

Experiments in learning by imitation - Grounding and Use of Communication in Robotic Agents

Aude Billard* , Kerstin Dautenhahn†

*Dept. of Artificial Intelligence, U. of Edinburgh, 5 Forrest Hill, Edinburgh EH1 2QL

†Department of Cybernetics at the University of Reading

PO Box 225 Reading, RG6 6AY United Kingdom

audeb@dai.ed.ac.uk, K.Dautenhahn@cyber.reading.ac.uk

Contact:

Tel: +44-131-650-4491

Fax: +44-131-650-6899

Abstract

Social behaviour and in particular social learning are key mechanisms for the cohesion and evolution of primate societies. Similarly, social skills might be desirable for artificial agents who are expected to interact with other natural or artificial agents. We view learning, communication and imitation as important capabilities to possess by social artificial agents and study how these skills can be designed and used by physically embodied autonomous robots. We study grounding and use of communication among heterogeneous agents. In particular, we investigate the role of social interactions for sharing of context and building of joint attention among communicative agents. Grounding and use of communication is investigated through simulations within a group of autonomous agents. Results show that social behaviour benefit the agents in two circumstances: (1) agents capable of following one another, and in this way imitating each other's

movements, develop faster and better a common understanding of the language; (2) furthermore, the agents' capability of communicating with one another via a common vocabulary benefits to the group and to each agent individually as it speeds up the transmission of information. We use a connectionist model, based on Hebbian associative learning, for the learning of the word-signal pairs. This work follows robotic experiments [6, 5, 7] in which a physical autonomous robot was taught a vocabulary to describe its perceptions of objects, movement, inclination and orientation. The robot was taught either by a human instructor or by another robot. The teacher-learner robot experiments were based on an imitative strategy whereby the learner robot followed the teacher robot. The work of this paper demonstrates scaling up of this movement imitative strategy for transmitting a vocabulary across a group of robotic agents, i.e. from a teacher agent to several learner agents. In particular, it shows that imitative behaviour is necessary for the grounding of the agents' proprioceptions and speeds up the grounding of exteroceptions. These studies stress the importance of behavioural social mechanisms in addition to general cognitive abilities of associativity for grounding communication in embodied agents. In particular, it shows that a simple movement imitation strategy is an interesting scenario for the transmission of a language, as it is an easy means of getting the agents to share a common context of perceptions, which is a prerequisite for a common understanding of the language to develop. It is thus suggested that a behaviour-oriented approach might be more appropriate than a pure cognitivist one which is dominating in related studies of the mechanisms involved in grounding communication.

Keywords: Social learning and behaviour; grounding communication; embodied and situated agents.

1 The social aspect of communication

Social learning theory studies the mechanisms by which an agent learns from its interactions with conspecific agents. Examples of social learning can be found in the animal kingdom, and especially

in primate societies, whereby, for instance, one agent learns new skills by the observation and imitation of another agent's behaviour. At present there is no widely agreed definition of imitation existing in the animal and human psychology literature. The variety of definitions ranges from strong definitions, which require an aspect of novelty in the imitated behaviour to very weak definitions which have no clear boundaries to forms of non-imitative social learning like contagion or social facilitation. Often researchers refer to Thorpe's definition of 'true imitation' which is the 'copying of a novel or otherwise improbable act or utterance, or some act for which there is clearly no instinctive tendency' [53]. However, such a definition leads automatically to the question of what can be counted as a 'novel' behaviour. On the other hand of the spectrum we find definitions like one given by Davis who refers to a behavioural skill which leads to 'some sort of similarity in behaviour among two or more individuals' [15]. Matching behaviour which involves unlearned species-typical behaviour (like contagion: e.g. synchronised courtship, flocking and herding behaviour) is then difficult to distinguish from imitation. For following behaviour (e.g. rats learning to follow a trained rat to food in a maze) the term 'matched-dependent behaviour' is often used. Thus, from the literature it is not quite clear when the term imitation should be used properly (see [25] for an extensive discussion). Applying the term imitation to artificial systems like physical or simulated agents poses a number of additional problems. In [41] Nehaniv and Dautenhahn discuss imitation in terms of mappings between dissimilar bodies, which can but need not involve learning. Thus, for the purpose of this paper we use imitation or imitative behaviour in the sense of having two agents, a model and an imitator, and both agents are able to actively maintain a following relationship. The capacity to follow is designed and assumed to be 'instinctive'. Since our experiments address learning by imitation instead of learning to imitate, the imitative behaviour (following) is not novel.

Imitation capabilities have been studied e.g. in macaques (see the well known example of the Japanese macaques of Koshima Island whose habit of washing potatoes develop from the obser-

vation and imitation of the individual finding of one macaque [26]), in parrots and mynah birds for reproducing songs [42] and of course in humans (e.g [9, 37, 55]). Similarly, recent work showed that artificial agents can also benefit from mutual interactions [14], e.g. for improving performance in collaborative tasks (e.g. [27, 31]), and from interactions with humans, e.g. for learning complex motor skills (e.g. [30, 52]). We view learning, communication¹ and imitation as important capabilities to possess by social robotic agents, and our previous work studied how these skills can be designed and used by physical autonomous agents. In this paper, we study grounding and use of communication among simulated agents, and, in particular, we investigate the role of social interactions for the sharing of a common context and the building of joint attention between the communicative agents.

Recent studies on the development of communication have addressed the problem essentially from an evolutionary perspective, (1) either through theoretical models based on biological studies of brain evolution (e.g. [16, 46]), or on sociological studies of primates and human societies (e.g. [2, 18]), (2) or through computer simulations (e.g. [28, 34, 45, 43, 51]). For these studies, the *symbol grounding problem*² [22] is solved, once the necessary cognitive abilities have evolved. However, few of these studies consider the influence of behavioural and social factors on the development of communication, exceptions are [18, 45]. A common trend among the above mentioned simulation studies is to give a very simplified physical description of the communicative agents and their environment. The communicative agents are described only in terms of their cognitive (by opposition to behavioural) abilities that enable production and reception of the communicative signals. We

¹Communication refers here to symbolic communication, characterised by the transmission of symbolic signals or symbols, that is arbitrary physical signals whose meaning follows from social convention among the communicative agents.

²The expression *symbol grounding problem* was first introduced by [22] to refer to the question of how the symbols or the representations of the world acquire their *meaning* for a computational agent.

distinguish between behavioural and cognitive capabilities. The agent’s behaviour relates to regularities of the agent’s dynamic interaction with its environment, as can be detected by an external observer. Behavioural capabilities are evaluated by having the agent interacting in an environment. Cognition generally refers to rational processes involved in the acquisition, organisation and use of knowledge, thus, cognitive capabilities refer to internal processes inside an agent. Cognitive and behavioural capabilities in animals are closely coupled and dependent on one another. However, in artificial systems the distinction can be made much more explicit, since models which are focusing on cognitive capabilities are often neglecting or strongly simplifying agent-environment dynamics (e.g. assuming complete or global information of the world and other agents).

The cognitive functions involved in grounding of communication have been simulated e.g. as the input/output of an artificial neural network [28], as a matching process [51] or as a probability function [43]. The agents in these simulations are disembodied, they do not occupy a physical space (they have no body, no sensors or actuators, and generally occupy not more than a single point in the space) and the result of their actions is atemporal (an action and its result occur in one time step). But most importantly, each agent has a perfect and identical perception of the environment features, based on an abstract model of the world.

By contrast, in our approach, grounding of communication is a process inherently *situated* [10] and dependent on the agent’s *embodiment* [10], as it is based on the agent’s individual and subjective world representation³ which it constructs through its interactions with the environment [13]. We say that two agents are *communicating* once they have developed a similar interpretation of a set of arbitrary signals in terms of their own sensor perceptions, that is once they have achieved a similar *categorisation* [22] of sensor perceptions and have successfully attached them with the same set of

³The agent’s representation of the world consists of the categorisation of its perceptions, which is done through updating the connections of its neural network (its controller), see section 3.

arbitrary signals. In [5, 7] we carried out physical experiments in which an autonomous mobile robot was taught by a human instructor and showed that it is not necessary for two agents to be of the same type to be able to communicate. It is not the means by which the agents perceive their environment that matters but rather that they can make similar perceptual distinctions and that these occur in close temporal relationship. In our experiments, we use an imitative strategy, namely mutual following of two agents⁴, to create a perceptual context common to learner and teacher agents, upon which the learner grounds its understanding of the teacher’s words. The set-up does not require that both agents use the same sensory data or even the same sensors in order to characterise a word. A concrete situation is interpreted on the basis of the learner’s own sensory-motor measurements.

In the computational linguistic studies we mentioned earlier ([28, 34, 43, 51]), grounding of communication is regarded as a computational problem that can be solved solely by means of combinatorial analysis. For these authors, categorisation of sensor perceptions into concepts results from a process of statistical elimination among all possible meaning-object pairs, where the most likely pairs, i.e. the most frequently observed, are chosen. However, combinatorial analysis alone is not always sufficient to discard all irrelevant information, as it is often difficult to present a sufficiently high number of relevant pair-meaning examples compared to irrelevant ones [16]. There are also numerous situations in which one feature does not appear (naturally) without another one, e.g. the eyes, mouth, nose and other human face features are bound to appear together with the whole face. In this case, combinatorial analysis would fail to attach two different concepts to the eyes and the face respectively, as there could be no example in which each of these features appears alone. Humans overcome this problem by using attentional mechanisms provided

⁴Following behaviour is an implicit means of imitation, as when the first robot follows the second robot, it replicates the second robot’s movements (moving, stopping, turning) in the plane.

either by the speaker/teacher (pointing, increasing the tone of voice, linguistic deixis) or by the listener/learner (focus of gaze in the direction of the speaker's gaze or the direction pointed by the speaker's finger). Attentional mechanisms act as a cognitive process which restricts the number of observations before combinatorial analysis. However, there is more to communication than just a single cognitive process. There is an interactive process between the two communicative agents. For those who study the development of language in children, it is clear that "to be effective early language learning must take place in a social setting .. [where] .. turn taking, mutual gaze and pointing are social devices .. [used for] .. establishing a joint attention [between speaker and listener] that creates a meaningful social setting necessary for the development of language" [21]. Other works implemented such attentional mechanisms as processes distinct from the learning mechanisms, e.g. Steels & Vogt's pointing strategy [50] and Yanco & Stein's action-selection mechanism [57]. By contrast, we develop a single cognitive architecture which enables both associative learning, selective attention from parsing of continuous sensory information, and the creation of a mutual binding between the two agents by means of mutual phototaxis (as an underlying social relationship). In agreement with Deacon's suggestion [16] that "the acquisition and use of symbols requires considerable facility for conditional associative learning, including an efficient short-term memory for sequences and combinations, and an ability to easily and rapidly produce new combinations", we developed a Dynamical Recurrent Associative Memory Architecture (DRAMA) [7] which provides short-term memory of events, learning of temporal sequences and fast retrieval of the learned combinations by means of Hebbian mutual associativity. Note that recent neurobiological studies support such an Hebbian approach to account for the brain processing involved in discrimination and categorisation of words (see [48] for a good review of these studies). In addition, following Deacon's suggestion that "symbolic processes and sensory-motor attentional process may utilise the same neurological computations for different purposes" [16], in our architecture the same

learning and retrieving algorithms are used for generating the symbolic associations (input from a radio sensor) and the sensor-motor associations, which are then used for controlling the robot's movements (motor activity) and communication (transmission of radio signals).

