

An investigation of the evolutionary origin of reciprocal communication using simulated autonomous agents

Elio Tuci

Received: 10 March 2009 / Accepted: 19 August 2009 / Published online: 3 September 2009
© Springer-Verlag 2009

Abstract How does communication originate in a population of originally non-communicating individuals? Providing an answer to this question from a neo-Darwinian epistemological perspective is not a trivial task. The reason is that, for non-communicating agents, the capabilities of emitting signals and responding to them are both adaptively neutral traits if they are not simultaneously present. Research studies based on rather general and theoretically oriented evolutionary simulation models have, so far, demonstrated that at least two different processes can account for the origin of communication. On the one hand, communicative behaviour may first evolve in a non-communicative context and only subsequently acquire its adaptive function. On the other hand, communication may originate thanks to cognitive constraints; that is, communication may originate thanks to the existence of neural substrates that are common to the signalling and categorising capabilities. This article provides a proof-of-concept demonstration of the origin of communication in a novel-simulated scenario in which groups of two homogeneous (i.e. genetically identical) agents exploit reciprocal communication to develop common perceptual categories and to perform a collective task. In particular, in circumstances in which communication is evolutionarily advantageous, simulated agents evolve from scratch social behaviour through acoustic interactions. We look into the phylogeny of successful communication protocol, and we describe the evolutionary phenomena that, in early evolutionary stages, paved the way for the subsequent development of reciprocal communication, categorisation capabilities and successful cooperative strategies.

Keywords Communication · Artificial evolution · Artificial neural networks · Autonomous simulated agents

1 Introduction

Tracing animal communication back to its origins from a neo-Darwinian epistemological perspective is neither a trivial nor an obvious task (see [Dawkins and Krebs 1978](#); [Maynard-Smith and Harper 1995](#); [Hauser 1996](#)). A scientific account of how a population of communicating agents originates from a population of non-communicating agents has to determine under which circumstances communicating is evolutionarily advantageous ([Maynard-Smith and Price 1973](#)). However, there are further issues that need to be investigated to understand the origins of communication other than those related to its adaptiveness. In particular, a neo-Darwinian account of the origin of communication has to provide evidence showing how agents that do not emit signals, and are not capable of appropriately responding to them, can acquire both capabilities through natural selection.

Communicating agents must possess the required mechanisms to emit appropriate signals and develop appropriate reactions to them. Natural evolution mainly works by favouring the emergence of traits that increase the chance of survival and reproduction. However, both the capability to emit signals and the capability to respond to them are adaptively neutral if not considered in a communicative context. That is, in a population of agents that are not capable of appropriately responding to signals, the reproductive chance of individuals that are capable of emitting signals does not differ from that of individuals that are not capable of emitting signal. Similarly, in a population of agents that are not capable of emitting appropriate signals, the reproductive chance of individuals that are capable of responding to signals does not

E. Tuci (✉)
ISTC-CNR, Via San Martino della Battaglia, No. 44, 00185 Rome, Italy
e-mail: elio.tuci@istc.cnr.it

differ from that of individuals that cannot respond to signals. The fact that emitting and responding to signals, when taken in isolation, are adaptively neutral traits, leaves us with the problem of explaining how the capability of responding to signals evolves if no one is signalling, or how the capability of emitting signals evolves if no one is responding.

Evolutionary Robotics (hereafter, referred to as ER) models are rather recent methodological tools that have been complementing classic biological/ethological modeling methods to study the evolutionary origins of individual and social behaviour (Nolfi and Floreano 2000). ER is based on the use of artificial evolution to find sets of parameters for artificial neural networks that guide agents to the accomplishment of their objectives, while avoiding dangers. By using ER models, scientists can determine the characteristics of the agents and of their environment that facilitate the evolution of certain behavioural capabilities and their underlying mechanisms (see Pfeifer and Scheier 2001; Cangelosi and Parisi 2002). Several ER models focus on issues concerning the operational principles of communicative behaviour due to the fundamental role, which communication plays in fostering cooperation and collaboration amongst artificial and natural organisms. Other ER models, described in Quinn et al. (2003), Marocco and Nolfi (2007), Mirolli and Parisi (2008), and Ampatzis et al. (2008), focus on the evolutionary origins of communication, trying to explain how communication may originate in a population of non-communicating agents. Following this line of investigation, this study illustrates a theoretically oriented ER model focused on the origin of acoustic communication. In particular, we provide a proof-of-concept demonstration of the origin of communication in a scenario in which genetically related agents must “talk to each other” to correctly categorise their world, and to cooperatively act on the basis of the chosen category.

Generally speaking, the term categorisation can be referred to the ability of mapping continuous signals received by sensors into discrete categories whose members resemble each other more than members of other categories. Categorical perception represents one of the most fundamental cognitive capacities displayed by natural organisms, and it is an important pre-requisite for the exhibition of several other cognitive skills (see Harnad 1987). In our experimental scenario, categorisation refers to the capability of an agent to gather individual perceptual experience and to share this experience with its partner to develop a common perspective on the current environment. In other words, since in our scenario individual perception is not sufficient to perform the categorisation task, the agents must share their experiences by interacting through sound signals. Whether or not the agents correctly categorise the environment can be inferred through the effectiveness of the collective strategy they are required to perform in order to accomplish their task. It is important to note that, at the beginning of their

evolutionary history, the agents do not possess any built-in mechanism to explicitly represent the environmental categories or other mechanisms that explicitly regulate acoustic interactions. The underlying structures for categorisation and communication have to evolve from scratch.

We demonstrate that communication originates in populations of non-communicating agents in which individuals exploit social interactions to develop common perceptual categories. We look into the phylogeny of a successful communication protocol, and we describe the evolutionary phenomena that, in early evolutionary stages, paved the way for the subsequent development of reciprocal communication, categorisation capabilities and successful cooperative strategies. Although the evolved behavioural and communication strategies may be limited to the peculiarities of this case study, we provide insights into the origin of reciprocal communication without addressing the evolution of any particular communication system of any given species.

