

# The Facilitatory Role of Linguistic Instructions on Developing Manipulation Skills

Gianluca Massera - *Institute of Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy* - [gmassera@istc.cnr.it](mailto:gmassera@istc.cnr.it)

Elio Tuci - *Institute of Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy* - [elio.tuci@istc.cnr.it](mailto:elio.tuci@istc.cnr.it)

Tomassino Ferrauto - *Institute of Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy* - [tomassino.ferrauto@istc.cnr.it](mailto:tomassino.ferrauto@istc.cnr.it)

Stefano Nolfi - *Institute of Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy* - [stefano.nolfi@istc.cnr.it](mailto:stefano.nolfi@istc.cnr.it)

## Abstract

In this paper, we show how a simulated humanoid robot controlled by an artificial neural network can acquire the ability to manipulate spherical objects located over a table by reaching, grasping, and lifting them. The robot controller is developed through an adaptive process in which the free parameters encode the control rules that regulate the fine-grained interaction between the agent and the environment, and the variations of these free parameters are retained or discarded on the basis of their effects at the level of the behaviour exhibited by the agent. The robot develops the sensory-motor coordination required to carry out the task in two different conditions; that is, with or without receiving as input a linguistic instruction that specifies the type of behaviour to be exhibited during the current phase. The obtained results shown that the linguistic instructions facilitate the development of the required behavioural skills.

## I. INTRODUCTION

In this paper, we describe a series of experiments in which a simulated iCub robot acquires through an adaptive process the ability to reach, grasp, and lift a spherical object. The robot develops the sensory-motor coordination required to carry out the whole task in two different conditions; that is, with or without receiving as input linguistic instructions that specify the type of behaviour that should be exhibited during the current phase. These are binary input vectors associated with elementary behaviours that should be displayed by the robot during the task. The main objective of this study is to investigate whether the use of linguistic instructions facilitates the acquisition of a sequence of complex behaviours. The long term goal of this research is to verify whether the acquisition of elementary skills guided by linguistic instructions provides a scaffolding for more complex behaviours.

The first theoretical assumption behind this work is that the activity of developing robots displaying complex cognitive and behavioural skills should be carried out by taking into account the empirical findings in psychology and neuroscience which show that there are close links between the mechanisms of action and those of language. As shown in [1], [2], [3], [4], [5] action and language develop in parallel, influence each other, and base themselves on each other. If brought into the world of robotics, the co-development of action and language skills might enable the transfer of properties of action knowledge to linguistic representations, and vice versa, thus enabling the synthesis of robots with complex behavioural and cognitive skills [6], [7].

The second theoretical assumption behind this work is that behavioural and cognitive skills in embodied agents are emergent dynamical properties which have a multi-level and multi-scale organisation. Behavioural and cognitive skills arise from a large number of fine-grained<sup>1</sup> interactions occurring among and within the robot body, its control system, and the environment [8]. Handcrafting the mechanisms underpinning these skills may be a hard task. This is due to the inherent difficulty in figuring out from the point of view of an external observer, the detailed characteristics of the agent that, as a result of the interactions between the elementary parts of the agent and of the environment, lead to the exhibition of the desired behavior. The synthesis of robots displaying complex behavioural and cognitive skills should instead be obtained through an adaptive process in which the detailed characteristics of the agent are subjected to variation and in which variations are retained or discarded on the basis of their effects at the level of the overall behaviour exhibited by the robot situated in the environment [8]. Therefore, the role of the designer should be limited to the specification of the utility function, that determines whether variations should be preserved or discarded, and eventually to the design of the ecological conditions in which the adaptive process takes place [9], [10], [8].

## II. BACKGROUND AND LITERATURE REVIEW

The control of arm and hand movements in human and nonhuman primates and in robots is a fascinating research topic actively investigated within several disciplines including psychology, neuroscience, and robotics. However, the task to model in detail the mechanisms underlying arm and hand movement control in humans and primates and the task of building robots able to display human-like arm/hand movements still represents an extremely challenging goal [11]. Moreover, despite the progress achieved in robotics

<sup>1</sup>The granularity refers to the extent to which the robot-environmental system is broken into small parts and to the extent to which the dynamics of the system is divided into short time periods. The term fine-grained interactions thus refer to interactions occurring at a high frequency between small parts.

through the use of traditional control methods [12], the attempt to develop robots with the dexterity and robustness of humans is still a long term goal. These difficulties can be explained by considering the need to take into account the role of several aspects including the morphological characteristics of the arm and of the hand, the bio-mechanics of the musculoskeletal system, the presence of redundant degrees of freedom and limits on the joints, non-linearity (e.g., the fact that small variations in some of the joints might have a strong impact on the hand position), gravity, inertia, collisions, noise, the need to rely on different sensory modalities, visual occlusion, the effects of movements on the next experienced sensory states, the need to coordinate arm and hand movements, the need to adjust actions on the basis of sensory feedback, and the need to handle the effects of the physical interactions between the robot and the environment. The attempt to design robots that develop their skills autonomously through an adaptive process permits, at least in principle, to delegate the solutions to some of these aspects to the adaptive process itself.

The research work described in this paper proposes an approach that takes into account most of the aspects discussed above, although often by introducing severe simplifications. More specifically, the morphological characteristics of the human arm and of the hand are taken into account by using a robot that reproduces approximately the morphological characteristics of a 3.5 year-old in term of size, shape, articulations, degrees of freedom and relative limits [13]. Some of the properties of the musculo-skeletal system have been incorporated into the model by using muscle-like actuators controlled by antagonistic motor neurons. For the sake of simplicity, the segments forming the arm, the palm, and the fingers are simulated as fully rigid bodies. However, the way in which the fingers are controlled, enable a certain level of compliance in the hand. The role of gravity, inertia, collision, and noise are taken into account by accurately simulating the physic laws of motion and the effect of collisions (see Section IV for details of the model).

One of the main characteristics of the model presented in this paper is that the robot controller adjusts its output on the basis of the available sensory feedback directly updating the forces exerted on the joints (see [14] for related approaches). The importance of the sensory feedback loop has been emphasised by other works in the literature. For example in [15] the authors describe an experiment in which a three-fingered robotic arm displays a reliable grasping behaviour through a series of routines that keep modifying the relative position of the hand and of the fingers on the basis of the current sensory feedback. The movements tend to optimise a series of properties such as hand-object alignment, contact surface,

finger position symmetry, etc.

In this work, the characteristics of the human brain that processes sensory and proprio-sensory information and control the state of the arm/hand actuators are modelled very loosely through the use of dynamical recurrent neural networks. The architecture of the artificial neural network employed is not inspired by the characteristics of the neuroanatomical pathways of the human brain. Also, many of the features of neurons and synapses are not taken into account (see [16], for an example of works that emulate some of the anatomical characteristics of the human brain). The use of artificial neural networks as robot controller provides several advantages with respect to alternative formalisms, such as robustness, graceful degradation, generalisation and the possibility to process sensory-motor information in a way that is quantitative both in state and time. These characteristics also make neural networks particularly suitable to be used with a learning/adaptive process in which a suitable configuration of the free parameters is obtained through a process that operates by accumulating small variations.