This work follows previous experiments in which we studied grounding of communication in a two-agent (teacher-learner) set-up [5, 6]. A learner robot followed a teacher agent, which was in one case a robot [6] and in another case a human instructor [5], around in a given environment. The learner robot was taught to differentiate between locations in the arena [6] (being on the plane, climbing up and climbing down a hill) and between different objects [5] (boxes, coloured patches on the floor) by associating the correct perceptual combinations with the different radio signals emitted by the teacher. Successful results demonstrated, on the one hand, the efficiency of the learning architecture at associating structurally different sensor patterns under a considerable amount of noisy data. On the other hand, these experiments showed the validity of the proposed imitative strategy, namely mutual following of the two agents, as a means of transmission of a vocabulary from one agent to a second heterogeneous agent (a human or a morphologically different robot).

In this paper, we report on three different studies, which are all carried out in a simulated environment. In the first part, we study how the imitative teaching scenario, used in the above described physical experiments in [5, 6], scales up to grounding communication among a group of robots. In the second part, we evaluate how the imitative behaviour, namely following of the teacher agent, benefits the agent by improving its learning performance. Finally, we implement a case study in which using the communication system which was learned in the first set of simulation, improves the learning performance of the communicative agents in another learning task. Here, the agents use the vocabulary to transmit to each other object locations, hence speeding up the learning of these locations by each agent individually. Note that we will use the term *robot* as well as the term *agent* in the rest of the paper to refer to the simulated mobile agents. The term

robot is used because the agents in the simulations represent physical robotic agents, as used in previous experiments. They have similar sensor and actuator capabilities (sensitivity and range) as that of the physical robots and use the same control mechanism (see section 4). The aim of the simulation studies was to carry out experiments with a group of robots, what we could not do in a physical environment (as we could not have more than two physical robots). We are however aware that simulations are not the same as physical experiments and have addressed this point previously in [7, 5]. Current work of one of the author investigates the implementation of the simulations reported here in a real set-up of several Khepera robots, see section 7.

The rest of this paper is divided as follows. In section 2 we explain the scenario of the experiments. In section 3 we describe the agents' control architecture and in section 4 we present the experimental set-up of the experiments. We report on the results of the experiments in section 5 and discuss in section 6 the contribution of these experiments to the study of the symbol grounding problem in situated robotic agents. We conclude the paper (section 7) with a short summary of the results of our studies.

2 The experimental scenario

We carry out two sets of simulation studies in which we investigate first grounding and then use of communication in a group of nine robots.

2.1 Learning the vocabulary

In the first set (section 5.1), the agents learn a vocabulary to differentiate between coloured patches and to describe their locations in terms of distance and orientation, i.e. in polar coordinates, relative to a homing point. The vocabulary is transmitted from a teacher agent T0, which has a complete knowledge of the vocabulary from start, to eight learner agents, which have no knowledge of the

vocabulary at the start of the experiments. Once a student becomes confident enough in its learning of the correct signal-meaning correlation (that is after having observed a sufficient number of times the same signal-sensor measurement pair), it becomes teacher in its turn. The teacher agent T0 does not learn, that is its definition of the vocabulary remains unchanged during the whole experiment. The learner agents, however, carry on learning even when they begin to teach. Thus, a ‘bad’ teacher, that is a learner agent which has become confident in incorrect signal-meaning pairs, can become a ‘good’ teacher later in the experiment if it can update again its confidence on the correctness of its word-meaning pairing under correct external teaching (and reversely a good teacher can become a bad teacher under opposite circumstances). The level of confidence under which the learner becomes a teacher is varied in the experiments and its effect on the learning performance of the whole population investigated.

Transmission of the vocabulary from teacher agent to learner agent occurs as part of an imitative strategy, namely mutual following, between the two agents. Each agent follows the other by means of phototaxis, namely tracking lights which are attached to the robots. Because following is mutual, it results in a smooth binding between the two agents. The agents seldom loose sight of each other, because when the learner runs slower than the teacher waits for it, and they find each other more easily as they are both looking for one another. While the two agents wander randomly in the environment and follow each other, the teacher sends radio signals, i.e. ‘words’, to describe its external perceptions, i.e. observation of a coloured patch, or internal perceptions, namely odometry (relative measure of distance from homing point) and orientation relative to the homing point. The learner attaches a meaning to the teacher’s signals in terms of its own sensor-motor perceptions, that is its vision perception (coloured patches’ detection) and its measure of odometry and orientation. While bounded by the following process, learner and teacher agents are set in a position from which they share a common context of both external (face the same direction)

and internal perceptions (perform the same movement, travel the same distance and on the same ground). This implicit similarity between the two agents' perceptions is what enables the learner to make sense of the teacher's words, as the teacher talks only of what it senses, unaware of the actual learner's perceptions. It is thus an unsupervised learning mechanism used in a teacher-learner set-up. We use a connectionist⁵ model for the learning of 'word'-observation pairs, where incorrect associations, due to mismatched agent's observations, are discarded compared to correct ones by a process of statistical elimination depending on their relative frequency of occurrence (cf. section 3). Further simulation studies are carried out (section 5.1.2), which investigate the importance of the following strategy for the success of the learning by comparing the learning curve of the agents with and without this following capability.

This study follows previous experiments which we carried out using two teacher-learner autonomous robots. There, the learner robot was taught by the teacher robot a vocabulary to describe its external perceptions of objects [7, 5] and internal perceptions of movement [8], inclination [6] and orientation [8]. The teaching scenario was based on the imitative-following strategy, as described above. The study reported here validates the imitative teaching scenario by showing that it scales up successfully to transmitting a vocabulary across a group of agents.

2.2 Using the vocabulary

In a second set of simulations (section 5.2) we study how the vocabulary learned in the first set of experiments can be used beneficially by the same group of nine agents. The experiments start with all agents knowing a vocabulary of 31 words for describing the nine different colours

⁵Our choice for using a connectionist architecture was driven by consideration relative to its implementation, see explanations given in [7]. This type of architecture was advantageous as compared to other learning technique, because it could provide fast computation and the ability for spatio-temporal association and learning of time series, as required by the task. Note that there is no biological plausibility in the model and in its implementation.

of the coloured patches present in the environment and their locations in terms of 14 distance measurements (measured by odometry) and 8 angle measurements (using its compass measurements and the odometry calculus). As the agents wander randomly around the environment, they learn the locations of the coloured patches by associating the sensor perception of the particular colour input with the coordinates of the agent's own position. When one agent has learned the location of a coloured patch, it can further transmit the coordinates of the location via the communication channel (using the learned vocabulary) to all other agents. Each agent can thus learn the coloured patches' locations from listening to the speaker agent's talking and without actually coming across the particular location itself. This results in a speeding up of the learning of the whole population, which we study using two different scenarios for the information transmission. The transmission occurs as soon as one agent comes across a coloured patch location (one-to-many, long distance communication) or only when two agents come across one another (one-to-one, short-distance communication).

Note that in the scenarios which we used in our experiments (i.e. learning of a vocabulary for colour types and locations in polar coordinates, transmission of the vocabulary through mutual following of teacher and learner agents and use of the vocabulary for transmitting information about coloured patches locations) was inspired by the communication system used by the Honey bees, following Moukas & Hayes' robotic implementation [40]. Honey bees communicate through a 'language' composed of dance patterns that form a vocabulary to describe different flower types and locations, the location being measured in terms of the distance, measure of energy required to reach it, and of the direction to follow from the nest. New born bees 'learn'⁶ to ground each dance pattern with a particular sensor perception of sugar, energy and orientation. It is suggested

⁶Note that the general meaning of each signal whether describing the colour type or location is the same in all bee society, only the actual sensor value associated to it varies.

that the newborn bees learn the meaning of the dance by associating their memory of the dance patterns and of the scent of the flower (of which traces had been brought back by the dancer bee) with the location of the particular flower, when they later discover it [19, 36]. Note, however, that we make no claim as to our system being a valid model of the Honey bees' communication system; we only stress its analogy with this system.

3 Control Architecture

Learning of the vocabulary, i.e. grounding of the teacher's signals in the learner robot's sensor-actuator state results from an association process across all the robot's sensor-actuator states, which is produced by a Dynamic Recurrent Associative Memory Architecture (DRAMA). A complete mathematical description of the architecture is given in [7]. In the following, we briefly summarise its main properties as given in [7], and then describe the training and retrieving algorithms which we use in the experiments reported in this paper.

The complete control system of our agents is composed of a set of *event recognition* modules (one for each sensor and actuator) and an associative module that contains the DRAMA architecture (see figure 1). Sensor and actuator information is encoded in binary (0/1) bit-strings. The event detector modules act as a selective mechanism on the sensor and actuator information. Once a variation in one sensor measurement or actuator state occurs, i.e. an event has happened, it triggers the corresponding event detector module and the new information is passed further to the associative module (in the experiments a one bit variation was considered to be sufficient to trigger the module). The associative module (DRAMA architecture) consists of a fully recurrent network with self connections to each unit. The structure of the network is dynamically updated each time a unit has been activated by an input from the event detector. To each connection in the network are associated two parameters: a *time parameter* and a *confidence factor*. Time parameters and

confidence factors are updated following Hebbian rules, providing an associative type of learning; the time parameters record the time delay between units' activation while the confidence factors keep a memory of the frequency of units' coactivation. The self connections to the units provide a short-term memory of the activation of the unit, by sending back the activity to the unit. The duration of the memory is limited, as the activity decreases by a fixed ratio after each passage along the recurrent connection. The unit activity stops (becomes zero) when the recurrent activity is lower than a fixed threshold. The short-term memory of a unit's activity enables associations between patterns of unit activation that have been delayed in time, a specificity required for solving our particular learning task (the following strategy implies that radio signals have to be associated with sensor measurements which are received with a varying time delay).

