# Imitation: a means to enhance learning of a synthetic proto-language in an autonomous robot.

Aude Billard*

*Robotics Laboratory, University of Southern California, SAL 230, Los Angeles 90089, U.S.A

billard@usc.edu institutions with particular

September 9, 1999

**Abstract**

This paper addresses the role of imitation as a means to enhance the learning of communication skills in autonomous robots. A series of robotic experiments are presented in which autonomous mobile robots are taught a synthetic proto-language. Learning of the language occurs through an imitative scenario where the robot replicates the teacher's movements. Imitation is here an implicit attentional mechanism which allows the robot imitator to share a similar set of proprio- and exteroceptions with the teacher. The robot grounds its understanding of the teacher's words, which describe the teacher's current observations, upon its own perceptions which are similar to those of the teacher. Learning of the robot is based on a dynamical recurrent associative memory architecture (DRAMA). Learning is unsupervised and results from the self-organization of the robot's connectionist architecture. Results show that the imitative behavior greatly improves the efficiency and speed of the learning. Moreover, without imitation, learning of symbolic expressions to describe the robot's proprioceptions is not possible.

These experiments are meant to be an illustration of the following proposal: *the sharing of a similar perceptual context between imitator and imitatee creates a meaningful social context onto which language, that is the development of a common means of symbolic communication, can develop.* This work follows a non-nativist approach to language development ([1], [2], [3]) which stresses the importance of social cues, such as coordinated behavior (of which imitation is one instance), as the precursor to language development in infants. Furthermore, these experiments show that association of arbitrary symbols to *meaning* is possible in embodied agents other than humans, and

that it can be done by a simple, compared to animals, neural architecture. In particular, it stresses the fact that the word-meaning association is strongly dependent on constraining the temporal uncertainty of the association. This is achieved by the creating a spatio-temporal bounding between the teacher-learner agents through the imitative strategy.

# 1 Robotic models of imitation:

Psychologists, ethologists and roboticists find a common interest in the study of high-level social skills for robots, which can be used for gaining a better understanding of the mechanisms behind similar skills in animals ([4], [5], [6]) or for improving the robot's performance in interaction with humans or robots ([7], [8], [9], [10], [11]). Relevant to this manuscript are studies which investigate the use of imitation[1] through human-robot or robot-robot interactions.

While there is still some debate as to what observed behavior the term "imitation" refers to, there is also a need for clarification as to how many different cognitive processes studies on animal and human imitation refer to. The attempt of roboticists is to make the light between the different cognitive processes by implementing separate models for each of them and commenting on the emergent behaviors each model leads to. From a programming point of view, the ability to imitate amounts to solving at least three problems: 1) how to get the agent to observe the example, that is how to constrain the agent's attention to observing only the relevant stimuli [13], 2) how to get the agent to reproduce the observed action pattern, which amounts to defining a mapping process from the agent's sensor (visual) stimuli (observation of the other actions) to the agent's actuators [14, 15], 3) what to learn from the reproduction/imitation process, e.g. learning action sequence patterns ([16], [17], [18], [19]) or patterns of perception-action sequences [20], [21], [7], [22].

There is an intricate relationship between learning and imitating. One might better separate between them if one thinks separately of the three following behaviors: the ability of imitating, that of learning to imitate and that of learning by imitation. Biologically inspired models of the ability to imitate have been developed ([14], [9]) and applied to experiments in which the robot could mirror the movements of the head and of the arms of a human. Learning to imitate, that is how to instantiate the visuo-motor mapping, was addressed to some extend by Demiris [14] and Schaal [23] who proposed algorithms for comparing and choosing among a set of predictors of visual patterns (recognition of visual primitive) which were already connected to motor patterns. The work

[1] I follow Davis's general definition of the term *imitation*. He refers to a behavior skill which leads to "some sort of similarity in behavior among two or more individuals'[12].

presented in this paper is concerned with learning by imitation. In particular, it addresses the role of imitation as a means to enhance the learning of communication skills in autonomous robots. Here, the ability of imitating is built-in.

**Imitation as a means to direct attention:**

Learning by imitation has been used in different experiments for teaching the robot new motor skills ([24], [7], [21], [22], [25], [26], [19]. While the robot replicates the demonstrator's movements, it learns to adjust its own motor parameters and perceptuo-motor mapping to replicate at best the actions of the demonstrator. An important advantage of such an approach to roboticists is that it saves the tedious programmer's work of tuning the robot's motor control parameters.

A key idea underlying these works is that imitation can be an indirect and efficient means of directing the robot's actions. Having the robot replicating the demonstrator's actions leads the robot to make similar proprioception (of movements) which further enables the robot to learn new motor skills without explicit (symbolic or internally prespecified) teaching.

I want to stress this point further: imitation allows imitator and imitatee to share similar sets of both proprio- and exteroceptions.[2] While replicating the imitatee's actions, the imitator makes similar proprioceptions, not only of movement, but also of inclination and of relative energy consumption. Moreover, the imitator also makes similar exteroceptions to that of the imitatee, as both agents share simultaneously the same physical space. Note, however, that the similarity can never be perfect as two physical agents, whether robotics or humans, are never the same; noise in sensor reception makes two simultaneous sensor information different, and never share exactly the same "view" of the environment.

This paper and the experiments which I present are meant to be an illustration of the proposal that *the sharing of a similar perceptual context between imitator and imitatee creates a meaningful social context onto which language, that is the development of a common means of symbolic communication, can develop.* The idea that social cues, such as coordinated behavior (of which imitation is one instance), can be the precursor to language development in infants is not new to psycho-linguistics, and shall be presented in more detail in Section 2.

The remaining of the paper presents a series of robotic experiments in which autonomous mobile robots are taught a proto-language; *proto* relates to the fact that the robot is taught a symbolic communication system which shares only some of the features associated with human language. In these experiments, a robot learns a lexicon

[2] The terms proprio- and exteroception relate, respectively, to internal and external perceptions of the agent.

to name its proprio and exteroceptions and it learns to make proto-sentences to describe its own actions as well as that of the teacher. Learning of the language occurs through an imitative scenario where the robot replicates the teacher's movements.

The behavioral and learning abilities of the robot are enabled by a single connectionist architecture, DRAMA [16]. It is a general mechanism whose computation capacity for fast and robust learning of time series and of spatio-temporal regularities have been demonstrated theoretically [16, 27] and through experimental work [28, 29]. Section 3 gives a brief presentation of the algorithm, while section 4 summarizes the different sets of robotic experiments. Section 5 discusses the experiments' results and their contribution to the above proposal (that social biases create a meaningful perceptual context which can enhance the development of a symbolic form of communication). Section 6 concludes the discussion by proposing further robotics experiments to pursue this approach and to investigate outstanding questions not addressed by this study.