Newborn babies display a rough ability to perform reaching, which evolves into effective reaching and grasping skills by 4/5 months, into adult-like reaching and grasping strategies by 9 months, up to precision grasping by 12/18 months [17], [18], [19]. Concerning the role of sensory modalities, the experimental evidence collected on humans indicates that young infants rely heavily on somatosensory and tactile information to carry out reaching and grasping action and they use vision to elicit these behaviours [20]. However, the use of visual information (employed to prepare the grasping behaviour or to adjust the position of the hand by taking into account the shape and the orientation of the object) starts to play a role only after 9 months from birth [21]. On the basis of this, we provide our robot with proprioceptive and tactile sensors and with a vision system that only provides information concerning the position of the object but not about its shape and its orientation. Moreover, we do not simulate visual occlusions on the basis of the assumption that the information concerning the position of the object can be inferred in a relatively reliable way even when the object is partially or totally occluded by the robot's arm and hand.

In accordance with the empirical evidence indicating that early manipulation skills in infants are acquired through self-learning mechanisms rather than by imitation learning [16], the robot acquires its skills through a trial and error process during which random variations of the free parameters of the robots' neural controller (which are initially assigned randomly) are retained or discarded on the basis of their effect at the level of the overall behaviour exhibited by the robot in interaction with the environment. More precisely, the effect of variations is evaluated using a set of utility functions that determine the extent to

which the robot manages to reach and grasp a target object with its hand, and the extent to which the robot succeeds in lifting the object over the table. The use of this adaptive algorithm and utility functions leaves the robot free to discover during the adaptive process its own strategy to reach the goals set by the experimenter. This in turn allows the robot to exploit sensory-motor coordination (i.e., the possibility to act in order to later experience useful sensory states) as well as the properties arising from the physical interactions between the robot and the environment. In [22] it is shown how this approach allows the robot to distinguish objects of different shapes by self-selecting useful stimuli through action, and in [23] it is shown how this approach allows for the exploiting of properties arising from the physical interaction between the robot body and the environment for the purpose of manipulating the object.

Finally, in this work we shape the ecological conditions in which the robot has to develop its skills by allowing the robot to access linguistic instructions that indicate the type of behaviour that should be currently exhibited by the robot. We do not consider any other form of shaping, such as, for example, the possibility to expose the robot to simplified conditions in some of the trials (in which, for example, the object to be grasped is initially placed within the robot's hand) although we assume that other forms of shaping might favour the developmental process as well.

### III. EXPERIMENTAL SETUP

Our experiments involve a simulated humanoid robot that is trained to manipulate a spherical object located in different positions over a table in front of the robot by reaching, grasping, and lifting it. More specifically the robot is made up of an anthropomorphic robotic arm with 27 actuated degrees of freedom (DOF) on the arm and hand, 6 tactile sensors distributed over the inner part of the fingers and palm, 17 proprioceptors encoding the current angular position of the joints of the arm and of the hand, a simplified vision system that detects the relative position of the object (but not the shape of the object) with respect to the hand and 3 sensory neurons that encode the category of the elementary behaviours that the robot is required to exhibit (i.e., reaching, grasping, or lifting the sphere). The neural controller of the robot is a recurrent neural network trained through an evolutionary algorithm for the ability: (i) to reach an area located above the object, (ii) to wrap the fingers around the object, and (iii) to lift the object over the table. The condition in which the linguistic instructions are provided has been compared with the condition in which the linguistic instructions are not provided. For each condition, the evolutionary process has been repeated 10 times with different random initialisations. The robot and the robot/environmental interactions

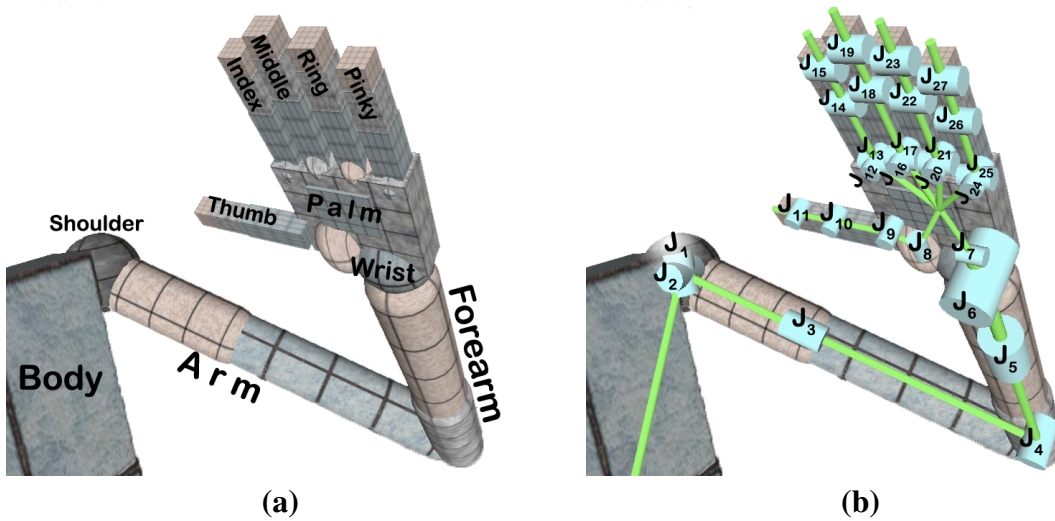


Figure 1. (a) The robot structure and (b) its kinematic chain. Cylinders represent rotational DOFs where its main axis indicates the corresponding axis of rotation; the links amongst cylinders represents the rigid connections that make up the arm structure.

have been simulated by using Newton Game Dynamics (NGD, see: [www.newtondynamics.com](http://www.newtondynamics.com)), a library for accurately simulating rigid body dynamics and collisions. For related approaches, see [23], [22], [24].

In section IV, we describe the structure and the actuators of the arm and hand. In section V, we describe the architecture of the robot controller and the characteristics of the sensors. In section VI, we describe the adaptive process that has been used to train the robot. In section VII, we describe the results obtained, and, finally, in section VIII, we discuss the significance of these results and our plans for the future.