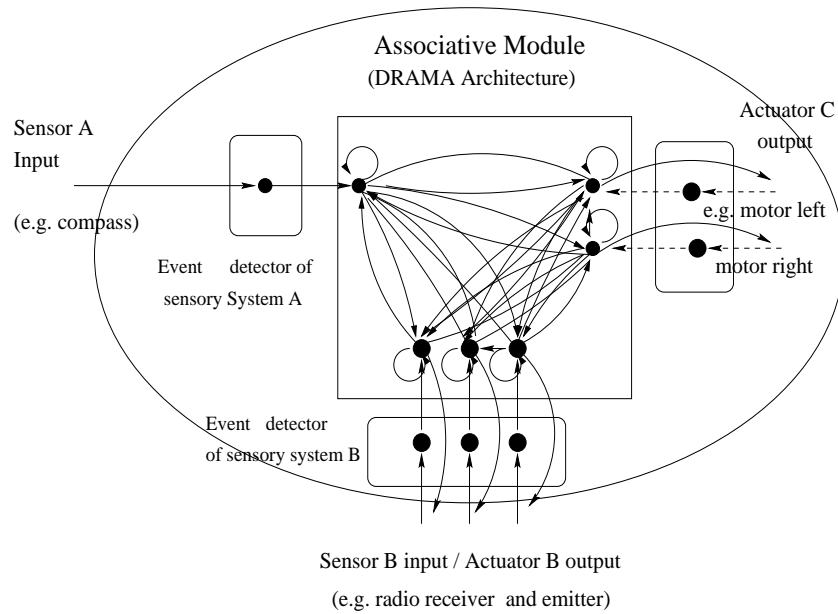


Figure 1: Schematic representation of the control system of the learner and teacher agents with three sensor-actuator systems, e.g. compass, radio and motors. Note that in the experiments the motors do not input to the associative architecture as there is no external mechanisms for changing their activity, hence the dotted arrows in the figure. Only the output of the architecture to the motors is used to directing the robot's movements.

3.1 Basic behaviours

In the experiments we report here, DRAMA is used for both learner and teacher agents. Each robot starts with a set of basic behaviours for wandering, obstacle avoidance and following. The behaviours are predefined by setting in advance the values of the connection parameters between infra-red sensors, bumpers, light sensors and the agents' motors. Obstacle avoidance uses infra-red and bumpers. Following consists of phototaxis using the light detectors (cross-connectivity between light sensors and motors). Wandering is a default input unit to both motors, always active, that leads the robot to move straight when no other movement is produced by the two other behaviours. Similarly, the teacher robot's knowledge of the vocabulary (in the simulations) is predefined by setting the value of the connections between relevant sensor information and the radio emitter. Details of the implementation can be found in [4].

At each time step, i.e. one processing cycle, the robot's behaviour is determined by retrieving the network outputs to the actuators (motors and radio emitter) given the current sensor input. The teacher 'speaks' (emits radio signals) only when it sees the learner.

3.2 Training algorithm

Learning is dynamical and is performed continuously by updating the connection parameters to account for variations measured in the sensor-motor state. That is, at each cycle the current sensor and actuator information is compared (in the event detectors) to the information measured in the previous cycle. If a variation is noticed, the new information is passed on to the associative memory where it is memorised for a fixed duration, during which it is associated to any incoming event. It is important to understand that the same network (the DRAMA architecture) is used for learning (i.e. generating the symbolic associations between radio sensor and other sensors) and control of the agent's behaviours (i.e. motor activity and radio signal transmission).

The success of the learning is determined in the experiments by comparing the proportion of correct and incorrect signal-object associations. Learning is successful when correct associations have been done more often than incorrect ones. Incorrect associations result from a mismatch between teacher and learner perceptions, where the learner robot associates the teacher’s radio signal with an incorrect perception (which does not correspond to that of the teacher). This mismatch is due to the imperfect following of the two agents. Note that there is no noise/imprecision in signal transmission, i.e. the signal is perfectly received. Incorrect associations can also be due to incorrect emission of the signal, as it is the case in the experiments (section 5.1) in which we vary the threshold of confidence under which the learner robots can speak. It happens that a learner robot (which has become teacher) emits incorrect signals (relative to the definition of the vocabulary), when it is allowed to speak before it had made the correct signal-object correlations.

The comparison between correct and incorrect correlations consists of calculating the ratio between the values of the confidence factor parameters⁷ attached to the correctly and incorrectly activated connections. The confidence factor parameter gives a measure of frequency of co-activation of the two units linked by the connection, that is, e.g., of the frequency of correlation between one signal (one activated radio unit) and an object’s features (pattern of activated units in colour, angle and odometry sensors). In the stage of learning the vocabulary (section 5.1), we consider only the connections between the radio sensor units and the units of colour (coloured patches’ feature), angle and odometry sensors (location), while in the stage of learning the locations of the coloured patches (section 5.2), we consider only the connections from the colour sensor to the angle and odometry sensors. Two units are said to be *correlated* if the value of the confidence factor associated with their mutual connection is higher than half the maximal value of confidence

⁷In the experiments presented here, we did not analyse the time parameters, which are mainly useful for the recording of sequences, see [7].

factors for all connections leading to each of these units. Note that the network’s connections are bidirectional and asymmetric and that we require correct correlation in both directions. Learning of the signal-object associations is successful when all radio units (one unit stands for one signal) are correctly connected to the corresponding colour-angle-odometry unit combinations (describing the corresponding objects), that is once all the connections between radio units and these other sets of units satisfy the above mentioned correlation criteria. Choosing the threshold on correlation to be half the maximal activation means that we accept a proportion of up to 50% of noisy data, where the noise consists of incorrect signal-sensor associations.

Note finally that there is no a-priori limitation on the number of words the robot could learn; this depends on the number of inputs (i.e. radio input units in the experiments) to the network, which fixes the maximal capacity of the network (i.e. the number of patterns which can be stored). This capacity has been evaluated to be of the square of the number of network units (i.e. N^2 , where N is the number of units, see [7]). The system could thus possibly scale up to learning a vocabulary of much bigger size. In [3, 4], we report on an experiment with a mini-sized humanoid robot (doll-shaped), which is taught by a human instructor English proto-sentences to describe its interactions with the instructor. This experiment exploited DRAMA’s ability for learning sequences of inputs, which allows learning of the sequence and combination of the words forming the sentences. The experiments reported in this paper exploit only DRAMA’s ability at binary association for word-meaning pairs.

3.3 Retrieval

The phases in which the robot uses its understanding of the vocabulary to learn new colour patches’ locations (section 5.2) is obtained by retrieving the correlated units activity in the colour, odometry and angle sensors given the particular radio signal. The robot’s emitter sends three radio signals for

describing the patch type (colour) and its location in terms of angle and odometry measurements. The three radio signals retrieve a specific unit activity in colour, angle and odometry sensors. Learning of the so transmitted coloured patches' locations results from updating the connection between the co-activated units in these sensors.

A retrieval mechanism based on the notion of retrieved and real sensor states was implemented to allow the robots to use their understanding of the vocabulary to learn new colour patches' locations from hearing another robot's saying so. The robots' predefined understanding of the vocabulary is set-up by fixing from the start their networks' connections between the radio sensor and the colour, angle and odometry sensor, such as to represent a vocabulary of 31 words to label 9 colours, 14 odometry levels and 8 angle measurements. The robots communicate by exchanging information about objects' locations. One robot's message is composed of three radio signals (1 byte with 1 bit activated) for describing the object's colour and its location in terms of angle and distance measurements. When one agent receives a radio signal sent by a second agent (which produces a pattern of activity in the radio sensor units), it decodes it by retrieving the correlated activity in units corresponding to the colour, angle and odometry sensors. That is, the three radio signals retrieve a specific unit activity in each of the set of units attached to colour, angle and odometry sensors. Learning of location of the coloured patches transmitted in this way results then from updating the connections between the co-activated units in the colour, angle and odometry sensors (but not the radio sensor, see explanations below).

The sensors' activities which are retrieved for the colour, odometry and angle sensors are stored in a *retrieved vector state* (one for each sensor), which are different from the *real vector states*, which contains the current sensor measurement. The retrieved and real vector states of the sensors can be thought of as two sets of network units, parallel to each other and receiving outputs from all other units in the network (as shown in figure 2). The retrieved vector state does not input

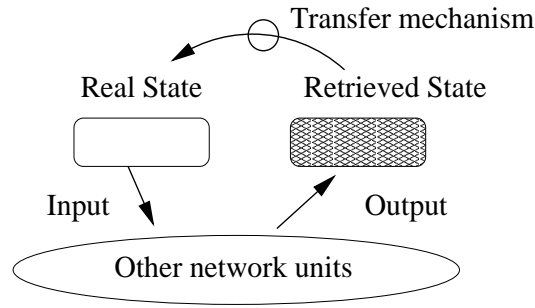


Figure 2: The control system with retrieved and real sensor states.

to the network. Thus, the retrieved sensor activity in the experiment (which is placed in the retrieved sensor state) does not affect directly the robot's behaviour; that is, it does not participate in controlling the motors and radio activity (by retrieving the *real* unit activity in these sensors). However, in order to control the robots' speaking, i.e. to activate the radio emitter to transmit the objects' locations, a mechanism is implemented, which transfers the retrieved state into the real one. This transfer occurs either when the robot discovers an object or when it meets another robot and engages in an answer-question dialogue (see description of the results below).

Learning of the objects' locations results then from association among retrieved sensor states and association among real sensor states. It is important to note that association among retrieved sensor units is done separately from association among real sensor units. Thus, in the experiment, the listener agent can either learn from the retrieved sensor units' activity (which correspond to the sensor perceptions 'transmitted' by the speaker robot) by associating mutually the active units in each sensor system or it can learn from the real sensor units' activity (its current sensor measurements), but it does not associate mutually its retrieved sensor state (speaker's measurement) with its real sensor state (its current measurement). By definition, radio measurements are never associated with other sensor measurements in this experiment. When they are produced by the speaker, they do not produce an input activity in the speaker's radio sensor but rather an output

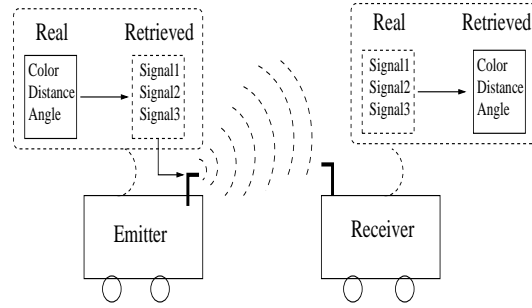


Figure 3: Transmission of the location of an object. The emitter robot first translates the *real* (as measured when it discovered the object) sensor vector states containing the description of the object features (color, angle, distance) into the three corresponding (*retrieved*) radio sensor states (signal 1-2-3). It then transfers the *retrieved* radio states into real ones; that is, it emits the three signals. The receiver robot picks these signals up by storing them in its *real* radio sensor state. It then translates these three signals by retrieving the corresponding *retrieved* sensor states for color, angle and distance of the object. This information is further used by the receiver robot to learn the location of the object.

activity, which is then not forwarded to the DRAMA associative module (see beginning of section 3 and [7] for explanations); thus, no association occurs. When they are received by the learner, only the retrieved sensor states they produce are associated. Association between real sensor states was defined so that all except the radio sensor state could be associated. Figure 3 illustrates the transfer from real to retrieved sensor states during a two robots' transmission of an object's location.