In what follows, we first present a review of previous study in ER focused on the origin of communication (see Sect. 2). Then, we describe the cooperative task investigated in this research study (see Sect. 3). In Sect. 4, we describe the agents' morphological structure, their control systems and the evolutionary algorithm used to design them. In Sect. 5, we illustrate the results of a series of post-evaluation analysis. Discussion, conclusions and directions for future study are presented in Sect. 6 and 7.

2 Comparison with previous studies

The issue of how artificial agents can develop social behaviour and communication skills has already been addressed in several research studies. In this section, we describe those works in which, like in ours, the structure of both communicative and non-communicative behaviour are automatically designed by an evolutionary algorithm and in which evolutionary analyses revealed interesting aspects about the origins of communication. Consequently, we do not consider those interesting studies on communication in multi-agent systems in which the mechanisms for social interactions are hand-coded (e.g. see Balch and Arkin 1994; Cao et al. 1997; Fong et al. 2002; Støy 2001; Steels 2003). Moreover, we do not review those ER models, in which the mechanisms for social interactions are entirely evolved, but the focus of that is on the ontogeny than on issues concerning the phylogeny of communication (e.g. Di Paolo 2000; Baldassarre et al. 2003; Trianni and Dorigo 2006; Tuci et al. 2008; Williams et al. 2008).

The study described in Quinn (2001) can be considered the first ER model to look at issues concerning the origin of communication in autonomous agents. Two years later, that work was extended in Quinn et al. (2003). In both works,

teams of two or three homogeneous robots were required to move in an arbitrarily chosen direction, while remaining close to each other at a distance smaller than the range of their infrared sensors. Neural mechanisms designed through artificial evolution allowed the agents to engage in social interactions which resulted in the emergence of roles such as leader/follower. In [Quinn \(2001\)](#) and [Quinn et al. \(2003\)](#), the authors found that behaviours that have a communicative value resulted from the adaptation of other elementary behaviours, which did not have a communicative function (i.e., obstacle avoidance). Thus, they concluded that behaviour for social coordination may evolve in a non-communicative context, and only subsequently acquire its adaptive function.

In [Marocco et al. \(2003\)](#), the authors described a categorisation task in which a simple artificial arm made of three segments is required to distinguish a sphere from a cube by remaining close to the first object and moving away from the second one. Each agent receives tactile sensations as well as a linguistic input provided by previous generation agents that have already interacted and named the objects by using two binary outputs. That work shows that the evolutionary conditions, which facilitate the emergence of effective communication are those in which each arm receives a linguistic input from its parent rather than from another individual of the previous generation, and in which naming begins after the agents have already evolved basic categorisation capabilities. The results of that work induced the authors to emphasise two distinctive aspects: first, they claimed that a kinship relation between a speaking parent and a listening offspring facilitates the origin of a common “language” in a population of originally non-communicating agents. This issue is further discussed in [Floreano et al. \(2007\)](#). Second, they took up a thesis already formulated in [Cangelosi and Parisi \(1998\)](#), by suggesting that the ability of categorising perceptual states constitutes the grounding for the evolution of effective communication protocols. The idea that the capability of categorising perceptual states can facilitate the establishment of links between linguistic production and comprehension abilities is the message of the work described in [Marocco and Nolfi \(2007\)](#) and [Mirolli and Parisi \(2008\)](#). In particular, [Mirolli and Parisi \(2008\)](#) showed a model in which communication amongst artificial agents originates because the signals produced by an organism tend to reflect the way, in which the organism categorises its experience. In other words, links in the brain of an organism between the mechanisms for categorisation and those for communication were proved to represent a bias (referred to as ‘producer-bias’ in [Mirolli and Parisi \(2008\)](#)) towards the production of useful and reliable signals.

In the work of [Amptz et al. \(2008\)](#), the authors described a set of experiments in which artificial evolution is used as a methodological tool to engineer robot

neuro-controllers capable of guiding groups of robots in a categorisation task by producing appropriate actions. In particular, two autonomous mobile robots required to perform an individual task exploit their signalling system to develop a simple form of cooperation. Categorisation is the result of how robots’ sensory inputs unfold in time, and, more specifically, of the integration over time of sensory input. Although not explicitly rewarded by the fitness function, and not necessarily required to perform the task, communication originates as it enhances group performance, revealing a ‘hidden’ benefit for social behaviour. This benefit is related to obtaining robust and fast decision-making mechanisms. More generally, the authors showed how processes requiring the categorisation of noisy dynamical information might be improved by social interactions mediated by communication.

With respect to the research works mentioned above, our model is a general and theoretically oriented study on the origin of communication by using acoustic signals. We do not address the evolution of any particular communication system of any given species. As in [Marocco et al. \(2003\)](#), [Amptz et al. \(2008\)](#) and in [Mirolli and Parisi \(2008\)](#), we target communication skills that allow the agents to perform categorisation tasks. However, our simulation differs from those mentioned above in that we investigate the origin of communication in a scenario in which the agents need to ‘talk to each other’ to correctly categorise elements of their world, and to cooperatively act on the basis of the chosen category. Communication is reciprocal, since both agents need to recognise and distinguish environmental features, to influence each other’s behaviour through the emission of signals and to properly respond to the perceived signals. Social interactions are evolutionarily advantageous as long as they allow agents to properly categorise elements of the environment in which they are currently located. In this sense, our work is similar to the work described in [Quinn \(2001\)](#) and [Quinn et al. \(2003\)](#), in which both agents engaged in mutually influencing ritualised interactions. However, our work differs from that of [Quinn \(2001\)](#) and of [Quinn et al. \(2003\)](#), in that agents need to recognise distinctive perceptual states to bring forth adaptive social interactions.

By synthesising through artificial evolution collective strategies tailored to this adaptive problem, we illustrate the operational principles of successful communication protocols. Moreover, we describe the selective pressures and the evolutionary dynamics that produced the best evolved communication protocols.