## 2 The role of imitation in language learning:

Learning of a language is perceived by many and, in particular, by linguists as a demonstration of highly complex cognitive capabilities of humans. It is often thought that this is one of the main characteristics which distinguishes us from other animals[3]. Language is very complex because of its important number of grammatical rules, its large vocabulary, its verbs conjugations, etc. It is therefore astonishing that infants master it so quickly (that is too quickly for it to be the result of pure association across all perceptual cues), given the low redundancy of single utterances and the high variability of other perceptions. To answer this problem, two main approaches have often been contrasted. The first approach, lead by Chomsky [36], is nativist. It argues for innate cognitive capacities for understanding and producing a language. In particular, it speculates the existence of an innate, i.e. the product of a long evolutionary process [37], *universal* grammar underlying all human languages. Learning of a language consists then of instantiating the parameters of the sets of built-in rules.

The second, non-nativist approach favors the idea that infants' language development results from complex learning skills, applied to general cognition including that of language. These skills are the product of evolution [34, 38, 39] and of development [1, 40, 41]). This approach assumes that humans are given general learning abilities

[3] This is a very controversial subject. For instance, people argue that some type of apes might be capable of simple forms of language. Their argumentation is based, e.g., on the successful experiment of Savage-Rumbaugh [30] at teaching chimpanzees a basic lexicon, on the studies of parrots' amazing ability of reproducing human speech [31] and that of whales complicated songs [32]. For some insight into this debate, see e.g. [33, 34, 35].

for combinatorial association, which are then applied to map words and meanings and to produce combinations of words such as to satisfy any grammatical rules (rather than only a universal one) [34, 42, 43].

The two above approaches give only a cognitivist account of language development. However, cognitive functions of associativity are not sufficient for learning a language. In order to account for the infants' facility to learn a language, a behaviorist[4] approach has been proposed which looks for external, socially driven, attentional processes which would shorten the time required for the associative process to converge by reducing the number of considered perceptions. It is interesting to note that the query of how language is learned can be traced back to St. Augustine, who viewed learning of words as an association of sound input with visual cues, driven by two social cues, namely gaze tracking and pointing (see quotation in Messer, p. 59 [44]). Closer to our time, Vygotsky's behaviorist position (followed later by Threvarten[5] has often been contrasted to Piaget's cognitivist position [41]). By opposition to Piaget who concentrates on the infants' internal and asocial cognitive processes, Vygotsky stressed the importance of social interactions alongside the infants' cognitive development.

In short, I see two views to language development. The first one — nativist — assumes it to be the results of biologically predefined abilities and independent on the development of other skills. The second one – non nativist — views language development as the results of the brain cognitive development as a whole. In this case, two variations can be distinguished: a Piaget's approach which considers the infant's (asocial, egocentric) interactions with the world as the main source of language development; a Vygotsky's approach which views the infant's social interactions with other humans as a key factor for language development. Similarly to [45], the approach taken in this paper can be considered close to a Vygostky's approach. In the following, I summarize what I think are the key aspects in social interactions underlying language development.

### Social interactions behind language development

Imitation has been attributed three different roles in infants' language development: in the motor control of speech, in the infants' internal cognitive development, and in the infants' social interactions. In the latter, imitation is as a social factor which guide the infants' cognitive development.

Vocal and verbal imitation is the process whereby the infant repeats part of the speech output produced by the caregiver (partial reproduction of the sound pattern in the case of vocalization and of words or sentences in the case

[4] *Cognitive* and *behavioral* capabilities are here distinguished. The former relates to the agent's internal computational processes and the latter relates to the interaction of these internal processes' outputs and inputs, the agent's actions and perceptions, with the environmental dynamics.

[5] There are several other approaches to the idea of social bias behind language development, see [44] for a review.

of verbal imitation). Practicing vocal utterances is thought to be a means of reinforcing the network responsible for the sequential activation of articulation muscles [46].

Infants' ability for imitating others is thought to play a role in the infant's cognitive development. Vocal imitation, as a rehearsal of sound utterances, might reinforce the process of meaning-sound matching [46, 47]. *Deferred* imitation, that is the ability to reproduce the observed actions within a long delay (hours, days) after their presentation, requires the ability for internalizing the visual observation of a situation. It is a first step towards symbolic representation and therefore might be a precursor to language [48].

Gestual and behavioral imitation, displayed on activities other than speech, has a social function [2] . It is a means to coordinate one's behavior on that of conspecifics, which can be used for engaging in body communication with peers ([1], [2], [3]). This last role of imitation in language learning is particularly relevant to the study presented in this chapter. In the following, I expand the ideas underlying this approach.

**Turn taking as a precursor to verbal communication:**

"*Social communication is fundamental to the development of language and cognition, permitting the establishment of a partnership within which communication takes place*" [1]. This partnership consists of being an active participant in a turn taking game. Indeed, verbal communication would suffer if one would not respect the social rule of turn taking in speech [44]. Therefore, alongside learning of a language, the child should also learn certain social rules. Moreover, learning of these social rules might play a role in the actual development of the infant's understanding of language. By learning the rhythmic of speech patterns (conversation), the infant might build up expectancies for certain form of speech which would simplify its learning of speech-meaning associations, where the meaning is given by the perceptual context.

The joint activity of the mother (or other child caretaker) and the child participates in the *scaffolding* process of learning a language. Scaffolding relates to a step by step teaching process which directs the child's learning process."*This process is local, task-directed and focuses the child's attention on relevant aspect of the task*" [49]. The child must first learn to communicate, that is, to correctly respond in a turn taking game. This implies understanding the specificity of speech patterns (interrogative, imperative, declarative speech) and learning to produce the correct response. The child's prelinguistic understanding of these communication patterns is described by [50, 51], who observed that children develop an ability for "*synchrony of movements with adult speech patterns, imitation of facial expressions and selective attention to aspects of speech*" [51]. Further, the child develops the ability to synchronize his visual attention to that of the mother, by following the mother's direction of gaze, and to produce vocalizations in synchrony to the mother's speech. Nadel and others [2] study the development of infants' ability for turn taking games, from being the imitator to being the imitatee. She observes that this developmental step is a marker for the child's development of more complex form of verbal communication exchanges[2].

**Imitation as a means to coordinate behavior:**

The synchrony of activity between the child and his mother results in their shared attention to specific visual and auditive patterns. The child develops his understanding of the mother's speech by associating a meaning to the mother's utterance in terms of his visual perceptions. The importance of a "*co-ordination of joint activities involving mutual direction of attention*" [52] between the mother and child for the grounding of the child's understanding of language was first pointed out by [49] and then further expanded by diverse authors (see [52, 53, 2]). In summary: "*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*"[53].