## IV. ROBOT STRUCTURE

### A. Arm Structure

The arm consists mainly of three elements (the arm, the forearm, and the wrist) connected through articulations placed into the shoulder, the arm, the elbow, the forearm and wrist (see figure 1).<sup>2</sup>

The joints  $J_1$ ,  $J_2$  and  $J_3$  provide *abduction/adduction*, *extension/flexion* and *supination/pronation* of the arm in the range  $[-140^\circ, +100^\circ]$ ,  $[-110^\circ, +90^\circ]$  and  $[-110^\circ, +90^\circ]$ , respectively. These three degrees of freedom (DOFs) acts like a ball-and-socket joint moving the arm in a way analogous to the human shoulder joint.  $J_4$ , located in the elbow, is a hinge joint which provides *extension/flexion* within the  $[-170^\circ, +0^\circ]$  range.  $J_5$  twists forearm providing *pronation/supination* of the wrist (and the palm) within  $[-100^\circ, +100^\circ]$ .  $J_6$  and  $J_7$  provide *flexion/extension* and *abduction/adduction* of the hand within  $[-40^\circ, +40^\circ]$  and  $[-100^\circ, +100^\circ]$  respectively (see figure 1).

<sup>2</sup>Details about arm and hand dimensions are available at the supplementary web page <http://laral.istc.cnr.it/esm/linguisticExps>.

## B. Arm Actuators

The arm joints ( $J_1, \dots, J_7$ ) are actuated by two simulated antagonist muscles implemented accordingly to Hill's muscle model [25], [26]. More precisely, the total force exerted by a muscle is the sum of three forces  $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$  which depend on the activity of the corresponding motor neuron ( $\alpha$ ) on the current elongation of the muscle ( $x$ ) and on the muscle contraction/elongation speed ( $\dot{x}$ ) which are calculated on the basis of the following equations:

$$\begin{aligned}
 T_A &= \alpha \left( -\frac{A_{sh} T_{max} (x - R_L)^2}{R_L^2} + T_{max} \right) \\
 A_{sh} &= \frac{R_L^2}{(L_{max} - R_L)^2} \\
 T_P &= T_{max} \frac{\exp\left\{K_{sh} \left(\frac{x - R_L}{L_{max} - R_L}\right)\right\}^{-1}}{\exp\{K_{sh}\} - 1} \\
 T_V &= b \cdot \dot{x}
 \end{aligned} \tag{1}$$

where  $L_{max}$  and  $R_L$  are the maximum and resting lengths of the muscle,  $T_{max}$  is the maximum force that can be generated,  $K_{sh}$  is the passive shape factor, and  $b$  is the viscosity coefficient.

The active force  $T_A$  depends on the activation of muscle  $\alpha$  and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero regardless of the activation  $\alpha$ . At the resting length  $R_L$ , the active force reaches its maximum that depends on the activation  $\alpha$ . The red curves in figure 2 show how the active force  $T_A$  changes with respect to the elongation of the muscle for some possible values of  $\alpha$ . The passive force  $T_P$  depends only on the current elongation/compression of the muscle (see the blue curve in figure 2).  $T_P$  tends to elongate the muscle when it is compressed less than  $R_L$  and tends to compress the muscle when it is elongated above  $R_L$ .  $T_P$  differs from a linear spring for its exponential trend that produces a large opposition to muscle elongation and little to muscle compression.  $T_V$  is the viscosity force. It produces a force proportional to the velocity of the elongation/compression of the muscle.

The parameters of the equation are identical for all 14 muscles controlling the seven DOFs of the arm and have been set to the following values:  $K_{sh} = 3.0$ ,  $R_L = 2.5$ ,  $L_{max} = 3.7$ ,  $b = 0.9$ ,  $A_{sh} = 4.34$  with the exception of parameter  $T_{max}$  which is set to  $3000N$  for joint  $J_2$ ,  $300N$  for joints  $J_1$ ,  $J_3$ ,  $J_4$ , and  $J_5$ , and  $200N$  for  $J_6$  and  $J_7$ .

Muscle elongation is computed by linearly mapping the angular position of the DOF, on which the muscle acts, into the muscle length range. For instance, in the case of the elbow where the limits are  $[-170^\circ, +0^\circ]$ , this range is mapped onto  $[+1.3, +3.7]$  for the agonist muscle and  $[+3.7, +1.3]$  for the

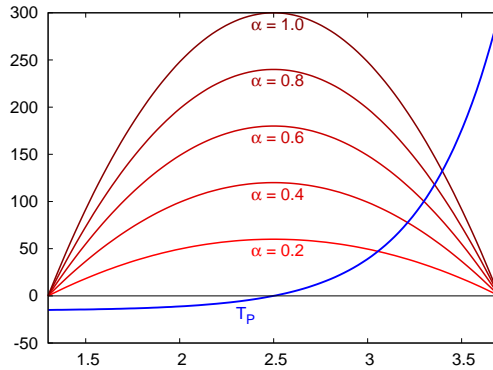


Figure 2. An example of the force exerted by a muscle; the graph shows how the force exerted by a muscle varies as a function of the activity of the corresponding motor neuron and of the elongation of the muscle for a joint in which  $T_{max}$  is set to  $300N$ .

antagonist muscle. Hence, when the elbow is completely extended (angle 0), the agonist muscle is completely elongated (3.7) and the antagonist muscle is completely compressed (1.3), and vice versa when the elbow is flexed.

### C. Hand Structure

The hand is attached to the robotic arm just after the wrist (at joint  $J_7$  as shown in figure 1). One of the most important features of the hand is its compliance. In details, the compliance has been obtained setting a maximum threshold of  $300N$  to the force exerted by each joint. When an external force acting on a joint exceeds this threshold, either the joint cannot move further, or the joint moves backward due to the external force.

The robotic hand is composed of a palm and 15 phalanges that make up the digits (three for each finger) connected through 20 DOFs,  $J_8, \dots, J_{27}$  (see figure 1).

Joint  $J_8$  allows the opposition of the thumb with the other fingers and it varies within the range  $[-120^\circ, +0^\circ]$ , where the lower limit corresponds to thumb-pinky opposition. The knuckle joints  $J_{12}$ ,  $J_{16}$ ,  $J_{20}$  and  $J_{24}$  allow the *abduction/adduction* of the corresponding finger and their ranges are  $[0^\circ, +15^\circ]$  for the index,  $[-2^\circ, +2^\circ]$  for the middle,  $[-10^\circ, +0^\circ]$  for the ring, and  $[-15^\circ, +0^\circ]$  for the pinky. All others joints are for the *extension/flexion* of phalanges and vary within  $[-90^\circ, +0^\circ]$  where the lower limit corresponds to complete flexion of the phalanx (i.e., the finger closed).

