An Experiment on Behaviour Generalisation and the Emergence of Linguistic Compositionality in Evolving Robots

Elio Tuci, Tomassino Ferrauto, Arne Zeschel, Gianluca Massera, and Stefano Nolfi

***Abstract*—Populations of simulated agents controlled by dynamical neural networks are trained by artificial evolution to access linguistic instructions and to execute them by indicating, touching or moving specific target objects. During training the agent experiences only a subset of all object/action pairs. During post-evaluation, some of the successful agents proved to be able to access and execute also linguistic instructions not experienced during training. This owes to the development of a semantic space, grounded in the sensory motor capability of the agent and organised in a systematised way in order to facilitate linguistic compositionality and behavioural generalisation. Compositionality seems to be underpinned by a capability of the agents to access and execute the instructions by temporally decomposing their linguistic and behavioural aspects into their constituent parts (i.e., finding the target object and executing the required action). The comparison between two experimental conditions, in one of which the agents are required to ignore rather than to indicate objects, shows that the composition of the behavioural set significantly influences the development of compositional semantic structures.**

***Index Terms*—Compositional Semantics, Behaviour Generalisation, Evolutionary Robotics, Artificial Neural Networks.**

I. Introduction

Recent research on action and language processing in humans and animals clearly demonstrates the strict interaction and co-dependence between language and action [e.g., 1, 2, 3, 4, 5].

For example, in [3] the authors describe a seminal psychological study showing that the execution of actions (e.g., bringing something close to or far away from to the body) facilitates/disrupts the comprehension of concurrently presented sentences which imply similar/opposite actions (e.g., sentence direction toward/away from the body). According to the authors, the results of this study show that understanding a sentence invokes the same cognitive mechanisms as those used in planning and executing actions. On the neurophysiological side, the authors in [6] performed a study in which by means of single-pulse transcranial magnetic stimulation, either the hand or the foot/leg motor area in the left hemisphere was stimulated in distinct experimental sessions, while participants were listening to sentences expressing hand and foot actions. The results of the study show that processing verbally presented actions activates different sectors of the motor system, depending on the effector used in the described action. The authors conclude that certain action words modulate areas of the brain concerned with performing those actions.

Developmental psychology studies based on emergentist and constructivist approaches also support a view of cognitive development strongly dependent on the contribution of various cognitive capabilities [e.g., 7, 8, 9]. These studies demonstrate the gradual emergence of linguistic constructions built through the child's experience with her social and physical environment. This is in line with the cognitive linguistic assumption that linguistic categorisation involves the same principles and mechanisms that also underlie non-linguistic cognition [see 10, 11].

In recent years, a fruitful exchange of ideas between roboticists and cognitive linguists has begun to develop. On the one hand, more and more language-related research in robotics embraces key ideas of the usage-based language model developed in cognitive linguistics [12, 13]. Several roboticists explicitly acknowledge this framework as their main theoretical inspiration on the language side [e.g., 14, 15, 16, 17]. On the other hand, it is becoming progressively more common for cognitive linguists to draw on insights and suggestions from works on computational modelling [see e.g., 18, 19]. This is especially evident in the field of language acquisition, where computational modelling has become a prominent aspect of the research agenda of various scientists [see 20, 21, 22, for recent reviews].

In this paper, we describe a further robotic model designed to look at aspects related to the emergence of compositional semantic structures in simulated agents. Our results demonstrate how the agents, trained to execute several actions by responding to linguistic instructions, can generalise their linguistic and behavioural skills to never experienced instructions through the production of appropriate behaviours. The analysis of the best agents and the comparison of different experimental conditions, in which the representation of the linguistic instructions is the same but in which the behavioural set is varied, demonstrates how the emergence of compositional semantics is affected by the presence of behavioural regularities in the execution of different actions. Post-evaluation tests also unveil further details of the behavioural and linguistic strategies used by agents equipped with compositional semantics to accomplish the task.

The paper is structured as follow. Section II reviews the

E. Tuci, T. Ferrauto, G. Massera, and S. Nolfi are with the ISTC-CNR, Via San Martino della Battaglia, n. 44, 00185 Rome, Italy, e-mail: {elio.tuci, tomassino.ferrauto, gianluca.massera, stefano.nolfi}@istc.cnr.it (see http://laral.istc.cnr.it/elio.tuci/).

A. Zeschel is with the University of Southern Denmark, Institute of Business Communication and Information Science, Alsion 2, 6400 Sønderborg, Denmark, email: zeschel@uni-bremen.de

most relevant works in the literature and in particular those described in [see 23, 24, 25], which have been particularly inspiring for our work. Section III describes the task investigated in this research work and the agents' morphological structure. In Section IV, Section V, and Section VI, we describe the agent's control system, the evolutionary algorithm and the fitness function used to design it. In Section VII, we illustrate the results of a series of post-evaluation analyses. In Section VIII, we express some reflections on potential connections between empirical studies of child language learning and robotic models trying to indicate fruitful directions for future work. Conclusions are presented in Section IX.

II. Background

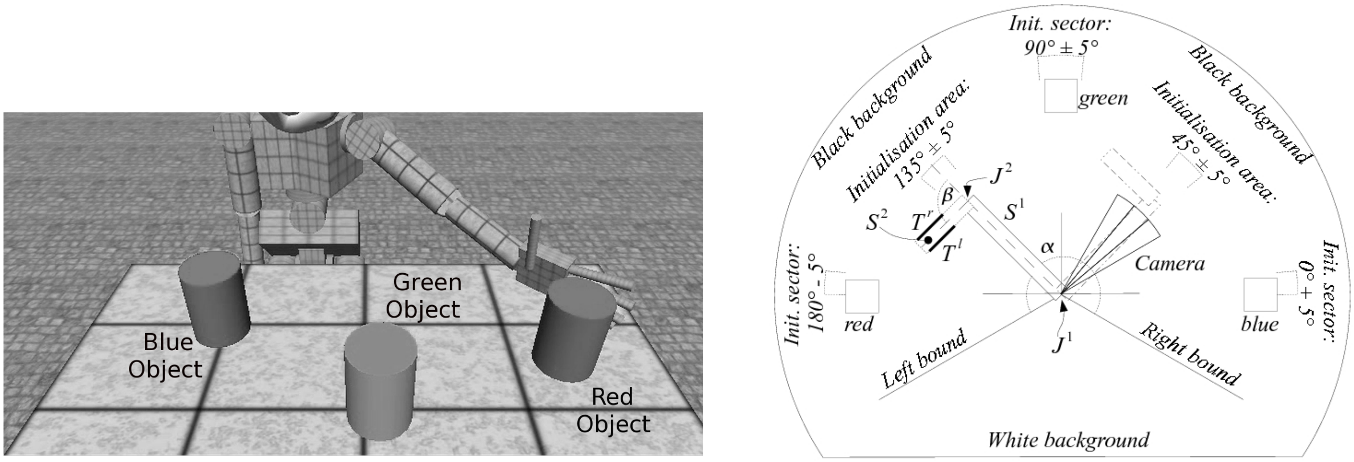
By the term "compositional semantics", we refer to a functional dependence of the meaning of an expression on the meaning of its parts. Compositional semantics in natural language refers to the human ability to understand the meaning of spoken or written sentences from the meaning of their parts, and the way in which these parts are put together. For example, the meaning of an unknown sentence like "Susan likes tulips" can be understood by learning the following three sentences: "Julie likes daisies", "Julie likes tulips", and "Susan likes daisies". In this example, the meaning of the original sentence is achieved through compositional semantics by generalising the meaning of single words from a known (already learnt) to an unknown (yet to be learnt) context.

During the cognitivist era, compositionality was supposed to be underpinned by concatenative processes in which the tokens of an expression's constituents (and the sequential relations among them) are preserved in the expression itself [see 26]. The difficulties shown by classic symbolic AI in accounting for general associations between semantic representations and sensory-motor profiles, and in particular in accounting for the acquisition of linguistic semantics through behavioural experiences, determined a paradigm shift in which an alternative perspective on compositionality emerged [see 27, for a critical perspective on classic AI]. In the last decade of the previous century, the connectionist approach to cognition proposed the idea of functional compositionality; that is compositional semantics systems in which the tokens of an expression's constituents (and the sequential relations among them) are not preserved in the expression itself [see 28]. Various connectionist models proved that artificial neural networks can be employed to physically instantiate functional compositional semantic structures [see 29].

More recently, autonomous (real or simulated) robots have been used to investigate how a form of language can emerge and evolve in a population of robots interacting between themselves and with the physical environment [30, 31, 32, 33]. Moreover, several works have investigated how a robot can acquire a language by interacting with a human user. For example, in [34], the authors designed robotic experiments with robots that, in addition to react to language commands issued by the user are also able to acquire both the meaning of new linguistic instructions and new behavioural skills on the fly, by grounding the new commands in pre-existing motor skills. In [35] the authors designed robots able to cooperate and to share attention with a human user in a restricted experimental setting. This is achieved by allowing the robot to observe the goal-directed behaviour exhibited by the user and to adopt her plan. In [36], the author designed a developmental learning architecture that allows a robot to progressively expand its behavioural repertoire while interacting with a human trainer that shapes its behaviour. In [37], the authors studied how new, higher-order behavioural abilities can be autonomously built upon previously-grounded basic action categories, acquired through language-mediated interactions with human users.