3 The task

Each simulated agent has a cylindrical body of 5.8 cm radius, with infrared, ambient light and floor sensors, a microphone and a traction system made of four wheels. The agents’ world

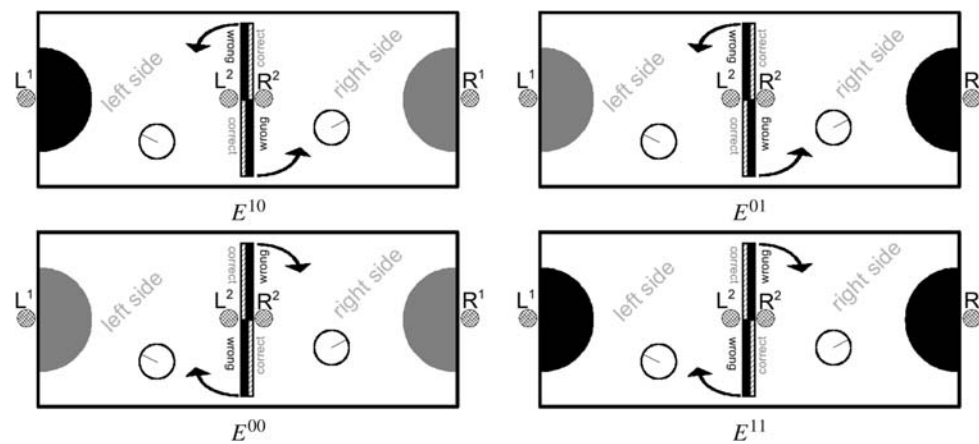


Fig. 1 The four environments E^{10} , E^{01} , E^{00} and E^{11} . The *small filled circles* labeled L^1 , L^2 , R^1 and R^2 indicate the position of the lights. The revolving door is indicated by the central bar orthogonal to the long arena walls. The *black parts* of the revolving door indicate the wrong arms (i.e. those that the agents should not touch), the *white parts* with

diagonal lines are the correct arms (i.e. those that, if simultaneously pushed, make the door rotate). In each environment, the *arrows* indicate the direction in which the door revolves. The two *empty circles* on the *white floor* represent the agents and their headings

is a rectangular arena (120 cm \times 50 cm) divided into two equal sides by a revolving door. The latter is a bar that rotates clockwise or anticlockwise if simultaneously pushed by both agents. The white arena floor is characterised by two painted zones placed close to the short arena walls (see Fig. 1). The direction of rotation of the revolving door is determined by the combination of colours of the painted zones, with each single colour (black and grey) associated with both clockwise and anticlockwise rotational movement.

Two agents, initially placed one on the left and one on the right side of the arena, are required to repeatedly swap sides by opening the revolving door without touching the current wrong arms (see Fig. 1). In a single trial, an agent can only perceive the colour of one painted zone. Therefore, in order to accomplish their task, the agents should first explore their sides to find out the colour of the painted zones. Then, they should ‘talk to each other’ using sound signals to find out the direction of rotation of the revolving door. Finally, they are required to employ an effective cooperative pushing strategy to exert forces only on the current correct arms of the revolving door. We chose to use a communication system as simple as possible, i.e. sound signals that do not vary in frequency not in amplitude. This minimal model of sound has already been ported on real robots without having to simulate aspects concerning the propagation of sound and sound interference (see Ampatzis et al. 2008). Moreover, this minimal sound model does not provide the agents any built-in mechanisms for the distinction between self- and non-self-produced sound. Whether or not the capability to distinguish self-produced from non-self-produced signals is required by the agents to solve this task, is an issue entirely left to evolution.

As shown in Fig. 1, there are four lights in the arena: L^1 , L^2 , R^1 and R^2 . L^1 and L^2 can only be seen by an agent

located in the left side of the arena, while R^1 and R^2 can only be seen by an agent located in the right side of the arena. When L^1 is turned on L^2 is turned off and vice versa; the same for R^1 and R^2 . The arena floor is white except in two semicircular zones of 15 cm radius located in front of L^1 and R^1 , respectively, in which the floor can be either black or grey. The agents can experience four different combinations of black and grey zones. The type of environment, in which the agents are located, is determined by the combination of colours of the painted zones. More specifically, the environments are labeled E^{xy} with $x, y \in \{0, 1\}$, where x corresponds to the colour of the floor in the proximity of L^1 and y to the colour of the floor near R^1 . Grey corresponds to 0 and black to 1. The four types of environment are E^{10} , E^{01} , E^{00} and E^{11} . The revolving door, orthogonal to the long arena walls, rotates 90° clockwise or anticlockwise, if simultaneously pushed by both agents in the correct manner. In fact, the correct direction of rotation depends on the type of environment. The agents have to exert forces to make the door rotate (a) anticlockwise, if located in E^{10} or in E^{01} (b) clockwise, if they are located in E^{00} or in E^{11} (see the arrows in Fig. 1).

At the beginning of the first trial and in those that follow an unsuccessful one, the agents are randomly placed in the proximity of the revolving door one in each arena side. A trial begins with L^1 and R^1 turned on, and it can last for up to 100 simulated seconds (1,000 time steps). A trial is terminated earlier (a) in the case, an agent crashes with the arena walls, (b) if the agents fail to swap sides within the time limit and (c) if both agents manage to travel (by rotating the door) into the opposite side of the arena up to a distance of 24 cm from the door hub before the time limit. Trials terminated earlier due to (a) or (b) are considered unsuccessful. In trials following a successful one, the agents are not repositioned. Each trial

differs from the others in the initialisation of the random number generator, which influences the agents' starting position and orientation anytime the agents are positioned, and the noise added to actuation of motors and sensor readings.

The sequence of desired actions that each agent is demanded to carry out during a trial can be decomposed into two phases. At the beginning of the first phase, L^1 and R^1 are turned on, the revolving door is oriented orthogonally to the long arena walls and the colour of the floor in the proximity of L^1 and R^1 is set according to the type of environment that characterises the trial. During this phase, the agents are required to find the painted zone in their arena side. L^1 and R^1 facilitate the search of the painted zones, since they can be used by the agents as beacons. For the agent located on the left and on the right side, the first phase finishes as soon as the time spent on the painted zone is bigger than T' and T'' , respectively. T' and T'' are randomly and independently chosen between 4 and 8 s.