### D. Hand Actuators

The joints are not controllable independently of each other, but they are grouped. The same grouping principle used for developing the iCub hand [13] has been used. More precisely, the two distal phalanges



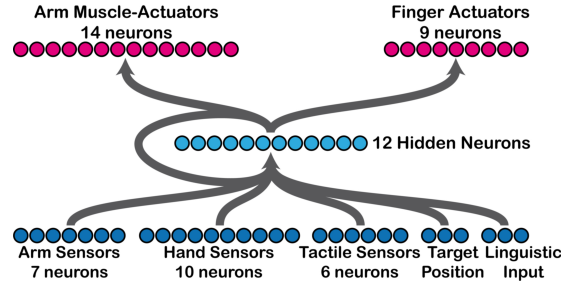


Figure 3. The architecture of the neural controllers. The arrows indicated blocks of fully connected neurons

of the thumb move together as do the two distal phalanges of the index and the middle fingers. Also, all *extension/flexion* joints of the ring and pinky fingers are linked as are all the joints of *abduction/adduction* of the fingers. Hence, only 9 actuators move all the joints of the hand, one actuator for each of the following group of joints:  $\langle J_8 \rangle$ ,  $\langle J_9 \rangle$ ,  $\langle J_{10}, J_{11} \rangle$ ,  $\langle J_{13} \rangle$ ,  $\langle J_{14}, J_{15} \rangle$ ,  $\langle J_{17} \rangle$ ,  $\langle J_{18}, J_{19} \rangle$ ,  $\langle J_{12}, J_{16}, J_{20}, J_{24} \rangle$  and  $\langle J_{21}, J_{22}, J_{23}, J_{25}, J_{26}, J_{27} \rangle$ . These actuators are simple motors controlled by position.

## V. NEURAL CONTROLLER

The architecture of the neural controllers varies slightly depending on the ecological conditions in which the robot develops its skills. In the case of the development supported by linguistic instructions, the robot is controlled by a neural network which includes 29 sensory neurons, 12 internal neurons with recurrent connections and 23 motor neurons. In the case without the support of linguistic instructions, the neural network lacks the sensory neurons dedicated to the linguistic instructions. Thus, it is composed of 26 sensory neurons instead of 29. The sensory neurons are divided into four blocks.

The Arm Sensors encode the current angles of the 7 DOFs located on the arm and on the wrist normalised in the range  $[0, 1]$ .

The Hand Sensors encode the current angles of hand's joints. However, instead of feeding the network with all joint angles of the hand, the following values are used:

$$\left\langle a(J_8), a(J_9), \frac{a(J_{10}) + a(J_{11})}{2}, a(J_{13}), \frac{a(J_{14}) + a(J_{15})}{2}, a(J_{17}), \frac{a(J_{18}) + a(J_{19})}{2}, a(J_{21}), \frac{a(J_{22}) + a(J_{23})}{2}, a(J_{12}) \right\rangle$$

where  $a(J_i)$  is the angle of the joint  $J_i$  normalised in the range  $[0, 1]$  with 0 meaning fully extended and 1 fully flexed. This way of representing the hand posture mirrors the way in which the hand joints are actuated (see section IV-D).

The Tactile Sensors encode how many contacts occur on the hand components. The first tactile neuron corresponds to the palm and its activation is set to the number of contacts normalised in the range  $[0, 1]$

between the palm and another body (i.e., an object or other parts of the hand). Normalisation is performed using a ramp function that saturates to 1 when there are more than 20 contacts. The other five tactile neurons correspond to the fingers and are activated in the same way.

The Target Position Sensors can be seen as the output of a vision system (which has not been simulated) that computes the relative distance in cm of the object with respect to the hand over three orthogonal axes. These values are fed into the networks as they are without any normalisation.

The Linguistic Instruction Sensors is a block of three neurons each of which represents one of the commands *reach*, *grasp* and *lift*. Specifically, the vector  $\langle 50, 0, 0 \rangle$  corresponds to the linguistic instruction “*reach the object*”,  $\langle 0, 50, 0 \rangle$  corresponds to the linguistic instruction “*grasp the object*” and  $\langle 0, 0, 50 \rangle$  corresponds to the linguistic instruction “*lift the object*”. The way in which the state of these sensors is set is determined by equation 4 explained below.

Note that the state of the Linguistic Instruction and Target Position Sensors varies on a larger interval than the other sensors in order to increase the relative impact of these neurons. Indeed, control experiments in which all sensory neurons were normalized within the  $[0, 1]$  interval led to significantly lower performance (result not shown).

The outputs  $H_i(t)$  of the Hidden Neurons are calculated on the basis of following equation:

$$\begin{aligned} y_i(t) &= \sigma \left( \sum_{j=1}^{29} w_{ji} I_j(t) + \beta_i \right) \\ H_i(t) &= \delta_i \cdot y_i(t) + (1 - \delta_i) \cdot y_i(t - 1) \end{aligned} \quad (2)$$

where  $I_j(t)$  is the output of the  $j^{th}$  sensory neuron,  $w_{ji}$  is the synaptic weight from the  $j^{th}$  sensory neuron to the  $i^{th}$  hidden neuron,  $\beta_i$  is the bias of the  $i^{th}$  hidden neuron,  $\delta_i$  is the decay-factor of the  $i^{th}$  hidden neuron, and  $\sigma(x)$  is the logistic function with a slope of 0.2.

The output neurons are divided into two blocks, the Arm Muscle Actuators and the Finger Actuators. All outputs of these neurons are calculated in the same way using the following equation:

$$O_i(t) = \sigma \left( \sum_{j=1}^{12} w_{ji} H_j(t) \right) \quad (3)$$

where  $H_j(t)$  is the output of hidden neuron  $j$  as described in 2,  $w_{ji}$  is the synaptic weight from the  $j^{th}$  hidden neuron to the  $i^{th}$  output neuron and  $\sigma(x)$  is the logistic function with slope 0.2. With respect to the hidden neurons, the output neurons do not have any bias or decay-factor.

The Arm Muscle Actuators output sets the parameter  $\alpha$  used in equation 1 to update the position of

the arm as described in section IV-B while the Finger Actuators output sets the desired *extension/flexion* position of the nine hand actuators as described in IV-D. The state of the sensors, the desired state of the actuators, and the internal neurons are updated every  $10\text{ ms}$ .

This particular type of neural network architecture has been chosen to minimize the number of assumptions and to reduce, as much as possible, the number of free parameters. Also, this particular sensory system has been chosen in order to study situations in which the visual and tactile sensory channels need to be integrated.

## VI. THE ADAPTIVE PROCESS

The free parameters of the neural controller (i.e., the connection weights, the biases of internal neurons and the time constant of leaky-integrator neurons) are set using an evolutionary algorithm [27], [28].

The initial population consists of 100 randomly generated genotypes, which encode the free parameters of 100 corresponding neural controllers. In the conditions in which Linguistic Instruction Sensors are employed (hereafter, referred to as Exp. A), the neural controller has 792 free parameters. In the other condition without the Linguistic Instruction Sensors (hereafter, referred to as Exp. B) there are 756 free parameters. Each parameter is encoded into a binary string (i.e., a gene) of 16 bits. In total, a genotype is composed of  $792 \cdot 16 = 12672$  bits in Exp. A and  $756 \cdot 16 = 12096$  bits in Exp. B. In both experiments, each gene encodes a real value in the range  $[-6, +6]$ , but for genes encoding the decay-factors  $\delta_i$  the encoded value is mapped in the range  $[0, 1]$ .

