The Evolution of Behavioural and Linguistic Skills to Execute and Generate Two-word Instructions in Agents Controlled by Dynamical Neural Networks

Elio Tuci, Tomassino Ferrauto, Gianluca Massera and Stefano Nolfi

Institute of Cognitive Sciences and Technologies, ISTC-CNR, Rome, Italy
{elio.tuci, tomassino.ferrauto, gianluca.massera, stefano.nolfi}@istc.cnr.it

Abstract

This paper illustrates an agent-based simulation model focused on the acquisition of linguistic skills. Populations of simulated agents controlled by dynamical neural networks are trained by artificial evolution to perfom two tasks: the behaviour-production task which consists in accessing and executing linguistic instructions; and the behaviour-recognition task which consists in linguistically recognising behaviours. During training the agent experiences only a subset of all linguistic instructions/behaviours. Trained agents successfully acquire an ability to perform both tasks. Moreover some of the successfull agents proved to be able to access and execute also linguistic instructions not experienced during training. However, none of the successfull agents manage to linguistically recognise behaviours corresponding to the execution of linguistic instructions not experienced during training. We conclude by speculating on potential factors that may have inhibited the agents from developing fully compositional semantics structures.

Introduction

The main objective of this study is to design neural mechanisms to allow autonomous agents to develop the linguistic skills necessary to perform both a behaviour-production task and a behaviour-recognition task. The behaviour-production task requires the agents to access linguistic instructions and to correctly execute them. The instructions are made of two parts: a part that defines the type of action, and a part that defines the object on which to perform the action. The behaviour-recognition task requires the agents to observe their own behaviours during the successful execution of each linguistic instruction and to generate the corresponding linguistic instruction (i.e., the object label and the action label).

Successful agents will be further post-evaluated to learn more about the semantics structures underpinning their linguistic skills. We will look at how the development of behavioural and linguistic skills required for the comprehension and the generation of the linguistic instructions changes the way in which the agents represent linguistic labels and attach meaning to them. For example, in the behaviour-production task, we are interested in whether, and eventually at which point in the learning phase, the agents perform the task by exploiting a flexible conceptual system in which object labels and action labels are parsed in a way that even never experienced object-action pair can be conceived as a recombination of previously experienced linguistic elements. In the behaviour-recognition task, we are also interested in whether, and eventually when, the capability of recognising the linguistic instructions associated with the perceived behaviours is underpinned by a compositional semantic system. Owing to this system, previously unexperienced behaviours are seen to be made of elementary behavioural units corresponding to already experienced elementary linguistic labels.

The broad objective of this study is to capture and to systematically investigate, through the use of simulated agent-based modelling, phenomena related to language learning observed in humans. Models of embodied (physical or simulated) agents focused on the study of phenomena related to language learning have become more significant with recent psychological and neuroscientific evidence of close links between the mechanisms of action and those of language (Glenberg and Kaschak, 2002; Gallese, 2008). This is because embodied and situated agent-based models represent a suitable methodological platform to test or to generate various hypothesis concerning the relationship between the development of motor and linguistic skills (Hutchins and Johnson, 2009). In recent years, various types of agent-based models have been employed to generate proof-of-concept demonstrations on how language-like symbolic systems can be acquired by artificial agents through interactions with a physical and/or social environment (e.g., Cangelosi and Parisi, 2002; Steels, 2002; Roy, 2002; Cangelosi and Riga, 2006).

Particularly inspiring for our work is a series of articles specifically focused on the acquisition of a compositional semantics (Sugita and Tani, 2005, 2008). That is, a compositional system grounded on the agent's sensory-motor skills (see Harnard, 1990, for the meaning of grounding in language learning). In (Sugita and Tani, 2005, 2008), the authors investigate this issue on tasks that require the shift from rote knowledge to systematised knowledge. This work has

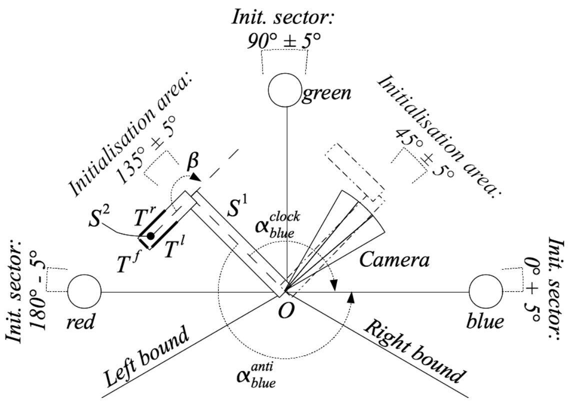

Figure 1: The agent structure and its world. The vision system of the agent is drawn only with respect to the arm initialised on the right initialisation area.

contributed evidence for a dynamical perspective on compositional semantic systems, an alternative perspective to the one in which neural correlates of language are viewed as atomic elements semantically associated to basic units of the linguistics systems (see also Van Gelder, 1990, on this issue).