4 Experimental setup

Simulation studies are carried out in a rectangular arena measuring 700 by 700 units (1 unit = 1 cm), which contains nine objects, coloured patches, which can be distinguished by their different colours. Figure 4 shows a graphical representation of the simulated environment with the nine robots. The objects are represented as big rectangles of different colours. These are coloured patches lying on the ground over which the robots can run. In the figure, we see the teacher robot

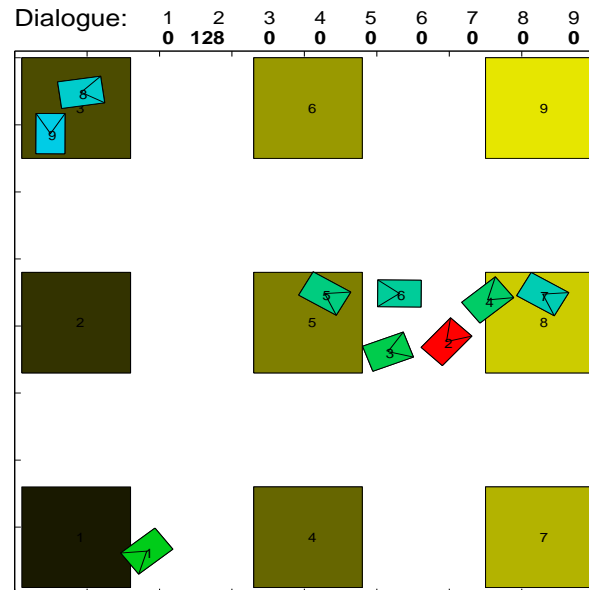


Figure 4: Simulation studies are carried out in a 2-D simulator, in which the robots are represented as rectangles of 30 by 20 units, a triangle determining the front. Coloured patches are represented as big rectangles whose number corresponds to the colour type. There are nine types of coloured patches with one location for each type. In the figure, we see the teacher robot (in dark grey) between five of the eight learner robots (in light grey), moving between objects 5 and 8.

(in dark grey) between five of the eight learner robots (in light grey), moving between objects 5 and 8. On top of the figure, the result of the robots' speaking is written for each of the nine robots. The teacher robot (2nd column) outputs '128'. This refers to the activation of the 8th radio unit in the robot's network, which is the radio encoding for the label of the object, across which the three robots are currently running. None of the learner robots is speaking, i.e. they all output '0', as they have not yet seen the object and thus have not yet associated the teacher's signal with the object's features.

There are nine⁸ robots, which are equipped with colour vision (nine colours) to distinguish

⁸The choice of using nine robots and nine coloured patches is purely arbitrary. The point is to demonstrate transmission of communication among a group of agents, i.e. composed of at minimum three agents. These numbers correspond in fact to the maximal number of elements for which simulations could be carried out in a reasonable

Table 1: Sensors sensitivity

Sensor type	Sensitivity range	Bit-string Encoding
Infra-red	180 ⁰ and 0-45cm	8 quadrants/bits
Light front	180 ⁰ and 0-45cm	8 quadrants of 22.5 ⁰ (8 bits)
Compass	bearings of 45 ⁰	8 bits
Angle	bearings of 45 ⁰	8 bits
Odometry	7 levels: dist. of 25cm	7 bits
Colour sensor	9 types	9 bits
Radio	whole arena	31 bits

between each coloured patch, infra-red vision to see the walls, light vision in front and back for mutual following by phototaxis and a radio transceiver to communicate. They also have a compass which measures bearings of 45 degrees and an odometry sensor which is incremented at each wheel rotation, hence giving a notion of travelled distance. Given the compass and odometry sensors information, the robots calculate their positions relative to the centre of the arena. They can then determine their position in the arena (the angle (virtual) “sensor”), which is divided in 8 quadrants. Infra-red and light detectors are associated with a cone of vision of 180 degrees which is segmented into eight quadrants. The measurement of the sensor is given by an 8-bit string where each bit corresponds to the value measured in each of the eight quadrants (e.g. infra-red=(11000000) stands for an infra-red activation of the first two quadrants). The range of sensitivity of the sensors are given in table 1. Each coloured patch is defined by a different colour. The behaviours of all robots are calculated by the same routines, that is there is one network (DRAMA architecture) per robot and the same retrieving and training algorithms are applied sequentially to them for determining the behaviour and learning of each of the robots independently⁹. In order to produce a more realistic amount of time, each complete set of simulations (10 runs) requiring a week of CPU time.

⁹Code is written in C and is processed serially. The C programs for the simulations were run on Ultra 1 Model 140s SPARCstations and the graphics representation was made using the MATLAB environment.

simulation¹⁰, the following behaviour is made imperfect. An agent is able to determine the position of the other agent with respect to its own by measuring the difference of intensity measured by its eight frontal light detectors. It can determine the other’s position with a precision of 20 degrees. In addition, noise¹¹ is introduced in the calculation of the robot’s movements in order to represent the imprecision measured in the real robot’s movements. The imprecise alignment and following of the robots leads to differences in their respective travelling and hence differences in their sensor measurements. Noise, i.e. incorrect teaching-observation associations in our experiment is due to this mismatch between teacher and learner perceptions, which results from the imprecise robots’ alignment (thus the noise is contained between 0 and 20%).

5 Results

In this section, we report first on simulation studies in which we study transmission of a vocabulary among a group of nine agents which then use the vocabulary for transmitting information about object locations.

5.1 Learning the vocabulary

In a first set of simulations (section 5.1.1), we study how a common understanding of the vocabulary can spread among a group of agents when starting with one teacher agent and when each successful learner agent can in turn become another teacher. In a second set of simulations (section 5.1.2), we study the influence of the following strategy on the learning performance for different learning

¹⁰The aim of the simulation studies was to further investigate the validity of the following teaching scenario, which had been first implemented in a physical set-up. It was therefore important to make the simulation as close as possible to the physical experiments. A simulated environment had to be used as it was not possible to have a real set-up of nine mobile robots.

¹¹This noise consists of adding or subtracting a random number to the distance travelled in one wheel rotation.

tasks.

5.1.1 Social learning in a group of agents

We carry out a set of 10 runs (1 run lasting for 400,000 cycles) in which we study the speed of learning of a vocabulary of nine words by eight learner robots, given 1 teacher robot. The nine words refer to nine different coloured patch types, defined by nine different colours. The coloured patches are spread homogeneously in a square area as shown in figure 4. The learner robots can become teachers, that is they can emit signals for ‘naming’ the coloured patches once they have reached a sufficient level of confidence (specified by a threshold) in their word-colour associations. We ran 10 simulations with a different value of the threshold (from 10 to 100) for each. Each threshold value corresponds to considering the agent sufficiently confident in a particular word-colour association when it has observed this combination for at least 10, 20, 30, ...,100 times. This refers to the actual value of the confidence factors for these word-colour connections, as the confidence factors are increased by a value of 1 at each co-activation of the corresponding units. Because the agents can make incorrect associations due to mismatched observations of teacher and learner agents, correct learning of the word-colour pairs is not immediate but results from the repetition of the teaching process until incorrect associations can be discarded from the correct ones by virtue of their relative frequency of occurrence. Therefore, choosing too low a threshold on confidence factors can allow the learner agent to become a teacher before it has made a sufficient number of correct correlations compared to incorrect ones. In this case the robot would emit incorrect signals given a colour perception, hence letting the learner agent following it make an incorrect association. This would increase the global amount of noise, i.e. of incorrect association episodes, and could lead to the failure of the experiment (i.e. unsuccessful learning for all agents) if the number of incorrect associations outnumbers the number of correct ones. This is demonstrated by the results of the

experiments, where we observe that learning is unsuccessful as long as the threshold is lower than 50.

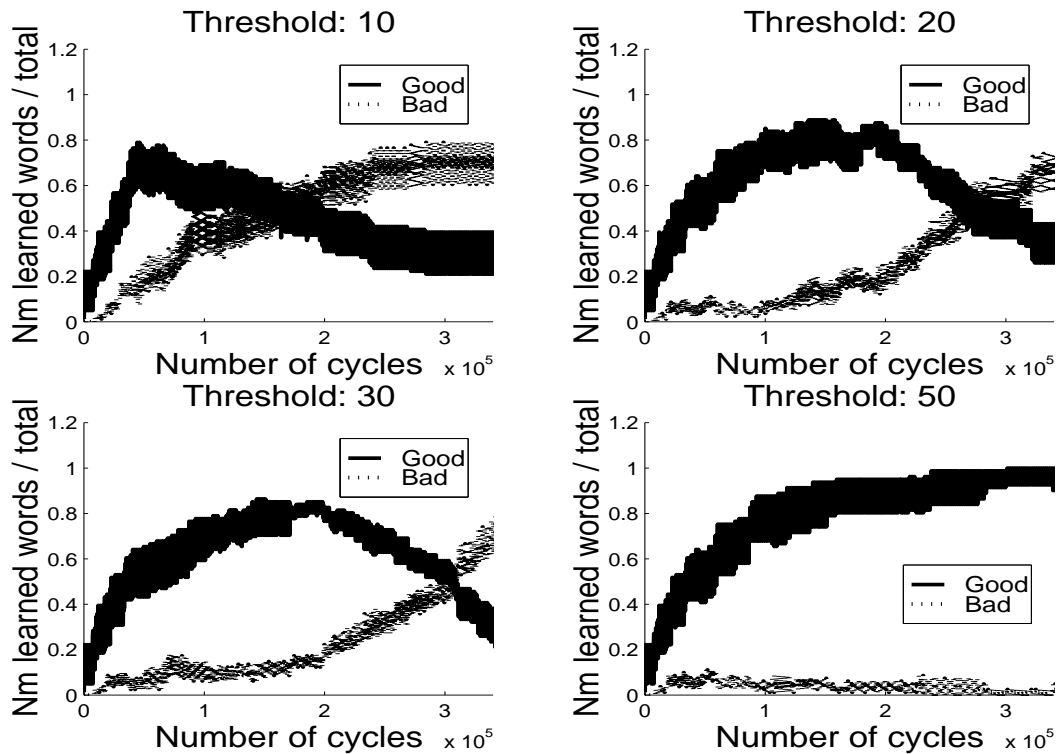


Figure 5: Learning curve (mean value for all agents) for four different thresholds (10,20,30 and 50) on the radio unit activation.