In [23, 24, 25], the authors investigate the issue of grounding compositional semantic structures in an agent's sensory-motor skills in tasks that require the shift from rote knowledge to systematised knowledge. In particular, in [23, 25] a robot learns to execute actions in response to linguistic instructions consisting in two-words sentences. The robots neural controller comprises a behavioural and a linguistic module. The behavioural module is trained through a learning-by-demonstration method in which the sensory-motor states experienced while the robot is moved by the experimenter, through tele-operation or kinesthetic teaching, are used as a training set. The linguistic module is trained to predict the next word of a two-word linguistic instructions in which the words are provided to the agent sequentially. In [25] both the behavioural and the linguistic module are trained only on a subset of all possible linguistic instructions resulting from the combination of all possible objects with all possible actions. In [23], the linguistic module is trained only on a subset of all possible linguistic instructions whereas the behavioural module is trained to execute all the possible instructions. In all three studies [23, 24, 25], the agent proves capable of performing actions associated with linguistic instructions not experienced during training. The authors claim that behavioural and/or linguistic generalisation is achieved by "conceiving something not experienced as a recombination of learnt examples" [see 23, for details]. The contribution of these works is in bringing evidence for a dynamical perspective on compositional semantic systems, alternative to the one in which neural correlates of language are viewed as atomic elements semantically associated to basic units of the linguistics system. The authors show that compositional systems can be underpinned by neural structures in which the neural correlates of the linguistic instructions are dynamically self-organised topological properties of the neural substrate, induced by similarities among sensory-motor sequences. Each instruction (i.e., action plus object) is represented in a two-dimensional semantic space by a single point which lies in a grid-like geometrical structure in which one dimension refers to actions and the other to objects. The geometrical arrangement of neural correlates that emerged during the simultaneous training of the behavioural and linguistic modules, allows the agent to successfully respond to non-experienced linguistic instructions.

In this paper, we describe a series of simulations in which a robot is required to perform a task very similar to the one described in [23]. As in [23], our goal is also to investigate the emergence and the underlying properties of a functionally compositional semantic system in a task that requires the shift



(a) (b)

Fig. 1: (a) An image of the simulated iCub and its word. (b) A schematic representation of the agent structure and its world in the 2D simulator. The vision system of the agent is drawn only with respect to the arm initialised on the right initialisation area. α refers to the angular position of S^1. β refers to the angular position of S^2 with respect to S^1. See text for further details.

from rote knowledge to systematised knowledge. However, we look at the problem with different methods that, as we will see, lead to a qualitatively different type of solution. In our case, a neural controller is trained to execute a subset of possible linguistic instructions through an evolutionary method in which the robot is rewarded for the ability to achieve a certain goal without specifying the sequence of movements through which this goal should be realised. As shown in Section VII, this allows the robot to co-develop linguistic skills to access the meaning of the instructions and behavioural skills to execute them.

III. The agent structure and the task

The experimental scenario concerns a humanoid iCub robot [see 38] placed in front of a table with a red, green, and blue object as shown in Figure 1a. The robot is trained to execute seven actions on the object by responding to linguistic instructions formed by all the possible combinations of the three action words "INDICATE", "TOUCH", and "MOVE" and the three object word "Red", "Green", and "Blue" with the exception of the sentences "TOUCH Green object" and "MOVE Blue object". After training, the robot is then tested on the two *non-experienced* sentences to assess whether it produces the appropriate corresponding behaviours even though it had neither experienced these sentences before nor received training on the two corresponding behaviours. To reduce the computational costs associated to the simulation of such a complex robot, we carried out our experiments on a simpler experimental 2D scenario involving a two-segments arm described below. We then port the obtained results on a simulated iCub by controlling the robots hand position on the basis of the current position of the end-effector of the simplified arm through the inverse kinematic software described in [39]. The best evolved controllers have been successfully ported on the iCub simulator[1].

In the simple two-dimensional simulated world, an agent is composed of an arm with two segments referred to as S^1 (100 cm) and S^2 (50 cm), and two degrees of freedom (DOF). Each DOF comprises a rotational joint which acts as the fulcrum and an actuator. The first actuator causes S^1 to rotate clockwise or anticlockwise around joint J^1, with the movement restricted in the right ($-30°$) and the left ($210°$) bound. The other actuator causes S^2 to rotate clockwise or anticlockwise around joint J^2 within the range [90°, 0°] with respect to S^1 (see Figure 1b). Friction and momentum are not considered.

In the environment there are three objects of different colours (i.e., a blue, a green, and a red object). The objects are placed 150 cm from J^1 with their centre placed anywhere on the chord delimiting their corresponding Init. sector (see Figure 1b). The objects do not move unless pushed by the arm. The agent is equipped with a linear camera with a receptive field of 30°, divided in three sectors, each of which has three binary sensors (C_i^B for blue, C_i^G for green, and C_i^R for red, with $i \in [1, 2, 3]$ sectors). Each sensor returns 1 if the blue/green/red object falls within the corresponding sector. If no coloured object is detected, the readings of the sensors are set to 0 (i.e., the camera perceives a black background). The camera and S^1 move together. The experimental set up is built in a way that at each time step there can be only one object in the camera view.

The agent has means to perceive whenever S^1 reaches the right or the left bound through the activation of the camera sensors. That is, when S^1 reaches the right bound C_1^B, C_1^G,

[1]Movies and further methodological details concerning the porting can be found at http://laral.istc.cnr.it/esm/tuci-etal-IEEE_TAMD2010/.

and C_1^R are set to 1 (i.e., the first camera sector perceives a white background). When S^1 reaches the right bound C_3^B, C_3^G, and C_3^R are set to 1 (i.e., the third camera sector perceives a white background). Finally, two binary touch sensors (i.e., T^r, T^l) are placed on the right, and left side of S^2. Collisions between the agent and an object are handled by a simple model in which whenever S^2 pushes the object the relative contact points remain fixed.

To assess whether the composition of the behavioural set affects the developmental process and the generalisation capabilities of the agents, we run two sets of evolutionary experiments. In the **With-Indicate** experimental condition, the task consists in the execution of the following instructions: TOUCH Blue object ($Inst^T_{blue}$), TOUCH Red object ($Inst^T_{red}$), MOVE Green object ($Inst^M_{green}$), MOVE Red object ($Inst^M_{red}$), INDICATE Blue object ($Inst^{IN}_{blue}$), INDICATE Green object ($Inst^{IN}_{green}$), and INDICATE Red object ($Inst^{IN}_{red}$). In the **With-Ignore** experimental condition, the action INDICATE is substituted with the action IGNORE. Thus, $Inst^{IG}_{blue}$ refers to IGNORE Bkue object, $Inst^{IG}_{green}$ refers to IGNORE Green object, and $Inst^{IG}_{red}$ refers to IGNORE Red object. For both evolutionary conditions, the linguistic instructions experienced during training are referred to as *experienced* instructions, while the instructions TOUCH Green object ($Inst^T_{green}$) and MOVE Blue object ($Inst^M_{blue}$), never experienced during training, are referred to as *non-experienced* instructions (see also Table I). The object-label and the action-label are given to the agent concurrently and for the entire duration of a trial.

TOUCH and MOVE require the agent to rotate S^1 and S^2 until S^2 collides with the target object. TOUCH requires an agent to remain in contact with the target object with the right side of S^2 (that is, by activating the touch sensor T^r) for an uninterrupted interval of 100 time steps. During this interval, S^1 must not rotate. MOVE requires an agent to rotate S^1 more than 35° while S^2 is touching the object with its right side. The rotation of S^1 while S^2 is touching the object determines the movement of the object. INDICATE requires an agent to rotate S^1 until the angular distance between S^1 and the object is less than 30°. INDICATE is correctly executed only if S^1 remains at less than 30° from the target object for more than 100 time steps. IGNORE requires the agent to look at anything except the target object. The agent has to move away from positions in which the target object falls within its visual field. During the execution of INDICATE and IGNORE, an agent must not collide with any object. During the execution of TOUCH and MOVE, an agent must not collide with the non-target objects (i.e., the objects not mentioned in the current linguistic instruction).

After training, all the agents are evaluated for their capability to access *experienced* and *non-experienced* linguistic instructions and to execute the corresponding behaviours.

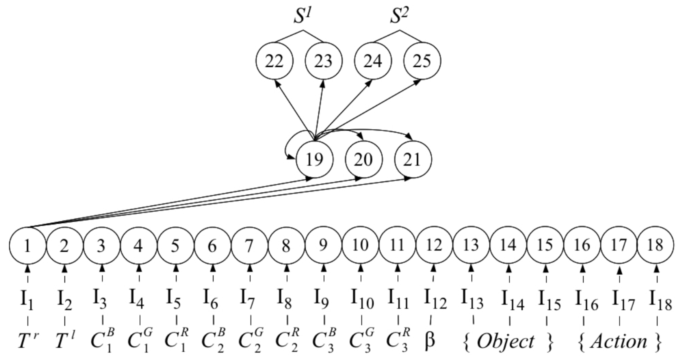


Fig. 2: The neural network. Continuous line arrows indicate the efferent connections for the first neuron of each layer. Underneath the input layer, it is shown the correspondences between sensors/linguistic instructions, the notation used in equation 1 to refer to them, and the sensory neurons.

TABLE I: The linguistic instructions. In grey the *non-experienced* instructions, that is, those not experienced during training. The table also shows the notation used in equation 1 to refer to each bit of the linguistic instructions.

