

Transmitting Communication Skills Through Imitation in Autonomous Robots*

Aude Billard and Gillian Hayes

University of Edinburgh, Dept. of Artificial Intelligence
5 Forrest Hill, EH1 2QL Edinburgh
E-mail: audeb@dai.ed.ac.uk

Abstract. Sharing a common context of perception is a prerequisite in order for several agents to develop a common understanding of a language. We propose a method, based on a simple imitative strategy, for transmitting a vocabulary from a teacher agent to a learner agent. A learner robot follows and thus implicitly imitates the movements of a teacher robot. While bounded by mutual following, learner and teacher agents are set in a position from which they share a common set of perceptions, that they can then correlate with an arbitrary set of signals, a vocabulary. The teacher robot leads the learner robot through a series of situations in which it teaches it a vocabulary to describe its observations. The learner robot *grounds* the teacher's 'words' by associating them with its own perceptions. Learning is provided by a dynamical recurrent associative memory, an Artificial Neural Network architecture. The system is studied through simulations and physical experiments where the vocabulary concerns the agents' movements and orientation. Experiments are successful, learning is fast and stable in the face of a significant amount of experimental noise. This work suggests then 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 the same physical context.

1 Introduction

Communication is a social skill and as such would be desirable for autonomous agents that are expected to interact with other agents or humans ([5], [6]). In particular, collaborative tasks would be easier if the agents could inform each other of what they are doing ([8]). For communication to be successful, the agents must share a common interpretation of the language they use. Whilst a fixed protocol could be designed to this end, we expect the system to be more robust and adaptable if the agents would develop their understanding of the language through a learning process. This work studies a method for transmitting a language from a teacher agent to a learner agent.

Recent simulation studies of language acquisition tend to address essentially the question of the origin of language and how it can evolve among a popu-

* In lecture Notes in Artificial Intelligence, A. Birk & J. Demiris publishers, Springer-Verlag editors, 1998.

lation (e.g. [9], [12], [1], [10]). In our approach, we focus on the learning and understanding of a language at the level of the individual rather than at the level of the society. We propose a method, in terms of the agent’s behavioural capacities, whereby a meaning can be associated with an abstract signal, i.e. a ‘word’. To our knowledge, this physical approach to the problem of the language understanding has very seldom been questioned nor investigated.

Sharing a common context, and so a common set of perceptions, is a prerequisite for the development of a common understanding of a language by different agents. In order for the learner agent to successfully interpret the teacher’s words, it has to be able to make the same observations to which the teacher is referring. Sharing the same spatial and temporal context of perceptions is often not enough and attentional mechanisms must be used to restrict the set of perceptions to a relevant subset shared by both agents. In related robotics studies, this reduction of the information was enabled by pointing to the referred-to objects [11] or by using global knowledge of the learner’s perception [14]. These methods are costly in terms of pre-programming and not easily adaptable to different agents or tasks. In this work, we use a simple imitative method, namely mutual following of teacher and learner agents. While bounded by mutual following, learner and teacher agents face the same direction and thus share a similar if not identical (due to different sensor sensitivity) set of external perceptions. This set of perceptions forms a subset of all the possible observations the agents could make from the same location. In addition, while following the teacher, the learner agent replicates the teacher’s movements and thus shares similar internal perception of movement and energy consumption, while restricting itself to a subset of possible actions.

Learning to communicate means in our experiments that one agent, the learner tries to achieve a similar ‘interpretation’ of a situation to that of a second agent, the teacher. While followed by the learner, the teacher robot leads the learner robot through a series of situations in which it teaches it a vocabulary to describe its observations. In this context, the ‘language’ consists of a simple vocabulary, where the ‘words’ are (radio) signals that label specific configurations of sensor-actuator state. The teacher sends distinct signals to discriminate between different movements and orientations. The learner robot *grounds* the teacher’s ‘words’ by associating them with its own perceptions, i.e. sensor measurements and motor states.

Other works used movement imitation for teaching a robot about movement related tasks ([4], [7]). Here, we propose to use this method for teaching a language. Movement imitation, as a means to learning a language, is interesting compared to a conventional supervised learning method, as used e.g. in [14], for at least two reasons: Firstly, it enables the learning agent to share a perception of the world similar to the teacher’s. Secondly, it enables the teacher to direct the learner’s learning. Teaching can concern both proprio and exteroceptive perceptions. By leading the learner to specific locations, the teacher can teach it a vocabulary about objects or situations. By directing the learner’s movements, the teacher can teach it a vocabulary about its internal states (position, move-

ments or energy consumption). In the following strategy we propose, movement imitation is only implicit and not the result of a particular cognitive mechanism.

The rest of this paper is divided as follows. Section 2 describes the controller architecture of teacher and learner agents. Section 3 presents the experimental procedure of the simulation studies and physical experiments. Section 4 reports and discusses the results of these experiments, on which we conclude this paper in section 5.