This study complements previous research on the development of compositional semantics by looking at circumstances in which the development of linguistic skills concerns both the domain of language comprehension and language production. The analysis of the obtained results indicatates that the agents successfully develop a semantic space, grounded on their sensory motor capability and organised in a way that enable linguistic compositionality and generalisation in the case of behaviour generation but not in the case of behaviour recognition. That is, the recognition of behaviour through the production of linguistic instruction seems to be acquired by rote knowledge. We conclude by speculating on potential factors that may have inhibited the agents from developing fully compositional semantics structures.

The task and the agent

Each agent lives in a two-dimensional world and 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. One actuator causes S^1 to rotate clockwise or anticlockwise around point O, with the movement restricted within the right ($-30°$) and the left ($210°$) bound. The other actuator causes S^2 to rotate within the range $[-90°, 90°]$ with respect to S^1. Friction and momentum are not considered (see Fig. 1). In the environment there are three rounded objects of different colours (i.e., a blue, a green, and a red object). The objects are placed at

Table 1: The linguistic instructions. In grey the non-regular instructions, that is, those not experienced during training.

	MOVE $Inst^M_o$					
	Object			Action		
	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}	I_{20}
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_{15}	I_{16}	I_{17}	I_{18}	I_{19}	I_{20}
Blue	1	1	0	1	0	1
Green	1	0	1	1	0	1
Red	0	1	1	1	0	1

	INDICATE $Inst^I_o$					
	Object			Action		
	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}	I_{20}
Blue	1	1	0	1	1	0
Green	1	0	1	1	1	0
Red	0	1	1	1	1	0

150 cm from point O with their centre placed anywhere on the chord delimiting their corresponding Init. sector (see Fig. 1). 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^B_i for blue, C^G_i for green, and C^R_i for red, with $i \in [1, 2, 3]$ sectors). Each sensor returns 1 if the blue/green/red object falls with the corresponding sector. 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. If no coloured object is detected, the readings of the sensors are set to 0. The agent is also equipped with right and left bound binary sensors (B^r and B^l) which activate (i.e., their reading is set to 1) whenever S^1 reaches the right or the left bound, respectively. Finally, three binary touch sensors (i.e., T^r, T^f, T^l) are placed on the right, front, 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.

Agents are trained on both a behaviour-production task and on a behaviour-recognition task. The behaviour-production task consists, for the agents, in the execution of the following instructions (which will be referred to in the remaining part of the paper as *regular 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 ($Inst^I_{blue}$), INDICATE GREEN object ($Inst^I_{green}$), and INDICATE RED object ($Inst^I_{red}$, see also Table 1). 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. During the execution of INDICATE, 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).

The behaviour-recognition task consists, for the agents, in recognising and correctly labelling own behaviours perceived through sequences of α, β duplet. Each duplet corresponds to the angular rotation of the two segments of the arm. In particular, α corresponds to the normalised clockwise angle from S^1 to the axis from O to the lower end position of the blue object Init. sector. β corresponds to the normalised relative rotation of S^2 with respect to S^1 (see Fig. 1). The duplets are recorded during the successful execution of the behaviours at the behaviour-production task.

We run two different series of simulations (referred to as Exp. A and Exp. B) which differ in the training schema. In Exp. A, the agents are evaluated on the behaviour-recognition task only if they successfully perform all the regular instructions during the behaviour-production task. In Exp. B, each agent performs the behaviour-recognition task as soon as it successfully executes at least one regular instruction at the behaviour-production task. In this case, the behaviour-recognition task is limited only to those regular instructions successfully executed at the behaviour-production task. After training, all the agents are evaluated for their capability to access regular and non-regular linguistic instructions and to execute the corresponding behaviours and also for their capability to label behaviours corresponding to the execution of regular and non-regular instructions.