	MOVE $Inst^M_o$					
	Object			Action		
	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}
Blue	1	1	0	0	1	1
Green	1	0	1	0	1	1
Red	0	1	1	0	1	1

	TOUCH $Inst^T_o$					
	Object			Action		
	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}
Blue	1	1	0	1	0	1
Green	1	0	1	1	0	1
Red	0	1	1	1	0	1

	INDICATE $Inst^{IN}_o$ - IGNORE $Inst^{IG}_o$					
	Object			Action		
	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}
Blue	1	1	0	1	1	0
Green	1	0	1	1	1	0
Red	0	1	1	1	1	0

IV. The Agent Controller

The agent controller is composed of a continuous time recurrent neural network (CTRNN) of 18 sensor neurons, 3 inter-neurons, and 4 motor neurons [40]. At each time step sensor neurons are activated using an input vector I_i with $i \in [1, .., 18]$ corresponding to the sensors readings. In particular, I_1 and I_2 are the readings of touch sensors T^r and T^l, respectively; I_3 to I_{11} are the readings of the camera sensors; I_{12} is refers to the normalised angular position of S^2 with respect to S^1 (i.e., β); I_{13} to I_{18} are the linguistic input and their value depend on the current linguistic instruction. I_{13}, I_{14}, and I_{15} identify the object, I_{16}, I_{17}, and I_{18} identify the action to execute (see Fig. 2).

The inter-neuron network is fully connected. Additionally, each inter-neuron receives one incoming synapse from each sensory neuron. Each motor neuron receives one incoming synapse from each inter-neuron. There are no direct connections between sensory and motor neurons. The states of the motor neurons are used to control the movement of S^1 and S^2 as explained later. The states of the neurons are updated using the following equation:

$$\Delta y_i = -y_i + gI_i; \text{ for } i \in \{1,..,18\}; \quad (1)$$

$$\tau_i \dot{y}_i = -y_i + \sum_{j=1}^{21} \omega_{ji}\sigma(y_j + \beta_j); \text{ for } i \in \{19,..,21\}; \quad (2)$$

$$\Delta y_i = -y_i + \sum_{j=19}^{21} \omega_{ji}\sigma(y_j + \beta_j); \text{ for } i \in \{22,..,25\}; \quad (3)$$

with $\sigma(x) = (1 + e^{-x})^{-1}$. In these equations, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i the decay constant, g is a gain factor, I_i the intensity of the perturbation on sensory neuron i, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β_j the bias term, $\sigma(y_j + \beta_j)$ the firing rate (hereafter, f_i). All sensory neurons share the same bias (β^I), and the same holds for all motor neurons (β^O). τ_i and β_i with $i \in \{19,..,21\}$, β^I, β^O, all the network connection weights ω_{ij}, and g are genetically specified networks' parameters. At each time step the angular movement of S^1 is $2.9H(f_{22} - 0.5)sgn(0.5 - f_{23})$ degrees and of S^2 is $2.9H(f_{24} - 0.5)sgn(0.5 - f_{25})$ degrees, where H is the Heaviside step function and sgn is the sign function. Cell potentials are set to 0 when the network is initialised or reset, and equation 2 is integrated using the forward Euler method with an integration time step $\Delta T = 0.1$.

V. The Evolutionary Algorithm

A simple generational genetic algorithm is employed to set the parameters of the networks [41]. At generation 0, a random population of 100 vectors is generated by initialising each component of each vector to a value chosen uniformly random in the range $[0, 1]$. Each vector comprises 84 real values (i.e., 75 connection weights ω_{ji}, 3 decay constants τ_i, 5 bias term β and 1 gain factor g shared by all the sensory neurons). Hereafter, using terms derived from an analogy with biological systems, a vector is referred to as genotype and its components as genes.

Generations following the first one are produced by a combination of selection with elitism and mutation. For each new generation, the three highest scoring genotypes ("the elite") from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection from the 50 best genotypes of the old population. New genotypes, except "the elite", are produced by applying mutation. Mutation entails that a random Gaussian offset is applied to each gene, with a probability of 0.4. The mean of the Gaussian is 0, and its standard deviation is 0.1. During evolution, all genes are constrained to remain within the range $[0, 1]$. That is, if due to mutations a gene falls below zero, its value is fixed to 0; if it rises above 1, its value is fixed to 1.

Genotype parameters are linearly mapped to produce network parameters with the following ranges: $\beta^I \in [-4, -4]$, β^O in$[-5, -5]$, β_i in $[-5, -5]$ with $i \in \{19,..,21\}$, $\omega_{ij} \in [-8, 8]$, with $i \in \{1,..,18\}$, and $j \in \{19,..,21\}$, $\omega_{ij} \in [-10, 10]$, with $i \in \{19,..,21\}$, and $j \in \{19,..,25\}$, gain factor $g \in [1, 13]$. Decay constants τ_i with $i \in \{19,..,21\}$, are firstly linearly mapped into the range $[-1.0, 2.0]$ and then exponentially mapped into $\tau_i \in [10^{-1.0}, 10^{2.0}]$. The lower bound of τ_i corresponds to the integration step-size used to update the controller; the upper bound, arbitrarily chosen, corresponds to about 4% of the maximum length of a trial.

VI. The Fitness Function

During evolution, each genotype is translated into an arm controller and evaluated more than once for all the object-action *experienced* instructions by varying the starting positions. The agents perceive *experienced* instructions and they are required to execute the corresponding behaviours. Agents are evaluated 14 times initialised in the left and 14 times in the right initialisation area, for a total of 28 trials. For each initialisation area, an agent experiences all the *experienced* linguistic instructions twice. The *non-experienced* linguistic instructions $Inst^M_{blue}$ and $Inst^T_{green}$ are never experienced during the training phase. At the beginning of each trial, the agent is randomly initialised in one of the two initialisation area, and the state of the neural controller is reset. A trial lasts 25 simulated seconds ($T = 250$ time steps). A trial is terminated earlier in case the arm collides with a non target object. In each trial k, an agent is rewarded by an evaluation function which seeks to assess its ability to execute the desired action on the target object.

A. **With-Indicate**

In **With-Indicate**, the fitness F^{tot}_k attributed to an agent in trial k is the sum of three fitness components F^1_k, F^2_k, and F^3_k. F^1_k rewards the agent for reducing the angular distance between S^1 and the target object. F^2_k rewards the agent for performing the required action on the target object. F^3_k rewards the agent for extending S^2 when it is perceiving the target object and it is required to touch or to move it.

$$F^{tot} = \tfrac{1}{K}\sum_{k=1}^{K} F^{tot}_k; \quad \text{with } K = 28;\ F^{tot}_k = F^1_k + F^2_k + F^3_k; \quad (4)$$

F^1_k, F^2_k, and F^3_k are computed as follows:

$$F^1_k = \max\left(0, \frac{d^i - d^f}{d^i} \cdot P^1_k, \mathbb{1}_{d^f < 4.6^\circ}\right); \quad (5)$$

where d^i and d^f are respectively the initial (i.e., at $t = 0$) and final (i.e., at the end of the trail k) angular distances between S^1 and the target object and $\mathbb{1}_{d^f < 4.6^\circ}$ is 1 if $d^f < 4.6^\circ$, 0 otherwise. P^1_k is the penalty factor. It is set to 0.6 if the agent collides with a non target object, otherwise to 1.0. The angle between S^1 and the target object o can be measured *clockwise* (α^{clock}_o) or *anticlockwise* (α^{anti}_o). In equation 5, d^i and d^f are the minimum between the clockwise and anticlockwise distance, that is $d = \min\left(\alpha^{clock}_o, \alpha^{anti}_o\right)$.

$$F^2_k = \begin{cases} \dfrac{\textit{steps-on-target}}{\textit{max-steps-on-target}} \cdot P^2_k; & \text{for TOUCH or INDICATE} \quad (6a) \\ \dfrac{\Delta\theta}{\textit{max-angular-offset}} \cdot P^2_k; & \text{MOVE} \quad (6b) \end{cases}$$

where *max-steps-on-target* $= 100$, $P_k^2 = 0$ if $F_k^1 < 1$ otherwise $P_k^2 = 1$, *max-angular-offset* $= 34.4°$. For the action INDICATE, *steps-on-target* refers to the number of time steps during which $F_k^1 = 1$, and S^2 does not touch the target object. For the action TOUCH, *steps-on-target* refers to the number of time steps during which $F_k^1 = 1$, S^2 touches the target object by activating the touch sensor T^r, and S^1 does not change its angular position. $\Delta\theta$ is the angular displacement of the orientation of S^1 recorded while $F_k^1 = 1$, and S^2 is touching the target object by activating the touch sensor T^r.

$$F_k^3 = 1.0 - \frac{\beta}{0.5\pi}; \tag{7}$$

with β corresponding to the angular position of S^2 with respect to S^1. F_k^3 is computed only when the target object is falling within the visual field of the agent and in those trials in which the agent is required to touch or to move the target object. If the current linguistic instruction requires the agent to indicate an object and during the time of a trial in which the agent is not perceiving the target object $F_k^3 = 0$. A trial is terminated earlier if *steps-on-target* = *max-steps-on-target* during the execution of INDICATE or TOUCH and when $\Delta\theta$ = *max-angular-offset* during the execution of MOVE.

B. **With-Ignore**

With-Ignore differs from **With-Indicate** only in the computation of F_k^1 and F_k^2 during the execution of the linguistic instructions IGNORE Blue object $Inst_{blue}^{IG}$, IGNORE Green object $Inst_{green}^{IG}$, and IGNORE Red object $Inst_{red}^{IG}$. During the trials in which an agent is required to IGNORE an object $F_k^1 = 1$ if at the end of the trial the target object does not fall within the visual field of the agent, otherwise $F_k^1 = 0$.

$$F_k^2 = \frac{\textit{steps-out-of-target}}{\textit{max-steps-out-of-target}} \cdot P_k^2; \text{ for IGNORE} \tag{8}$$

where *max-steps-out-of-target* $= 100$, and *steps-out-of-target* refers to the number of time steps during which $F_k^1 = 1$, and S^2 does not touch the target object.