2 Controller Architecture

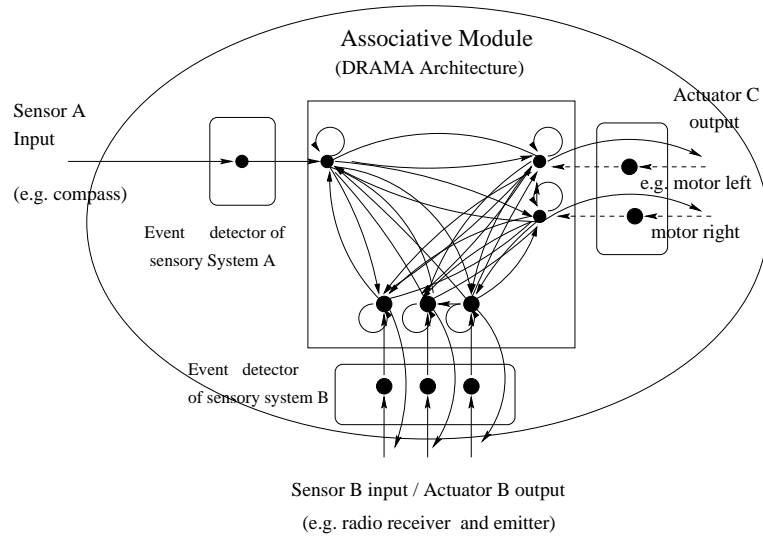


Fig. 1. Schematic representation of the control system of the learner and teacher agents with three sensor-actuator systems, e.g. compass, radio and motors. Learning is dynamical and performed continuously during the experiment. At each processing cycle, the learner robot's sensor-actuator information is compared to the measurements collected the previous cycle. Whenever a variation in the measurement is detected (by the *event recognition* module), the new measurement is presented to the associative architecture (DRAMA) to be associated with all simultaneous and previously recorded measurements in the same or other sensor-actuator system. DRAMA is a fully recurrent Neural Network. Its recurrent structure provides a short term memory of the measurements. Long term memory is obtained by updating the internal connections following Hebbian rules.

Learning to communicate is implemented in our experiments as part of a general multiple sensory association process, produced by a Dynamic Recurrent Associative Memory Architecture (DRAMA), previously described in [2]. We

here briefly summarise its main properties, together with a description of the global control system of the agent:

- 1) It combines two connectionist model types, associative memory (Willshaw type) and recurrent networks, in one single network structure for achieving learning of temporally coincident patterns.
- 2) The network is fully connected with self-recurrent connections on each node.
- 3) Each connection in the network is associated with a time parameter and a confidence factor. The time parameters represent the temporal correlation between events while the confidence factors keep a memory of the frequency of a pattern occurrence.
- 4) Time parameters and confidence factors are updated following Hebbian rules, providing an associative type of learning. The closer the temporal occurrence of the events, the stronger the correlation.
- 5) Encoding of inputs and outputs is binary.
- 6) The self-recurrent connections on the units provide a short-term memory of one event per unit, whose maximal duration is fixed.
- 7) Long term memory of sequential and simultaneous correlations is provided by updating the connections between different units.
- 8) Data retrieving depends on the value of the time parameters and confidence factors associated with the connections. The time parameters play the role of a first threshold on the time of occurrence of the event; only the events 'older' than the memorised temporal correlation are considered. The unit output is activated when the sum of confidence factors of the different active inputs passes a minimal threshold.
- 9) The structure of the network is dynamically updated, at each processing cycle. Learning is performed continuously during the experiments, that is we do not distinguish between learning and testing phases. Testing, i.e. retrieving the output occurs at each cycle in order to direct the robot's actions (moving and emission of radio signal).

Figure 1 presents a schematic description of the learning architecture. The learning process used in the experiments is described in the following section.

2.1 Learning Process

Learning is dynamical and performed continuously during the experiment. At each processing cycle, the learner robot's sensor-actuator information is compared to the measurements collected the previous cycle. Whenever a variation in the measurement of one sensor or actuator system is detected by the corresponding *event recognition* module, the new measurement is presented to the associative architecture (DRAMA) to be associated with all simultaneous and previously recorded measurements in the same or other sensor-actuator system. Sensor data are collected as bit-strings (different length for each sensor); in the experiments we considered a one-bit variation in the measurement as an event. Learning occurs only when an *event* has been detected. Events are memorised in the self-recurrent connections of the network input units for a fixed period of time (100 cycles, i.e. about 2.5 minutes of real time on the robot)² or until

² This was preferred to an infinite event memorisation, where e.g. the memory of the event is erased only when a new event is detected, in order to avoid misleading associations. The experimental set-up was done so that the teaching of a word oc-

a new event occurs. In figure 1, we present a schematic view of the learning architecture. The sensor-actuator systems A, B and C in the figure could be interpreted, for instance, as the motors, compass and radio systems respectively. The learning process works then as follow: If, at time t , one event is detected by the sensory system A and one by the sensory system B, they will be associated with each other. If then, at time $t + n$ ($n < \text{memory length}$), a new event is detected by the system C, it will be associated with the two previously recorded events in systems A and B. Association consists of 1) increase the confidence factors of connections between active units (following an Hebbian rule) by a factor inversely proportional to the temporal delay between activation of each unit 2) update the time parameters to record this temporal delay (mean value over all examples).