# 3 Robotic experiments

This work wishes to contribute to a non-nativist approach which favors social biases behind language learning. It presents robotics experiments in which the robot's ability to imitate another robot or a human teacher enhances the efficiency with which the robot learns a proto-language.

Experiments are carried out with two types of robots, wheeled based vehicle robots and a doll-shape robots. For a technical description of the robots' hardware, see [16, 54]. The robots' imitative behavior consists in the ability to follow another robot and to mirror the movements of a human instructor. The imitative behavior and the learning of the robot are controlled by a single artificial neural network architecture DRAMA (Dynamical Recurrent Associative Memory Architecture). A complete mathematical description of DRAMA can be found in [16]. Below is a short summary of DRAMA's functioning.

**The robot's controller**

The robot's controller in the experiments is composed of two parts: a set of *event recognition* modules for detecting variations in the robot's sensor-actuator state and a *learning module*, the DRAMA architecture, which associates temporally the changes across all the sensor-actuator modalities of the robot. At each processing cycle, the sensor-actuator vector state is measured. When a variation (i.e. more than one bit change) in one sensor or actuator input has been measured, the new information is forwarded to the associative architecture (DRAMA) to be correlated with all simultaneous and previously recorded events in other sensor-actuator systems.

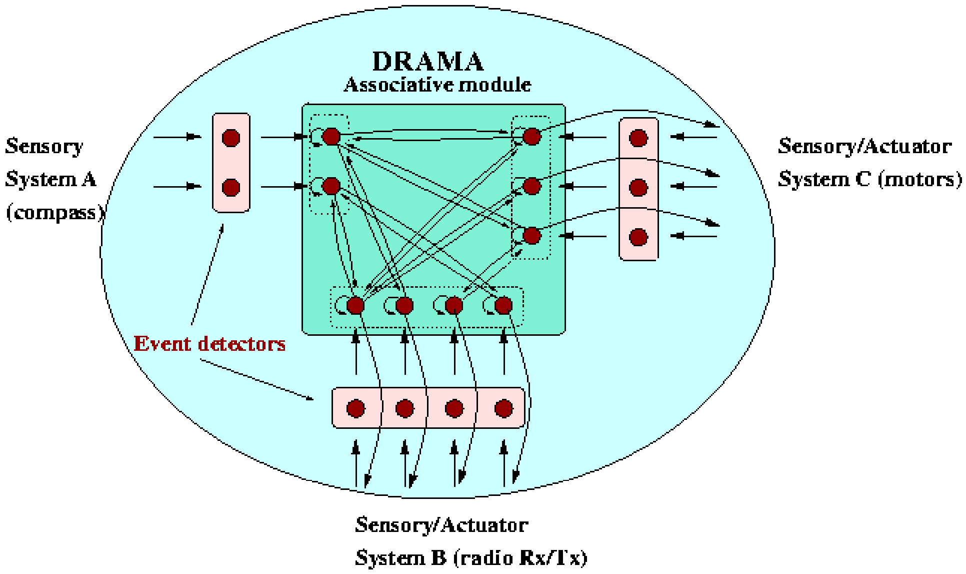


Figure 1: Schema of the connectionist architecture DRAMA which controls the imitative and learning skills of the robot.

DRAMA is a fully connected recurrent neural network, without hidden units. The self-connections on the units provide a short term memory of the units' activation. Consequently, sensor and actuator information (i.e. their corresponding network units' activation) is memorized for a fixed duration; the memory duration is determined by the decay rate of unit activation along the self-connection on the unit, see equation 1. This allows association to be made between time delayed presentations of two inputs, hence learning of time series. In the experiments, the robot learns sequences of actions (dance patterns), i.e. consecutive activation of different actuator states. It also learns words to label its perceptions, i.e. it associates labels (radio signal or combination of typed keys on a keyboard) with other consecutive, simultaneous or precedent perceptions. Finally, it learns to combine sequentially the labels to form proto-sentences and to associate these combinations with a meaning in terms of other perceptual inputs.

Figure 1 shows a schematic representation of the robot's controller with three sensor systems as inputs. The number of sensors and actuators used in the experiments vary from 10 (for the vehicle robots which have 2 bumpers, 2 infra-red sensors, 2 light detectors, 1 compass, 1 radio transceiver and 2 motors) to 19 (for the doll robot which has an 8-keys keyboard, 6 touch switches on the body, 4 infra-red sensors and 3 motors). The number of nodes of DRAMA varies consequently from 18 up to 169, each sensor being represented by a different number of nodes depending on the granularity of its encoding.

Long term memory of consecutive activation of two units in DRAMA consists of updating the parameters of the connection linking these two units following pseudo-Hebbian rules (see equations 3 and 4). Similarly to time-delay neural networks, each connection in the DRAMA network has two associated parameters: a time parameter ($tp$) and a confidence factor ($cf$). Time parameters and confidence factors are positive numbers; they record respectively the time delay between two units' activations and the number of their co-activations.

**Unit activation function**

$$y_i(t) = F(x_i(t) + tp_{ii} \cdot y_i(t-1) + \sum_{j \neq i} G(tp_{ji}, cf_{ji}, y_j(t-1))) \tag{1}$$

$F$, the transfer function, is the identity function for input values less than 1 and saturates to 1 for input values greater than 1 ($F(x) = x$ if $x \leq 1$ and $F(x) = 1$ otherwise) and $G$ is the retrieving function whose equation is given below in equation 2 and explained in the following paragraph. The indices notation used in the equations should be interpreted as follows: $cf_{ji}$ is the confidence factor of the connection leading from unit $j$ to unit $i$.

$$G(tp_{ji}, cf_{ji}, y_j(t-1)) = A(tp_{ji}) \cdot B(cf_{ji}) \tag{2}$$

$$A(tp_{ji}) = 1 - \theta(|y_j(t-1) - tp_{ji}|, e)$$

$$B(cf_{ji}) = \theta(cf_{ji}, \frac{\max_{y_j>0}(cf_{ji})}{T})$$

$\max_{y_j>0}(cf_{ji})$ is the maximal value of confidence factor of all the connections between activated units $j$ and unit $i$, which satisfy the temporal condition encoded in $A(tp_{ji})$. The function $\theta(x, H)$ is a threshold function that outputs 1 when $x >= H$ and 0 otherwise.