VII. Results

For each experimental condition (**With-Indicate**, and **With-Ignore**), we run ten evolutionary simulations for 6000 generations, each using a different random initialisation. Recall that our objective is to generate agents that are capable of successfully accessing and executing *experienced* linguistic instructions. Moreover, we are interested in investigating whether agents develop semantic structures that are functionally compositional. Agents endowed with functionally compositional semantics should be able to successfully access and execute *experienced* linguistic instructions and to generalise their linguistic and behavioural skills to *non-experienced* instructions (i.e., linguistic instructions never experienced during training). We run two different series of simulations to test whether a different training bears upon the development of the required mechanisms for compositional semantics.

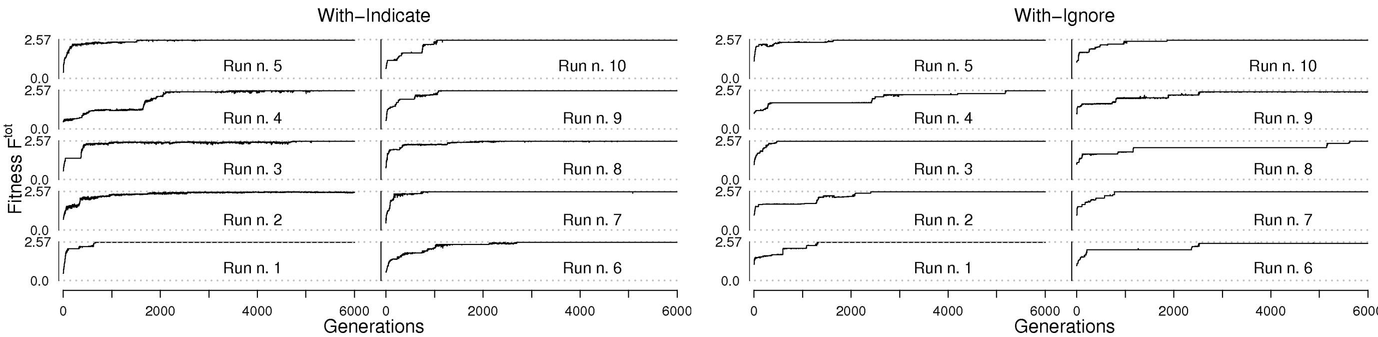
Figure 3 shows the fitness of the best agent at each generation of ten evolutionary Runs per condition. All the curves reach a stable plateau with fitness either firmly fixed or progressing with small oscillation around the maximum score (i.e., $F^{tot} \simeq 2.57$). There are Runs in which the agents reach the maximum fitness very quickly (e.g., Run n° 1 condition **With-Indicate**, or in Run n° 2 condition **With-Ignore**) other in which it takes longer (e.g., Run n° 4 condition **With-Indicate**, or in Run n° 3 condition **With-Ignore**). For all the Runs, before reaching the last fitness plateau, we have periods of very rapid fitness growth induced by the acquisition of new skills to access and execute either entire linguistic instructions or just single linguistic labels. These periods are always followed by either long or short fitness plateaus characterised by rather small oscillations. Just by looking at the fitness curves, we can say that, at the end of the simulation, most of the best agents in both conditions looked capable of correctly solving the linguistic task. However, to estimate the effectiveness and robustness of some of the best evolved agents, with respect to the initial position of the arm, we post-evaluated them for a larger number of trials.

A. First post-evaluation test: Performances on experienced and non-experienced linguistic instructions

In the first post-evaluation test, the best 5 agents of each generation, from generation 4000 to generation 6000, of each evolutionary Run in both conditions, have been repeatedly post-evaluated in each *experienced* and *non-experienced* linguistic instruction. We decided to test the best 5 agents instead of the best one of each generation, because, during evolution, the agents have been ranked according to their fitness, which does not take into account the agent capability to access and execute *non-experienced* linguistic instructions. Recall that *non-experienced* linguistic instructions have not been presented during evolution. Thus, with respect to the capability to access and execute *non-experienced* linguistic instructions, the best agent of each generation may not represent the most effective solution that appeared at each evolutionary time. Overall, 100000 agents per condition have been post-evaluated (i.e., 5 agents, times 2000 generations, times 10 Runs).

During this post-evaluation test, each agent is required to execute 80 times each of the nine instructions (40 trials with the agents randomly initialised in the right initialisation area and, 40 trials in the left one, see also Figure 1b). The position of the objects is also randomly varied as explained in Section III. In each trial k, an agent can be either successful or unsuccessful. It is successful if $F_k^{tot} = 1$, otherwise it is unsuccessful (see equation 4, Section VI for details on F_k^{tot}). At the end of the post-evaluation test, an agent capability to solve the linguistic and behavioural task is represented by nine scores, one for each linguistic instruction. Recall that each score ranges from 0 to 80, and it represents the number of times an agent is successful at the execution of the corresponding linguistic instruction.

It is worth noting that, the results of this test gave us a rather heterogeneous picture, with performances that, even for a single agent, vary remarkably from one linguistic instruction to the other. We felt that readings and interpreting these data by only concentrating on general trends, it would have



(a) (b)

Fig. 3: Graphs showing the fitness curves of the best agent at each generation of ten evolutionary Runs in condition (a) **With-Indicate**; (b) **With-Ignore**.

significantly impoverished the message or this research work. Therefore, we chose a way of representing the results which gives the reader a coherent and exhaustive, although a bit articulated, synthesis of what the post-evaluated agents are capable of doing at the linguistic task. In particular, for each condition, the performances of the agents are compared with respect to four different sub-tasks. For each sub-task, the comparison were accomplished by grouping the 100000 agents in eleven different categories. We first describe what the sub-tasks are and then we explain the meaning of each category.

Sub-task I takes into account only the seven scores recorded during the execution of the *experienced* linguistic instructions.

Sub-task II takes into account the seven scores recorded during the execution of the *experienced* linguistic instructions plus the score recorded during the execution of the *non-experienced* linguistic instruction MOVE Blue object.

Sub-task III takes into account the seven scores recorded during the execution of the *experienced* linguistic instructions plus the score recorded during the execution of the *non-experienced* linguistic instruction TOUCH Green object.

Sub-task IV takes into account all the nine scores (i.e, seven of them for the *experienced* instructions plus two for the *non-experienced* instructions).

For each sub-task, the agents are allocated to one of eleven possible categories (from Cat^0 to Cat^{10}). For a given sub-task, an agent is assigned to Cat^n with $n \in [0,..,10]$, if its lowest score among those considered for that particular sub-task, is within the interval $(80\frac{n-1}{10},..,80\frac{n}{10}]$. Cat^0 comprises all agents that failed to complete a single trial out of 80 attempts on at least one of the instructions. The higher the category, the better the overall performance of the agent. For example, Cat^6 subsumes those agents for whom the lowest score among those considered in a given sub-task is within the interval $(40, 48]$. Cat^7 subsumes those agents for whom the lowest score among those considered in a given sub-task is within the interval $(48, 56]$, etc. Let's consider an agent whose performances at the post-evaluation test are represented by the following nine scores vector (80, 80, 80, 80, 80, 80, 80, 52, 67), in which the first seven scores refer to the performances while executing *experienced* instructions, the eighth score refers to the performance while executing the *non-experienced* instruction TOUCH Green, and the ninth score refers to the performance while executing the *non-experienced* instruction MOVE Blue object. This agent would be assigned to the following categories: i) category Cat^{10} as far as it concerns sub-task I; ii) category Cat^9 as far as it concerns sub-task II; iii) category Cat^7 as far as it concerns sub-task III, and sub-task IV.

Table II shows the number of post-evaluated agents for each category and for each sub-task. These results can be summarised in the following:

- for both conditions, more than half of the post-evaluated agents (about 60% of the agents in **With-Indicate**, and about 66% of them in **With-Ignore**), are perfectly capable of accessing and executing the seven linguistic instruction *experienced* during evolution (see sub-task I, Cat^{10}, condition **With-Indicate**, and **With-Ignore**). This is expected from what was previously observed in the fitness curves shown in Figure 3.
- for both conditions, only a very small number of post-evaluated agents is perfectly capable of accessing and executing all the *experienced* plus one single *non-experienced* linguistic instruction, no matter which one of the two we consider (see Table II, sub-task II, and III, Cat^{10}, condition **With-Indicate**, and **With-Ignore**). The great majority of the agents in sub-task II and III completely fails to access and execute exactly the single *non-experienced* linguistic instruction included in the corresponding sub-task. This has been confirmed by further checks on the data. However, it can also be inferred from the fact that the same agents that are in Cat^{10} for sub-task I tend to be in Cat^0 for sub-tasks II and III.
- for both conditions, only a tiny fraction of the post-evaluated agents is perfectly capable of accessing and executing both the *experienced* and *non-experienced* linguistic instructions (see Table II, sub-task IV, Cat^{10}, **With-Indicate**, and **With-Ignore**).

From these results, it clearly emerges that only a tiny fraction of the post-evaluated agents is capable of accessing

TABLE II: Results of post-evaluation tests showing, for each evolutionary condition, the number of agents for each performance category and for each sub-task. The total number of post-evaluated agents per condition is 100000 (i.e., 5 agents, times 2000 generations, times 10 Runs).