The 20 best genotypes of each generation are allowed to reproduce by generating five copies each. Four out of five copies are subject to mutations and one copy is not mutated. During mutation, each bit of the genotype has a 1.5% probability to be replaced with a new randomly selected value. The evolutionary process is repeated for 1000 generations.

### A. Fitness Function

The agents are rewarded for reaching, grasping and lifting a spherical object of radius  $2.5\text{ cm}$  placed on the table in exactly the same way in both Exp. A and Exp. B. Each agent of the population is tested 4 times. Each time the initial position of the arm and the sphere change. Figure 4 shows the four initial positions of the arm and of the sphere superimposed on one another. For each initial arm/object configuration, a random displacement of  $\pm 1^\circ$  is added to each joint of the arm and a random displacement of  $\pm 1.5\text{ cm}$  is added on the  $x$  and the  $y$  coordinates of the sphere position. Each trial lasts  $6\text{ sec}$  corresponding to 600

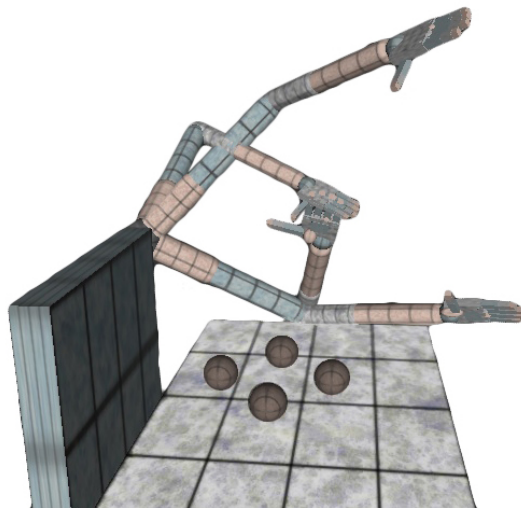


Figure 4. Initial positions of the arm and the sphere over imposed; the joints  $J_1, \dots, J_4$  are initialised to  $\langle -73, -30, -40, -56 \rangle$ ,  $\langle -73, -30, -40, -113 \rangle$ ,  $\langle -6, +30, -10, -56 \rangle$  and  $\langle -73, -30, +45, -113 \rangle$ ; the initial sphere positions are  $\langle -18, +10 \rangle$ ,  $\langle -26, +18 \rangle$ ,  $\langle -18, +26 \rangle$  and  $\langle -10, +18 \rangle$ .

simulation steps. The sphere can move freely and it can eventually fall off the table. In this case, the trial is stopped prematurely.

The fitness function is made up of three components:  $FR$  for reaching,  $FG$  for grasping and  $FL$  for lifting the object. Each trial is divided in 3 phases in which only a single fitness component is updated. The conditions that define the current phase at each timestep and consequently which component has to be updated are the following:

$$\begin{aligned}
 r(t) &= 1 - e^{(-0.1 \cdot ds(t))} \\
 g(t) &= e^{(-0.2 \cdot graspQ(t))} \\
 l(t) &= 1 - e^{(-0.3 \cdot contacts(t))} \\
 Phase(t) &= \begin{cases} reach & r(t) > g(t) \vee g(t) < 0.5 \\ grasp & otherwise \\ lift & g(t) > 0.7 \wedge l(t) > 0.6 \end{cases} \quad (4)
 \end{aligned}$$

where  $ds(t)$  is the distance from the centre of the palm to a point located  $5\text{ cm}$  above the centre of the sphere.  $graspQ(t)$  is the distance between the centroid of the fingertips-palm polygon and the centre of the sphere.  $contacts(t)$  is the number of contacts between the fingers and the sphere. The shift between the three phases is irreversible (i.e. the reach phase is always followed by the reach or grasp phases and the grasp phase is always followed by the grasp or lift phases).

Essentially, the current phase is determined by the values  $r(t)$ ,  $g(t)$  and  $l(t)$ . When  $r(t)$  is high (i.e., when the hand is far from the object) the robot should reach the object. When  $r(t)$  decreases and  $g(t)$  increases (i.e., when the hand approaches the object from above) the robot should grasp the object. Finally, when  $l(t)$  increases (i.e., when the number of activated contact sensors are large enough) the robot should lift the object. The rules and the thresholds included in equation 4 have been set manually on the basis of our intuition and have not been adjusted through a trial and error process. In Exp. A, the phases are used to define which linguistic instruction the robot perceives.

The three fitness components are calculated in the following way:

$$\begin{aligned}
 FR &= \sum_{t \in T_{Reach}} \left( \frac{0.5}{1 + ds(t)/4} + \frac{0.25}{1 + ds(t)} (fingersOpen(t) + palmRot(t)) \right) \\
 FG &= \sum_{t \in T_{Wrap}} \left( \frac{0.4}{1 + graspQ(t)} + \frac{0.2}{1 + contacts(t)/4} \right) \\
 FL &= \sum_{t \in T_{Lift}} objLifted(t)
 \end{aligned}$$

where  $T_{Reach}$ ,  $T_{Wrap}$  and  $T_{Lift}$  are the time ranges determined by equation 4.  $fingersOpen(t)$  correspond to the average degree of extension of the fingers, where 1 occurs when all fingers are extended and 0 when all fingers are closed.  $palmRot(t)$  is the dot product between the normals of the palm and the table, with 1 referring to the condition in which the palm is parallel to the table and 0 to the condition in which the palm is orthogonal to the table).  $objLifted(t)$  is 1 only if the sphere is not touching the table and it is in contact with the fingers, otherwise it is 0.

The total fitness is calculated at the end of four trials as:  $F = \min(500, FR) + \min(720, FW) + \min(1600, FL) + bonus$ , where *bonus* adds 300 for each trial where the agent switches from reach phase to grasp phase only, and 600 for each trial where the agent switches from reach to grasp phase and from grasp to lift phase.

During the reach phase the agent is rewarded for approaching a point located 5 *cm* above the centre of the object with the palm parallel to the table and the hand open. Note that the rewards for the hand opening and the rotation of the palm are relevant only when the hand is near the object (due to  $0.25/(1 + ds(t))$  factor); in this way the agent is free to rotate the palm when the hand is away from the sphere allowing any reaching trajectory.