**Training rules:**

$$cf_{ji}(t) = cf_{ji}(t-1) + a \tag{3}$$

$$tp_{ji}(t) = \frac{tp_{ji}(t-1) \cdot \frac{cf_{ji}}{a} + \frac{y_j(t)}{y_i(t)}}{\frac{cf_{ji}}{a} + 1} \tag{4}$$

The DRAMA network has general capacities for learning of spatio-temporal regularities and time series. A formal analysis of these properties can be found in [27, 16]. The network can learn series of the type 1) $ABA$, cyclic sequence; 2) $ABCDE$ and $HJCKAE$, crossing sequences on state $C$ and $A$; 3) $ABCDEFCDG$, sequence with a loop on states $CD$ and divergence on state $D$. These properties are very relevant to learning of a language, which is made of multiple combinations of the same primary items (words) and where sentences are often recursive combinations of the same instances. For instance, $ABCDEFCDG$ could represent the sentence 'I think that you should recognize that you are wrong' and $ABA$ could represent '(To be) (or not) (to be)'. $ABCDE$ and $HJCKAE$ could be "My father works at home", "I will work for my pleasure".

DRAMA's property for extracting spatio-temporal invariance is exploited in the experiments of section 4.1 where vehicle robots learn word-meaning pairings. Learning of time series is used in experiments of section 4.2 where the doll robot is taught sequences of actions and proto-sentences (combination of words of the lexicon) .

Built-in behaviors can easily be defined by presetting particular sensor-actuator connections. In the experiments, the robot's imitative behavior of following (section 4.1) and of mirroring (section 4.2) the teacher agent's movements result from the presetting of particular connections between light sensors (sensitive to visual and infra-red light) and the robot's motors. Recall of the DRAMA units' outputs (following equation 1) is used to direct the robot's movements (retrieval of the predefined sensor-motor connectivity). Results of the learning can then be immediately exploited. For instance, in section 4.2., recall of learning of the motor sequences results in the robot's immediate repetition of the sequence. Similarly, recall of the word-perceptions associations (section 4.2) results in the robot emitting the correct word or combination of words given a particular combination of perceptions.

# 4 Experiments:

Two main sets of experiments were carried out. In the first set (section 4.1), a learner robot is taught by a teacher robot a lexicon to describe its proprio- and exteroceptions. In the second set (section 4.2), a doll robot is taught by a human experimenter. It learns a lexicon and how to combine the words of the lexicon to produce proto-sentences in order to describe its interactions with a human teacher.

In the following, I shortly summarize the main results of these experiments. I refer in places to detailed reports of each of these experiments.

### Imitative scenario:

The experiments are based on an imitative scenario whereby the learner robot replicates the movements of the teacher agent. In the first set of experiments (section 4.1), the learner robot follows closely the teacher robot and, therefore, implicitly replicates the teacher's movements in the plane (see figure 2 left). Following proceeds from mutual phototaxis of the two agents. Because tracking is mutual, it results in a smooth binding between the two agents. The agents seldom lose sight of each other, as, if the learner runs slower, the teacher waits for it.

In the second set of experiments (section 4.2), the doll robot mirrors the demonstrator's movements of the arms (up and down lifts) and of the head (left and right turns). The robot's mirroring behavior results from a predefined coupling between motors and infra-red sensors (IRs). A pair of IRs is attached to the robot's ears and detects the demonstrator's head movements by measuring the reflection of the two IR emitters on the demonstrator's glasses, see figure 2 right. A second pair of IR receivers placed on the robot's chest detects the demonstrator's hand movements by measuring the IR emission of the sensors which the demonstrator holds in her hands.

Figure 2: The experiments are based on an imitative scenario, namely following behavior (left) and mirroring of arm and head movements (right), whereby the learner robot replicates the movements of the teacher.

**Learning of a lexicon:**

Four experiments were carried out in which a learner robot was taught by a teacher robot a small vocabulary (composed of four to eight words). Two types of wheeled robots were used (FischerTechnik and Lego based robots), with different sensors and different morphologies. Learning of the vocabulary consists for the robot to associate correctly radio signals (the labels) with different sensory inputs. While following the teacher robot, the learner robot picks up radio signals emitted by the teacher robot. These signals are labels for either the teacher's external perceptions (observation of a box or a light bulb [28]) or of its internal perceptions of movement (move, stop, turn left or right)[55], of inclination (up, down, plane)[56] and of orientation (North, South, West, East)[16].

The robot learns the meaning of the teacher's signals by associating those with its owns perceptions. While they are bounded by the following process, learner and teacher agents are set in a position from which they share a common context of both external (facing the same direction) and internal (performing the same movement, traveling the same distance and over the same ground) perceptions. 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 learner's actual perceptions. It is thus an unsupervised teaching strategy. Learning of word-observation pairs results from the statistical associative process provided by the DRAMA architecture.

Signal and meaning are usually not perceived simultaneously. There is thus an important temporal uncertainty in the presentation of the signal-meaning pattern. Moreover, in addition to the radio signal and the relevant sensor input (the word's meaning), the robot perceives several other irrelevant stimuli (provided by its other sensors). The robot learns the correct word-meaning pairs by extracting the correct spatio-temporal invariance

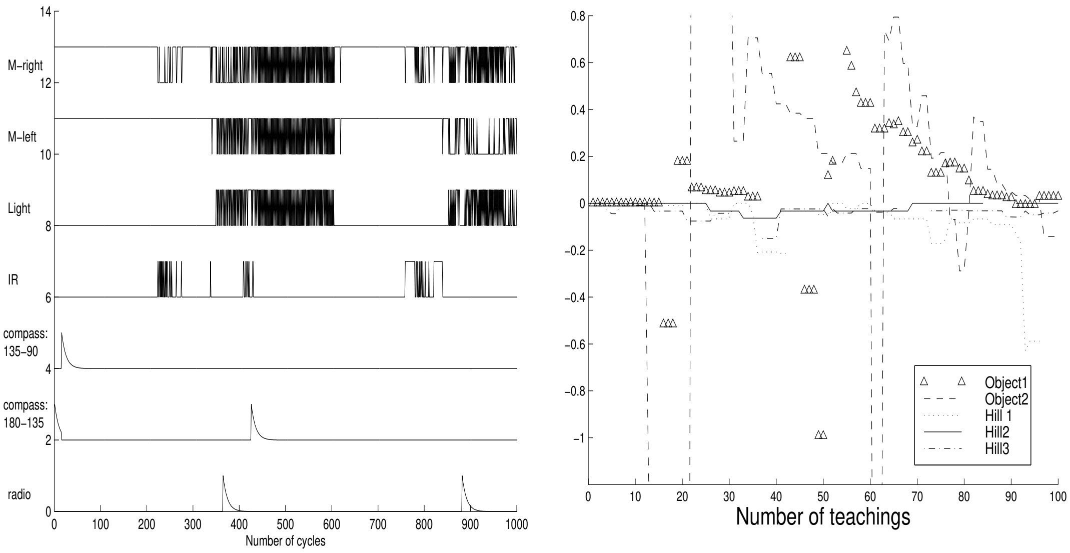


Figure 3: **Left** Robot's neural activity resulting from the sensor-actuator activation during a run. **Right** Variation of the time parameter values for each correct connection signal-object

in the continuous pattern of sensor-actuator state. Incorrect associations, which arise from a mismatch between the agents' observations, are discarded compared to correct ones by a process of statistical elimination depending on their relative frequency of occurrence (see section 3). In the experiments, the robot faced up to 30% and 50% imprecision respectively in the temporal and spatial dimensions of its perceptual information. Under these conditions, learning was shown to be successful.