	With-Indicate			
	Sub-task I	Sub-task II	Sub-task III	Sub-task IV
Cat^0	9408	75200	70787	90263
Cat^1	1545	3962	5840	3313
Cat^2	578	1252	2477	1092
Cat^3	823	1314	2174	889
Cat^4	1458	1703	2016	939
Cat^5	3558	2161	8217	2430
Cat^6	2483	2004	1493	346
Cat^7	2780	2061	922	197
Cat^8	5020	1668	957	174
Cat^9	12116	1906	995	135
Cat^{10}	60231	6769	4122	222
Total	100000	100000	100000	100000
	With-Ignore			
Cat^0	8127	87238	92457	98516
Cat^1	15	3502	2439	643
Cat^2	26	1220	1069	218
Cat^3	102	989	1021	220
Cat^4	275	890	928	160
Cat^5	15733	3836	1363	178
Cat^6	451	382	208	15
Cat^7	822	215	145	6
Cat^8	2049	231	141	10
Cat^9	6107	302	121	8
Cat^{10}	66293	1195	108	26
Total	100000	100000	100000	100000

and executing all the linguistic instructions, independently from the initial position of the arm. However, since the number of agents in condition **With-Indicate**, Cat^{10}, sub-task II, III and IV, is significantly different from the number of agents in condition **With-Ignore**, Cat^{10}, sub-task II, III and IV (pairwise Wilcoxon test with 99% confidence interval), we conclude that condition **With-Indicate** facilitates the evolution of agents capable of accessing and executing both *experienced* and *non-experienced* linguistic instructions. In other words, evolutionary pressures to evolve a behavioural repertoire to execute the INDICATE behaviour seem to facilitate the development of compositional semantics. In the next Section, we will further investigate this issue and present a closer look at what makes condition **With-Indicate** more suitable to the evolution of compositional semantic structures.

Obviously, it is important to emphasise the fact that the evolutionary conditions detailed in previous Sections, and in particular those in condition **With-Indicate**, generate the neural mechanisms required by the agents to go beyond what was experienced during evolution. Nevertheless, the fact remains that in either condition, the agents capable of generalising their skills are only a tiny fraction of the agents capable of successfully accomplishing the evolutionary task. This can be explained by the fact that: (i) evolution only seldom produced agents fully capable of generalising their skills; (ii) the selective process does not differentiate between compositional and non-compositional agents since they tend to produce equally good performance with respect to the conditions in which they are evaluated. We noticed that agents capable of generalising appear only in six Runs out of ten, and they are never more than one or two per generation[2]. When they appear, they generally have the highest fitness recorded at that particular generation, which almost always is the highest possible fitness. However, they tend to appear when there are already many more agents with the same fitness in the population that are nevertheless not capable of generalising their linguistic and behavioural skills to *non-experienced* linguistic instructions. The selection mechanism, which can not distinguish on the basis of the fitness alone, agents capable of generalising from those not capable of generalising, tends to favour the latter, to the detriment of the former, simply because the latter are more frequent in the population.

A final point of minor significance is that generalisation capabilities with respect to the MOVE Blue object instruction are more frequent than that with respect to the TOUCH Green object instruction. That is, for both conditions, the number of agents in Cat^{10} sub-task II is significantly different from the number of agents in Cat^{10} sub-task III (pairwise Wilcoxon test with 99% confidence interval). Although we have no empirical explanation for this finding, we know that the action MOVE, which requires the agents to rotate both arms around their joints, is an action that, in evolutionary terms, appears earlier than the capability to TOUCH an object, which requires the agents to stop rotating both arms. At the beginning of the evolution, when the agents' linguistic and behavioural skills are rather simple, it pays more to be able to rotate the arms in order to approach the target objects, rather than to be able to stop a not existing yet rotation of the arms. This evolutionary progression of the behavioural skills may explain why the *non-experienced* instruction which requires to MOVE a target object turns out to be more easily accessible and executable than the *non-experienced* instruction that requires to TOUCH a target object.

B. *Compositionality: Operational principles*

What kind of operational principles do agents employ to be able to access and execute both *experienced* and *non-experienced* instructions? What are the mechanisms underpinning compositional semantics? By visually inspecting the behaviour of some of the agents, we notice that, contrary

[2]Data not shown in the paper can be found at http://laral.istc.cnr.it/esm/tuci-etal-IEEE_TAMD2010/.

to the behaviour of the agents evolved in condition **With-Ignore**, the behaviour of compositional agents evolved in condition **With-Indicate** is the result of the combination of two types of elementary behaviour: an "INDICATE Red object" or "INDICATE Green object", or "INDICATE Blue object" behaviour produced during the first phase of the trial, eventually followed by a "TOUCH" or "MOVE" behaviour, in the second phase of the trial. During the first phase of the trial, regardless of the action to be performed on the object, the agents search the target object by rotating S^1 in order to reduce the angular distance between the target object and S^1, keeping S^2 bent as at start until the target object falls into the agent visual field. During the second phase of the trial, regardless of the target object, the agents rotate S^2 without moving S^1 if TOUCH is required. They rotate S^2 until this segment collides with the target object and then they start rotating S^1 again if MOVE is required. They keep S^1 pointing to the object and S^2 fully bent as at start if INDICATE is required. This qualitative analysis of the behaviour of compositional agents suggests that the agents have developed behavioural skills that, being independent from the particular nature of the linguistic instructions in which they are employed, can be used in contexts already experienced as well as in context not experienced during training.

From this observation, we hypothesised that compositional semantics is underpinned by simple mechanisms by which, during the first part of the trial, the agents regulate their actions on the basis of the object-label and not on the basis of the action-label, and viceversa, during the second part of the trial. This quite intuitive hypothesis suggests that, in any given trial, there may be a first temporal phase, which starts right at the beginning of the trial, in which agents access the part of the linguistic instruction that defines the target object (i.e., the object-label) and act in order to execute the appropriate search behaviour. During this phase, the other part of the linguistic instruction (i.e., the action-label) should not influence the agent's behaviour. The first phase would be followed by a second one, which begins roughly when the target object is visually found. In the second phase, the agents regulate their behaviour on the basis of the action-label only (i.e., the object-label does not have any influence) in case the instruction is TOUCH or MOVE. In the case of INDICATE, instead, the agents keep producing the same behaviour during the entire trial. On this account of compositionality, linguistic instructions not experienced during training (i.e., MOVE Blue object, TOUCH Green object), would be:

- accessed by exploiting the capability to extract from a *non-experienced* instruction already experienced linguistic labels (i.e., TOUCH, MOVE, Blue object, and Green object).
- executed by calling upon known elementary behaviours associated to or triggered by one or the other linguistic label.

In what remains of this Section, we show the results of two post-evaluation tests designed in order to verify whether the agents temporally and functionally decompose the linguistic and behavioural task into two sequential phases as suggested by our hypothesis. These tests are referred to as the *action-transition test* and the *object-transition test*. Both tests follow a similar logic. In the *action-transition test*, the action-label is changed during the course of a trial, while in the *object-transition test*, the object-label is changed during the course of a trial. In both tests, the change takes place in a single time step randomly chosen within a 10 time steps interval which starts at the time when the target object falls within an agent visual field. Based on our hypothesis, agents equipped with compositional semantics are expected to execute the second given action-label and neglect[3] the first given one, at the *action-transition test*. This is because the first given action-label is experienced during the first phase of a trial, when the attention of the agents should be focused on the object-label. At the *object-transition test*, these agents are expected to neglect the second given object-label. This is because this object-label is experienced during a time in which the agents already see the first given target. Consequently, they should pay attention only to the action-label.

The *action-transition test* and the *object-transition test* have been run on a pool of agents selected on their results at the first post-evaluation test (see Section VII-A). In particular, for each evolutionary condition (i.e., **With-Indicate**, and **With-Ignore**), we chose the agents that proved to be more than 75% successful at executing each *experienced* instruction. For the purposes of these tests, these agents have been further selected, and the following three categories have been created: i) *non-compositional* agents, referring to those agents that, at the first post-evaluation test, proved to be less than 10% successful at executing each of the *non-experienced* instructions; ii) *partially-compositional* agents, referring to those agents that, at the first post-evaluation test, proved to be more than 75% successful at executing only one of the two *non-experienced* instructions, and less than 10% successful at executing the other *non-experienced* instructions; iii) *fully-compositional* agents, referring to those agents that, at the first post-evaluation test, proved to be more than 75% successful at executing each of the *non-experienced* instructions.

For both tests, the agents are evaluated 80 times (i.e., 80 trials) on each transition. In half of the trials, the agents are randomly initialised in the right, and in half of the trials, in the left initialisation area. In each trial k, an agent can either succeed, if at the end of the trial $F_k^{tot} = 1$, or fail, if $F_k^{tot} < 1$. Following the logic of each test, the fitness components F_k^1, F_k^2, and F_k^3 are updated with respect to the execution of the second given action-label on the current target object, in the *action-transition test*, and with respect to the execution of the current action-label on the first given target object, in the *object-transition test*. For both tests, an agent's overall performance on each specific transition is considered a success if the agent successfully executes the transition in more than 60 out of 80 trials (i.e., 75% success rate). Since both tests are indiscriminately done on *non-compositional*, *partially-compositional*, and *fully-compositional* agents, we

[3]In this Section, we often take an anthropomorphic stance, by talking about agents that attend or neglect linguistic labels. This is purely for ease of exposition. It is not our intention to claim that the agents are cognitively rich enough to intentionally attend or neglect sensory stimuli.

removed from the two sets of possible transitions, those which, assuming our hypothesis holds, require a response that *non-compositional*, and *partially-compositional* agents are not capable of performing. That is, we remove those transitions which require a MOVE Blue object, or a TOUCH Green object response[4].