The change from the first to the second phase of the task is characterised by the fact that L^1 and R^1 turn off, while L^2 and R^2 turn on. The function of L^2 and R^2 is to show the agents in which direction they have to move to approach the revolving door. Each agent, with a phototactic behaviour, should get close to the revolving door and push it with its body until it is open enough to let the agent go through. As mentioned above, the agents have to simultaneously exert forces to make the door rotate (a) anticlockwise, if they are located in E^{10} or in E^{01} ; (b) clockwise, if located in E^{00} or in E^{11} . A trial successfully terminates when both agents manage to swap sides and to travel into the opposite side of the arena up to a distance of 24 cm from the door hub. At the end of a successful trial, L^2 and R^2 are turned off, L^1 and R^1 are turned on, the rotating door automatically returns to its rest position (i.e. orthogonal to the long arena walls) and a new trial begins.

Note that the sequence of events experienced by the agent on the left side is completely independent from the sequence of events experienced by the agent on the right side. For example, the change of state of L^1 and L^2 is exclusively determined by the behaviour of the agent on the left side. The change of state of R^1 and R^2 is exclusively determined by the behaviour of the agent on the right side. This means that the agents might start pushing the door at different times. Nevertheless, the door rotates only if both agents simultaneously push it in the correct way. Moreover, genuine cooperative successful strategies are those in which, during a trial, each agent touches only the correct arm of the revolving door. Agents should find out which is the correct arm to push by communicating through sound signals. Communication should allow the agents to complement their knowledge of their respective arena side by 'informing' each other on the colour of the painted zone in the opposite side. Communication is effective if by emitting sound signals the agents first

find out whether they are in an environment in which both painted zones are coloured in the same way or not, and second develop the cooperative pushing strategy that open the revolving door without touching its current wrong arms.

The agents can complete both phases of the task and swap arena side with simpler strategies in which communication and categorisation are not required. For example, one or both agents may push in a single trial both arms (i.e. the correct and the wrong one) of the revolving door until it rotates. This behaviour, referred to as trial-and-error, is not considered the correct way to solve the task and, during evolution, it is penalised by the fitness function (see Sect. 4.3).

4 Methods

4.1 The robot and its control structure

The agents and their world are simulated using software based on Open Dynamic Engine (see <http://www.ode.org/>), a 3D rigid body dynamics library that provides primitives for the implementation of detailed and realistic physics-based simulations. Our simulation models some of the hardware characteristics of the real *s-bot*, which are small wheeled cylindrical robots, 5.8 cm of radius, equipped with a variety of sensors, and whose mobility is ensured by a differential drive system (see Mondada et al. 2004). Our simulated agent has a differential drive motion provided by a traction system composed of four wheels: two lateral, motorised wheels and two spherical, passive wheels placed in the front and in the back. The four wheels are fixed to the cylindrical body that holds the sensors. In particular, agents make use of five infrared sensors IR_i , two ambient light sensors AL_i , one floor sensor FS , one loudspeaker SO and one omnidirectional microphone SI (see Fig. 2a). Infrared sensors have a maximum range of 15 cm; ambient light sensors have a maximum range of 60 cm and light levels change as a function of the agent's distance from the lamp.¹ The floor sensor, placed underneath the agent, detects the level of grey of the floor. It outputs the following values: 0—if the agent is positioned over a white floor; 0.5—if the agent is positioned over a grey floor; 1—if the agent is positioned over a black floor. The loudspeaker produces a binary output with $SO = 1$ corresponding to sound on, and $SO = 0$ to sound off. The omnidirectional microphone has no directionality and intensity features and it returns 1 if one or both agents are emitting sound, otherwise 0. A total of 10% uniform noise is added to IR_i and AL_i readings, and also to motor outputs and agents' position.

¹ The readings of the infrared and light sensors are taken from look-up tables which contain sampled information from the real *s-bot*. More details can be found in Vicentini and Tuci (2006).

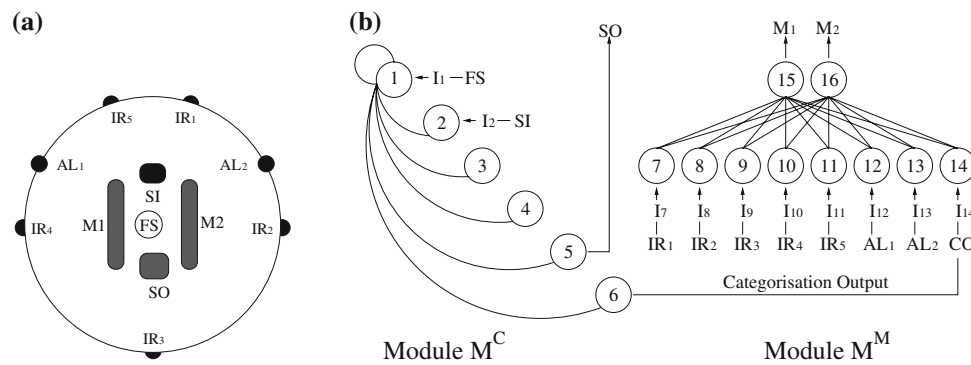


Fig. 2 **a** Depiction of the agent with motors and sensors. IR_i , with $i = 1, \dots, 5$ are the infrared sensors; AL_1 and AL_2 are the ambient light sensors; FS is the floor sensor; SI is the omnidirectional microphone; SO is the loudspeaker; M_1 and M_2 are the left and right motor, respectively. **b** The control architecture: module M^C and module M^M . For M^C only the efferent connections of one neuron are drawn. At each

time step, the categorisation output CO emitted by M^C is input into M^M . The correspondences between agent’s sensors and input neurons are indicated in each module. The labels on the continuous line arrows refer to the notation used in Eq. 1 to indicate the readings of the corresponding sensors

Each agent controller is composed of two modules referred to as M^C and M^M (see Fig. 2b). The modularisation is hand-coded to facilitate the evolution of successful behavioural strategies. M^C is a six neuron fully connected Continuous Time Recurrent Neural Network (CTRNN; see also Beer and Gallagher 1992). This module takes input from FS and SI and it outputs the state of SO and CO (i.e. the categorisation output, with $0 < CO < 1$). M^C is devoted to (i) communication, by controlling the acoustic signalling system; and to (ii) decision making, by operationally defining, through the values of the categorisation output CO , the agent decision concerning the nature of the environment.