Learning is unsuccessful as long as the threshold is lower than 50. The lines *good* and *bad* represent the proportion of correct and incorrect respectively signal-object associated pairs over the total number of taught pairs.

We show in figure 5 the learning curve for four different choices of minimal levels of confidence out of the ten investigated. We observe that the number of incorrect correlations increases in the three first examples, so as to become more frequent than correct ones, leading finally to unsuccessful learning. Note that the curves for threshold values greater than 50 have similar shape to that shown in figure 5. The simulation with threshold 50 was carried out for 20000 more cycles and it was shown that the curve did not decrease after having reached the maximal value of 1. The value of 50 for the threshold under which the learning is successful reflects the proportion of noise, i.e. incorrect

associations due to incorrect matchings of the teacher-learner perceptions in the particular set-up. This value would then vary from one experiment to another. In the present case the increase of the noise which leads to unsuccessful learning is due to incorrect teaching provided by the learner robots. The more robots are speaking, the more noise is contained in the teaching. By definition, each robot can hear only other robots which are in a distance of 1.5 its body size. This means that each robot can be taught by at most four other robots (when placed in each quadrant around it). In fact, each robot is usually taught by two other learners, as the robots tend to quickly form long chains. If one robot is a bad teacher then its incorrect teaching spreads quickly to other robots. We make three observations: 1) There is in average 20% of noise, i.e. 20 % of associations are incorrect; this is due to incorrect matching of the teacher-learner perceptions (as evaluated in the two agents, teacher-learner scenario [5, 7]); 2) Each object's description (angle and distance measurements) overlap with at least two other objects' descriptions (which makes the association more difficult, as the correct set of features is less distinct, see discussion of the architecture's properties in [7]); 3) Each robot can be taught at most by two other learners. Given the three above facts, the noise increase due to one bad learner is enormous and chances are small that correct learning would result in this case. Correct learning is thus ensured when all learners speak only when they have correctly learned the vocabulary. Each robot is taught each signal about 5 times during its passage across one object (the teacher repeats the signal several times). Correct learning of the vocabulary requires a statistics of at least 10 passages over each object (i.e. 10 different teaching episodes in which the robot approaches the object from a different direction). Therefore, correct learning should occur after about $10 \cdot 5 = 50$ associations, i.e. a value of 50 for the confidence factors associated with the signal-object features.

The simulations in which learning was unsuccessful were stopped, while the curve of bad association was still increasing. If we had let the simulations run longer we could have checked whether

the curve would converge to the maximal value of 1. In this case, this would have meant that all robots had finally converged to a common definition of the vocabulary, while this definition would have been different from that initially taught by the teacher robot. The observation of a shift in the meaning of the words of the vocabulary for the agent is similar to the ‘emergence of a dialect’ as in the studies of [1, 51] in their simulations of the development of language. We did not make this analysis mainly because the aim of our experiment was to study correct transmission of a fixed pre-defined vocabulary (and also because these simulations were extremely long to run), as opposed to [1, 51] who studied the emergence and variation of a lexicon as an effect of its transmission. It would of course be very interesting to carry out similar studies in the future, using the physical simulation and the DRAMA architecture. In particular, it would be valuable to compare our results with results obtained in [1, 51] where no physical or behavioural description of the agents were given. We could then determine the role played by these physical factors in the transmission of a language. This point will be further discussed in section 6.

5.1.2 The benefit of following

We carried out a set of ten runs (1 run lasting for 100000 cycles) in which we study the importance of the imitative teaching scenario, namely mutual following, for a good transmission of the vocabulary. For these studies we use only three agents, 1 teacher and two learners. One of the learners does not possess the predefined capability of following another agent by means of phototaxis. Instead, when it meets one of the two other agents it simply avoids it as an obstacle and carries on its random wandering. It can however learn, similarly to the other learner agent. If it is close enough to the speaking agent, it can receive its signals. It can then associate them with all other sensor measurements it made during a short time delay window (plus minus the short-term memory duration, c.f. section 3) before and after the signals’ reception.

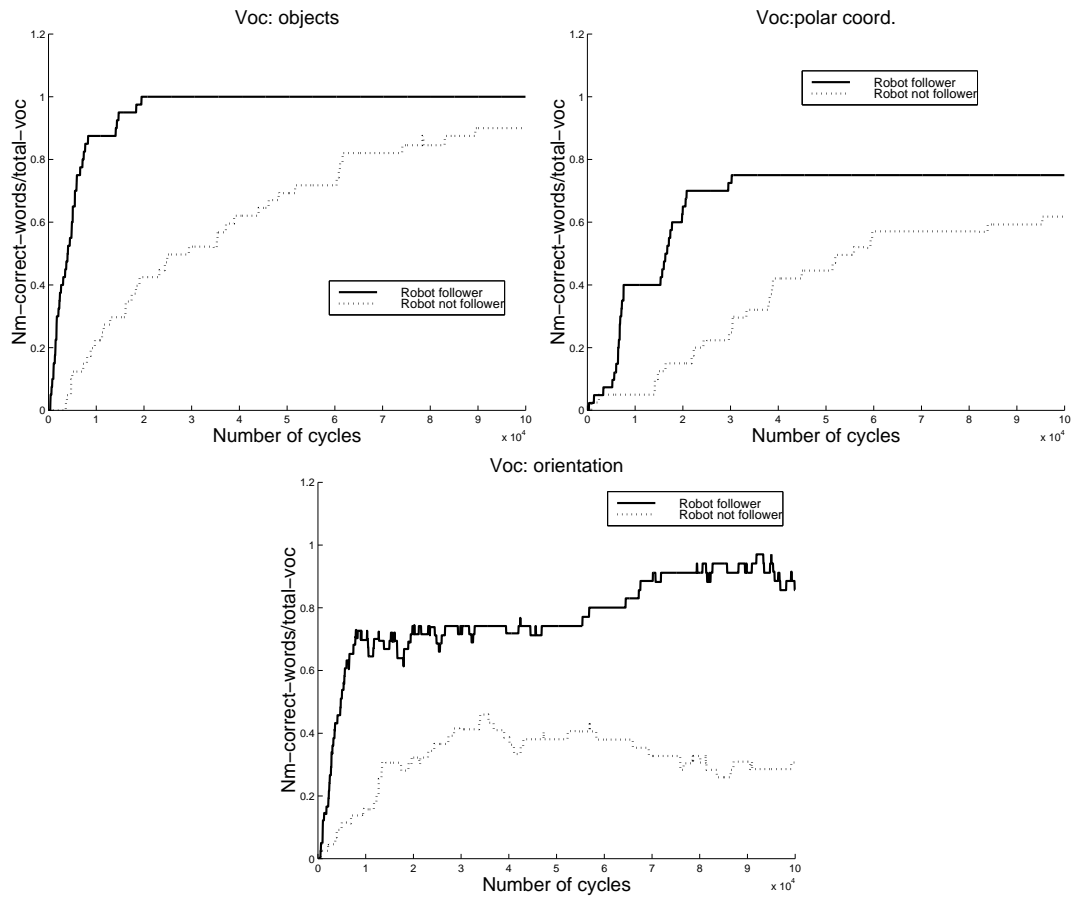


Figure 6: Learning curves for follower and non follower learner agents in three different learning tasks: learning a vocabulary for objects, for scaling values of polar coordinates and for orientation relative to a compass.

Figure 6 shows the learning progress of the two learner agents (with and without phototaxis capabilities) when learning three different types of vocabularies: 1) a vocabulary to describe the nine objects (coloured patches) of the environment, 2) a vocabulary to label scaled values of polar coordinates referring to the objects' locations and 3) a vocabulary for the four quadrants of a compass which measures the robot's individual orientation. Learning each of these vocabularies corresponds to associating radio signals, i.e. the teacher's words, with respectively 1) nine different colour measurements (objects' features), 2) five different measurements along the odometry scale and eight different measurements of angle, position in the arena with respect to 8 quadrants (objects'

Table 2: Comparison of learning success of follower and non-follower agent

	objects		polar coord.		orientation	
	Foll	Non-foll	Foll	Non-foll	Foll	Non-foll
Nm-learned-words/total-voc.	1.0	0.6	0.5	0.5	0.6	0.0

polar coordinates), and 3) four different compass measurements. The Y-axis represents the ratio between correctly and incorrectly learned words. In table 2, we show the mean values of this ratio over the 10 runs for each experiment and for each learner robot.

Results show that the non-follower agent is less successful on average and slower at learning the vocabulary concerning the coloured patches and the polar coordinates, and that it was always unsuccessful at learning the vocabulary concerning the orientation relative to a compass. These results imply that the ability of following improves the grounding of exteroceptions, as done when naming the coloured patches detected by different colour perceptions and when naming its position relative to global polar coordinates. But it is especially important for grounding proprioceptions, as done when naming its relative orientation. Being close enough spatially is often sufficient for the agents to share a common context of external perceptions and then to successfully ground the vocabulary onto the same sensor perceptions. In the experiments the non-follower agent learns the vocabulary concerning exteroceptions correctly because when it receives the teacher’s signal, it is often close enough to get a similar measurement of polar coordinates (the spatial scaling of the environment it learns to name is wide enough to allow two agents to share the same set of coordinates at one time) and to make a measure of colour (detection of coloured patch) shortly before or after meeting the teacher. The follower agent is, however, faster and more successful at learning because of its constant spatial closeness to the teacher agent. By contrast, it is not sufficient for the agents to be spatially close to one another for them to share a common set of

internal perceptions which would allow them to successfully ground these proprioceptions onto a common vocabulary. The imitator replicating the model’s actions by using the following strategy results in both agents performing the same movements relative to their body axes. It also allows it to make other similar internal perceptions which are a consequence of its actions, e.g. its orientation, inclination (see experiments reported in [6]), and odometry. In the experiments the follower learner agent implicitly imitates or replicates the teacher’s movements while following it, and consequently orients its body towards the same direction as the one pointed to by the teacher. In contrast, the non-follower agent which tries to avoid the teacher agent is not or seldom (and then only for a short period) oriented similarly to the teacher and is thus less likely to correctly associate the teacher’s signals. This accounts for the failing of its learning observed in all 10 simulations.

