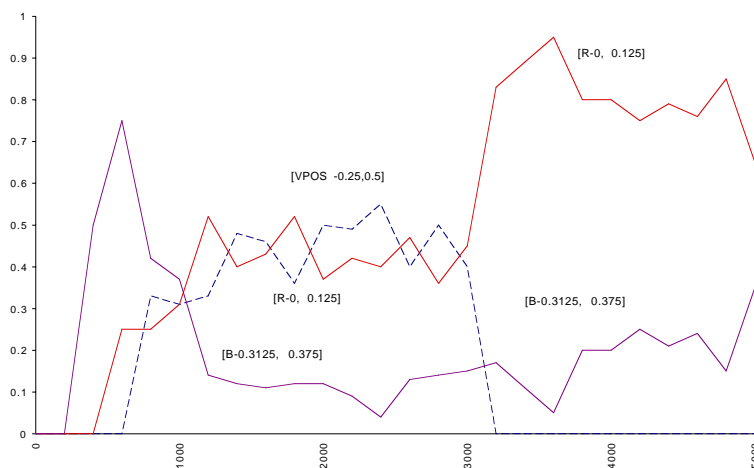


FIG. 5 – *Diagramme FS: Evolution de la fréquence de tous les sens utilisés pour la même forme "va" sur 5000 jeux. Une situation au jeu 3000 provoque la perte d'un des sens.*

ce phénomène: le mot "va" devient dominant pour l'expression de ce référent. Cette suppression de la synonymie est due à l'inhibition latérale et à la boucle de rétroaction entre le succès d'une association et son utilisation dans les jeux futurs.

Quand nous étudions les différents sens du mot "va", grâce au diagramme FS (figure 5), nous voyons de façon claire que même après 3000 jeux l'ambiguïté reste dans le langage. Trois sens stables de "va" ont émergé: [R-0,0.125], [B-0.3125,0.375], et [VPOS-0.25,0.5]. Ils sont tous aussi efficaces pour distinguer le sujet désigné par "va" et aucune situation pouvant lever cette ambiguïté ne s'est encore présentée.

Au jeu 3000, l'environnement produit une scène dans laquelle une des catégories, qui était distinctive pour l'objet désigné par "va" n'est plus acceptable. Concrètement, nous avons, en tant qu'expérimentateurs, déplacé l'objet en question jusqu'à une position très proche d'un autre objet: la catégorie faisant intervenir la position verticale ne permet plus de les distinguer. La figure 3 montre que le succès chute (plusieurs communications ont été des échecs), puis remonte rapidement. Les échecs se produisent car "va" ne permet plus de sélectionner le bon objet pour les agents pour qui "va" est associé à [VPOS-0.25,0.5]: ils doivent adopter un nouveau sens

pour "va" compatible avec la nouvelle situation. Le diagramme FM de la figure 5 montre que le sens [VPOS-0.25,0.5] a disparu. Les autres sens, correspondants à des informations sur les couleurs, sont encore possibles et non pas été affectés par le déplacement de l'objet.

#### 4. Interactions entre agents humains et agents artificiels

La plateforme expérimentale est construite pour que des humains puissent jouer le rôle des agents artificiels, c'est-à-dire participer aux jeux de langage en contrôlant le corps d'une des Talking Heads. Lorsqu'un humain est locuteur, il choisit mentalement un objet du tableau et tape au clavier un mot pour le désigner. Si l'interlocuteur n'a pas deviné juste, il donne une information extra-linguistique en dirigeant la caméra vers l'objet choisi. Lorsqu'il est interlocuteur, il doit, à partir du mot indiqué par le locuteur, deviner le sujet de l'interaction. Il est donc possible de faire interagir de façon transparente : deux agents artificiels, deux agents humains ou un agent artificiel avec un agent humain.

Les premières expériences faisant intervenir des populations mixtes d'agents artificiels et humains montrent qu'effectivement les agents créent un langage plus proche du nôtre. Le tableau ci-dessous montre la propagation d'un mot anglais ("GREEN") dans une population de 4 agents artificiels et d'un agent humain.

<i>Jeu</i>	<i>Locuteur</i>	<i>Interlocuteur</i>	<i>Sujet</i>	<i>Association créée</i>
154	a5 (humain)	a3 (artificiel)	triangle vert	a3 : ("GREEN") + [G-0.5,1.0]
156	a5 (humain)	a1 (artificiel)	triangle vert	a1 : ("GREEN") + [VPOS-0.4375,0.5]
167	a3 (artificiel)	a4 (artificiel)	triangle vert	a4: ("GREEN") + [R-0.0,0.5]
211	a4 (artificiel)	a2 (artificiel)	carré vert	a2: ("GREEN") + [VPOS-0.5,1.0]

Nous voyons que les agents attribuent dans un premier temps des sens très différents au mot ("GREEN"), mais le lexique et les ontologies des agents convergent progressivement vers une situation où chaque agent associe au mot ("GREEN") un sens similaire indépendant de la position. Ceci est dû aux dynamiques de desambiguïsation présentées dans la partie précédente et au fait que, contrairement aux agents artificiels, l'agent humain est peu adaptatif du point de vue lexical (il insiste pour imposer son mot) et du point de vue ontologique (il ne remet pas en cause le sens du mot ("GREEN")).