The agent controller and the evolutionary algorithm

The agent controller is composed of a continuous time recurrent neural network (CTRNN) of 22 sensor neurons, 8 inter-neurons and 10 output neurons (Beer and Gallagher, 1992). During the behaviour-production task, at each time step, sensor neurons from 1 to 20 are activated using an input vector I_i with $i \in [1, .., 20]$ corresponding to the sensors readings indicated in Fig. 2, and the input to sensor neuron 21 and 22 is set to 0. During the behaviour-recognition task,

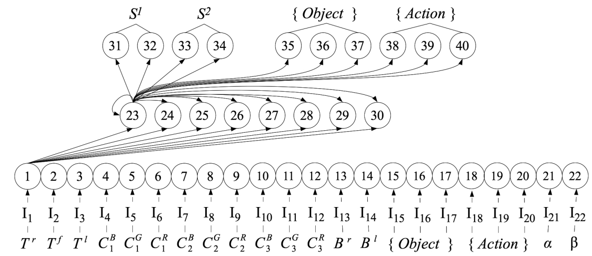

Figure 2: The neural network. Continuous line arrows indicate the efferent connections for the first neuron of each layer. Underneath the input layer, it is shown the correspondences between sensors/linguistic instructions, the notation used in equation 1a to refer to them, and the sensory neurons.

at each time step, the input to sensor neurons 1 to 20 is set to 0, and sensor neurons 21 and 22 are activated using an input vector I_i with $i \in [21, 22]$ corresponding to the α, β generated by successfully executing the linguistic instructions at the behaviour-production task.

The inter-neuron network is fully connected. Additionally, each inter-neuron receives one incoming synapse from each sensory neuron. Each output neuron receives one incoming synapse from each inter-neuron. There are no direct connections between sensory and output neurons. The states of the output 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 equations:

$$\frac{\Delta y}{\Delta T} = \begin{cases} \Big(- y_i + gI_i \Big) \dfrac{1}{\Delta T}; & \text{(1a)} \\ \Big(- y_i + \displaystyle\sum_{j=1}^{30} \omega_{ji} \sigma(y_j + \beta_j) \Big) \dfrac{1}{\tau_i}; & \text{(1b)} \\ \Big(- y_i + \displaystyle\sum_{j=23}^{30} \omega_{ji} \sigma(y_j + \beta_j) \Big) \dfrac{1}{\Delta T}; & \text{(1c)} \end{cases}$$

for $i \in \{1, .., 22\}$ in eq. 1a, for $i \in \{23, ..., 30\}$ in eq. 1b, for $i \in \{31, .., 40\}$ in eq. 1c, and 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 output neurons (β^O). τ_i and β_i with $i \in \{23, ..., 30\}$, 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 $2.9H(f_{31} - 0.5)sgn(0.5 - f_{32})$ degrees and of S^2 is $2.9H(f_{33} - 0.5)sgn(0.5 - f_{34})$ degrees, where H is the Heaviside step function and sgn is the sign function.

A generational genetic algorithm is employed to set the parameters of the networks (Goldberg, 1989). The population contains 100 genotypes. Generations following the first one are produced by a combination of selection with elitism, recombination and mutation. For each new generation, the five highest scoring individuals from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection from the 70 best individuals of the old population. Each genotype is a vector comprising 340 real values. At the beginning of the evolutionary process, each gene is chosen randomly from a uniform distribution in the range $[0, 1]$. Cell potentials are set to 0 when the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration time step $\Delta T = 0.1$.

The fitness function

During evolution, each genotype is translated into an arm controller and evaluated more than once for all the object-action regular instructions by varying the starting positions. The agent fitness is computed on both the behaviour-production task and the behaviour-recognition task ($F^{total} = F^{production} + F^{recognition}$, see below for details).

The behaviour-production task

During the behaviour-production task, the agents perceive regular 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 2 times all the regular linguistic instructions. The 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 12 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. The final fitness $F^{production}$ attributed to an agent is the sum of two fitness components F^1_k and F^2_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.

$$F^{production} = \frac{1}{28}\sum_{k=1}^{28}(F^1_k + F^2_k); \tag{2}$$

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

$$F^1_k = \max\left(0, \frac{d^i - d^f}{d^i} \cdot P^1_k, \mathbb{1}_{d^f < 4.6^\circ}\right); \tag{3}$$

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, which is set to 0.6 if the agent collides with a non target object, to 1.0 otherwise. The angle between S^1 and the target object o can be measured *clockwise* (α^{clock}_o) or *anticlockwise* (α^{anti}_o). In equation 3, d^i and d^f are the minimum between the clockwise and anticlockwise distance, that is $d = \min\left(\alpha^{clock}_o, \alpha^{anti}_o\right)$.