5.2 Simulations: Using the vocabulary

We carried out 3 sets of simulation studies (10 runs each) in which the nine robots use the vocabulary, learned in the experiment of the previous section, to transmit to each other the location (polar coordinates) of the nine objects (coloured patches) in their environment. The experiment starts with all agents knowing a vocabulary of 31 words for describing the 9 different colours of the coloured patches present in the environment and their locations in terms of 14 distance measurements (odometry levels) and 8 angle measurements. As the agents wander randomly around the environment, they learn the locations of the coloured patches by associating the sensor perception of the particular colour input with the coordinates of their own position. When one agent has learned the location of a coloured patch it can further transmit it via the communication channel (using the learned vocabulary) to all other agents. Each agent can thus learn the coloured patches’ locations from listening to the speaker agent’s talk and without actually coming across the particular location itself. This results in speeding up the learning of the whole population.

We study two different scenarios for the information transmission. The transmission occurs as soon as one agent comes across a coloured patch location (one-to-many, long distance communication) or only when two agents come across one another (one-to-one, short-distance communication).

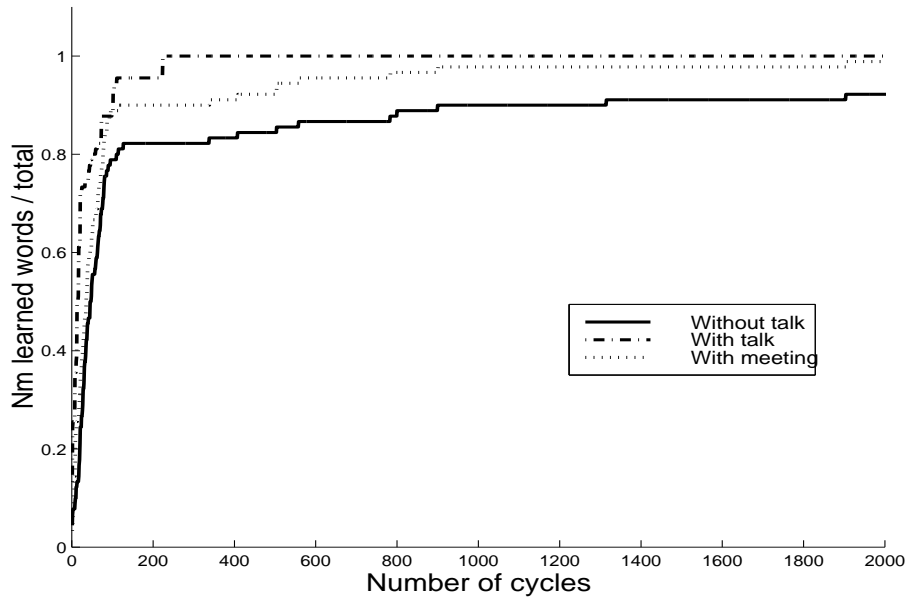


Figure 7: Speed of learning of the locations of the colour patches in three cases (mean value over all nine robots): as a result of each robot's individual exploration (-), by listening to another robot's speech which talks as soon as it finds a new location (-.-), through dialogue conversation when meeting another robot (....). Data are mean values over 10 runs. Maximal standard deviation of all three curves is 0.1. There is no overlap between any two curves.

In figure 7 we compare the results for three different learning scenarios (called Experiments 1, 2 and 3). In the first case the agents learn the locations of the coloured patches by making the associations when they travel over a coloured patch; in this case learning of the patches' locations results from each agent's individual search. In Experiments 1 and 3 the agents use their knowledge of the vocabulary to transmit to each other information about each patch location. In these two experiments learning results from both individual search and from social learning (interaction with other agents). An agent informs the other agent about a patch location by emitting a set of three

signals, one corresponding to the colour type, and two corresponding to the ‘words’ for its location in terms of distance and orientation. The other listener agents learn the new location by associating together the corresponding three sensor stimuli which have been activated by the reception of the signals (as explained in the previous subsection). In Experiment 2 the agents transmit the location of the coloured patch as soon as they discover it, that is when they travel on it. The speaker robot’s signal can be received by all robots in the whole arena (long distance communication). In Experiment 3 the robots’ transmissions of information occur only when two robots meet (short distance communication). That is, when two robots are close enough to ‘see’ each other, they engage in a conversation; each robot speaks in turn, the robot with lowest numbering (robots are numbered arbitrarily as 1,2,..to 9) first, and asks the other for all coloured patches whose location it does not know yet. That is, one robot sends a signal for the colour type (question) which activates in the receiver agent’s network the corresponding patch location (if known). The retrieval process proceeds in two stages: first the corresponding sensor stimuli for angle and odometry are retrieved, and then by transitivity of the associations the corresponding radio signals are retrieved which are further emitted by the robot as its ‘answer’ to the other robot’s query. The dialogue ends as soon as the agents become separated after 20 cm (e.g. when one agent turns in an opposite direction for avoiding an obstacle or another agent), or when they have enumerated all the objects.

We observe that learning of the whole group is faster when the robots can transmit to each other their current knowledge, that is the learning curve converges faster in the second and third experiments than in the first one. Learning is also faster when the transmission can go from one robot to all robots (second experiment), rather than from one to one (third experiment). This is not a surprising result. Surely, learning with one-to-many communication is faster than with one-to-one communication. One-to-one communication of the objects’ locations speeds up learning of the locations compared to an individual search; this is especially due to the fact that the agents

do not stop moving to speak to each other; that is, they engage in a question-answer dialogue whenever moving close to one another. Therefore, the dialogue does not interrupt their random search behaviour and, thus, communicating with one another does not delay their individual search behaviour. Dialogues can therefore only enhance the robot's knowledge about locations.

6 Discussion of results and theoretical implications

This paper reported on simulated experiments in which we demonstrated the validity of a connectionist learning architecture and a teacher-learner scenario based on imitative skills for transmitting a vocabulary among heterogeneous agents. In the first set of studies, we showed that the scenario scales up successfully for transmitting the vocabulary among a group of robots. Further, the importance of the imitative strategy for a successful transmission of the vocabulary was pointed out, as it was shown that agents lacking this capability would be slower at learning or would simply fail. The capacity of imitating the other is particularly important when the word to be learned concerns the agent's proprioceptions, e.g. orientation, as it allows the two agents to share similar internal perceptions, e.g. pointing in the same direction. An experiment was implemented to show an example of a situation in which it is advantageous and useful for the agents to communicate. Symbolic communication is used by a group of robots for transmitting to each other the coordinates of objects locations. This speeds up the robots' learning of these locations, as compared to the robots learning the locations only when finding one by chance during the robots' random walk. The rest of this section discusses the contribution of these studies and their results on our understanding of the symbol grounding problem as faced by embodied situated robotic agents.

We built our approach on the assumption that grounding of communication creates constraints not only on the cognitive capabilities of the agent but also and especially on its behavioural capacities. In the introduction, we distinguished between behavioural and cognitive capabilities, where

the former do not result only from the outcome of the agent's cognitive processes, but from the interaction of these processes' outputs, i.e. the agent's actions, with the dynamics of the agent's environment. In [12] we discuss the coupling of internal and external dynamics (i.e. behavioural and cognitive processes) and their roles in social understanding. Agents with poor cognitive capabilities cannot be expected to exhibit overwhelming behavioural complexity, and vice versa. A balanced design of cognitive and behavioural capabilities of the agent on the one hand, and environmental complexity on the other hand, is hard to implement but seems to be most promising. This is increasingly important in the case of social agents, where the environment is not static and fully predictable but consists of other agents.

With respect to cognitive capacities we defined key features for the learning capacities, such as a selective mechanism for the discretisation of sensor perception and a short-term memory of perceptual events and associative capacities. We implemented these in our experiments by using the DRAMA architecture. In the introduction, we pointed out the importance of a basic behavioural social relationship between the two communicative agents to act as an attentional mechanisms for eliminating irrelevant information which can not always be discarded by means of combinatorial analysis only. We thus provided our agents with a basic ability of mutual following which creates a spatial and temporal binding between the two agents necessary for their sharing of a similar set of perceptions. However, although we used a simple protocol, namely mutual phototaxis, for achieving this basic 'social' interaction, this proved to be a powerful 'external' (behavioural instead of cognitive) attentional mechanism.

Similar experiments on grounding radio signals in robots' sensor capabilities were carried out previously by Yanco & Stein [57] and Steels & Vogt [50] who used respectively reinforcement learning and evolutionary techniques. We showed in [7] that our model is more efficient and faster than the above models to learn a vocabulary of the same size and type. Moreover, our model

(the associative memory and the following scenario) has the advantage, as compared to these methods, to be less restricted in the sensor stimuli the robots could talk about. In Yanco & Stein’s work [57], the vocabulary consisted only of the robot’s actions because the learning algorithm was based on an action-selection mechanism. In Steels & Vogt [50], the vocabulary concerned only the robots’ external perceptions as these were the only perceptions they could share. By contrast, the mutual following strategy we use in our work allows the two agents to share a common context, namely sharing external (facing the same direction) and internal perceptions (performing the same movement, travelling the same distance and on the same ground). When aligning itself behind the teacher the learner agent naturally points in the direction the teacher ‘looks’ at. While we avoid the problem of the mirror effect¹² occurring in pointing, we however introduce another problem lying in the fact that the learner cannot see over the teacher’s body. In our experiments this is partly solved by the two agents moving and the learner eventually reaching the teacher’s previous position. This obstruction problem also occurs in Steels & Vogt’s pointing model [50], as in their experiments the two robots spoke also about objects on their sides which one of the agent could not see. In addition, because the learning mechanism we use is based on mutual associations between inputs from any sensor or actuator systems of the agent, the vocabulary can concern any (proprio and extero) perceptions of the agent. We reported here on experiments where the agents talked about external perceptions of objects and internal perceptions of orientation. In previous work, we reported on experiments where the learner agent was taught a vocabulary concerning its inclination [6] and its actions [8].

¹²In Steels & Vogt’s experiments, the teacher robot points ‘verbally’ by indicating the quadrant in which to look at relatively to itself; however, there is a left right symmetry between the two agents in their definition of quadrants for all objects lying between them; this effect was, however, not considered by the authors and might explain why the agents could never agree on more than a third of the vocabulary (only a third of the same ‘quadrant’ space can be seen by both agents).

However, a disadvantage of our following strategy over a more precise pointing mechanism, is that we are restricted to speaking only about static events or events that last long enough for the learner to eventually see them. In particular, a ‘finger’ pointing mechanism which is more accurate than our ‘body’ pointing mechanism, would be more appropriate in an environment where features cannot be spatially separated (as in the eyes and face example which we described in the introduction). Another disadvantage of our mutual following strategy is that it restricts the number and type of movements that can be taught to only those of motion. This limitation could however be overcome by using a more complex imitative strategy, as for instance the imitative scenario developed in [17] or in [30]. Then, it should be possible to teach concepts relative to movements of many more body actuators and also of more complex sequences of movements (compare with experiments in [3]).