Figure 3 left shows an example of the robot's neural activity resulting from the sensor-actuator activation. These data are the results of experiments in which the robot was taught four words to describe the four directions of compass (North/South/West/East)[6], see [16]. The three lines starting from the bottom show the receptions of the teacher's radio signals and the consecutive measurements of compass variations. In order to correctly associate the radio and compass stimuli, the robot has to discard the activity of its other sensors (namely the two light and infra-red sensors) and actuators (the two motors), which bears no direct correlations with the radio signals. Figure 3 on the right shows the time parameter ($tp$) variation between two time steps. These data are the results of experiments in which the robot was taught five words to describe objects (boxes) and hills (inclined objects)[28]. The graphs show strong and continuous fluctuations during the teaching. This reflects the important temporal variability of the word-meaning pattern.

[6] An on-line animated demonstration of this experiment can be found at http://www-robotics.usc.edu/ billard/Drama.html.

Table 1: 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. | $1.0 \pm 0.2$ | $0.6 \pm 0.2$ | $0.5 \pm 0.1$ | $0.5 \pm 0.2$ | $0.6 \pm 0.1$ | $0.0 \pm 0.1$ |

### The importance of the imitative skills

Further, simulations were carried out to investigate scaling up of the teacher-learner scenario to transmit a vocabulary among a group of robots, from a teacher robot to a group of learner robots [29]. Simulations were used to validate the model, the learning architecture and the imitative teaching scenario, by generating a statistically significant number of data. 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 [57]).

In order to determine the importance of the imitative/following strategy for the success of the learning, simulations were carried out with two types of learner agents, some with and some without the ability of following. The agents had to learn three different types of vocabularies: 1) a vocabulary to describe nine objects (colored patches) in 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 measured the robot's individual orientation. Table 1 shows the mean value of the ratio (number of correctly learned words over total vocabulary) over the 10 runs for each experiment and for each learner robot.

Results showed that the non-follower agents are less successful on average and slower at learning the vocabulary concerning global variables, such as the name of the objects and that of the measure of the objects' location in global coordinates. However, the non follower agents are in all cases unsuccessful at learning the vocabulary concerning the agent's relative perception of orientation (relative to its compass). In this experiment, the capacity of following the other is particularly important when the word to be learned concerns the agent's proprioceptions, e.g. of orientation, as it allows the two agents to share a similar internal perception of it, as they point in the same direction.

### Learning of proto-sentences:

The second set of experiments was carried out with a small doll-shaped robot, called Robota, see [27, 20] for a complete report. As mentioned in section 3, the robot can imitate (mirror) the arms and head movements of a demonstrator (see figure 2 right). In a experiment, the robot was taught to perform different sequences of actions (consecutive movements of arms and head) by replicating those performed by the human demonstrator and to label these action sequences with different 'names'. In these experiments, rather than radio signals, a name consisted of a combination of keys typed on a keyboard connected to the robot (the eight keys of the keyboard are interpreted by the DRAMA controller as another set of switch inputs, in addition to that on the robot's body). In a second experiment, the robot was taught combinations of words (combination of keys, where one key is one word) which formed English proto-sentences (e.g. "I touch left arm", "You move head right"), to describe its actions and perceptions of touch on different parts of its body. Just to mention it, tests were done with children of 5 and 6 years old, who taught the robot words to label different parts of its body and simple actions sequences [54]. Children took a real pleasure to having the robot responding to their interaction. This tests suggest that the imitative and communicative behaviors of Robota makes it an interesting toy for children.

Table 2: Results of teaching of dance patterns and sentences

| **Experiment type** | **Time (min)** | **Nm of trials** | **Nm of patterns** |
|---|---|---|---|
| Action sequences | $15 \pm 4$ | $1.2 \pm 0.5$ | $8 \pm 0$ |
| Word sequences | $12 \pm 10$ | $2 \pm 1.5$ | $6 \pm 2$ |

Table 2 shows the results of the experiments on teaching sequences of movements (1st line) and on teaching proto-sentences (2nd line). Each experiment was carried out by five different experimenters. Data are mean value and standard deviation of time before learning, the number of trials per taught pattern (before learning) and the number of taught patterns. In the first experiment, the robot was taught eight actions patterns, which consisted of different sequences of head and arm movements. The experimenter was free to choose the number of movements of each dance he wanted to teach. As a result, none of the teaching was similar (there were dance patterns common to two teaching episodes but the complete sets of dance patterns were different). In addition, as each experimenter would move at different speed, one could evaluate the performance of the DRAMA network to adapt to variation in the input timing.

In the second experiment, the robot was taught four to eight different sentences to describe its actions and its perceptions of touch on the body. The experimenter was free to choose the eight words combination (one word per keyboard key) it wanted to use in the sentences. Examples of taught proto-sentences are *I move right arm, I turn head left, You touch mouth, You touch left/right foot.* Table 3 shows the progress of the robot's learning alongside one of the teaching. The robot was taught five complete sentences (middle column) for describing its different

Table 3: Example of sentence teaching

| **Activated sensor** | **Taught sentence** | **Retrieved sentence** |
|---|---|---|
| Move right arm | *I move right arm* | |
| Touch left arm | *You touch left arm* | |
| Move right arm + touch left arm | | *arm* |
| Touch right foot | *You touch right foot* | |
| Touch left foot | *You touch left foot* | |
| Touch right foot + left feet | | *You touch foot* |
| Touch mouth | *You touch head* | |
| Touch right arm | | *You touch* |

motions and touch perceptions (left column). It learned the meaning of each word from extracting the invariances across all teachings (e.g arm in line 3, you touch and foot in line 6 in table 3). The robot learns also the position of each word relative to one another in a sentence by extracting temporal invariance in the words consecutive occurrences. Finally, the robot can use its learning to infer the correct combination of word for describing a touch on the right arm (line eight of the example), for which it had yet not been taught a sentence.

These experiments demonstrated that the DRAMA architecture could allow learning of a basic 'proto-language', which shares some properties with natural language: 1) each word (key in our experiment) can carry a specific meaning; 2) words can be combined and the combination can be given a different meaning while not losing the meaning of each word taken separately; 3) different combinations of the same words can be given different meanings, the meaning of each combination being determined by the order of appearance of each word in the combination; 4) conceptual meanings of each word can be learned implicitly by only presenting them as part of complete sentences, which can then be used to infer new word combinations; 5) precedence between words appearance in the combination is learned and can be used to infer the correct order when constructing a new word combination.