2.2 Experimental implementation

Figure 2 shows a schematic description of the control system in the present implementation. In the experiments we report here, the architecture is used for both learner and teacher agent. Basic behaviours of wandering, obstacle avoidance and following are predefined by setting in advance the values of the connection parameters between infra-red sensor, bumpers, light sensors and the agents' motors. Figure 2 shows how obstacle avoidance behaviour with the frontal infra-red sensor was predefined: connections from the frontal infra-red sensor to the motor left are set with predefined values, resulting in the robot turning to the right (activation of the left motor only) when detecting an obstacle by infra-red reflection. Mutual following of the two agents is defined as a phototaxis behaviour performed by both agents with light detectors ([2]) (each agent carrying the relevant sensors on its front (learner) or on its back (teacher)), i.e. asymmetric connections are set between these left-right sensors and right-left motors in order to result in phototaxis behaviour. The robot's behaviour is then determined by retrieving the motor outputs at each cycle given the current sensor input. The motor activity is determined by the activity of the units corresponding to the motor system. The teacher robot's knowledge of the vocabulary is predefined by setting the value of the connections between relevant sensor information and the radio emitter (actuator). Similarly to what is done for the motor activation, the teacher 'speaking' (emission of radio signals) is directed by retrieving the radio output given the current sensor-actuator state. The teacher 'speaks' only when it sees the learner. Finally, learning occurs dynamically as a consequence of the continuous update of the two connection parameters to account for the variations in the sensor-actuator states. Figure 2 shows an example of connection update after training. Training of the network results in associating different compass measurements with different radio signals: 4 signals (external teaching) are associated with measurement of the 4 quadrants (North, South, East, West). Self-organisation of the network results also in associating different compass

curred at the latest one or two minutes following the event which the word should be associated with.

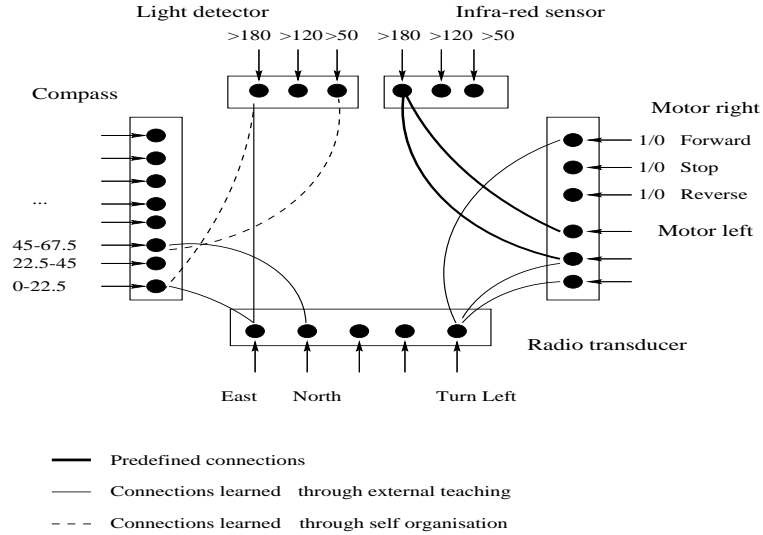


Fig. 2. Schematic representation of the processing of the sensor-actuator information through the DRAMA architecture as implemented in the experiments. The motor activity is determined by retrieving the activity of the units corresponding to the motor system. The motor activity is encoded in a 3-bit string. Bit 1 determines the state of activity of the motor (*active/not* = 1/0), bit 2 encodes for the direction of activity (*forward/reverse* = 1/0) while the third bit determines the speed of activity (*full/half* = 1/0). Basic obstacle avoidance behaviour is predetermined by setting the connection (straight lines) between a high infra-red value detected in infra-red detector I1 and the left motor. As a result, the robot turns to the right in front of an obstacle by activating its left motor forward and keeping the right one inactive. Training of the network results in associating different compass measurements (East, North) with different radio signals (external teaching provided by the teacher agent) and with different light measurements (when facing the north and so the windows, the robot measures an increase in the global lighting).

measurements with different light measurements (we observed in the experiments that when facing the north and so the windows, the robot measures an increase in the global lighting).