We mentioned in the introduction that previous studies on the development and evolution of communication differ from our work mainly by the absence or the simplification of the physical description of the agents and their environment. However, although these studies used disembodied a-behavioural¹³ agents, their results have sometimes been interpreted in terms of the agents’ behavioural skills. E.g., Oliphant & Batali [43] show that imitation is not a factor enhancing the evolution of communication. This result seems to contradict ours as, in our simulations, the capacity of following and so imitating another robot’s movement is advantageous. The problem lies in our different definition of imitation or imitative behaviour. The agents’ imitative capability in Oliphant & Batali’s work is not a behavioural skill which would involve at minimum 1) a definition of the agent’s actuators, 2) a mechanism responsible for the agent’s observation of the second agent’s actions, and 3) a mapping mechanism which interprets the observed actions into the agent’s own

¹³We refer here to the discussion of the introduction section where we pointed out that in these studies, the agents have no body, no sensors or actuators; they generally occupy not more than a single point in space; they have usually no other actions than that of sending and receiving a signal, the result of their actions being atemporal.

set of actions [14, 17, 38]. The same critic applies to Smajuk & Zanutto's [49] neural model of conditional learning in which imitation is only feeding the observer agent's network with data given by the observed agent's network, without use of any transfer mechanism. In Oliphant & Batali's work imitation is for an agent to match its statistical function of word-meaning associations, a purely cognitive functionality, with the one of a second agent; in other words, one agent learns the language by strictly reproducing each word-pair association as produced by the other agent. The authors conclude that this 'imitative' learning procedure is not advantageous for the development of communication as it only allows for copying a currently existing communicative system and does not allow for its modification and further improvement. These authors' work and their results relative to the influence of imitative behaviour on the development of communication is thus not comparable to ours as we are not approaching the problem from the same angle. In our model imitative behaviour is not the learning mechanism *per se* but acts alongside associative learning to constrain association of relevant data.

We follow a behaviour-oriented (rather than a pure cognitivist) approach to the problem of grounding communication as we investigate the influence of social and behavioural aspects onto the development of communication. Our approach differs from previous studies on the influence of sociality onto the development of communication (e.g. [18, 44, 45, 56]), as we give a complete spatial and temporal description of our agents' behaviour. This allowed us to point out the influence on the success of the learning of environmental factors (see [5]), such as the teaching of objects' featural description and relative dispersion in the environment in relation to two parameters of the learning architecture (short-term memory duration and long-term memory capacity). In addition, we showed the importance of having a spatial and temporal synchronisation between the two communicative agents to allow sharing of a common perceptual context (the experiments reported here and earlier in [5]). The role of context is a key issue in sociality and social understanding. In

particular, studies of (first or second) language acquisition demonstrated the importance of social and verbal context for inferring the correct ‘meaning’ of spoken words or gestural signals ([21]). Our approach to the symbol grounding problem confirmed these ideas and showed that a simple movement imitation strategy is an interesting scenario for the transmission of a language as it is an easy means of getting the agents to share a common context of perceptions. Nevertheless, however similar the two agents’ perceptions could be, they can never be exactly the same and therefore the agent’s learning capacities should be sufficiently complex to compensate for these differences. Not only should the agent be able to associate temporally delayed patterns, but this under a great amount of noisy and spurious data, two performance criteria which have been shown to be satisfied by the learning architecture (see [7]).

This brings us to stress the following points. Whether used for grooming [18] or for transmitting information, communication is an interactive process between the two communicative agents and as such it is a social interaction. Communication does not exist without the physical means of its production and reception. Whatever the level of interpretation chosen for the communicative signals, it is about the physical perception the communicative agents have of their world. As Steven Harnad’s puts it [23]: “even at such abstract cognitive heights, [referring to] the highest level of abstraction of natural language when our interactions with objects are based only on the interactions between names and descriptions, [...] embodiment is never escaped, for the power of names and propositions is completely parasitic on the meanings of those names, and those must all eventually be grounded in the sensorimotor interactions with the kinds of objects they designate, and the sensorimotor invariants on the basis of which the names are assigned”. This requires very complex cognitive processes of segmentation of sensor information, sequence processing and spatial and temporal mapping. A model of the evolution of communication in terms of the agents’ cognitive capabilities should then encapsulate a description of **all** the required cognitive functions. Similarly,

if we consider the progression of brain evolution (e.g. comparing fish to apes and humans), the 'brain' of an individual animal has always evolved as a whole, as a complete organ; although different areas have differentiated and specialised e.g. when comparing 'primitive' with 'advanced' vertebrates (compare reptiles and mammals), different parts of the body plan of a species and therefore different parts of the nervous system have not evolved independently from each other and from the rest of the body. However, the computational models of the evolution of communication have usually ignored most of the so called 'low-level' cognitive functions, such as sensory perception, episodic memory, and focus of attention.

Note, finally, that the evolution of an animal's cognition cannot be separated from the agent's behaviour, as these cognitive processes have been shaped by the constraints created by the agent's interactions with its environment. The body of an animal is a functionally and physiologically well integrated system. Thus, perception, action, communication and cognition are intrinsically interrelated in an embodied system. Our perception of the world is linked to our way of interpreting the world and talking about events in and actions upon this world. Such a system, or behavioral account of cognitive processes, is closely related to the *enactivist* position [35, 54], which describes a living system in terms of the structural coupling of its cellular and molecular components (*autopoiesis*) and of its interaction with its environmental medium: "the behaviour of a living system is not something that the living system does, not something that the medium specifies of its own, the behaviour arises and takes place in the relation living system/medium" [35]. Similar positions are discussed in the area of embodied artificial intelligence (EAI), see [47] for a discussion of implications of embodiment for cognitive theories. As Erich Prem points out, cognitive science and linguistic research has strongly focused on formal aspects and neglected e.g. the notion of time, situatedness and interaction dynamics in animal cognition. 'Getting the interaction dynamics right' is according to Prem a key principle of EAI research. This statement and the position to which it

relates is supported by our experimental results on robot communication grounded in social interaction dynamics. This brings us, then, to suggest that a behaviour-oriented approach [12], as e.g. an enactivist position, might be more appropriate than a pure cognitivist one for describing the cognitive processes involved in addressing the symbol grounding problem.

7 Conclusion and future perspectives

We proposed a teaching scenario based on movement imitation for transmitting a vocabulary among a group of agents and implemented it in simulated experiments. Results showed that social behaviour benefit the agents in two circumstances; 1) agents capable of following one another and so imitating each other's movements develop faster and better a common understanding of the language; 2) the agents benefit from the so developed capability of communicating with one another via a common vocabulary as it speeds up the transmission of information.

An important point to note is that the system we described in this paper, namely the learning architecture and the following strategy, is not restricted to the particular implementation we presented here. For instance, the capacity of the connectionist architecture (DRAMA) allows learning of a vocabulary of bigger size (of order 2 of the memory size) and learning of words combinations (viewed as pattern sequences) [7]. In [3] we report on the implementation of DRAMA in experiments in which a small humanoid robot was taught regular combinations of words which formed 'proto' sentences to describe patterns of actuator sequences and of sensor perceptions. Note also that the architecture is not restricted to be used only with simple mobile robots, as it does not encapsulate constraints on the sensor-actuator mode and could therefore be implemented in robots provided with finer sensor capabilities and actuators with more degrees of freedom. The grounding of symbols using an imitative strategy, as described in the paper, could, for instance, be extended to 'label' behaviours defined as a sequence of sensor-actuator states [33]. Recent work has developed

robotic platforms which can perform such complex set of actions, e.g [11, 20, 32].

Current work by Aude Billard at the Laboratory of MicroInformatics (LAMI) at the Swiss Institute of Technology at Lausanne (EPFL) investigates further some aspects of the simulation studies reported here. Experiments are carried out in which a group of robots learn a vocabulary and the objects' location (similarly to the first and second set of studies reported here), but in a continuous fashion rather than in two separated phases. The robots interact in a biologically inspired environment, where robots and objects appear and disappear ('are born' and 'die') continuously. The study evaluates the influence of parameters, such as the range of communication and the rate of death and birth of robots and objects on the robots' learning performance. Further, the simulation will be implemented in a real set-up of ten Khepera [39] robots.

Our work showed the importance of behavioural capacities alongside cognitive ones for addressing the symbol grounding problem. Behavioural mechanisms which act as external attentional processes are required alongside general cognitive abilities of associativity. We proposed a connectionist architecture in which the behavioural and cognitive processes are produced by the same general ability of spatio-temporal associativity. We used a simple movement imitative strategy to create a spatial and temporal synchronisation between the communicative agents, which provides the required attentional mechanisms. These studies and their results lead us to suggest that a behaviour-oriented approach might be more appropriate than a pure cognitivist one which is dominating in linguistics and cognitive science studies on the cognitive processes involved in grounding communication.

Acknowledgement

Many thanks to anonymous reviewers for their precise comments which helped greatly to improve the writing of this paper. Facilities for the experiments were provided by the Department of

Artificial Intelligence of the University of Edinburgh. Aude Billard was supported by a personal grant from the Swiss National Science Foundation.

References

- [1] Arida T. & Taylor C. E., (1996), 'A Simple Model for the Evolution of Communication', The Fifth Annual Conference on Evolutionary Programming, 1996.
- [2] J. Aitchison (1996), 'The seeds of speech', Cambridge University Press.
- [3] Billard, A. (1999), 'DRAMA, a connectionist architecture for on-line learning and control of autonomous robots: Experiments on learning of a synthetic proto-language with a doll robot'. *Industrial Robot Journal*, 26:1 (due out on 23rd Jan 1999).
- [4] Billard A. (1998), 'DRAMA, a connectionist model for robot learning: Experiments on grounding communication through imitation in autonomous robots', PhD thesis, Dept. of Artificial Intelligence, University of Edinburgh, Submitted on the 30th of Sept. 1998.
- [5] Billard A. & Dautenhahn K. (1998), 'Grounding communication in autonomous robots: an experimental study', in *Robotics and Autonomous Systems*, special issue on 'Quantitative Assessment of Mobile Robot Behaviour', vol. 24:1-2., pp. 71-79, Edited by U. Nehmzow, M. Recce, and D. Bisset.
- [6] Billard A. & Dautenhahn K. (1997), 'Grounding communication in situated, social robots', Proc. of TIMR 97, Towards Intelligent Mobile Robots Conference, (Technical Report Series UMCS-97-9-1), Dept of Computer Science, Manchester University.
- [7] Billard A. & Hayes G. (1999), 'DRAMA, a connectionist architecture for control and learning in autonomous robots'. *Adaptive Behaviour* journal, vol. 7:1.
- [8] Billard A. & Hayes G. (1997), 'Transmitting Communication Skills through Imitation in Autonomous robots'. Proceedings of EWLR97, Sixth European Workshop on Learning Robots, Brighton, July 97. *Lecture Notes in Artificial Intelligence*, A. Birk & J. Demiris (editors), Springer-Verlag.