$$F^2_k = \begin{cases} \dfrac{\textit{steps-on-target}}{\textit{max-steps-on-target}} \cdot N \cdot P^2_k; & \text{for TOUCH or INDICATE} \quad (4a) \\ \dfrac{\Delta\theta}{\textit{max-angular-offset}} \cdot N \cdot P^2_k; & \text{MOVE} \quad (4b) \end{cases}$$

where *max-steps-on-target* $= 100$, $P^2_k = 0$ if $F^1_k < 1$ otherwise $P^2_k = 1$, *max-angular-offset* $= 34.4°$, $N = 2$ for TOUCH and MOVE, and $N = 1$ for INDICATE. For the action INDICATE, *steps-on-target* refers to the number of time steps during which $F^1_k = 1$, 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 $F^1_k = 1$, 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 $F^1_k = 1$, and S^2 is touching the target object by activating the touch sensor T^r. 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.

The behaviour-recognition task

During the behaviour-recognition task, the agent is evaluated for labelling its behaviours corresponding to the successful execution of each of the regular instructions. That is, the arm of the agent is moved so as to display a behaviour previously exhibited during the behaviour-production task by the agent itself, and it is asked to produce the corresponding linguistic instruction (without receiving it as input).

In Exp. A, an agent moves on to the behaviour-recognition task only if it successfully completes all the trials of the behaviour-production task (i.e., $F^{production} > 2.57$) . In Exp. B, an agent moves on to the behaviour-recognition task as soon as it successfully completes at least one trial at the behaviour-production task (i.e., $\exists k|(F^1_k + F^2_k) > 2.57$). The behaviour-recognition task comprises only the trial/s successfully executed at the behaviour-production task. In other words, in Exp. A, the evolution of the mechanisms to accomplish the behaviour-recognition task follows the evolution of the mechanisms to successfully execute the behaviour-production task. In Exp. B, the evolution of the mechanisms for the behaviour-production task and the behaviour-recognition task evolve simultaneously, since it suffices for an agent to successfully complete a single trial of the behaviour-production task to move on to the behaviour-recognition task.

In each trial k, the functions F_k^{obj} and F_k^{act} reward the agents for matching with the firing rate of the output neurons 35, 36, 37, 38, 39, and 40 the six digit regular instruction that triggered the currently experienced successful behaviour. F_k^{obj} and F_k^{act} are computed as follow:

$$F^{recognition} = \frac{1}{28}\sum_{k=1}^{K}(F_k^{obj} + F_k^{act});$$

$$F_k^3 = \frac{\sum_{t=T-5}^{T}\left(2^{-2\cdot rank_{k,t}} + \frac{2(1-f_{k,t}^{r}) + \sum_{i\in W_{k,t}} f_i}{4}\right)}{2\cdot 5}; \quad (5)$$

with $F_k^{obj} = F_k^3$ with $W_{k,t}$ the subset of output neurons defining the object label (i.e., neurons 35, 36, and 37) whose activation should be 1, $f_{k,t}^{r}$ the firing rate of the neuron defining the object label whose activation should be 0, $rank_{k,t}$ the rank of $f_{k,t}^{r}$ when the output neurons defining the object label are ranked in ascending firing rate order. F_k^{act} is computed as F_k^{obj} considering the output neurons defining the action label (i.e., neurons 38, 39, 40). $F_k^3 = 0$ if $(F_k^1 + F_k^2) < 2.57$ (i.e. if the behaviour at trial k has not been correctly executed).

Results

For each experimental condition (Exp. A, Exp. B), we run ten evolutionary simulations for 10000 generations, each using a different random initialisation. Recall that our objective is to generate agents that are capable of successfully performing both the behaviour-production task and the behaviour-recognition task. Moreover, we are interested in investigating whether successful agents develop semantic structures that are functionally compositional. Agents endowed with a functionally compositional semantics should be able to access and execute linguistic instructions never experienced during training (i.e., from non-regular instructions to the execution of the corresponding behaviours). They may also be able to linguistically describe a behaviour never performed/experienced during training (i.e., from the perception of behaviours never executed during training to the generation of non-regular instructions). We run two different series of simulations (i.e., Exp. A and Exp. B) to see whether a different training bears upon the development of functionally compositional neural structures.