M^M is a feed-forward artificial neural network made of eight sensory neurons and two motor neurons. M^M takes input from IR_i , with $i = 1, \dots, 5$, AL_1 , AL_2 and CO . It sets the speed of the agent’s wheels. M^M is devoted to (i) phototaxis, by moving the agent towards the light currently on; and to (ii) the door opening behaviour. This latter behaviour can be accomplished by a group without touching the current wrong arm, only if each agent is able to act differently with respect to the corresponding central light. In other words, the agents have to keep the central light on their left when located in E^{10} and in E^{01} , and on their right when located in E^{00} and in E^{11} (see Fig. 1). The differentiation of actions with respect to the central light can only be achieved if the values of the categorisation output CO , emitted by module M^C , while the agent is approaching the revolving door, varies to allow module M^M to discriminate the two environmental categories. This is because M^M is a reactive module that would systematically bring the robot either on the left or on the right of the central light, in the absence of input signals that co-vary with the environmental category.

The neurons’ state of an agent’s controller is updated using the following equations:

$$\frac{\Delta y_i}{\Delta t} = \begin{cases} \frac{1}{\tau_i} \left(-y_i + \sum_{j=n'}^{m'} \omega_{ji} \phi(y_j + \beta_j) + g^C I_i \right); & \text{for } i = 1, 2; \\ \frac{1}{\tau_i} \left(-y_i + \sum_{j=n'}^{m'} \omega_{ji} \phi(y_j + \beta_j) \right); & \text{for } i = 3, \dots, 6; \\ \frac{1}{\Delta t} (-y_i + g^M I_i); & \text{for } i = 7, \dots, 14; \\ \frac{1}{\Delta t} \left(-y_i + \sum_{j=n''}^{m''} \omega_{ji} \phi(y_j + \beta) \right); & \text{for } i = 15, 16; \end{cases} \quad (1)$$

with $n' = 1$, $m' = 6$, $n'' = 7$, $m'' = 14$ and $\phi(x) = (1 + e^{-x})^{-1}$. In these equations, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i the decay constant, ω_{ji} is the strength of the synaptic connection from neuron j to neuron i , β_i the bias term, $\phi(y_i + \beta_i)$ the output of neuron i , g^C and g^M the gain factors, and I_i is the intensity of the sensory perturbation on sensory neuron i . The parameters ω_{ji} , τ , β and g are genetically encoded. Cell potentials are set to 0 any time the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of $\Delta t = 0.1$. The output of the 5th neuron sets the state of the agent’s sound actuator SO , according to the following rule:

$$SO = \begin{cases} 1 & \text{if } \phi(y_5 + \beta_5) \geq 0.5; \\ 0 & \text{otherwise;} \end{cases} \quad (2)$$

The output of the 6th neuron corresponds to the state of the agent’s categorisation output (i.e. $CO = \phi(y_6 + \beta_6)$). The values of $\phi(y_{15})$ and $\phi(y_{16})$, linearly scaled into $[-6.5s^{-1}, 6.5s^{-1}]$, are used to set the agent motors output M_1 and M_2 , respectively.

4.2 The evolutionary algorithm

A simple generational genetic algorithm is employed to set the parameters of the networks (Goldberg 1989). At generation 0, a random population of 80 vectors is generated by initialising each component of each vector to a value chosen uniformly random in the range [0,1]. Each vector comprises 67 real values. Hereafter, using terms derived from an analogy with biological systems, a vector is referred to as genotype and its components as genes. The first 18 genes of each genotype are used to set the parameters of M^M (i.e. 16 connection weights ω_{ji} , 1 bias term β and 1 gain factor g both shared by all the sensory neurons). The other 49 genes are used to set the parameters of M^C (i.e. 36 connection weights ω_{ji} , 6 decay constants τ_i , 6 bias terms β_i , and 1 gain factor g).

Generations following the first one are produced by a combination of selection with elitism, recombination 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 64 best genotypes of the old population. New genotypes, except 'the elite', are produced by applying recombination and mutation. Each new genotype has a 0.3 probability of being created by combining the genetic material of two parents. During recombination, one crossover point is selected. Genes from the beginning of the genotype to the crossover point is copied from one parent, the other genes are copied from the second parent. Mutation entails that a random Gaussian offset is applied to each gene, with a probability of 0.15. 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: (i) for M^C , biases $\beta_i \in [-4, -4]$, weights $\omega_{ij} \in [-8, 8]$, gain factor $g \in [1, 13]$; (ii) for M^M , bias $\beta \in [-4, -2]$, weights $\omega_{ij} \in [-8, 8]$, gain factor $g \in [1, 13]$; (iii) for M^C , decay constants are first linearly mapped into the range $[-1.0, 1.5]$ and then exponentially mapped into $\tau_i \in [10^{-1.0}, 10^{1.5}]$. 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 1/3 of the maximum length of a trial.

4.3 The fitness function

During evolution, a generation corresponds to the evaluation of a set of 80 different genotypes. At the beginning of its evaluation, each genotype is translated into a controller, and

cloned in each agent of a group. Each group is made of two agents. The group is evaluated three times in each environment to yield a total of 12 trials. The sequence order of the environment experienced by the agents—randomly chosen at the beginning of each generation—has a bearing on the overall performance of the group since the agents' controllers are reset only at the beginning of the first trial.

The final fitness FF attributed to a group is computed as follows:

$$FF = F^1 + F^2 + F^3; \tag{3}$$

F^1 rewards the agents for performing phototaxis by facilitating the evolution of mechanisms that allow the agents to move towards the painted zones during the first phase of the task (i.e. when L^1/R^1 are on), and towards the revolving door during the second phase of the task (i.e. when L^2/R^2 are on). F^2 rewards the agents for opening the revolving door by facilitating the evolution of collective pushing strategy. F^3 rewards the agents for generating categorisation outputs which vary as little as possible within environments that require to push the revolving door in the same direction, and as much as possible between environments that require to push the revolving door in different directions. The aim of F^3 is to facilitate the emergence of mechanisms which allow modules M^C to generate categorisation outputs that, by discriminating the two environmental categories, induce the modules M^M to tune the agents pushing strategies to the current direction of rotation of the revolving door.