In summary: these experiments allowed us 1) to verify the computational ability of the DRAMA architecture for learning redundant combinations and sequences of inputs, such as to form the basis of a proto-language; 2) to show that a mode of interaction with a robot, which is based on imitation and synthetic communication, could be acceptable to humans (to both infants and adults who participated in the experiments); And that it could be an easy means of teaching a robot (in the capabilities of young children).

Note that only simple sentence examples were used so far, in which the words could easily be tied to the taught concepts. Moreover, the 'language' the robot was to learn was regular; that is, the robot's learning task was to recognise temporal regularities in the words' ordering across the taught sentences and to correlate the words' usage with its sensors and actuators' activity. As such, these experiments were a first step towards demonstrating the validity of the system (the learning architecture and the imitative strategy) for teaching a robot a symbolic communication system, such as a regular language. The fact that the DRAMA architecture is comparable in function to a Hidden Markov model or other recurrent neural networks [27], which are models currently used in techniques of Natural Language Processing, suggests that the model could scale up successfully to learning a regular language. However, it remains to be shown how the system could scale up to learning a complete language with grammar structure and irregularities.

A second doll has now been constructed, which uses a more powerful microcontroller. The sensor abilities of the doll have been increased. The keyboard has been replaced by a commercial speech processor. The doll is provided with a colored CCD camera and pyroelectric sensor for detecting the demonstrator's movements. Current experiments are testing the system for learning more complex speech patterns, as made possible by the speech recognizer system. Updates on this work's progresses can be found at http://www-robotics.usc.edu/ billard/poupees.html

# 5 Discussion:

This paper reviewed a series of experiments in which a robot was taught basic elements of a synthetic proto-language. The robot grounded a lexicon concerning its perceptions and learned to combine the words of the lexicon to describe situations. The experiments were based on an unsupervised learning strategy where the robot grounded meaning into another agent's utterances in terms of its own perception. The teacher agent had not access to the robot's internal state and, thus, did not direct its teaching depending on the robot's performance. The learning of the robot was, however, implicitly guided by the teacher, as the robot's movements and perceptual attention were co-ordinated to that of the teacher. This coordination resulted from the robot's ability of imitating the teacher's movements. Results showed that the imitative behavior improves greatly the efficiency and speed of the learning. Moreover, without imitation, learning of symbolic expressions to describe the robot's proprioceptions was not possible.

### Contribution to robotics

The work presented in this paper does not follow directly from any previous studies of robotics. Indeed, little work has yet been done in teaching a physical robot a synthetic form of communication. Closest works are those of Yanco and Stein [58] and Vogt and Steels [59], in which a robot learns a lexicon to describe sets of actions and perceptions respectively. The present study differs from those works in two main aspects:

1. The learning and behavioral capacities of the robots result from a single connectionist architecture, which has general ability for extracting spatio-temporal regularities in a dynamic environment. It is used, in the experiments, for other learning tasks in addition to that of learning the language, which involve learning time series of perception-action [16]. Thus, the model is more general than the compared studies, which used a learning mechanism, designed specifically for the particular language task.
2. The language which the robot is taught is not restricted only to a lexicon, where each word of the lexicon relates to a single specific perception, as in the two compared studies. Because of the general property of the DRAMA architecture for learning time series, the robot could be taught combinations of words of the lexicon to label combinations and sequences of perceptions.

I am a conscious, however, that these experiments are somewhat simple in terms of the number of words and combinations learned. The experiments so far were mainly aimed at illustrating the general capacities of DRAMA together with the imitative strategy for the learning of a simple proto-language. Further experiments are currently carried out with the aim to study the learning of more complex patterns of sensory inputs with more sophisticated robots (such as the second doll robot).

### Contribution to computational linguistics

The success of the experiments reported in this paper validated the proposed unsupervised imitative learning strategy for transmitting a basic synthetic mode of symbolic communication across heterogeneous agents. Moreover, these experiments showed that association of arbitrary symbols to meaning is possible in embodied agents other than humans; and that it can be done by a simple, compared to animals, neural architecture. In particular, it stresses the fact that the word meaning association is strongly dependent on constraining the temporal uncertainty of the association. Language learning occurs in a highly noisy and sensory rich environment. The experiments simulated only very partially some of this complexity, but as such it was already an important step compared to previous work in the domain.

In the robotics works on grounding of a lexicon, which were cited earlier ([58] and [59]), the spatial and temporal variability of the word-meaning pattern is almost inexistent. Learning occurs between only the relevant (to the learning) sensory channels, which are the motor and radio channels [58], and the infra-red, light and radio channels in [59]. This reduces strongly the possibility of incorrect signal-sensory input associations. There is no temporal ambiguity in the associations, as the experiments are static. As soon as the robot picks up a radio signal, it chooses (action-selection) an action to perform [58] or one of its current sensory input [59]. It gets reinforced (reinforcement learning[58], increase of statistic bias for the sensor[59]) on the validity of its choice.

The embodied aspect of grounding of meaning has also been neglected by simulated studies of language evolution (e.g. [60], [61], [62], [63], [64]). In these works, grounding of communication is regarded as a computational problem that can be solved solely by means of combinatorial analysis. For these authors, categorization 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.

In my opinion, combinatorial analysis is not sufficient: it becomes quickly intractable when the number of dimensions in the sensory space is high and when temporal delays are taken in consideration (as it should in any physical implementation). In this case, external attentional mechanisms would be required to reduce the uncertainty on the spatial and temporal dimensions of the sensory inputs. The introduction and second section of this paper stressed the importance of social interactions as an external attentional mechanism. I referred to studies of psycholinguistics which pointed out several social mechanisms which are possible attentional mechanisms. These are provided either by the speaker/teacher (pointing, increasing the tone of voice, linguistic deixis) and/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). These attentional mechanisms act as a cognitive process which restrict the number of observations before combinatorial analysis. However, there is more to this than just a single cognitive process. There is an interactive process between the two communicative agents, which requires a behavioral coordination between the two agents. In the experiments reported in this paper, this behavioral coordination between teacher and learner agents is achieved by creating a spatio-temporal bounding between the agents' behaviors through an imitative strategy.

In summary, this work brings a novel contribution to current research concerned with the modeling of language learning, as it addresses both the behavioral (agent-environment interaction) and cognitive aspects of the *symbol grounding problem* [65]. In particular, it suggests that providing the robot's with primary social abilities might enhance its development of more complex social skills. As such, this works is a tiny first step towards studies of the development of complex social skills for robots.