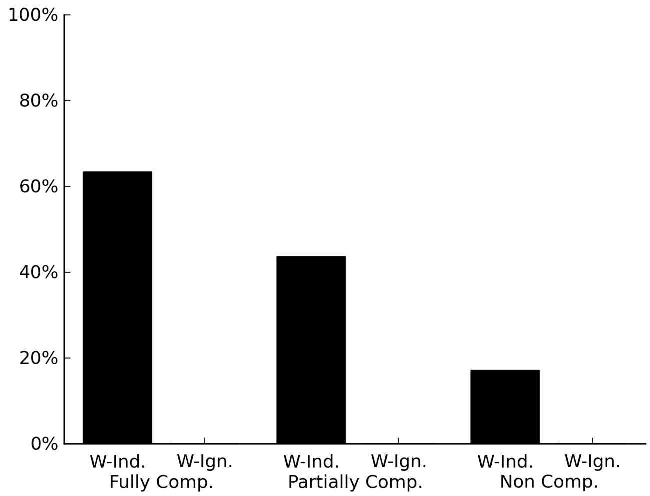
Figure 4a and 4b show the results of the *action-transition test* and of the *object-transition test*, respectively. In both graphs, each bar indicates the percentage of agents that managed to obtain a success rate higher than 75% in all possible transitions of the corresponding test. Black bars refer to the agents evolved in condition **With-Indicate**, white bars refer to the agents evolved in condition **With-Ignore**. Before commenting the results, the reader should be aware of the following. These are quite severe tests since they demands a high success rate on part of the agents on each experienced transition. If our hypothesis on the mechanisms underpinning compositionality holds, we expect *non-compositional* and *partially-compositional* agents to be very bad at least in one of the experienced transitions. This is because we assume that the test can be successfully performed only by agents possessing the capability to functionally and temporally decompose the linguistic and behavioural task into two sequential phases, and that this capability can only be found in *fully-compositional* agents. However, the agents may not need to fully decompose every single trial into two sequential phases in order to be able to successfully access and execute *non-experienced* instructions. In this sense, the test may demand more than what is required to be capable of behavioural and linguistic generalisation. Moreover, in these tests the agents' performance is influenced by whether the label change takes place exactly at the time when the agents switch the focus of their attention from the object-label to the action-label. For methodological convenience, we treated all the agents in the same way, by arbitrarily making this switch in a single time step randomly located in a 10 time steps interval that starts when the agents see the target object. Nevertheless, this may not fully comply with each agent's own strategy, causing failure even in those agents that can functionally and temporally decompose the task.

In spite of these limitations, these graphs tell us several important things. We first concentrate on the results of the *action-transition test*. Figure 4a indicates that the majority of *fully-compositional* agents evolved in condition **With-Indicate**, relies on strategies in which the action-label does not influence the agents' behaviour during the first phase of the task (see Figure 4a, black bar on the left). This suggests that the capability to neglect the action-label while searching for the target object is associated with the presence of compositional semantic structures, since it tends to be observed in *fully-compositional* agents. However, some of the *partially-compositional* and *non-compositional* agents in condition **With-Indicate** proved also capable of accomplishing their task without failing in any transition of the *action-transition test* (see Figure 4a, central and right black bars). Thus, the first conclusion we draw is that neglecting the action-label while reaching the target object is not sufficient to attain compositionality, since it does not allow those *partially-compositional* and *non-compositional* agents that possess it to access and execute *non-experienced* instructions.

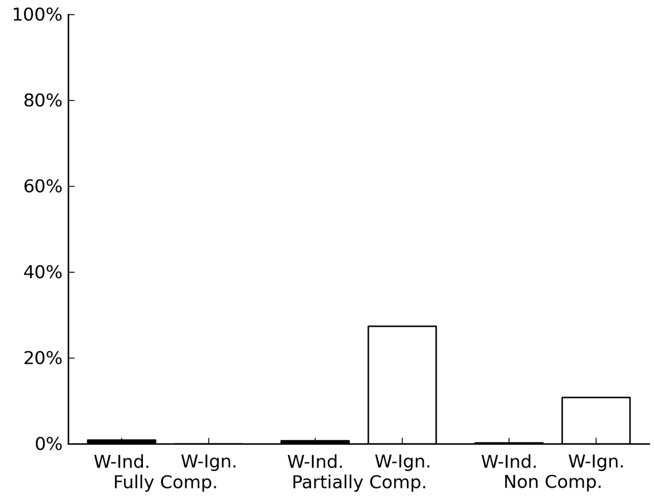
Figure 4a also shows that the capability to cope with the action-label change is completely absent in the agents evolved in condition **With-Ignore**. This result seems to suggest that the significant differences, illustrated in the previous Section, between the two evolutionary conditions in the generation of agents capable of accessing and executing *non-experienced* linguistic instructions, could be explained by the fact that solutions based on the combination of independent elementary behaviours are more rare in the **With-Ignore** condition. Thus, we further conclude that the condition **With-Indicate** seems to contain the evolutionary pressures that facilitate the emergence of compositionality by indirectly favouring those agents whose behaviour is not influenced by the action-label while they reach the target object .

Figure 4b, which refers to the *object-transition test*, tell us that the capability to neglect the object-label during the second phase of a trial, when the target object is already within an agent's visual field, is completely absent in agents evolved in condition **With-Indicate**, and in particular is completely absent in *fully-compositional* agents. Only some of the *partially-compositional* and of the *non-compositional* agents evolved in condition **With-Ignore** seem to be able to cope with the object-label change (see Figure 4b, central and right white bars). How do we explain these results? As far as it concerns the unexpected failure of the *fully-compositional* agents evolved in condition **With-Indicate**, we found out that, contrary to what hypothesised by us, the agents use the object-label during the second phase of the task to keep the target object within their visual field. We observed that, when the object-label does not match what is visually perceived, *fully-compositional*, *partially-compositional*, and *non-compositional* agents perform a search behaviour, loosing visual contact with the object indicated by the first given object-label. Thus, the object-label influences the agents' behaviour during both the first and second phase of a trial, by triggering the agents' response of searching and orienting toward the appropriate object. As far as it concerns the performance of the agents evolved in condition **With-Ignore**, we think that their successes at the *object-transition test* can be explained by considering the evolutionary circumstances in which they evolved. In particular, the action IGNORE can be accomplished by executing a common act for all the objects. Behavioural inspections have indeed demonstrated that *partially-compositional* and *non-compositional* agents evolved in condition **With-Ignore** and capable of coping with the object-label change, once required to IGNORE an object simply don't move at all from their position. This is a strategy which can be successfully applied to execute the action IGNORE regardless of the target object. This may have facilitated the emergence of mechanisms to be able to neglect

[4] In particular, in the *action-transition test*, the transitions experienced by the agents are those in which the second given action-label in combination with the object-label does not produce a *non-experienced* instruction. Similarly, in the *object-transition test*, the transitions experienced by the agents are those in which the first given object-label in combination with the action-label does not produce a *non-experienced* instruction



(a)



(b)

Fig. 4: Graphs showing the results of the (a) *action-transition test*; (b) *object-transition test*. In both graphs, each bar indicates the percentage of agents that managed to obtain a success rate higher than 75% in all possible transitions of the corresponding test. Black bars refer to the agents evolved in condition **With-Indicate**, white bars refer to the agents evolved in condition **With-Ignore**. See the text for the definition of *fully-compositional*, *partially-compositional*, and *non-compositional* agents.

the object-label while executing the required action. However, this is speculative and further analyses are required to test it.

Overall, these tests indicate that in *fully-compositional* agents obtained in condition **With-Indicate**, the "INDICATE Red object", "INDICATE Blue object", and "INDICATE Green object" behaviours are executed during the entire trial, as demonstrated by the fact that the agents are able to search for a new object and then keep indicating it when the object-label is modified during the second phase of the trial. The execution of the "INDICATE" behaviour during the second phase of the trial is not apparent in normal condition (i.e., when the position or the colour of the objects do not change) simply because the execution of this behaviour do not produce any movement. Thus, during the second phase of the trial, when the action label is "INDICATE", agents keep producing the same behaviour. When the action label is "TOUCH" or "MOVE", the agents perform the corresponding elementary behaviour that operates in parallel with the "INDICATE" behaviour. The key mechanism that enables compositionality, therefore, is the fact that the action-label does not affect the agents behaviour during the first part of the trial. In other words, "TOUCH" and "MOVE" behaviours constitute independent behavioural units realised through the execution of the same sequence of micro-actions irrespectively from the object-label. Moreover, we can now state that a different training bears upon the development of the required mechanisms for compositional semantics, and that condition **With-Indicate** facilitates the emergence of compositionality by favouring the emergence of the functional independence of the action-label from the behavioural experience of searching for the target object.

Indeed, by looking at the phylogeny of *fully-compositional* and *partially-compositional* agents in condition **With-Indicate**, we notice that in early stages of their evolutionary history, one of the first behavioural skill to appear is indeed related to the capability of these agents to systematically reduce the angular distance between S^1 and the target object regardless of what type of action the current linguistic instruction is demanding. For example, the ancestors of *fully-compositional* agents, when required to MOVE or to TOUCH an object, they successfully bring S^1 in correspondence of the target object, and they keep S^2 bent until the end of the trial, by systematically failing to execute the action MOVE and TOUCH. In other words, these agents proved to be capable of accessing the linguistic label that defines the object without being able to appropriately execute the linguistic label that defines the TOUCH and MOVE actions. The ability to handle these type of actions is developed later. This can be considered a further evidence that, since the early generation of evolution in condition **With-Indicate**, *fully-compositional* and *partially-compositional* agents learn to decompose the trial into two parts, in the first one of which their behaviour is entirely triggered by the object-label. It is important to note that the early appearance of the capability to decompose the task into two parts is not enforced by any means by the design of the fitness function, it emerges through the dynamics of evolution, and it is facilitated in condition **With-Indicate** by the presence of the instruction INDICATE. However, in the absence of further tests, we can not exclude that these phenomena are induced by design constraints, such as the fact that the segment S^1 and the vision system are bound together. This is because, this particular constraint makes it impossible for an agent to develop a visual search strategy without concurrently acting as required by the instruction INDICATE.