To summarise, at each processing cycle, given the current sensor input, the actuators' (motors, radio emitter) outputs are calculated. Then, the current sensor and actuator information is compared to the information measured in the previous cycle (in the event detectors). 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 connections, i.e. a single network, are used for learning, control (motor activity) and signal transmission (radio transmission). In addition, the same learning is used for generating the symbolic associations (radio-sensor) and other sensor-sensor and sensor-actuator associations, and the

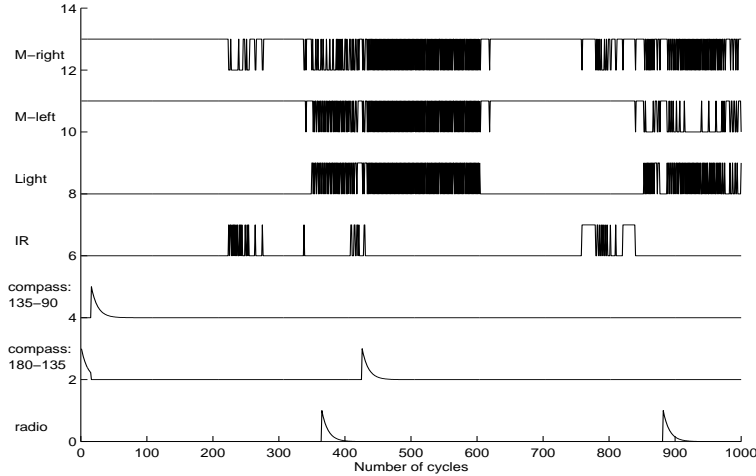


Fig. 3. Variation of activity of the units corresponding to the left and right motor, the compass, light and radio sensors, during 1000 processing cycles. We observe that activation of the infra-red detector unit (at times 210, 350, 430 and 780) produces an immediate deactivation of the right motor. The robot turns to the left when it faces an obstacle, as previously defined by the connection weights. Consequently, after a long rotation a new value is measured by the compass (at time 450). Light detection at time 380, i.e. detecting the other robot, produces a deactivation of left and right motors alternately (time 380-600). The robot aligns behind the other robot.

same retrieving algorithm then use these associations for both controlling the robot's movements and communicative behaviour. Figure 3 shows the variation of activity of the units corresponding to the left and right motors, the compass, radio and light sensors, during 1000 processing cycles³. We observe that activation of the infra-red detector unit (at times 210, 350, 430 and 780) produces an immediate deactivation of the right motor. The robot turns to the left when it faces an obstacle, as previously defined by the connection weights. Consequently, after a long rotation a new value is measured by the compass (at time 450). Light detection, corresponding to detection of the second robot, produces a deactivation of left and right motors alternately (time 380-600). The robot aligns behind the other robot. We also see the effect of the decrease of activation along the recurrent connections of the radio and compass units. The motor and infra-red units do not decrease because they are constantly activated by the new input. When the radio unit is activated, at cycles 380 and 890, it is associated with the following activation of the compass unit, just before its extinction.

³ Each sensor information is in fact given by more than one unit; in figure 3 we represent the maximal activation of the units corresponding to this system

3 The Experiments

In this section, we report on experiments carried out both in simulated and physical environments. Teachings concern the robot's movements (motor states) and orientation (compass measurements). Figure 4 presents a schema of the teaching scenario. The learner robot follows the teacher robot that wanders randomly in the arena. Whenever the teacher detects the learner behind it, it stops and sends a radio signal composed of the 'word' *stop* and a word for the *orientation* (*South*, *North*, *East*, *West*) given by its compass. Then it starts to move again and sends consecutively (after a delay of 200 cycles) a teaching signal containing the word *move*. The vocabulary consists at that point of 6 words: [*stop*, *move*, *South*, *North*, *East*, *West*]. In a second phase of the experiment, instead of teaching the compass direction three radio signals are sent to distinguish between three movements *stop*, *turn left* and *turn right*. Each signal is sent separately whenever the teacher performs the particular movement, as part of its wandering, while detecting the presence of the learner behind it. At the end of the experiment, the total vocabulary learned by the learner agent consists then of 8 words: [*stop*, *move*, *turn left*, *turn right*, *South*, *North*, *East*, *West*].