During the grasp phase, the centroid of the fingertips-palm polygon can reach the centre of the sphere

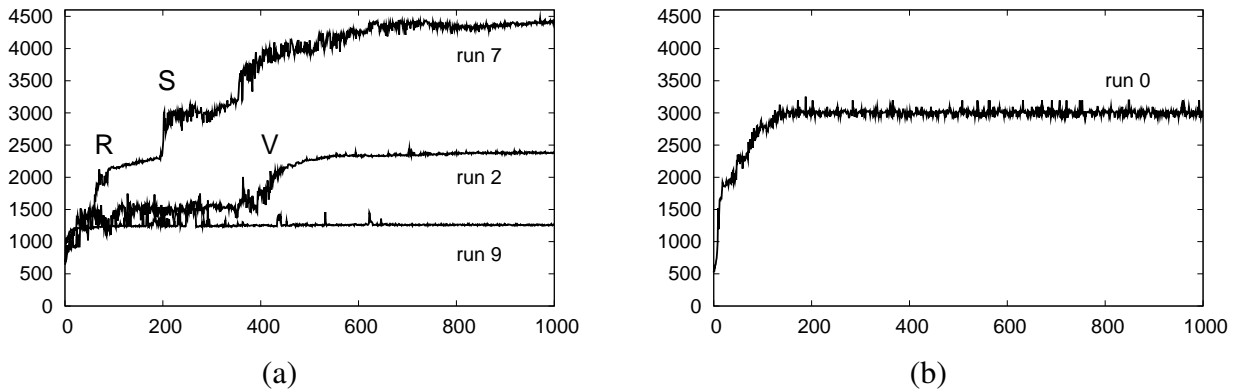


Figure 5. Fitness of the best agents at each generation of (a) run 2, run 7, and run 9 of Exp. A, and (b) run 0 of Exp. B.

only when the hand wraps the sphere with the fingers, producing a potential power grasp. During the lift phase, the reward is given when the agent effectively moves the sphere upward of the table.

## VII. RESULTS

For both Exp. A (with linguistic instructions) and Exp. B (without linguistic instructions), we run 10 evolutionary simulations for 1000 generations, each using a different random initialisation. Looking at the fitness curves of the best agents at each generation of each evolutionary run, we noticed that, for Exp. A, there are three distinctive evolutionary paths (see figure 5a). The most promising is run 7, in which the last generation's agents have the highest fitness. The curve corresponding to run 2 is representative of a group of seven evolutionary paths which, after a short phase of fitness growth, reach a plateau at  $F = 2000$ . The curve corresponding to run 9 is representative of a group of two evolutionary paths which are characterised by a long plateau slightly above  $F = 1000$ . Generally speaking, these curves progressively increase by going through short evolutionary intervals in which the fitness grows quite rapidly followed by a long plateau<sup>3</sup>. For Exp. B, all the runs show a very similar trend, reaching and constantly remaining on a plateau at about  $F = 3000$  (see figure 5b).

Due to the nature of the task and of the fitness function, it is quite hard to infer from these fitness curves what could be the behaviour of the agents during each evolutionary phase. However, based on what we know about the task, and by visual inspection of the behaviour exhibited by the agents, we found out how the agents behave at different generations of each evolutionary run. In Exp. A, the phases of rapid fitness growth are determined by the *bonus* factor, which substantially rewards those agents that successfully accomplish single parts of the task. The first fitness jump is due to the *bonus* factor associated

<sup>3</sup>The fitness curves of the runs not shown are available at the supplementary web page <http://laral.istc.cnr.it/esm/linguisticExps>.

to the execution of a successful reaching behaviour. This jump corresponds to the phase of fitness growth observed in run 7 in correspondence of label R figure 5a, and in run 2 in correspondence of label V figure 5a. The agents generated after these fitness jumps are able to systematically reach the object. Run 9 does not go through the first fitness jump, and the agents of this run lack the ability to systematically carry out a successful reaching behaviour.

The second fitness jump is due to the *bonus* factor associated with the execution of a successful grasping behaviour. Only in run 7 is it possible to observe a phase of rapid fitness growth corresponding to a second fitness jump (see label S figure 5a). The agents generated after this jump are able to successfully carry out reaching and grasping. Note also that, in run 7, the fitness curve keeps on growing until the end of the evolution. This growth is determined by the evolution of the capability to lift the object. Thus, in run 7, the best agents following generation 400 are capable of reaching, grasping, and lifting the object. The constant increment of the fitness is determined by the fact that the agents become progressively more effective in lifting the object. Run 2 does not go through a second fitness jump. The agents of this run lack the ability to systematically carry out a successful grasping behaviour.

In summary, only run 7 has generated agents (i.e., those best agents generated after generation 400) capable of successfully accomplishing reaching, grasping, and lifting.<sup>4</sup> The best agents of run 2, and of the other six runs that show a similar evolutionary trend, are able to systematically reach but not grasp the object and completely lack the ability to lift the object. The best agents of run 9, and of the other run that show a similar evolutionary trend, are not even able to systematically reach the object. In Exp. B, they are able to successfully reach and grasp the object, but not lift it.

#### A. Robustness & Generalisation

In this section, we show the result of a series of post-evaluation tests aimed at establishing the effectiveness and robustness of best agents' behavioural strategies of the four runs show in figure 5. In these tests, the agents, from generation 900 to generation 1000 of each run, are subjected to a series of trials in which the position of the object as well as the initial position of the arm are systematically varied. For the position of the object, we define a rectangular area ( $28\text{ cm} \times 21\text{ cm}$ ) divided in  $11 \times 11$  cells. The agents are evaluated for reaching, grasping and lifting the object positioned in the centre of each cell of the rectangular area. For the initial position of the arm, we use the four initial positions

<sup>4</sup>Movies of the behaviour and corresponding trajectories are available at the supplementary web page <http://laral.istc.cnr.it/esm/linguisticExps>.