The best agents of each generation in both experimental conditions have been post-evaluated by first running sets of 80 trials for each regular and non-regular linguistic instruction in which the agents are asked to perform the behaviour-production task. Hereafter, we refer to this first phase of the post-evaluation test as behaviour-production test. In half of the trials the agents are randomly initialised in the right initialisation area and half of the trials in the left one (see Fig 1). We considered those agents successful at the behaviour-production test (hereafter, referred to as *b-successful*) that manage to obtain a success rate higher than 80% in performing the behaviours corresponding to the execution of the regular linguistic instructions (i.e., those experienced during evolution). *b-successful* agents have been further classified into i) *b-non-compositional* agents, referring to those *b-successful* agents that proved to be less than 80% successful at performing the behaviour corresponding to the execution of both the non-regular instructions, $Inst_{blue}^{M}$ and $Inst_{green}^{T}$; ii) *b-partially-compositional* agents referring to those *b-successful* agents that proved to be more than 80% successful at performing the behaviour corresponding to the execution of only one of the two non-regular instructions, $Inst_{blue}^{M}$ or $Inst_{green}^{T}$; iii) *b-fully-compositional* agents re-

Table 2: Result of post-evaluation tests performed on the best agents of each generation for four runs of Exp. A, and for two runs of Exp. B. The tables show the number of successful agents at the behaviour-production task on regular linguistic instructions, and the percentage of them also successful on the non-regular instructions. The tables also show the number of successful agents at the behaviour-recognition task on regular linguistic instructions, and the percentage of them also successful on the non-regular instructions.

Exp. A				
run	1	8	9	10
Num. *b-successful*	5310	414	2079	6588
$Inst_{blue}^{M}$	0.00	0.00	6.54	37.66
$Inst_{green}^{T}$	0.00	0.00	33.57	0.00
$Inst_{blue}^{M}$ and $Inst_{green}^{T}$	0.00	0.00	0.00	0.00
Num. *l-successful*	0	0	0	0
$Inst_{blue}^{M}$	0.00	0.00	0.00	0.00
$Inst_{green}^{T}$	0.00	0.00	0.00	0.00
$Inst_{blue}^{M}$ and $Inst_{green}^{T}$	0.00	0.00	0.00	0.00

Exp. B		
run	5	7
Num. *b-successful*	3183	9613
$Inst_{blue}^{M}$	0.00	17.07
$Inst_{green}^{T}$	8.83	19.08
$Inst_{blue}^{M}$ and $Inst_{green}^{T}$	0.00	21.49
Num. *l-successful*	0	1753
$Inst_{blue}^{M}$	0.00	0.00
$Inst_{green}^{T}$	0.00	0.00
$Inst_{blue}^{M}$ and $Inst_{green}^{T}$	0.00	0.00

ferring to those *b-successful* agents that proved to be more than 80% successful at performing the behaviour corresponding to the execution of both the non-regular instructions, $Inst^{M}_{blue}$ and $Inst^{T}_{green}$.

During the second phase of the post-evaluation test, *b-successful* agents are asked to perform the behaviour-recognition task. That is, they are required to produce as output the regular and non-regular linguistic instructions that, during the behaviour-production test, triggered their successful behaviour. Hereafter, we refer to this second phase of the post-evaluation test as behaviour-recognition test. Recall that, behaviour-recognition test on non-regular instructions is performed only on *b-partially-* or *b-fully-compositional* agents. Moreover, recall that the agents perceive their successful behaviours through sequences of duplet α, β, recorded during successful post-evaluation trials of the behaviour-production test. As for the behaviour-production test, we considered those agents successful at the behaviour-recognition test (hereafter, referred to as *l-successful*) that manage to obtain a success rate higher that 80% in generating the regular linguistic instructions. Note that, the object label generated by the agent controller is considered "blue" if the neuron with the lowest firing rate is neuron 35, "green" if it is neuron 36, "red" if it is neuron 37. The action label generated by the agent controller is considered "touch" if the neuron with the lowest firing rate is neuron 38, "move" if it is neuron 39, "indicate" if it is neuron 40. *L-successful* agents have been further classified in i) *l-non-compositional* agents, referring to those *l-successful* agents that proved to be less than 80% successful at generating non-regular linguistic instructions, $Inst^{M}_{blue}$ and $Inst^{T}_{green}$; ii) *l-partially-compositional* agents referring to those *l-successful* agents that proved to be more than 80% successful at generating only one of the two non-regular instructions, $Inst^{M}_{blue}$ or $Inst^{T}_{green}$; iii) *l-fully-compositional* agents referring to those *l-successful* agents that proved to be more than 80% successful at generating both the non-regular instructions, $Inst^{M}_{blue}$ and $Inst^{T}_{green}$.