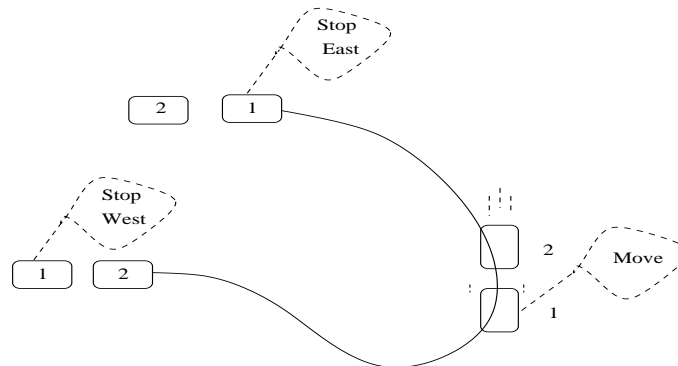


Fig. 4. The Teaching Scenario. A teacher robot (1) teaches a learner robot (2) a vocabulary to describe its movements and location. The learner, while following the teacher, associates the teacher's words with its own observations.

Learning consists of multiple associations between all the sensor and motor systems of the robot. More specifically, learning of the language amounts to associate different radio signals (bit-strings) with distinct sensory measurements and motor states. Association are made between all the sensors (light, infra-red, bumpers, compass and radio) and motor systems of the robot, that is associations are not only made between sensor and motor systems but also between the sensor systems themselves, e.g. the radio/compass and radio/light associations.

Learning is achieved once the confidence factors associated with the relevant sensory measurements show a strong correlation, easily distinguishable from the ones with other non relevant sensory measurements.

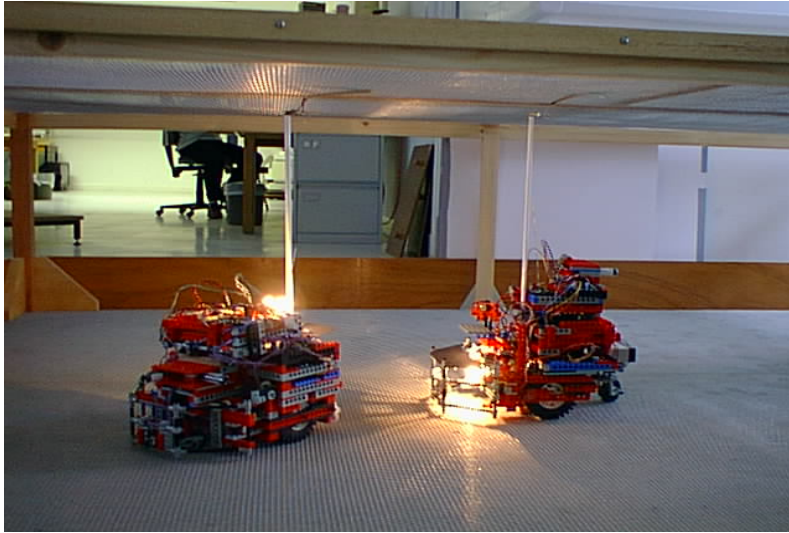


Fig. 5. Left Teacher (in front) and learner robots in their Dodgem environment.

Two autonomous LEGO robots (teacher and learner) are used for the experiments. Each robot is equipped with one frontal infra-red sensor and bumpers to avoid the obstacles (see figure 5). They also have two light detectors in the front (learner) or in the back (teacher) to enable them to follow each other. The range and sensitivity of the sensors are given in table 1. They have a radio transceiver, the means of transmission of communication signals. Each signal is encoded in one byte with only 1-bit activated (e.g. 'North' = (10000000), 'Stop' = (00000001)). That is, in our agents' 'language' 8-bit = 1 byte 'words'. The arena consists of a rectangular cage of 2.5m by 2m by 0.5m, in which the robots are continuously recharged similarly to the system used in the 'Dodgem' (bumping cars) game. Roof and bottom of the arena are electrified, creating a potential difference of 10V between them. The robots carry a long stick touching both ends of the cage from which they receive the current to power their battery and light bulb. We estimate that about 10 to 20% of the sensor measurements are noisy: 80% of the radio transmissions are correctly received (i.e. if a signal is received, then it is perfect; the noise corresponds to the case where an emitted signal has not been received), while the quadrants given by the compass are correctly detected in about 90% of the cases ('zigzagging', side movements, of the robots while following each other, and undesirable magnetic effects produced by

Sensor type	Sensitivity range	
	Physical Exp.	Simulations
Bumpers	0cm (touch contact)	None
Infra-red	cone of 15 ⁰ and 30-40cm	cone of 180 ⁰ and 45cm
Light	cone of 180 ⁰ and 40-50cm	cone of 180 ⁰ and 45cm
Compass	8 quadrants	8 bit-string (22.5 ^o)
Radio	30m	whole arena

Table 1. Table with the robots sensor sensitivities in physical and simulated environment.

the motors and the powering of the cage).