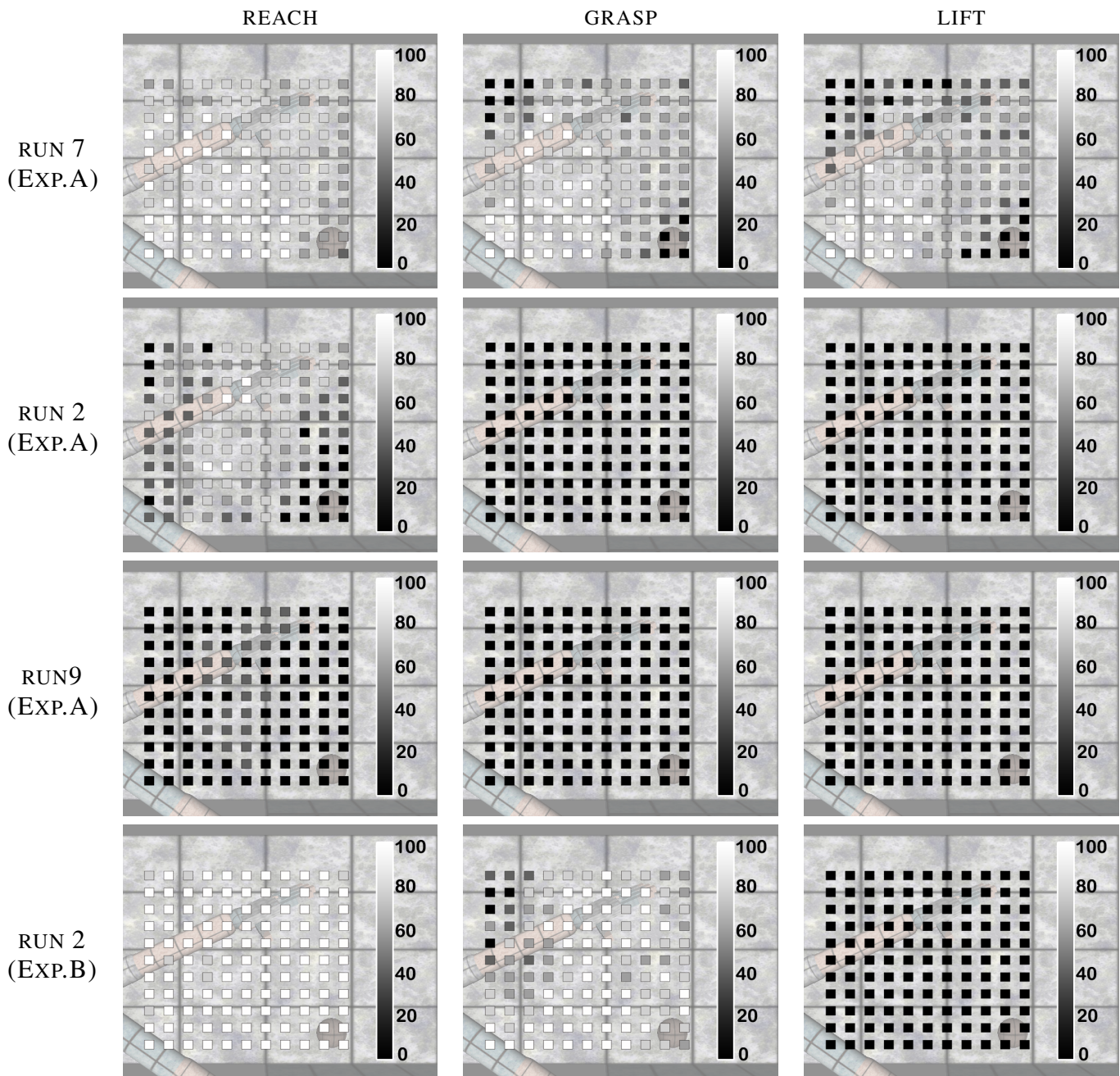


Figure 6. Results of post-evaluation tests on the robustness of reaching, grasping and lifting behaviour of the best agent at generation 1000 of run 7, run 2, and run 9 in Exp. A and run 0 in Exp. B. The cells in shades of grey indicate the percentage of successful trials (from 0% success rate in black, to 100% success rate in white), with the object located in the centre of each cell.

employed during evolution as prototypical cases (see figure 4). For each prototypical case, we generate 100 slightly different initial positions with the addition of a  $\pm 10^\circ$  random displacement on joints  $J_1$ ,  $J_2$ ,  $J_3$ , and  $J_4$ . Thus, this test is comprised of 48400 trials, given by 400 initial positions ( $4 \cdot 100$ ) for each cell, repeated for 121 cells corresponding to the different initial positions of the object during the test. In each trial, reaching is considered successful if an agent meets the conditions to switch from the *reach* phase to the *grasp* phase (see equation 4). Grasping is considered successful if an agent meets the conditions to



switch from the *grasp* phase to the *lift* phase (see equation 4). Lifting is considered successful if an agent manages to keep the object at more than 1 *cm* from the table until the end of the trial. In this section, we show the results of a single agent for each run. However, agents belonging to the same run obtained very similar performances. Thus, the reader should consider the results of each agent as representative of all the other agents of the same evolutionary run.

All the graphs in figure 6 show the relative position of the rectangular area and the cells with respect to the agent/table system. Moreover, each cell of this area is coloured in shades of grey, with black indicating 0% success rate, and white indicating 100% success rate. As expected from the previous section, the agent chosen from run 7 Exp. A proved to be the only one capable of successfully accomplishing all the three phases of the task. This agent proved capable of successfully reaching the object placed almost anywhere within the rectangular area. Its grasping and lifting behaviour are less robust than the reaching behaviour. Indeed, the grasping and lifting performances are quite good everywhere except in two small zones located in the top left and bottom right of the rectangular area in which cells are coloured black. The agent chosen from run 2 Exp. A proved to be capable of successfully performing reaching behaviour for a broad range of object initial positions, and completely unable to perform grasping and lifting behaviour. The agent chosen from run 9 Exp. A does not even manage to systematically bring the hand close to the object regardless of the object's initial position. The agent chosen from run 0 Exp. B, proved capable of successfully performing reaching and grasping behaviour but not lifting behaviour.

## VIII. CONCLUSION

In this paper, we showed how a simulated humanoid robot controlled by an artificial neural network can acquire the ability to manipulate spherical objects located over a table by reaching, grasping and lifting them. The agent is trained through an adaptive process in which the free parameters encode the control rules that regulate the fine-grained interaction between the agent and the environment, and the variations of these free parameters are retained or discarded on the basis of their effects at the level of the behaviour exhibited by the agent. This means that the agents develop their skills autonomously in interaction with the environment. Moreover, this means that the agents are left free to determine the way in which they solve the task within the limits imposed by i) their body/control architecture, ii) the characteristics of the environment, and iii) the constraints imposed by the utility function that rewards the agents for their ability to reach an area located above the object, wrap the fingers around the object, and lift the object. The analysis of the best individuals generated by the adaptive process shows that the

agents of a single evolutionary run manage to reach, grasp, and lift the object in an reliable and effective way. Moreover, when tested in new conditions with respect to those experienced during the adaptive process, these agents proved to be capable of generalising their skills with respect to new object positions never experienced before. The comparison of two experimental conditions (i.e., with or without the use of linguistic instructions that specify the behaviours that the agents are required to exhibit during the task) indicates that the agents succeed in solving the entire problem only with the support of linguistic instructions (i.e., in Exp. A). This result confirms the hypothesis that the possibility to access linguistic instructions, representing the category of the behaviour that has to be exhibited in the current phase of the task, might be a crucial pre-requisite for the development of the corresponding behavioural skills and for the ability to trigger the right behaviour at the right time. More specifically, the fact that the best agents of Exp. B succeed in exhibiting the reaching and then the grasping behaviour but not the lifting behaviour suggests that the linguistic instructions represent a crucial pre-requisite in situations in which the agent has to develop an ability to produce different behaviours in similar sensory-motor circumstances. The reaching to grasping transitions are marked by well differentiated sensory-motor states, which are probably sufficient to induce the agents to stop the reaching phase and to start the grasping phase, even without the support of a linguistic instruction. The grasping to lifting transition is not characterised by well differentiated sensory-motor states. Thus, in Exp. A, it seems to be that the valuable support of the linguistic instruction induces successful agents to move on to the lifting phase.