VIII. Discussion: perspectives for research on child language acquisition

Computational approaches to language learning are an intensely researched topic in several disciplines [for recent reviews, cf. 20, 21, 22]. As yet, however, there is still a marked gap between language learning research in cognitive robotics on the one hand and language acquisition studies in computational linguistics on the other. One reason for this is the different thrust of typical research in the two disciplines: in robotics, the focus is commonly on semantic issues to do with the grounding of individual linguistic symbols in agents' sensory-motor experience [42]. In computational linguistics, the focus is usually on structural issues to do with the induction of complex grammars from unrestricted text [43, 44]. In a nutshell, roboticists tend to concentrate on words as carriers of meaning (but neglect their combinatorial properties), while linguists tend to concentrate on their grammar (but neglect their meanings).

Given this apparent opposition, it is interesting to note that a currently influential theory of child language acquisition assumes both a phenomenological continuum and a developmental connection between these two seemingly complementary learning targets (i.e., meaningful "words" and meaningless "rules" in traditional terminology). In usage-based models of language learning, children are assumed to acquire linguistic "rules" (i.e., grammatical categories and constructional patterns thereof) through piecemeal abstractions over utterance-length concrete "words" [i.e., unanalysed holophrastic strings like "there+you+go" and "look+at+this" that are associated with a holistic communicative intention, see 9]. Learners' discovery of the internal structure of these units, coupled with the realisation that the segmented building blocks can be productively recombined within the abstracted constructional patterns, marks the crucial transition from finite lexicons to open-ended grammars. From this perspective, the above experiment is therefore concerned with the emergence of a genuine breakthrough on the way to language.

Needless to say, both the learning target and the learning architecture are substantially less complex here. However, most computational models of language acquisition do not purport to provide an accurate representation of the precise learning mechanisms and processes at work in human children. Rather, the more modest aim is usually to show that it is possible to solve a given task through learning at all (i.e., without innate domain-specific biases). In this way, computational models have made an important contribution to the debate over language learnability, innateness and the purported "poverty of the stimulus" [e.g., 45, 46]. However, none of the models in these debates is grounded in the way that human children's internal representation of language is. In other words, such research has focused on the combinatorial dimension of language alone, but has ignored the additional challenge of linking linguistic structures to the embodied conceptualisations that constitute their meanings. The present study takes steps towards closing this gap, and several of its findings can indeed be related to similar observations made in empirical studies of child language learning.

To better appreciate these connections, it will be helpful to translate aspects of the design into the terminology of usage-based models of child language acquisition. Agents' capacity to correctly access and execute a *non-experienced* linguistic instruction corresponds to their acquisition of an "item-based construction", for example, [move N] in the sense of [9]. As the term "item-based" implies, the generalisations that child language learners have acquired at this developmental stage do not apply across the board. For instance, they may begin to use grammatical marking on some verbs but not on others, indicating that the more inclusive generalisation that both items belong to the same overall category has not yet been formed. Empirical evidence for such item-specific effects in early language acquisition is abundant [cf. 9], and the theoretical vision of a transition from holophrastic units over networks of item-specific "islands" to ever more schematic grammars has also received support from different computational simulations of (non-grounded) language learning [47]. From this perspective, agents' differential performance on the two *non-experienced* instructions in the present experiment does not come as a surprise: also in child language acquisition, the transition from holophrases to compositional grammars is not instantaneous.

Similarly, also the second major finding of this study, that is the significant effect of learning condition (**With-Indicate** vs. **With-Ignore**) on agents' generalisation performance readily translates into a concept of usage-based models of child language learning: if the above assumptions about what makes the behaviour INDICATE more similar to MOVE and TOUCH than IGNORE are plausible, agents' poorer generalisation performance in condition **With-Ignore** would be said to be the outcome of a lower cue consistency (i.e., regularity of form-function mapping) of the category "Verb" in this condition. Furthermore, since such constellations of inconsistency, competition and syncretism are in fact taken to be the norm in natural language processing and learning, a look to usage-based acquisition models in linguistics could also suggest certain useful extensions of the present approach that might be worthwhile to explore in future work (e.g., studying agents' generalisation performance across more than one construction, with or without semantic similarity between actions and/or referents, with balanced or statistically skewed training input, etc.). In other words, we will investigate the characteristics that favour the emergence of compositional solutions (i.e., that ensure behavioural generalisation) and/or that reduce the chance to converge on non-compositional solutions. We will also investigate the possibility to scale the model with respect to the number and the complexity of the linguistic/behavioural repertoire of the agent.

IX. Conclusions

In this paper, we described a robotic model that allows a simulated robot to interact with three coloured objects (a Red, a Green, and a Blue object) located in its peripersonal space by executing three actions (INDICATE, TOUCH, and MOVE) during a series of trials. In each trial, the agent receives as input a linguistic instruction and is rewarded for the ability to exhibit the corresponding behaviour. The results of this study show that dynamical neural networks designed by artificial evolution allow the robot to access and correctly execute the linguistic instructions formed by all the possible combinations of the three action-labels and the three object-labels with the exception of the instructions "TOUCH Green object" and "MOVE Red object", which are *non-experienced* during training. Post-evaluation tests showed that some of the evolved agents generalise their linguistic and behavioural skills by responding to the two *non-experienced* instructions with the production of the appropriate behaviours.

Our study shows that behavioural and linguistic competences can co-evolve in a single non-modularised neural structure in which the semantics is fully grounded in the sensory-motor capabilities of the agents and fully integrated with the neural mechanisms that underpin the agent's behavioural repertoire. Owe to the use of artificial evolution, we leave the agents free to determine how to achieve the goals associated to each linguistic instruction. This allows the agents to organise their behavioural skills in ways that facilitate the development of compositionality thus enabling the possibility to display a generalisation ability at the level of behaviours (i.e., the ability to spontaneously produce behaviours in circumstances that have not been encountered or rewarded during training).

The comparison between two experimental conditions, in one of which the action-label INDICATE is substituted with the action-label IGNORE, shows that the composition of the behavioural set significantly influences the development of solutions that generalise to *non-experienced* instructions. Only individuals evolved in condition **With-Indicate** are characterised by a particularly successful linguistic and behavioural organisation based on the decomposition of the task into two phases, each of which can be associated with the execution of an elementary behaviour. In the first phase only the object-label bears upon the agents' behaviour by triggering the object search strategy. In the second phase, both the object-label and the action-label determine the agents' response. In particular, the object-label keeps an agent eliciting the same behaviour produced during the first phase (i.e., the agent keeps on searching/pointing the target object with the first segment of its arm). At the same time, the action-label triggers a different behaviour that consists in bending the second segment of the arm so to touch or move the object. The capability to decompose the task into two sequential phases as described above, and the use of elementary behaviours employed in different circumstances, are features that, although not sufficient *per se* to explain compositional semantics, they certainly facilitate its evolution.

The use of elementary behavioural skills to generate instructions denoting complex actions resembles the process described in [37], in which the ability to execute more complex linguistic commands, such as GRAB, is acquired by associating two or more previously acquired elementary behaviours (e.g., CLOSE-LEFT-ARM and CLOSE-RIGHT-ARM). However, in [37], the relation between complex and elementary behaviours is established through explicit teaching (i.e., through linguistic input such as: GRAB is CLOSE-LEFT-ARM and CLOSE-RIGHT-ARM). By contrast, in the experiments reported in this paper, behavioural decomposition emerge as a side effect of the need to acquire the ability to execute several related linguistic commands. Moreover, the way in which the agents accomplished the required functionality (i.e., by combining in sequence or in parallel relatively independent behavioural units) represents an important prerequisite for the emergence of compositionality. Therefore, leaving the agents as free as possible to organise how they produce the required skills might be advantageous since it might allow them to decompose the problem in a way that maximise skills re-use. This in turn implies that methods such as the evolutionary method that rewards the agent on the basis of the ability to achieve a given functionality without specifying in details the behaviour that should be produced might be advantageous with respect to alternative methods in that respect.

The results of our post-evaluation analyses also suggests us that there are further distinctive operational principles underpinning compositionality, other than those considered in this work, that are most probably related to the structural and functional characteristics of the agents' neural controller. In future work, we will specifically target these principles, trying to provide a clear description of their nature. Moreover, we mentioned that compositional agents tend to appear very rarely during evolution. It is our intention to work on the characteristics of the task to identify the elements that bear upon the evolutionary origins of agents equipped with compositional semantic structures. With respect to this issue, we think that it may be worth to vary linguistic features and behavioural aspects of the task. For example, in this simulation, the objects have fixed positions with respect to the agent (i.e., Red object on the left, Green object in front, and Blue object on the right of the agent). We wonder whether the necessity to evolved more robust exploration strategies, induced by the variability of the object position relative to the agent, facilitates or hinders the development of compositional structures. Moreover, we are interested in studying whether the use of more cognitively plausible coding schemes, in which the labels are perceived by the agent in a sequential order and just for a short interval of time, bears upon the emergence of compositional semantics. We are also interested in studying whether the development, during training, of a wider and more heterogeneous behavioural repertoire facilitates the emergence of more robust generalisation capabilities.

Acknowledgement

This research work was supported by the *ITALK* project (EU, ICT, Cognitive Systems and Robotics Integrating Project, grant n° 214668). The authors thank Cristina Burani and their colleagues at LARAL for stimulating discussions and feedback during the preparation of this paper.

References

[1] G. Rizzolatti and M. Arbib, "Language within our grasp," *Trends in Neuroscience*, vol. 21, pp. 188–194, 1998.