- [9] Bandura, A. Ross, D. & Ross, S.A. (1961), 'Transmission of aggression through imitation of aggressive models', *Journal of Abnormal and Social Psychology*, vol. 63, pp. 575-82.
- [10] Brooks R.A. (1991), 'Intelligence without reason', Proc. of IJCAI-91, International Joint Conference on Artificial Intelligence, pp.659-595, Sydney, Australia.
- [11] Brooks R.A. (1996), 'Behavior-based humanoid robotics', in Proc. 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 96, pp.1-8.
- [12] Dautenhahn K., (1997), 'I could be you - the phenomenological dimension of social understanding', *Cybernetics and Systems Journal*, 28(5), pp. 417-453, Special Issue on Epistemological Aspects of Embodied AI.
- [13] Dautenhahn K., (1997), 'Grounding Agent Sociality: The Social World is its Own Best Model', From Agent Theory to Agent Implementation (organiser Joerg Mueller and Paolo Petta) symposium. Published in Proceedings of 14th European Meeting on Cybernetics and Systems Research, Robert Trappl (editor), pp. 779-784.
- [14] Dautenhahn K., (1995), 'Getting to know each other – artificial social intelligence for autonomous robots', *Robotics and Autonomous Systems* 16, pp 333-356.
- [15] J. M. Davis, (1973), 'Imitation: a review and critique', Perspectives in Ethology, vol. 1, P. Klopfer eds., pp. 43-72, Plenum, New York.
- [16] Deacon, T.W. (1992), 'Brain-Language Coevolution', in *The Evolution of Human Language*, SFI Studies in the Sciences of Complexity, pp. 49-83, Proc. Vol. X, Eds. J.A. Hawkins and M. Gell-Mann, Addison-Wesley, 1992.
- [17] Demiris J. & Hayes G., (1996), 'A Robot Controller Using Learning by Imitation', Proceedings of the 2nd International Symposium on Intelligent Robotic Systems, Grenoble, France, July 1994.
- [18] Dunbar R.I.M. (1993), 'Coevolution of neocortical size, group size and language in humans'. *Behavioural and Brain Sciences*, 16:681-735.
- [19] Frisch, K. von (1967), *The dance Language and Orientation of Bees*, Bellknap, Cambridge, MA.

- [20] Fujita M. & Kageyama K. (1997), 'An Open Architecture for Robot Entertainment', in Proc. of the First International Conference on Autonomous Agents, Marina del Rey, CA, USA, W. Lewis Johnson (eds), pp. 435-442.
- [21] Harley, T.A. (1994), 'The Psychology of Language: from data to theory', chapt. 2-3, Erlbaum (UK) Taylor & Francis Publishers, c1995.
- [22] Harnad, S. (1990), 'The symbol grounding problem', *Physica D* 42:335-346.
- [23] Harnad, S. (1998), 'Beyond Object Constancy'. Institution of Electrical Engineers (IEE) Seminar on "Self-Learning Robots II: Bio-Robotics" (Digest 98/248: 2/1-2/3).
- [24] Henis, E. & Levinson, S., 'Language as part of sensorimotor behavior', *Embodied Language and Action*, 1995 AAAI Fall Symposium (Tech. Report FS-95-05). AAAI Press, Menlo Park, CA, USA; 1995; pp. 48-53.
- [25] C. M. Heyes & B. G. Galef (1996), 'Social Learning in Animals: The Roots of Culture', Academic Press.
- [26] Kawamura, S. (1963), 'The process of sub-culture propagation among Japanese Macaques'. In *Primates Social Behaviour*, ed. C.H. Southwick, Van Nostrand, New York, pp.82-90.
- [27] Kim, J. H. et al., (1997), 'Development of a multi-agent system for robot soccer game', Proceedings of IEEE International Conference on Robotics and Automation. IEEE, New York, NY, USA; 1997; 4 vol. pp. 626-31.
- [28] Kirby S. & Hurford J. (1997), 'The evolution of incremental learning', Proceedings of GALA'97, Groningen Assembly on Language Acquisition, Edinburgh, 4-6 April 1997.
- [29] Klingspor V., Demiris J., and Kaiser M., (1997), 'Human-Robot-Communication and Machine Learning', *Applied Artificial Intelligence Journal* 11, pp. 719-746, 1997.
- [30] Y. Kuniyoshi, M.Inaba and H. Inoue, (1994), 'Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance', IEEE Transactions on Robotics and Automation, vol.10, no.6, pp.799-822, Dec. 1994.

- [31] Mataric M., (1998), 'Using Communication to Reduce Locality in Distributed Multi-Agent Learning', *Journal of Experimental and Theoretical Artificial Intelligence*, special issue on Learning in DAI Systems, Gerhard Weiss, ed., 10(3), Jul-Sep, 1998, 357-369.
- [32] Demiris J. & Mataric M.J. (1998), 'Perceptuo-Motor Primitives in Imitation', in Working Notes, Autonomous Agents '98 Workshop on "Agents in Interaction - Acquiring Competences Through Imitation", Minneapolis/St. Paul.
- [33] Mataric M., (1998), Personal communication to Aude Billard.
- [34] MacLennan B. & Gordon M. Burghardt, (1993), 'Synthetic Ethology and the Evolution of Cooperative Communication', *Adaptive Behavior*, Vol. 2, No. 2, Fall 1993, pp. 161-187.
- [35] Maturana, H. (1988), 'Ontology of Observing: The Biological Foundations Of Self Consciousness And the Physical Domain Of Existence'. In: Conference Workbook: Texts in Cybernetics, *American Society For Cybernetics Conference*, Felton, CA. 18-23 October 98.
- [36] McFarland D. (1993), 'Animal Behaviour', Longman Publisher, 1993.
- [37] Meltzoff A, (1996), 'The human infant as imitative generalist: a 20-year progress report on infant imitation with implications for comparative psychology', in *Social Learning in Animals: The Roots of Culture*, Heyes C.M. and Galef B.G. (eds), New York: Academic Press.
- [38] Meltzoff A, (1996), 'Towards a developmental cognitive science. The implication of cross-modal matching and imitation for the development of representation and memory in infancy', in *The development and Neural Basis of Higher cognitive Functions*, Adele Diamond (ed.). Annals of the N. Y. Acad. of Sci., **608**.
- [39] F. Mondada, E. Franzi & P. Ienne (1993), 'Mobile Robot Miniaturisation: a Tool for Investigation in Control Algorithms', Proceedings of ISER'93, Kyoto, Japan.
- [40] Moukas, A. & Hayes G. (1996), 'Synthetic Robotic Language Acquisition by Observation', in Proc. of the Fourth Conference on Simulation of Adaptive Behaviour, Cape Cod, USA, September 1996.

- [41] C. Nehaniv & K. Dautenhahn (1998), 'Mapping between Dissimilar Bodies: Affordances and the Algebraic Foundations of Imitation', Proceedings European Workshop on Learning Robots 1998 (EWLR-7), Edinburgh, 20 July 1998, John Demiris and Andreas Birk (eds), pp. 64-72.
- [42] Nottebohm, F. (1976), 'Vocal tract and brain: a search for evolutionary bottlenecks', *Ann. N.Y. Acad. Sci.* **280**, pp. 643-649.
- [43] Oliphant, M. & Batali, J. (1997), 'Learning and the emergence of coordinated communication'. Center for Research on Language Newsletter 11(1).
- [44] Parisi, D. & Floreano, D. (1992), 'Prediction and imitation of linguistic sounds by neural networks'. In A. Paoloni (Ed.), Proceedings of the 1st Workshop on Neural Networks and Speech Processing. Roma: Fondazione Bordoni, 1992, 50-61.
- [45] Di Paolo E. (1998), 'Assessing the role of social development in the evolution of cooperation'. From Animals to Animats 5, Proceedings of the Fifth International Conference of The Society for Adaptive Behavior (SAB98)", Pfeifer, R. and Blumberg, B. and Meyer, J.-A. and Wilson, S.W. (eds), pp. 453-459, MIT Press.
- [46] Pinker S. (1994), 'The Language Instinct', Penguin books.
- [47] Prem E. (1998), 'The implications of embodiment for cognitive theories. A systems view', Proceedings of the 14th European Meeting on Cybernetics and Systems Research, EMCSR'98, pp.687-692.
- [48] Pulvermuller F., (1998) 'Words in the Brain's Language', *Behavioural and Brain Science Journal*, Cambridge University Press.
- [49] Smajuk, N.A. & Zanutto B.S., (1997), 'Escape, Avoidance, and Imitation: A Neural Network approach', *Adaptive Behaviour*, Vol.6:1, pp.63-129.
- [50] Steels L. & Vogt P. (1997), 'Grounding adaptive language games in robotic agents', Proceedings of the Fourth European Conference on Artificial Life, pp. 473-484, Brighton, U.K., July 97.
- [51] Steels, L., (1997) 'Language Learning and Language Contact', In: Daelemans, W. (ed.). Proceedings of the workshop on Empirical Approaches to Language Acquisition. Prague.

- [52] Tatsuno, J. et al., 1996, 'Human friendly teaching for industrial robots', Proc. of 5th IEEE International Workshop on Robot and Human Communication RO-MAN'96 Tsukuba. IEEE, New York, NY, USA; 1996; pp. 549–550.
- [53] W. H. Thorpe (1963), 'Learning and instinct in animals', London: Methuen.
- [54] Varela F., Thompson E. and Rosch E. (1993), 'The Embodied Mind', *American Journal of Psychology*, 106, p.121-6.
- [55] Whiten A. and Ham R. (1992), 'On the Nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research in Advances in the Study of Behaviour', 21, p. 239-83.
- [56] Worden R. (1998), 'The Evolution of Language from Social Intelligence', To appear in *The Evolution of Phonology and Syntax*, Hurford, Studdert-Kennedy & Knight (eds) Cambridge University Press.
- [57] Holly Y. & Stein L. A., (1993), 'An adaptive communication protocol for cooperating mobile robots', in *From Animals to Animats 2: Proceedings of the Second International Conference on the Simulation of Adaptive Behavior*, edited by J.-A. Meyer, H.L. Roitblat and S.W. Wilson. The MIT Press/Bradford Books, 1993, pp. 478-485.