Table 2 shows the results of post-evaluation tests on those evolutionary runs in which we recorded the presence of *b-successful* agents. First, only four out of ten runs in Exp. A, and two out of ten runs in Exp. B produced *b-successful* agents. Second, only run 7 in Exp. B produced agents that are both *b-successful* and *l-successful*. This result indicates that, given our methodological setup, it is extremely difficult to design the mechanisms to allow autonomous agents to perform both the behaviour-production task and the behaviour-recognition task as described in previous Sections. The experimental condition in which the mechanisms to perform the behaviour-production task and the behaviour-recognition task co-adapt simultaneously (i.e., Exp. B) seems to contain the necessary "ingredients" to accomplish the objective of this study. However, the fact that only one out of ten runs produced both *b-successful* and *l-successful* agents suggests that there are elements that severely hindered the evolution from generating the neural structured required by the agents to accomplish their objective. What are these elements? At this stage of our investigation, we have evidence to claim that the number of hidden neurons of the neuro-controllers has a bearing on the evolution of *b-successful* agents. In a previous study described in (Tuci et al., 2010), we have evolved agents to perform only the behaviour-production task in evolutionary circumstances identical to those illustrated in this study. In (Tuci et al., 2010), agents were controlled by neural controllers with only three hidden neurons. Almost all the evolutionary runs generated *b-successful* agents. It seems that smaller neural controllers corresponding to a smaller evolutionary search space facilitates the evolution of the mechanisms to accomplish the behaviour-production task. However, when employed in this study, three-hidden-neuron controllers proved to be insufficient to perform both the behaviour-production task and the behaviour-recognition task. We had to progressively increase the number of hidden neurons from three to eight to generate *b-successful* and *l-successful* agents. Further tests are certainly required to isolate other elements of our model that may have a strong bearing on the capability to generate *b-successful* and *l-successful* agents.

Table 2 also shows the results concerning compositionality. Only run n. 7 in Exp B produced agents that turned out to be *b-fully-compositional*. *b-partially-compositional* agents can be found in run 9 and 10 of Exp. A, and in both runs of Exp. B. None of the runs produced *l-partially-compositional* or *l-fully-compositional* agents. It is worth noting that the mechanisms to access non-regular instructions and to generate the corresponding behaviours do not underpin the inverse process, that is, from the perception of behaviours never executed during training to the generation of the corresponding non-regular instructions. This suggests that linguistic skills related to the capability to comprehend and to generate linguistic instructions in *b-fully-compositional* and *l-successful* agents are underpinned by different neural mechanisms. The mechanisms concerning the capability to be *b-fully-compositional* work as a functionally compositional semantic structure. The mechanisms concerning the capability to be *l-successful* allow the agents to learn by rote the association between the perception of sequences of α, β duplet and regular instructions.

Figure 3 show several graphs which tell us more about the evolutionary dynamics which led to the emergence of *b-successful* and *l-successful* agents in run 7 of Exp. B. These graphs show for each best agent of each generation of run n. 7 the percentage of success for each instruction of the behaviour-recognition test (see dotted, dashed, and continuous lines in Figure 3) as well as the generations in which the agents turned out to be *b-successful*, and the generation at which the agents turned out to be *b-fully-compositional*