[2] S. Cappa and D. Perani, "The neural correlates of noun and verb processing," *Journal of Neurolinguistics*, vol. 16, no. 2–3, pp. 183–189, 2003.

[3] A. Glenberg and M. Kaschak, "Grounding language in action," *Psychonomic Bulletin & Review*, vol. 9, pp. 558–565, 2002.

[4] F. Pulvermuller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge University Press, 2003.

[5] V. Gallese, "Mirror neurons and the social nature of language: The neural exploitation hypothesis," *Social Neuroscience*, vol. 3, pp. 317–333, 2008.

[6] G. Buccino, T. Riggio, G. Mellia, F. Binkofski, V. Gallese, and G. Rizzolatti, "Listening to action-related sentences modulates the activity of the motor system: A combined tms and behavioral study," *Cognitive Brain Research*, vol. 24, pp. 355–363, 2005.

[7] M. Bowerman and S. Levinson, *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 2001.

[8] B. MacWhinney, "The emergence of linguistic form in time," *Connection Science*, vol. 17, no. 3–4, pp. 191–211, 2005.

[9] M. Tomasello, *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University, 2003.

[10] G. Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press, 1987.

[11] R. Langacker, *Foundations of Cognitive Grammar*. Stanford, Calif.: Stanford University Press, 1987.

[12] A. Goldberg, *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press, 2006.

[13] R. Langacker, *Cognitive grammar: a basic introduction*. Oxford: Oxford University Press, 2008.

[14] P. Dominey, "From holophrases to abstract grammatical constructions: insights from simulation studies," in *Constructions in Acquisition*, E. Clark and B. Kelly, Eds. Stanford: CSLI Publications, 2006, pp. 137–162.

[15] E. Hutchins and C. Johnson, "Modelling the emergence of language as an embodied collective cognitive activity," *Topics in Cognitive Science*, vol. 1, pp. 523–546, 2009.

[16] S. Wermter, M. Page, M. Knowles, V. Gallese, F. Pulvermller, and J. Taylor, "Multimodal communication in animals, humans and robots: An introduction to perspectives in brain-inspired informatics," *Neural Networks*, vol. 22, no. 2, pp. 111–115, 2009.

[17] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. Nehaniv, K. Fischer, J. Tani, T. Belpaeme, G. Sandini, F. Nori, 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," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 3, pp. 1–28, 2010.

[18] R. W. Langacker, "A dynamic usage-based model," in *Usage-based Models of Language*, M. Barlow and S. Kemmer, Eds. Stanford: CSLI Publications, 2000, pp. 1–63.

[19] A. Goldberg, "Constructions work," *Cognitive Linguistics*, vol. 20, no. 1, pp. 201–224, 2009.

[20] J. Elman, "Computational approaches to language acquisition," in *Encyclopedia of Language and Linguistics*, 2nd ed., K. Brown, Ed. Oxford: Elsevier, 2006, vol. 2, pp. 726–732.

[21] F. Kaplan, P. Oudeyer, and B. Bergen, "Computational models in the debate over language learnability," *Infant and Child Development*, vol. 17, no. 1, pp. 55–80, 2008.

[22] B. MacWhinney, "Computational models of child language learning: an introduction," *Journal of Child Language*, vol. 37, pp. 477–485, 2010.

[23] Y. Sugita and J. Tani, "Learning semantic combinatoriality from the intercation between linguistic and behvioral processes," *Adaptive Behavior*, vol. 13, no. 1, pp. 33–52, 2005.

[24] ——, "Acquiring a functionally compositional system of goal-directed actions of a simulated agent," in *Proc. of the* 10^{th} *Int. Conf. on Simulation of Adaptive Behavior, SAB2008*, M. Asada, J. Hallam, J.-A. Meyer, and J. Tani, Eds. Springer Verlag, Berlin, GE, 2008, pp. 331–341.

[25] H. Arie, T. Endo, S. Jeong, M. Lee, S. Sugano, and J. Tani, "Integrative learning between language and action: A neuro-robotics experiment," in *Proceedings of the* 20^{th} *International Conference on Neural Networks*. Springer Verlag, Berlin, GE, 2010, p. In Press.

[26] J. Fodor and Z. Phylysyn, "Connectionism and cognitive architecture: A critical analysis," *Cognition*, vol. 28, pp. 3–71, 1988.

[27] S. Harnard, "The symbol grounding problem," *Physica D*, vol. 42, pp. 335–346, 1990.

[28] T. Van Gelder, "Compositionality: A connectionist variation on a classic theme," *Cognition*, vol. 14, pp. 355–384, 1990.

[29] J. Elman, "Finding structure in time," *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.

[30] A. Cangelosi and D. Parisi, Eds., *Simulating the evolution of language*. Springer Verlag, New York, 2002.

[31] S. Nolfi and M. Mirolli, Eds., *Evolution of Communication and Language in Embodied Agents*. Springer Verlag, Berlin, GE, 2010.

[32] L. Steels, "Experiments on the emergence of human communication," *Trends in Cognitive Sciences*, vol. 10, no. 8, pp. 347–349, 2006.

[33] E. Tuci, "An investigation of the evolutionary origin of reciprocal communication using simulated autonomous agents," *Biological Cybernetics*, vol. 101, no. 3, pp. 183–199, 2009.

[34] P. Dominey, A. Mallet, and E. Y. E., "Real-time spoken-language programming for cooperative interaction with a humanoid apprentice," *International Journal of Humanoid Robotics*, vol. 6, no. 2, pp. 147–171, 2009.

[35] P. Dominey and F. W. F., "The basis of shared intentions in human and robot cognition," *New Ideas in Psychology*, 2010, in Pres.

[36] J. Weng, "Developmental robotics: theory and experiments," *International Journal of Humanoid Robotics*, vol. 1, no. 2, pp. 199–236, 2004.

[37] A. Cangelosi and T. Riga, "An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots," *Cognitive Science*, vol. 30, no. 4, pp. 673–689, 2006.

[38] G. Sandini, G. Metta, and D. Vernon, "The icub cognitive humanoid robot: An open-system research platform for enactive cognition," in *50 Years of Artificial Intelligence*, M. Lungarella, F. Iida, J. Bongard, and R. Pfeifer, Eds.

Springer Verlag, Berlin, GE, 2007, pp. 358–369.

[39] U. Pattacini, F. Nori, L. Natale, G. Metta, and G. Sandini, "An experimental evaluation of a novel minimum-jerk cartesian controller for humanoid robots," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, p. In Press.

[40] R. D. Beer and J. C. Gallagher, "Evolving dynamic neural networks for adaptive behavior," *Adaptive Behavior*, vol. 1, no. 1, pp. 91–122, 1992.

[41] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.

[42] E. A. Di Paolo, "Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents." *Adaptive Behavior*, vol. 8, no. 1, pp. 27–48, 2000.

[43] N. Chater and C. Manning, "Probabilistic models of language processing and acquisition," *Trends in Cognitive Sciences*, vol. 10, no. 7, pp. 335–344, 2006.

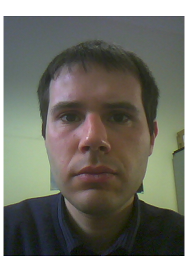
[44] Z. Solan, D. Horn, E. Ruppin, and S. Edelman, "Unsupervised learning of natural languages," *Proc. Natl. Acad. Sci.*, vol. 102, pp. 11 629–11 634, 2005.

[45] F. Reali and M. Christansen, "Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence," *Cognitive Science*, vol. 29, pp. 1007–1028, 2005.

[46] J. Lewis and J. Elman, "Learnability and the statistical structure of language: poverty of stimulus arguments revisited," in *Proceedings of the* 26^{th} *Annual Boston University Conference on Language Development*, 2001.

[47] W. Morris, G. W. Cottrell, and J. Elman, "A connectionist simulation of the empirical acquisition of grammatical relations," in *Hybrid neural symbolic integration*, S. Wermter and R. Sun, Eds. Springer Verlag, Berlin, GE, 2000.

Elio Tuci received a Laurea (Master) in Experimental Psychology from "La Sapienza" University, Rome (IT), in 1996, and a PhD in Computer Science and Artificial Intelligence from University of Sussex (UK), in 2004. His research interests concern the development of real and simulated embodied agents to look at scientific questions related to the mechanisms and/or the evolutionary origins of individual and social behaviour. He is currently a Lecturer in Developmental Robotics in the Department of Computer Science, Aberystwyth University, UK.



Tomassino Ferrauto is a PhD student at the University of Plymouth working under the supervision of Prof. A. Cangelosi and Dr. S. Nolfi. His main research interests are within the domain of evolutionary robotics, active perception, and language and behaviour co-development in robots.

Arne Zeschel studied English and German linguistics at the Universities of Hannover, East Anglia and Bremen. He received his Staatsexamen degree at the University of Hannover in 2001 and his PhD in English linguistics at the University of Bremen in 2007. His main research interests are cognitive linguistics, corpus linguistics and usage-based linguistic model building. He is currently working as a postdoctoral researcher on construction-based approaches to child language acquisition at the University of Southern Denmark.

Gianluca Massera is a PhD student at the Plymouth University working under the supervision of Prof. A. Cangelosi and Dr S. Nolfi. His research interests are within the domain of evolutionary robotics, active perception, and sensory-motor coordination in artificial arms.

Stefano Nolfi is research director at the Institute of Cognitive Sciences and Technologies of the Italian National Research Council (ISTC-CNR) and head of the Laboratory of Autonomous Robots and Artificial Life. His research activities focus on Embodied Cognition, Adaptive Behaviour, Autonomous Robotics, and Complex Systems. He authored or co-authored more than 130 scientific publications and a book on Evolutionary Robotics published by MIT Press.