F^1 is updated only during the first phase of the task, and is computed as follows:

$$F^1 = \frac{1}{ER} \sum_{e=1}^E \sum_{r=1}^R P'_{er} \left(1 - \frac{d_{er}}{D}\right); \tag{4}$$

where $E = 12$, $R = 2$ and d_{er} corresponds to the shortest distance, at the end of the first phase of trial e , between agent r and the perimeter of the painted zone on the agent arena side. $d_{er} = 0$ if the agent r is on the painted zone. $D = 60$ cm is used to normalise d_{er} to run between 0 and 1. P'_{er} is set to 0.2 if the agent r collides with the arena walls, otherwise it is set to 1.

F^2 corresponds to the aperture of the revolving door, at the end of trial e , normalised to run between 0 and 1. Formally speaking, F^2 is computed as follows:

$$F^2 = \frac{1}{E} \sum_{e=1}^E P''_e \left(\frac{2\theta_e}{\pi}\right); \tag{5}$$

where θ_e refers to the rotation of the revolving door at the end of the second phase of trial e . For $\theta_e = 0$, the revolving door is in its rest position, orthogonal to the long arena walls. For $\theta_e = \frac{\pi}{2}$, the revolving door is fully open, orthogonal to the short arena walls. The penalty factor P''_e is set to 0 if, during

trial e , either agent touches the door on the current wrong arm, otherwise it is set to 1.

F^3 concerns the categorisation outputs of both agents, and it is computed only if both agents manage to successfully complete the first phase of the task in each of the 12 evaluation trials. Otherwise, $F^3 = 0$. The fitness component F^3 is computed as follows:

$$F^3 = \left| C^{E^{clock}} - C^{E^{anti}} \right| \left(\frac{k - \sigma^{E^{clock}} - \sigma^{E^{anti}}}{k} \right); \quad (6)$$

$$C^{E^H} = \frac{1}{\sum_{e \in E^H} \sum_{r=1}^R (v_{er} - u_{er})} \sum_{e \in E^H} \sum_{r=1}^R \sum_{t=u_{er}}^{v_{er}} C O_{ert}$$

$$\sigma^{E^H} = \sqrt{\frac{1}{\sum_{e \in E^H} \sum_{r=1}^R (v_{er} - u_{er})} \sum_{e \in E^H} \sum_{r=1}^R \sum_{t=u_{er}}^{v_{er}} (C O_{ert} - C^{E^H})^2}$$

$$E^H = \begin{cases} E^{clock} & \text{if robots located in } E^{00} \text{ or } E^{11} \\ E^{anti} & \text{if robots located in } E^{10} \text{ or } E^{01} \end{cases}$$

where E^{anti} refers to environments in which the door rotate anticlockwise (i.e. $E^{anti} = E^{xy}$ with $x, y \in \{0, 1\}; x \neq y$) and E^{clock} refers to environments in which the door rotate clockwise (i.e. $E^{clock} = E^{xy}$ with $x, y \in \{0, 1\}; x = y$). C^{E^H} and σ^{E^H} correspond to the average group categorisation outputs and to the standard deviation in environments E^{00} and E^{11} for $E^H = E^{clock}$, and in environments E^{10} and E^{01} for $E^H = E^{anti}$. $C O_{ert}$ refers to the value of the categorisation output in trial e , of robot r , at time step t . For robot r , the categorisation outputs are considered during an interval that goes from the beginning of the second phase of the task (i.e. $t = u_{er}$) to the first collision with the revolving door (i.e. $t = v_{er}$). Hereafter, this interval is referred to as *approaching phase*. $k = 2$ is used to normalise $\left(\frac{k - \sigma^{E^{clock}} - \sigma^{E^{anti}}}{k} \right)$ to run between 0 and 1. The first part of Eq. 6 (i.e. $|x|$, the absolute value of x) rewards the agents for emitting categorisation outputs that differ as much as possible between environments of different categories (i.e. E^{anti} vs. E^{clock}). The other part of Eq. 6 rewards the agents for emitting categorisation outputs that differ as little as possible when the agents are located in environments of the same category (i.e. E^{anti} or E^{clock}).

The fitness function illustrated in Eq. 3 has been designed so that, during evolution, phototaxis and wall avoidance are the first behavioural capabilities to appear. Note also that in Eq. 3, there is nothing that directly refers to signalling behaviour. We intentionally omit to include selective forces that reward signalling behaviour to clean our model from preconceptions concerning what (i.e. semantics) and how (i.e. form) successful groups communicate. In this way, we explicitly leave to evolution the task to find the mechanisms underlying

successful communication protocols. However, the fitness function indirectly rewards those groups that possess adaptive communication systems, through the selective pressure produced by the fitness component F^3 . This is because the agents are rewarded for employing the categorisation output in a way that can only be achieved with the contribution of an adaptive sound signalling system. F^3 does not impose any specific associations between the categorisation output and the type of environment. Nevertheless, F^3 favours those groups that use solutions, which make possible for reactive modules M^M to bring the agents either left or right of the central lights L^2/R^2 . For example, highly rewarded strategies are those in which $C O$ values are close to the upper bound when the agents are located in E^{clock} , and to the lower bound when the agents are located in E^{anti} , or vice versa.

Summarising, each fitness component is in the range between 0 and 1, and the maximum fitness score a group can obtain is $FF = 3$. However, the final fitness of a successful group (i.e. a group in which the agents repeatedly swap arena side without touching the wrong arm of the revolving door) does not need to be exactly 3, but rather within the following interval: $1.8 < FF < 3$. This is because groups can be successful without necessarily obtaining the highest reward by the components F^2 and F^3 . For example, the revolving door does not need to be fully open for the agent to swap side. The two parts of Eq. 6 do not need to return exactly 1 for the agents to be able to adjust their pushing strategy according to the direction of rotation of the revolving door.

5 Results

Ten evolutionary simulations, each using a different random initialisation, were run for 6,000 generations. The evolutionary trend of the fitness of the best group at each generation is the first element that can be used to estimate whether a run produced successful groups or not. As shown in Fig. 3, four evolutionary runs produced best groups that are potentially successful.² In particular, two evolutionary runs produced best groups with the highest final fitness (i.e. run₁ and run₂ in Fig. 3a, b, respectively). Two evolutionary runs produced best groups whose final fitness is within the interval that may denote a successful strategy (i.e., run₃ and run₄ in Fig. 3c, d, respectively). Although all these best groups are potentially capable of accomplishing the task, the effectiveness and the robustness of their collective strategies have to be further estimated with more severe post-evaluation tests. In Sect. 5.1, we illustrate the results of post-evaluation tests

² Data concerning unsuccessful evolutionary runs, other data not shown in the paper, movies of successful post-evaluated groups, and other methodological details can be found at <http://laral.istc.cnr.it/elio.tuci/suppPagn/J13/suppMatJ13.html>.