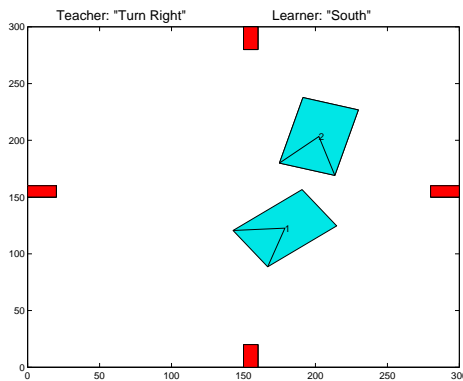


Fig. 6. Simulation studies are carried out in a 2-D simulator, in which two robots can interact in a rectangular arena. In the figure, we see the teacher robot (No 1) followed by the learner robot (2), while the teacher begins to turn to the right and sends the corresponding radio signal. The learner which follows it is still thinking 'move' as it has not begun to turn. The boxes on the side of the arena stand for the obstacles of the real arena.

Simulation studies are carried out in a 2-D simulator, in which two robots can interact in a given environment, here a rectangular arena measuring 300 by 300 units (see figure 6). There are 4 obstacles on the sides of the arena, represented by dark rectangles. The robots are represented as rectangles of 30 by 20 units, a triangle determining the front. The simulated robots are provided with the same sensor systems as the real robots (apart from not having bumpers): light detectors for mutual recognition, infra red sensor for obstacles avoidance, a radio transducer as the means of communication and a compass to measure bearings to the request of 22.5 degrees. 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 sensitivities of the sensors is given in table 1.

The robots' behaviour is calculated separately by the same routines as used in the physical robots. Code is written in C and is processed serially. At each cycle, each robot's routine measures the current sensor-actuator state of the robot. The robots' sensor perceptions are simulated by defining a specific measurement routine for each sensor. In order to reproduce reality more faithfully, the following behaviour is made imperfect. Following is mutual, each agent aligning towards the other one on the basis of its light detectors' measurement (each robot carries a bright light). Similarly to what happens in reality, an agent is able to determine the position of the other agent with respect to its own with a precision of 90 degrees. Therefore, the alignment of the two robots is imprecise. In addition, randomness is introduced in the calculation of the robot's movements, such as to represent the experimental imprecision.

We use a simulator in order to reproduce the physical experiments and study their successes and failures in a more reliable environment. The main advantage of simulations over physical experiments is that they are repeatable, faster (simulating a 1 hour experiment takes about 5 minutes) and do not suffer unexpected hardware breakdown. The disadvantages in terms of model faithfulness are, of course, well known. For a more complete discussion of this see [13]. Here simulation studies are carried out to show the stability and success of the learning in the particular set-up proposed, which is further demonstrated in the physical experiments.

4 Results

A set of 100 runs, each run simulating 10000 processing cycles of the robots (about 4 hours of real experiments) were carried out to test the speed and stability of the learning. The success or failure of learning in the experiment is determined by looking at the values of the connections' confidence factors between the radio sensor system and the rest of the sensor and actuator systems of the agent (light, infra-red, compass, motors). The experiment is said to be successful when only the connections between the correct radio unit, standing for the particular word and its associated sensor (compass) - actuator (motors) unit combination have a confidence factor value greater than the maximal one for all connections leading to these units. In other words, learning occurs if the number of correct associations (correct matching between radio signal reception and sensor-actuator measurement) is greater than the number of incorrect ones, i.e. if the noise (incorrect examples) is below 50% of the total number of examples in a run.

In a Willshaw type of associative memory as used here, there is no notion of convergence as in many other kinds of neural network, since weight update is only of one time step. The success of the learning can be measured at each time step as

the ratio between connection parameters. At each presentation of a new example (new teaching), the connections' confidence factors are updated. The study of the variation of the confidence factors values during the experiment informs us about the variation of the percentage of noise and consequently of the stability of the learning. Noise, in our case an incorrect example, is produced when teacher and learner are not perfectly aligned and therefore measure different compass orientations or perform different movements. Figure 7 left shows the mean variation of the confidence factor (cf) values along a run. The y axis represents the ratio between the cf value associated with the correct word and the maximal value of cf attached to all words for a given sensor measurement, i.e. $\text{cf}(\text{correct correlations}) / (\text{cf}(\text{correct correlations}) + \max(\text{cf}(\text{incorrect correlations}))$). The data are the mean value over all words and over all runs. A ratio greater than 0.5 means that the particular sensor measurement has been associated with the correct word more often than with other words. Figure 7 left shows that learning is stable and steadily increasing while more teachings are given. Once the cf value has crossed over the threshold of 0.5, i.e. the word has been learned, it stays over it with a standard deviation of 0.2 (see table 2).

	Simulations	Physic. exp
	Mean & Std	Mean & Std
1	0.62 ± 0.23	0.71 ± 0.18
2	0.80 ± 0.19	0.96 ± 0.10
3	21 ± 31	70 ± 16

Table 2. Comparison between results of simulations and physical experiments. From top to bottom: 1) Ratio between the confidence factor value of the correct word and the maximal value of confidence factor attached to all words for a given sensor measurement (mean value over all words and all runs). 2) Ratio between the number of words learned at the end of the run and the total vocabulary (mean value over all runs). 3) Number of teachings per run (mean value over the runs).