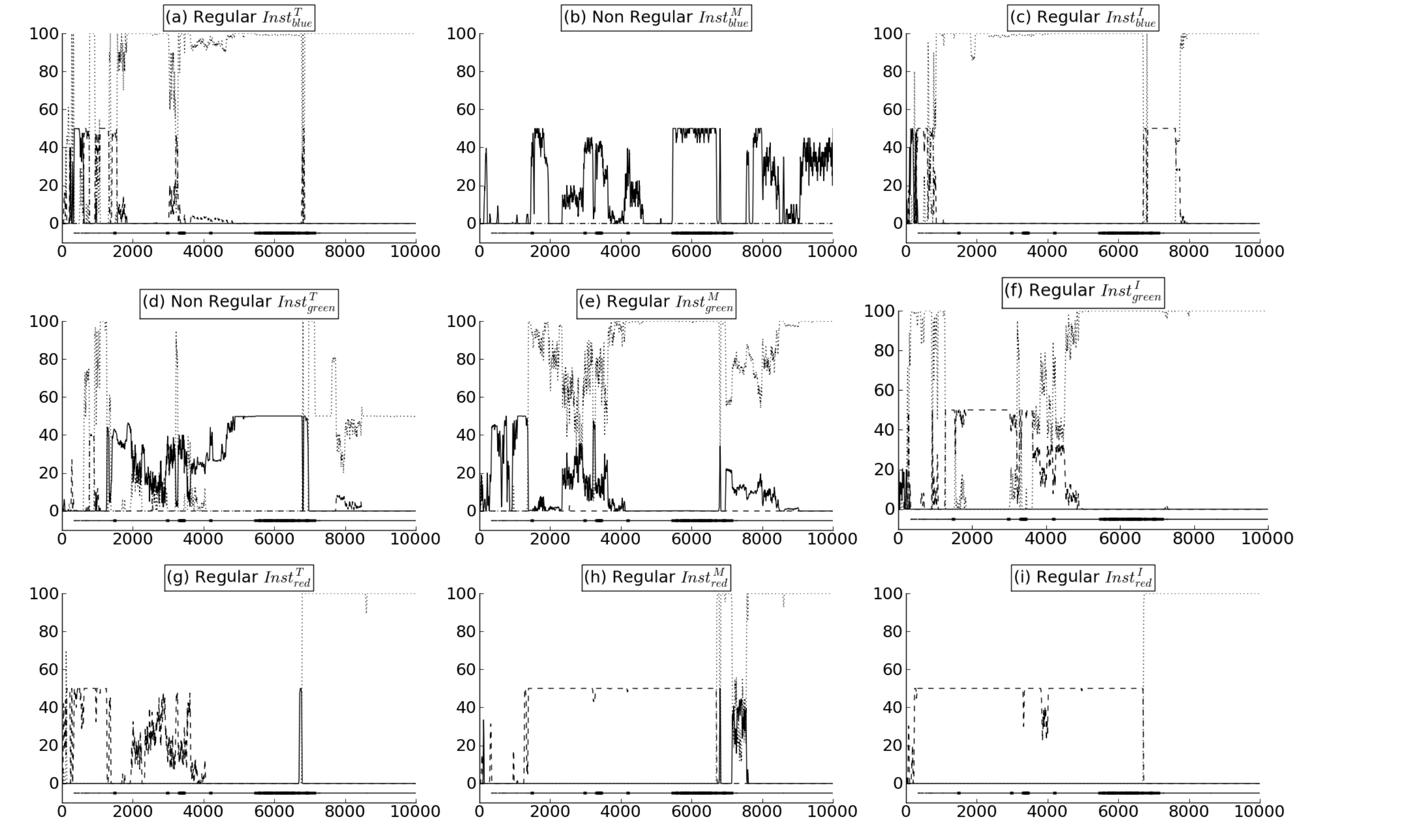

Figure 3: Graphs showing for each best agent of each generation of run n. 7 the percentage of success for each instruction of the behaviour-recognition test. Dotted lines refer to the percentage of success in generating the labels for both the object and the action. Continuous lines refer to the percentage of cases in generating the correct label for the object and the wrong one for the action. Dashed lines refer to the percentage of cases in generating the correct label for the action and the wrong one for the object. At the bottom of each graph, the thin horizontal continuous line indicates the generations in which the agents turned out to be *b-successful*. The tick horizontal line over-imposed on the thin one, indicates the generations in which the agents turned out to be *b-fully-compositional* (see text for details). Data are smoothed with a moving average of window size 20.

(see thin and thick horizontal lines below zero in Figure 3). First, we notice that *b-fully-compositional* agents keep on appearing and disappearing during evolution, while successful agents once generated, are almost never lost. These data suggest that compositionality is not automatically associated with, and is not a prerequisite for developing the capability of successfully performing the behaviour-production task. Second, *l-successful* agents appear very late in evolution. In particular, the agents seemed to have hard time to correctly label behaviours triggered by instructions concerning the red object (see continuous and dashed lines in Figure 3g, 3h, 3i). *l-successful* agents appear after generation 6000, definitely later than the appearance of *b-fully-compositional* agents (see dotted lines and the tick horizontal lines below zero in Figure 3a, 3c, 3e, 3f, 3g, 3h, 3i). This suggest that the emergence of a functionally compositional semantics is not determined by the evolution of the mechanisms to successfully perform the behaviour-recognition task. Third, the graphs concerning non-regular instructions tell us that the agents are not completely unable to deal with these circumstances. For example, as far it concerns $Inst^{M}_{blues}$ (see Figure 3b), several agents during evolution proved to be up to 50% successful in correctly labelling the object on which the action was performed. As far it concerns $Inst^{T}_{green}$ (see Figure 3d), up to generation 6000, the agents seemed to be more effective in labelling the object, while after generation 6000 they proved to be at least 50% effective in correctly labelling both the object and the action given the behaviour corresponding to the execution of this instruction.