# 6 Conclusion

This paper presented a series of robotic experiments in which autonomous mobile robots were taught a basic synthetic proto-language. The robot grounded a lexicon concerning its perceptions and learned to combine the words of the lexicon to describe its interactions with the teacher. Learning of the robot is unsupervised and results from the self-organization of the robot's connectionist architecture. The success of the experiments validated the proposed unsupervised imitative learning strategy for transmitting a basic synthetic mode of symbolic communication across heterogeneous agents.

This work wished to contribute to a non-nativist approach which favors social biases behind language learning. It was argued that external, socially driven, attentional processes are required aside associative processes for the grounding of word meaning. The robot experiments reinforced this claim by showing that imitative behavior, which implicitly constrains the robot's perceptions, could improve the robot's performance at learning a proto-language. In particular, results showed that imitative behavior was necessary for grounding the robot's proprioceptions, such as naming its relative orientation and inclination.

This works is a first step towards studies of the development of complex social skills for robots. It opens the way to further implementation of the model in experiments on teaching a robot a complete regular language, which includes a syntax and other grammatical rules (with the limitation that handling of irregularities and ambiguities might not be possible). Moreover, it suggests that providing a robot with primary social abilities might enhance its development of more complex social skills. Developing more complex imitative scenarios would allow transmitting more complex forms of communication. This could, for instance, be the development of the robot's capability for imitating facial expressions (as currently investigated by [66]), used then for transmitting symbolic expressions to label human states of emotions. Further, the robot's ability at replicating complex action sequences and at learning long perception-action sequences could be used to teach the robot more complex conceptual notions for describing its behavior. In this case, the robot's internal state would be more complex than simply the robot's sensor and actuator state (e.g., it would also include the value of different motivational factors) and, therefore, higher-level conceptual information (e.g., concerning goals, rewards, etc.) could be transmitted through the language. The latter direction is currently investigated by the author.

**Acknowledgements**

I want to thank very warmly Kerstin Dautenhahn and Gillian Hayes for their precious comments and guidelines all along this project. An enormous thanks to all the technicians of the department of Artificial Intelligence at the University of Edinburgh, whose precious help made this work possible. Facilities were provided by the department of Artificial Intelligence at the University of Edinburgh. Aude Billard was supported by personal grants from the LEMO company and the Swiss National Science Foundation.

# References

[1] A. F. Garton. Social interaction and the development of language and cognition. In *Essays in developmental psychology*. Laurence Erlbaum Associates, USA, 1992.

[2] J. Nadel, C. Guerini, A. Peze, and C. Rivet. *The evolving nature of imitation as a format for communication*, pages 209–234. Cambridge University Press, 1999.

[3] C. Threvarthen, T. Kokkinaki, and G. A. Fiamenghi Jr. *What infants' imitations communicate: with mothers, with fathers and with peers*, pages 61–124. Cambridge University Press, 1999.

[4] R.A. Brooks, C. Breazeal (Ferrell), R. Irie, C. C. Kemp, M. Marjanovic, B. Scassellati, and M. Williamson. Alternative essences of intelligence: lessons from embodied ai. In *AAAI*, 1998.

[5] K. Dautenhahn. I could be you - the phenomenological dimension of social understanding. *Cybernetics and Systems, Special Issue on Epistemological Aspects of Embodied AI*, 28(5):417–453, 1997.

[6] M.J. Mataric. Studying the Role of Embodiment in Cognition. *Cybernetics and Systems Vol. 28, No. 6, special issue on Epistemological Aspects of Embodied AI, pp. 457-470*, July 1997.

[7] K. Dautenhahn. Getting to know each other – artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

[8] V. Klingspor, J. Demiris, and M. Kaiser. Human-robot-communication and machine learning. *Applied Artificial Intelligence Journal, 11*, pages 719–746, 1997.

[9] M.I. Kuniyoshi. The science of imitation -towards physically and socially grounded intelligence. In *RWC Joint Symposium*, 1994.

[10] M J Mataric. Learning social behavior. *Robotics and Autonomous Systems*, 20:191–204, 1997.

[11] R. Pfeifer. Embodied system life. In *Proceedings of the International Symposium on System Life, Tokyo*, July 1998.

[12] J. M. Davis. Imitation: a review and critique. In P. Klopfer, editor, *Perspectives in Ethology*, volume 1, pages 43–72. Plenum, Ney York, 1973.

[13] B. Scassellati. Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. In Nehaniv C., editor, *Computation for Metaphor, Analogy, and Agents.* Springer-Verlag series, 1998.

[14] J. Demiris. *Movement imitation mechanisms in robots and humans.* PhD thesis, Dept. of Artificial Intelligence, University of Edinburgh, Mai 1999.

[15] J. Demiris and M.J. Mataric. Perceptuo-motor primitives in imitation. In *Working Notes, Autonomous Agents'98 Workshop on Agents In Interaction - Acquiring Competence Through Imitation, Minneapolis/St. Paul*, 1998.

[16] A. Billard and G. Hayes. Drama, a connectionist architecture for control and learning in autonomous robots. *Adaptive Behaviour Journal, Vol. 7:1*, 1999.

[17] G.M. Hayes and J. Demiris. A robot controller using learning by imitation. In *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems, Grenoble, France*, pages 198–204, July 1994.

[18] M.I. Kuniyoshi. Behavior matching by observation for multi-robot cooperation. In *International Symposium of Robotics Research, ISRR'95*, 1995.

[19] S. Schaal. Learning from demonstration. *Advances in Neural Information Processing Systems*, 9:1040–1046, 1997.

[20] A. Billard. 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*, 26:1, 1999.

[21] J. Demiris and G. Hayes. Imitative learning mechanisms in robots and humans. In *published in the Proceedings of the 5th European Workshop on Learning Robots,Bari, Italy*, pages 9–16. also as Research Paper No 814 at the dept. of Artificial Intelligence at the University of Edinburgh, 1996.

[22] P. Gaussier, S. Moga, J.P. Banquet, and M. Quoy. From perception-action loop to imitation processes: a bottom-up approach of learning by imitation. *Applied Artificial Intelligence*, 7:1, 1998.

[23] S. Schaal. Nonparametric regression for learning nonlinear transformations. In H. Ritter and O. Holland, editors, *Prerational Intelligence in Strategies, High-Level Processes and Collective Behavior.* Kluwer Academic Press, 1999.

[24] A. Billard and G. Hayes. Learning to communicate through imitation in autonomous robots. In Wulfram Gerstner et al., editor, *Proceedings of ICANN97, 7th International Conference on Artificial Neural Networks*, pages 763–768, 1997.