The slow increase of the slope (figure 7 left) at its beginning is due to the fact that we superpose the learning curves of 100 different runs and these do not begin at the same point. Each run starts with the agents in a new random position. The time before the first meeting and teaching varies then for each run. The slope of each single curve is however very steep. In most cases, the first teaching would be a correct example. The curve would then start from the maximal value 1, decreasing slightly afterwards under the effect of noise. This is illustrated by the very steep beginning of the curve after the first teaching.

A set of five physical experiments is reported here. Each physical experiment consisted of two teaching phases, in the first phase six words were taught (*stop, move, South, North, East, West*) and in the second 3 words (*stop, turn left, turn right*). Each experiment lasted for about 1 hour 30 minutes (45 minutes for each

phase) and was stopped when about 90 teachings had been done (the exact number of teaching events is imprecise because the radio transmission was imperfect and it was difficult to detect whether the radio signal had been received by the robot.). Table 2 compares results of simulations and physical experiments for learning of the first phase. Results are qualitatively similar. The mean values of confidence factor for all words at the end of each run are respectively 0.62 and 0.71. The standard deviation values represent the fluctuation of the cf values along the run. The qualitative similarity between physical and simulated results shows that the simulations studies gave a good account of the percentage of noise occurring in the physical experiments. Lines 2 and 3 of the table show respectively the ratio to total vocabulary of words that have been learned at the end of the run and the total number of teachings in a run (average values for all runs)⁴. Surprisingly, the results of the physical experiments are slightly better than the simulations. A reason may be the small numbers of the physical experiments. But the main reason is surely that learning in the physical experiments was helped by placing the robots one behind the other one from the beginning⁵. The chances of the two robots meeting and then teaching were therefore greatly increased. This may explain why the number of teachings and the number of learned words per run is higher in the physical experiments than in the simulations.

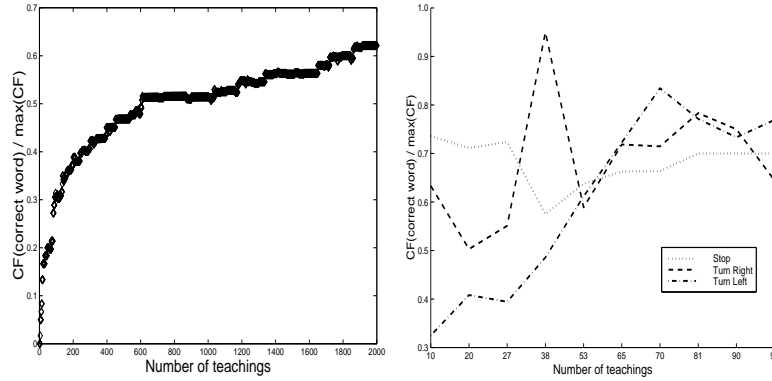


Fig. 7. Variation of confidence factor values along a run in simulation studies (left) and physical experiments (right). Y axis represents the ratio between the cf value of the correct word and the maximal value of cf attached to all words for a given sensor measurement. Data are mean value over all words and over all runs. A ratio greater than 0.5 means that the particular sensor measurement has been associated with the correct word more often than with other words.

⁴ The large value of standard deviation in line 3 is due to outlier values of 300 (teachings per run) in the data.

⁵ This was done in order to shorten the time of the experiments, the hardware being unlikely to stand more than one hour of continuous running of the robots.

In figure 7 right, we show the variation of the confidence factor values in the second phase of the experiments. We observe that the three words are correctly learned after 30 teachings, which corresponds to about 15 to 20 minutes of experiment, which is faster than in the simulations (figure 7 left). Again, this is probably due to our ‘helping’ the physical robots to align faster. However, the fluctuation of the values are more pronounced in the physical experiments, which means that the proportion of noise varies more in the physical world. The noise corresponds to the cases where the radio signals are associated with incorrect motor states or compass measurements. Because the movements of the physical robots are less smooth, incorrect alignment and imprecise following of the two robots is more pronounced and more frequent, which results in the robots facing and measuring different directions and producing different movements (e.g. turning in opposite directions). Note also that learning of the word *stop* is better and occurs already from the beginning, due to the fact that this word had already been taught during the first phase of the experiment (no erase of the memory between the two phases).