Conclusions

We have described a set of simulations which generated autonomous agents, controlled by a single non *a priori* modularised neuro-controller, capable of successfully executing both a language comprehension and a language production task. Post-evaluation tests revealed that, successful agents display a form of compositional semantics which allow them to access linguistic instructions not experienced during training and to execute the corresponding behaviours also no experienced during training. That is, we observed generalisation capabilities in the behaviour-production task. The same successful agents proved not capable of correctly labelling their own behaviours not experienced during training. That is, we did not observe generalisation capabilities in the behaviour-recognition task. Although at this stage we do not have enough empirical evidence to account for this result, we can definitely formulate a number of not mutually exclusive hypothesis that we will consider to identify future directions of work.

Why successful agents show generalisation capabilities at the behaviour-production task and no generalisation capabilities at the behaviour-recognition task? First, we can hypothesise that, the agents have enough computational resources (e.g., hidden neurons) to learn by rote the association between behaviours represented by sequences of α, β duplet and linguistic labels. Alternatively, it could be that the behaviour-recognition task did not produce sufficiently selective evolutionary pressures to generate the mechanisms required to shift from rote knowledge to a more flexible conceptual system. Second, from the agent point of view, the behaviour-production task and the behaviour-recognition task are mostly uncorrelated tasks. This becomes clear if we consider that the agent has two groups of input-output neurons: one (comprising input neurons 1 to 20 and output neurons 31 to 34) that is only used during the behaviour-production task; the other (comprising input neurons 21 and 22 and output neurons 35 to 40) that is only used during the behaviour-recognition task. Due to the different nature of the two input-output groups, the input received during the behaviour-recognition task is completely different from the motor output and from any other input experienced during the behaviour-production task. This may make it difficult for the agent to develop a coherent internal structure, common to the language comprehension and language production task. To try to cope with this problem we plan to explore two possibilities: one is to modify the agent body and neural architecture, the other is to slightly modify the task. As far as the agent is concerned, one possibility could be to change the way the output controlling the arm movement is encoded, so to have at least similar kinds of input and output signal. Another possibility could be to feed the α and β input neurons also during the behaviour-production task (as if the agent could "see" himself doing the task). On the task side, we plan to implement setups in which the two abilities have to be used together. For example, we could ask the agent to produce the correct linguistic instruction during the behaviour-production task. Even though this is a rather easy task (the correct instruction is already present in the input units), it could nonetheless favour the emergence of common structures underpinning both the language comprehension and language production task.

Acknowledgements

This research work was supported by the *ITALK* project (EU, ICT, Cognitive Systems and Robotics Integrating Project, grant $n°$ 214668). The authors thank Andrew Szabados and Domenico Parisi for helpful comments during the preparation of the paper.

References

Beer, R. D. and Gallagher, J. C. (1992). Evolving dynamic neural networks for adaptive behavior. *Adaptive Behavior*, 1(1):91–122.

Cangelosi, A. and Parisi, D., editors (2002). *Simulating the evolution of language*. Springer Verlag, New York.

Cangelosi, A. and Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots. *Cognitive Science*, 30(4):673–689.

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience*, 3:317–333.

Glenberg, A. and Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9:558–565.

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

Harnard, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Hutchins, E. and Johnson, C. (2009). Modelling the emergence of language as an embodied collective cognitive activity. *Topics in Cognitive Science*, 1:523–546.

Roy, D. (2002). Learning visually grounded wprds and syntax for a scene description task. *Computer Speech and Language*, pages 353–385.

Steels, L. (2002). Grounding symbols in through evolutionary language games. In Cangelosi, A. and Parisi, D., editors, *Simulating the evolution of language*, pages 211–226. Springer Verlag, London.

Sugita, Y. and Tani, J. (2005). Learning semantic combinatoriality from the intercation between linguistic and behvioral processes. *Adaptive Behavior*, 13(1):33–52.

Sugita, Y. and Tani, J. (2008). Acquiring a functionally compositional system of goal-directed actions of a simulated agent. In Asada, M., Hallam, J., Meyer, J.-A., and Tani, J., editors, *Proc. of the* 10^{th} *Int. Conf. on Simulation of Adaptive Behavior (SAB2008)*, pages 331–341. Springer Verlag.

Tuci, E., Ferrauto, T., Massera, G., and Nolfi, S. (2010). Co-development of linguistic and behavioural skills: compositional semantics and behaviour generalisation. In *Proceedings of the 11th International Conference on the Simulation of Adaptive Behaviour (SAB 2010)*, volume LNCS. Springer, Berlin, Germany. In Press.

Van Gelder, T. (1990). Compositionality: A connectionist variation on a classical theme. *Cognitive Science*, 14:355–384.