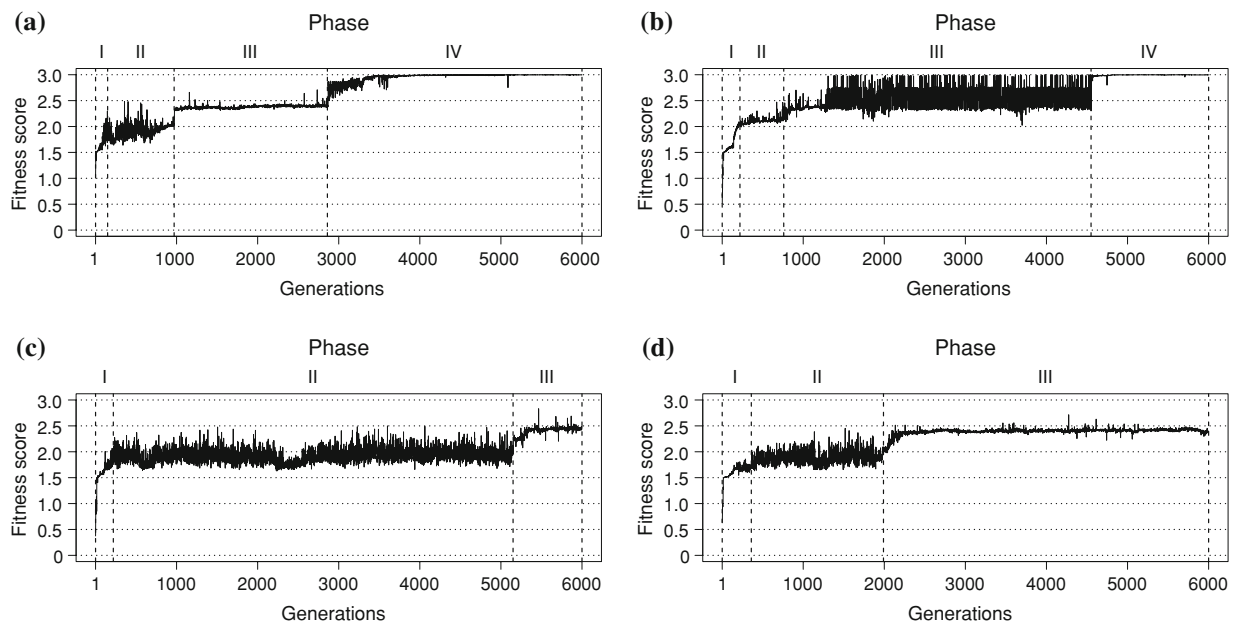


Fig. 3 Graphs showing the fitness of the best groups at each generation of the best four evolutionary runs: **a** run₁, **b** run₂, **c** run₃ and **d** run₄. Vertical dashed lines delimit the evolutionary phases

that tell us whether the best evolved groups are successful or not and what are the behavioural strategies they employ to accomplish the task. In Sect. 5.2, we talk about the evolutionary transitions which characterised the emergence of the best evolved strategies.

5.1 A first series of post-evaluation tests

From each run_{*i*}, we picked one of the best groups, hereafter referred to as *g_i*, chosen amongst those with the highest fitness. These best groups have been post-evaluated on all the possible four-trials sequences that differ in the ordering of the environments, without repetitions, and in which each environment appears at least once. Each sequence is repeated 100 times, with each repetition differently seeded to guarantee random variations in the initial position and orientation of the agents and in the noise added to sensors readings and motor outputs. Agents’ controllers are reset at the end of each four-trials sequence. This test, made of 2,400 four-trials sequences (i.e. 100 * *N!*, with *N* = 4), and a total of 9,600 trials, is referred to as test P. A post-evaluation trial is considered successful if the agents: (i) complete the first part of the task, (ii) move towards the revolving door, (iii) rotate the door without touching the current wrong arms and (iv) travel into the opposite side of the arena up to a distance of 24cm from the door hub, without colliding with the arena walls. If one or both agents collide with the current wrong arms of the revolving door, which trial is considered a failure of type *W*¹. Trials, in which one or both agents collide with the arena walls, are considered failure of type *W*². Failure

of type *W*³ refers to those trials in which the agents are not successful for other reasons than those considered in *W*¹ and *W*². For example, a failure of type *W*³ might concern a trial in which the agents do not manage to turn *L*¹ and *R*¹ off, or in which they do not manage to sufficiently rotate the revolving door within the time-limit.

Table 1 shows for each group and for each environment the percentage of success *S*, and the percentage of unsuccessful trials due to *W*¹, *W*² or *W*³ error. Data confirm that the groups with a success rate higher than 80% in all environments are those selected from run₁ and run₂ (i.e. *g*₁ and *g*₂, see Table 1, columns *S*). *g*₃ and *g*₄ are very good in carrying out the task in three out of four environments. For both groups, the performances in one type of environment are severely compromised by *W*¹ error. That is, at least one agent systematically touches the current wrong arm of the revolving door when located in *E*⁰⁰ for *g*₃, and in *E*¹¹ for *g*₄ (see Table 1, columns *W*¹). Post-evaluation tests carried out on best groups of other unsuccessful runs (i.e. run₅ to run₁₀), showed that these groups also fail primarily in *E*⁰⁰ and in *E*¹¹. This can be explained by the fact that while the two instances of the environment *E*^{clock} are two distinctive cases for the agents (i.e. either both agents experience grey or both black ground), the two instances of the environment *E*^{anti} are indeed a single case (i.e. one agent experiences grey ground while the other black ground). Thus, during evaluation, groups encounter the environmental circumstances referred to as *E*^{anti} twice as much as each of those referred to as *E*^{clock}. In term of fitness, in particular, with respect to component *F*², this means that a group that can successfully handle the combination