In summary, the experiments showed that 1) a number of words are learned and correctly associated; 2) the vocabulary is learned although the experimental conditions create incorrect teachings, which demonstrates the robustness of the learning architecture in the face of 20 percent of noise; 3) the learning is stable once a significant number of relevant correlations have been recorded. One may now question whether these experiments, because of their apparent simplicity (the vocabulary has only eight words), are a demonstration of the efficacy of the learning method. Although the sensory system and consequently the vocabulary are quite simple, the learning task is actually rather difficult. Because of their very limited sensory capabilities, the robots’ mutual recognition and following is very imprecise. Alignment of the two robots is seldom perfect. Consequently, the chances of their sensor measurements being identical (in particular their compass measurements) are quite slim and misleading associations may easily occur. On the contrary, if we had used more sophisticated sensors, giving a finer and thus more informative description of the environment, the number of taught words could have been greatly increased. On the other hand, however, we would have simplified the task of imitation by giving the robots more information on which to rely for adjusting their positions one towards the other. This leads us to think that the levels of difficulty of the learning task in the two cases are in some sense comparable. The significant data have to be extracted from a considerable amount of noise that is produced, in the first case, by misleading teachings due to an imprecise imitation and, in the second case, by a more rich source of information. Therefore, it seems reasonable to conclude that our experiments, although apparently simple, are relevant as a demonstration of the validity of the system.

5 Conclusion

This paper reported on experiments where a physical robot is taught a vocabulary concerning proprio and exteroceptive perceptions. Teaching is provided by a second heterogeneous robot using an unsupervised imitative teaching strategy, namely mutual following of the two agents. Grounding of the vocabulary occurs as part of a general associative process of all the robot's sensor-actuator measurements. Simulations and physical experiments are carried out which show successful and consistent results. The vocabulary is learned and learning is stable in the face of a significant amount of noise. Success of the experiments demonstrates the validity of the learning architecture for the particular task. This work suggests 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 the same context of perceptions, a prerequisite for a common understanding of the language to develop. This work may be a starting point towards using more complex imitative strategies for transmitting more complex forms of communication.

6 Acknowledgement

An enormous thanks to the technicians for their essential support. Many thanks to John Demiris and Auke Jan Ijspeert for their useful comments on this paper. Aude Billard is supported by a personal grant from the LEMO company and the Swiss National Research Fund.

References

1. T. Arita, C.E. Taylor: A Simple Model for the Evolution of Communication. The Fifth Annual Conference on Evolutionary Programming (1996).
2. A. Billard, G. Hayes: A Dynamical Recurrent Associative Memory Architecture for Learning Temporal Correlated Events in Autonomous Mobile Agents. To appear in *Adaptive Behaviour* journal, MIT Press (1998).
3. A. Billard, K. Dautenhahn: Grounding communication in autonomous robots: an experimental study. To appear in *Robotics and Autonomous Systems*, special issue on Quantitative Assessment of Mobile Robot Behaviour. Edited by U. Nehmzow, M. Recce, and D. Bisset (1998).
4. J. Demiris, G. Hayes: Imitative Learning Mechanisms in Robots and Humans. Proc. of the 5th European Workshop on Learning Robots, Bari, Italy (1994).
5. K. Dautenhahn: Getting to know each other: artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, vol. 16, pages **333-356** (1995).
6. V. Klingspor, J. Demiris and M. Kaiser: Human-Robot-Communication and Machine Learning. *Applied Artificial Intelligence Journal*, Vol. 11, to appear (1997).
7. Y. Kuniyoshi, M. Inaba and H. Inoue: Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance. In *IEEE Transactions on Robotics and Automation*, Vol. 10, No 6, Dec. 1994.

8. Mataric M. J.: Reducing Locality Through Communication in Distributed Multi-Agent Learning. To appear in the *Journal of Experimental and Theoretical Artificial Intelligence*, special issue on Learning in Distributed Artificial Intelligence Systems, Gerhard Weiss, ed., fall 1997.
9. B. MacLennan & Gordon M. Burghardt: Synthetic Ethology and the Evolution of Cooperative Communication, *Adaptive Behavior*, Vol. 2, No. 2, pp. 161-187, Fall 1993.
10. M. Oliphant: Formal Approaches to Innate and Learned Communication: Laying the Foundation for Language. Doctoral dissertation, Department of Cognitive Science, University of California, San Diego (1997).
11. L. Steels & P. Vogt: Grounding adaptive language games in robotic agents, Proceedings of the fourth European conference on Artificial Life, Brighton, U.K., July 1997.
12. L. Steels: The Spontaneous Self-organization of an Adaptive Language. In Muggleton, S. (ed.) *Machine Intelligence 15*. Oxford University Press, Oxford (1996).
13. M. C. Torrance: The Case for a Realistic Mobile Robot Simulator. In Working Notes of the AAAI Fall Symposium on Applications of Artificial Intelligence to Real-World Autonomous Mobile Robots, Cambridge, MA, October 1992.
14. H. Yanco & A. S. Lynn: 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, pp. 478-485 (1993).