Nous ne sommes aujourd'hui qu'à un stade préliminaire de l'étude des dynamiques dans des populations d'agents artificiels et humains. Nous travaillons sur une expérience à grande échelle au cours de laquelle des agents artificiels seront en contact continu avec des humains pendant plusieurs mois (Laboratorium, Anvers, été 1999). Cette expérience nous fournira le matériel nécessaire à une étude plus approfondie.

Ce type d'interactions mixtes ouvrent des perspectives intéressantes pour l'interaction homme-machine et le traitement automatique de la langue naturelle:

- Le type d'interaction, mis en jeu lorsqu'un humain interagit avec un agent artificiel par l'intermédiaire des Talking Heads, est un exemple d'interface homme-machine continuellement adaptative. L'humain et la machine développent coopérativement, au cours des interactions, un langage commun qui peut leur permettre de communiquer en vue de réaliser d'une tâche particulière. Si comme c'est le cas avec le développement d'un réseau étendu



de Talking Heads, les agents artificiels ont l'occasion d'interagir avec un nombre important de personnes différentes, un nouvel utilisateur humain peut bénéficier de l'expérience déjà acquise par les agents artificiels pour ce type d'interactions.

- Les Talking Heads peuvent également servir de base pour les premières expériences de *traduction ancrée*. L'installation des Talkings Heads dans différents pays du monde permet aux agents artificiels de développer des lexiques différents, adaptés aux agents humains avec qui ils sont amenés à interagir. Ces différents lexiques reposant tous sur une ontologie construite au cours d'interactions avec un environnement physique, il devient alors possible d'effectuer de la traduction lexicale automatique. Nous pensons que cette voie pourrait constituer une solution intéressante aux problèmes que les systèmes de traduction automatique rencontrent aujourd'hui.

## 5. Conclusions

L'approche que nous utilisons pour la formation du lexique est différente des méthodes plus traditionnelles inspirées de Quine (Quine, 1960). Quine suppose que les agents apprennent le sens des mots par des abstractions inductives successives à partir de situations au cours desquelles ils observent des relations particulières entre des objets et des mots. Les propriétés communes des référents constituent le sens d'un mot. Elles sont induites par l'étude des similarités sur de nombreux exemples. Ce point de vue est aussi sous-jacent dans les approches utilisant des réseaux de neurones pour l'acquisition d'un lexique (voir par exemple (Hutchins & Hazlehurst, 1995) ou (Regier, 1995)). Notre approche est, au contraire, inspirée de Wittgenstein (Wittgenstein, 1953): des agents inventent des mots et des sens au cours de jeux de langage, formulent différentes hypothèses sur le sens des mots utilisés par les autres et testent ces sens quand ils sont locuteurs. L'évolution vers une cohérence lexicale dans la population (où un mot est associé à un sens dominant et un sens est associé à un mot dominant) est un phénomène collectif dirigé par les réponses du système à de nouvelles situations où les différents sens d'un mot ne sont plus compatibles les uns avec les autres.

Une grande partie du travail d'analyse nécessaire à la compréhension des dynamiques complexes de la formation du lexique reste encore à étudier. De nouveaux outils pour analyser la coévolution de la formation des catégories et du lexique doivent encore être construits. Néanmoins, nous pensons que les résultats discutés dans cet article constituent un pas en avant important dans la mesure où nous montrons, pour la première fois, l'évolution d'un ensemble ouvert de sens et de mots dans un groupe d'agents autonomes en interaction dans un environnement physique.

## 6. Remerciements

Cette recherche a été menée au Sony Computer Science Laboratory à Paris. Le logiciel Babel (McIntyre, 1998) qui constitue le noyau de toute l'architecture de nos expériences et les routines visuelles de bas niveau ont été développées par Angus McIntyre.

## Références

HUTCHINS E. & HAZLEHURST B. (1995). How to invent a lexicon: the development of shared symbols in interaction. In N. GILBERT & R. CONTE, Eds., *Artificial Societies: The Computer Simulation of Social Life*. UCL Press.

- MCINTYRE A. (1998). Babel: A testbed for research in origins of language. In *Proceedings of COLING-ACL 98*, Montreal: ACL.
- OLIPHANT M. (1996). The dilemma of saussurean communication. *Biosystems*, **37**(12), 31–38.
- QUINE W. (1960). *Word and Object*. Cambridge Ma: The MIT Press.
- REGIER T. (1995). A model of the human capacity for categorizing spatial relations. *Cognitive Linguistics*, **6**(1), 63–88.
- STEELS L. (1997). Constructing and sharing perceptual distinctions. In M. VAN SOMEREN & G. WIDMER, Eds., *Proceedings of the European Conference on Machine Learning*, Berlin: Springer-Verlag.
- STEELS L. (1998). The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, (103), 1–24.
- STEELS L. & KAPLAN F. (1998). Spontaneous lexicon change. In *Proceedings of COLING-ACL 1998*, p. 1243–1249, Montreal: ACL.
- WITTGENSTEIN L. (1953). *Philosophical Investigations*. New York: Macmillan.