[25] S. R. Cooke, B. Kitts, R. Sekuler, and M. J. Mataric. Delayed and real-time imitation of complex visual gestures. In *Proceedings of the International Conference on Vision, Recognition, Action: Neural Models of Mind and Machine, Boston University*, May 28-31 1997.

[26] M.I. Kuniyoshi and I. Inoue. Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Transactions on Robotics and Automation, vol.10, no.6*, pages 799–822, 1994.

[27] A Billard. *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, U.K., 1998.

[28] A. Billard and K. Dautenhahn. Grounding communication in autonomous robots: an experimental study. *Robotics and Autonomous Systems, Special Issue on Scientific methods in mobile robotics*, 24:1-2, 1998.

[29] A. Billard and K. Dautenhahn. Experiments in social robotics: grounding and use of communication in autonomous agents. *Adaptive Behaviour Journal, special issue on simulation of social agents*, 7:3/4, 1999.

[30] E. S. Savage-Rumbaugh. *Language comprehension in ape and child.* Chicago : University of Chicago Press, 1993.

[31] Bruce R. Moore. Avian movement imitation and a new form of mimicry: Tracing the evolution of a complex form of learning. *Behavior*, 122:614–623, September 1992.

[32] AAS. Communication and behavior in whales. Washington DC, Selected Symposium, 1983. on-line references http://www.dai.ed.ac.uk/daidb/people/ homes/rbf/WHALES/warinentry.html.

[33] D. Bickerton. *Language & Human Behaviour.* UCL Press, 1990.

[34] T. Deacon. *Symbolic Species: the Co-Evolution of Language and the Human Brain.* General Science, 1997.

[35] E.O. Wilson. *Sociobiology:the new synthesis.* Social Sciences, 2nd edition, 1980.

[36] N. Chomsky. *Language and Mind.* Linguistics and philosophy. Harcourt Brace Jovanovich, Inc., 1968.

[37] S. Pinker. *The language instinct.* Penguin books, 1994.

[38] D.F. Armstrong, W. C. Stoekoe, and S. E. Wilcox. *Gesture and the nature of language.* Cambridge University Press, 1995.

[39] M. A. Arbib and J. C. Hill. *In Explaining Language Universals*, pages 56–72. Blackwell, 1988.

[40] J. Piaget. *Play, Dreams and Imitation in Childhood.* NewYork: Norton, 1962.

[41] C. Threvarthen. *Communication and cooperatipon in early infancy: A description fo primary intersubjectivity.* Cambridge University Press, 1979.

[42] S. Harnad. Grounding symbols in the analog world with neural nets. *Think, Special Issue on Connectionism versus Symbolism*, 2(1):12–78, 1993.

[43] F. Pulvermuller. Words in the brain's language. *Behavioural and Brain Science*, 1998.

[44] D. J. Messer. *The development of communication: from social interaction to language.* Wiley, England, 1994.

[45] K. Dautenhahn and A. Billard. Bringing up robots or - the psychology of socially intelligent robots: From theory to implementation. In *Proc. of Autonomous Agents (Agents '99) conference, Seattle, Washington, USA*, 1999.

[46] G. S. Speidel. *Imitation: a bootstrap for learning to speak*, pages 151–180. Springer Verlag, 1989.

[47] G. S. Speidel and M. J. Herreshoff. *Imitation and the construction of long utterances*, pages 181–198. Springer Verlag, 1989.

[48] A. N. Meltzoff and A. Gopnik. *On linking nonverbal imitation, representation and language: Learning in teh first two years of life*, pages 23–52. Springer Verlag, 1989.

[49] E.D. Lock. *Action, Gesture and Symbols: the emergence of language.* Academic Press, Harcourt Brace Jovanovich, 1978.

[50] J. Bremner and Gavin. *Infant development : recent advances.* Hove : Psychology Press, 1997.

[51] J. Bremner and Gavin. *Infancy.* Oxford : Basil Blackwell, 1988.

[52] J. S. Bruner and R. Watson. *Child's talk : learning to use language.* Oxford : Oxford University Press, 1983.

[53] T.A. Harley. *The Psychology of Language: from data to theory.* Erlbaum UK, Taylor & Francis Publishers, 1994.

[54] A. Billard, K. Dautenhahn, and G. Hayes. Experiments on human-robot communication with robota, an imitative learning and communicating doll robot. In K. Dautenhahn B. Edmonds, editor, *Proceedings of Socially Situated Intelligence Workshop, Zurich CH, as part of the Fifth Conference of the Society of Adaptive Behaviour.* Centre for Policy Modelling technical report series: No. CPM-98-38, 1998.

[55] A. Billard and G. Hayes. Transmitting communication skills through imitation in autonomous robots. In A. Birk and J. Demiris, editors, *Learning Robots: A Multi-Perspective Exploration.* LNAI Series, Springer-Verlag, 1998.

[56] A. Billard and K. Dautenhahn. Grounding communication in situated, social robots. In *Proceedings of TIMR 97, Towards Intelligent Mobile Robots Conference, Manchester.* Tech. Rep. Series, Dept. of Computer Science, Manchester University, ISSN 1361 - 6161. Report number UMCS-97-9-1, 1997. Longer version is published as technical report of the dept. of Artificial Intelligence, University of Edinburgh: Tech No 45.

[57] 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.

[58] H. Yanco and L. A. Stein. An adaptive communication protocol for cooperating mobile robots. In H.L. Roitblat J.-A. Meyer and S.W. Wilson, editors, *From Animals to Animats 2: Proceedings of the Second International Conference on the Simulation of Adaptive Behavior*, pages 478–485. The MIT Press, Bradford Books, 1993.

[59] L. Steels and P. Vogt. Grounding adaptive language games in robotic agents. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life, ECAL97*, pages 473–484. MIT Press, 1997.

[60] T. Arita and Y. Koyama. Evolution of linguistic diversity in a simple communication system. *Artificial Life*, 4:1:9–17, 1998.

[61] S. Kirby and J. Hurford. Learning, culture and evolution in the origin of linguistic constraints. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life, ECAL97.* MIT Press, 1997.

[62] B. MacLennan. Synthetic ethology: an approach to the study of communication. In J. D. Farmer C. G. Langton, C. Taylor and S. Rasmussen, editors, *Proceedings of the Second Artificial Life Workshop*, pages 631–658, 1991.

[63] E. Di Paulo. An investigation into the evolution of communication. *Adaptive Behavior*, 6:2, 1998.

[64] L. Steels. The spontaneous self-organisation of an adaptive language. *Machine Intelligence*, 15, 1996.

[65] S. Harnad. The symbol grounding problem. *Physica D 42*, pages 335–346, 1990.

[66] C. Breazeal and B. Scassellati. Infant-like social interactions between a robot and a human caretaker. *Submitted to a special issue of Adaptive Behavior on Simulation Models of Social Agents*, 1998.