In future work, we plan to verify whether these agents can be trained to self-generate linguistic instructions and use them to trigger the corresponding behaviours autonomously (i.e., without the need to rely on external instructions). In other words, we would like to verify whether the role played by linguistic instructions can be later internalised in agents' cognitive abilities [29], [30], [31]. Moreover, we plan to port the experiments performed in simulation in hardware by using the iCub robot and the compliant system recently developed [32]. Even though the iCub joints are stiff, the implementation of the muscle model used in this article is still possible. Two 6 axis force sensors placed on the arms and a module developed by the robotcub consortium allow the joints to react as if they were compliant. In this way, it is possible to move the joint applying a torque on its axis and thanks to the opensource aspect of the project, it would be possible to implement muscle actuation directly on the motor control boards.

## IX. ACKNOWLEDGEMENT

This research work was supported by the ITALK project (EU, ICT, Cognitive Systems and Robotics Integrating Project, grant no 214668). The authors thank their colleagues at LARAL for stimulating discussions and feedback during the preparation of this paper.

## REFERENCES

- [1] S. F. Cappa and D. Perani, "The neural correlates of noun and verb processing," *Journal of Neurolinguistics*, vol. 16, no. 2–3, pp. 183–189, 2003.
- [2] A. Glenberg and M. Kaschak, "Grounding language in action," *Psychonomic Bulletin & Review*, vol. 9, pp. 558–565, 2002.
- [3] O. Hauk, I. Johnsrude, and F. Pulvermuller, "Somatotopic representation of action words in human motor and premotor cortex," *Neuron*, vol. 41, no. 2, pp. 301–307, 2004.
- [4] F. Pulvermuller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge University Press, 2003.
- [5] G. Rizzolatti and M. A. Arbib, "Language within our grasp," *Trends in neurosciences*, 1998.
- [6] A. Cangelosi, V. Tikhanoff, J. F. Fontanari, and E. Hourdakis, "Integrating language and cognition: A cognitive robotics approach," *IEEE Computational Intelligence Magazine*, vol. 2, no. 3, pp. 65–70, 2007.
- [7] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. L. Nehaniv, K. Fischer, J. Tani, G. Sandini, L. Fadiga, B. Wrede, K. Rohlfing, E. Tuci, K. Dautenhahn, J. Saunders, and A. Zeschel, "Integration of action and language knowledge: A roadmap for developmental robotics," Tech. Rep., 2010.
- [8] S. Nolfi, "Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour," *Complexus*, vol. 2, no. 3–4, pp. 195–203, 2005.
- [9] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, no. 5504, pp. 599–600, 2001.
- [10] J. Weng, "Developmental robotics: Theory and experiments," *International Journal of Humanoid Robotics*, vol. 1, no. 2, pp. 199–236, 2004.
- [11] S. Schaal, "arm and hand movement control," in *Handbook of brain theory and neural networks*, 2nd ed., M. Arbib, Ed. MIT Press, Cambridge, MA, 2002, pp. 110–113.
- [12] M. Gienger, M. Toussaint, N. Jetchev, A. Bendig, and C. Goerick, "Optimization of fluent approach and grasp motions," in *Proceeding of the 8th IEEE-RAS International Conference on Humanoid Robots*. IEEE Press, 2008, pp. 111–117.
- [13] G. Sandini, G. Metta, and D. Vernon, "Robotcub: An open framework for research in embodied cognition," *International Journal of Humanoid Robotics*, 2004.
- [14] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," in *International Symposium on Robotics Research (ISRR2003)*, S. verlag, Ed., 2004, pp. 1–10.
- [15] J. Felip and A. Morales, "Robust sensor-based grasp primitive for a three-finger robot hand," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009.
- [16] E. Oztop, N. S. Bradley, and M. A. Arbib, "Infant grasp learning: a computational model," *Experimental Brain Research*, vol. 158, no. 4, pp. 480–503, 2004.
- [17] C. von Hofsten, "Eye-hand coordination in the newborn," *Developmental Psychology*, vol. 18, pp. 450–461, 1982.
- [18] ———, "Developmental changes in the organization of prereaching movements," *Developmental Psychology*, vol. 20, pp. 378–388, 1984.

- [19] —, “Structuring of early reaching movements: a longitudinal study,” *Journal of Motor behavior*, vol. 23, pp. 280–292, 1991.
- [20] P. Rochat, “Self-perception and action in infancy,” *Experimental Brain Research*, vol. 123, pp. 102–109, 1998.
- [21] M. K. McCarty, R. K. Clifton, D. H. Ashmead, P. Lee, and N. Goulet, “How infants use vision for grasping objects,” *Child Development*, vol. 72, pp. 973–987, 2001.
- [22] E. Tuci, G. Massera, and S. Nolfi, “Active categorical perception of object shapes in a simulated anthropomorphic robotic arm,” *IEEE Transaction on Evolutionary Computation*, In Press.
- [23] G. Massera, A. Cangelosi, and S. Nolfi, “Evolution of prehension ability in an anthropomorphic neurobotic arm,” *Front. Neurobot.*, vol. 1, pp. 1–9, 2007.
- [24] T. Buehrmann and E. A. Di Paolo, “Closing the loop: Evolving a model-free visually-guided robot arm,” in *Proceedings of the 9th International Conference on the Simulation and Synthesis of Living Systems*, J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. Watson, Eds. MIT Press, Cambridge, MA, 2004, pp. 63–68.
- [25] T. G. Sandercock, D. C. Lin, and W. Z. Rymer, “Muscle models,” in *Handbook of brain theory and neural networks*, 2nd ed., M. Arbib, Ed. MIT Press, Cambridge, MA, 2002, pp. 711–715.
- [26] R. Shadmehr and S. P. Wise, *The Computational Neurobiology of Reaching and Pointing: A Foundation for Motor Learning*. MIT Press, Cambridge, MA, 2005.
- [27] S. Nolfi and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge, MA, 2000.
- [28] X. Yao and M. M. Islam, “Evolving artificial neural network ensembles,” *IEEE Computational Intelligence Magazine*, vol. 3, no. 1, pp. 31–42, 2008.
- [29] L. S. Vygotsky, *Thought and language*. MIT Press, Cambridge, MA, 1962.
- [30] —, *Mind in society*. Harvard University Press, Cambridge, MA, 1978.
- [31] M. Mirolli and D. Parisi, “Towards a vygotskian cognitive robotics: The role of language as a cognitive tool,” *New Ideas in Psychology*, vol. In Press, Corrected Proof, pp. –, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/B6VD4-4X00P73-1/2/5eb2e93d3fc615eea3ec0f637af6fc89>
- [32] V. Mohan, J. Zenzeri, P. Morasso, and G. Metta, “Equilibrium point hypothesis revisited: Advances in the computational framework of passive motion paradigm,” pp. 1–3.