Table 1 Results of post-evaluation tests P showing for groups g_1 , g_2 , g_3 and g_4 , and for each environment the percentage of trials: (i) terminated successfully, see column S , (ii) in which the agents failed since one or both of them touched the current wrong arms of the revolving door, see column W^1 , (iii) in which the agents failed since one or both of them collided with the arena walls, see column W^2 and (iv) in which the agents failed for reasons other than those mentioned at (ii) and (iii), see column W^3

	S	W^1	W^2	W^3
E^{10} (%)				
g_1	82.21	4.08	1.12	12.58
g_2	97.87	0.46	0.00	1.67
g_3	100.00	0.00	0.00	0.00
g_4	98.42	0.00	1.58	0.00
E^{01} (%)				
g_1	81.87	4.62	0.67	12.83
g_2	97.62	0.46	0.00	1.92
g_3	100.00	0.00	0.00	0.00
g_4	98.33	0.04	1.62	0.00
E^{00} (%)				
g_1	97.00	0.08	1.25	1.67
g_2	99.75	0.00	0.00	0.25
g_3	0.00	100.00	0.00	0.00
g_4	96.00	0.08	1.33	2.58
E^{11} (%)				
g_1	89.79	3.42	1.08	5.71
g_2	99.42	0.08	0.00	0.50
g_3	99.92	0.08	0.00	0.00
g_4	0.00	99.29	0.71	0.00

grey/black zones immediately gets approximately twice the score of a group that can only accomplish the task in a single instance of E^{clock} . This represents a strong selective pressure that favours the emergence of groups that are successful in E^{anti} . This can also explain why, in every evolutionary runs, the capabilities required to solve the task in E^{anti} appear earlier in evolution than the capabilities required to solve the task in E^{clock} . The extent to which this ‘bias’ favours/hinders the evolution of fully successful groups has to be established with further tests.

In order to understand more about the nature of collective strategies of groups g_1 , g_2 , g_3 and g_4 , we run simple tests in which we let the groups undergo the four-trials sequence E^{10} , E^{01} , E^{00} and E^{11} , without resetting the agents controllers. At each time step, we recorded, for each agent, the sound output SO , and the categorisation output CO . This analysis revealed that both successful groups share the same strategy. That is, during the four-trials sequence, these groups move back and forth between two states: a NS -state in which there is no sound in the arena because none of the agents is emitting, and a S -state in which there is sound as either one or

both agents are emitting. When the group is in NS -state both agents set their CO to the lower bound. When the group is in S -state both agents set their CO to the upper bound. If during the *approaching phase*, the group is in NS -state, both agents tend to keep L^2 and R^2 on their respective right sides, and push the revolving door in a clockwise direction. If during the *approaching phase*, the group is in S -state, both agents tend to keep L^2 and R^2 on their respective left sides, and push the revolving door in an anticlockwise direction.

In what remains of this section, we look at how agents of g_1 and g_2 set the sound output and the categorisation output in each environment, and we illustrate the principles that regulate the transitions between NS -state and S -state.

At the beginning of each trial in any environment, the agents do not emit sound signals and the value of their CO is close to the lower bound. When the groups are located in E^{00} , the perception of grey ground does not produce any change. The group remains in NS -state throughout the entire trial and the agents end up correctly pushing the revolving door in a clockwise direction. When the groups are located in E^{10} or in E^{01} , the agent that perceives black ground emits, for few seconds, intermittent bursts of sound. The other agent is induced to signal as soon as it perceives sound while on grey ground.³ This agent keeps the group in S -state by emitting for the entire *approaching phase*. The presence of sound in the arena makes both agents set their CO to the upper bound. Thus, while approaching the revolving door, they tend to keep L^2 or R^2 on their respective left sides, and they end up correctly pushing the revolving door in a anticlockwise direction. When the groups are located in E^{11} , the presence of two black zones does not produce any long lasting change. Both agents emit sound due to the perception of black floor, and both of them temporarily set their CO to the upper bound. However, both agents stop emitting very quickly, and their CO return to values close to the lower bound. Thus, during the *approaching phase*, the group remains in NS -state. The agents keep L^2 and R^2 on their respective right sides as they do when located in E^{00} , and they end up pushing the revolving door in a clockwise direction. Note that long lasting signalling behaviour tends to last less than the time required to complete the second phase of the task. Thus, a group is in S -state only during part of the second phase of the task, while the agents are approaching the revolving door. Since S -state never lasts longer than the end of the current trial, the agents possess the required plasticity to successfully cope with any differently ordered four-trials sequence.

³ Note that, both in trials following a successful one, and after repositioning, both agents arrive on the painted zone almost at the same time. This is the reason why, when successful groups are located in E^{10} or in E^{01} , the emission of intermittent bursts of sound by the agent on black takes place when the other agent is on grey ground.

In order to summarise, our analysis revealed that signalling behaviour varies with respect to the combination of colours of the painted zones, and that each agent regulates its categorisation output in response to the perception of sound. In particular, during the first phase of the task before reaching the painted zone, none of the agents is emitting. The perception of grey ground does not trigger any signalling response. The perception of black ground triggers the emission of intermittent bursts of sound that lasts no more than 2 or 3 s. The perception of sound, while on grey ground triggers the emission of a continuous tone that lasts for about 40 s. An agent sets its categorisation output to its upper bound if it perceives sound, and to its lower bound if it perceives no sound. Thus, in E^{10} and E^{01} , successful groups perform a large part of the second phase of the task in *S-state*. In E^{00} , successful groups perform the entire trial in *NS-state*. In E^{11} , successful groups are in *S-state* only during the last part of the first phase of the task, and perform the entire second part of the task in *NS-state*, like in E^{00} .

Groups g_3 and g_4 exploit similar mechanisms as g_1 and g_2 . They vary the sound output in response to the perception of painted zones, and the categorisation output in response to sound. In contrast to what observed in successful groups, group g_3 is by default in *S-state*, with both agents emitting sound at the beginning of each trial. Signalling is temporarily interrupted by the perception of black ground. When located in E^{11} , both agents stop emitting during the entire *approaching phase*. The group switches from the *S-state* to the *NS-state*, and both agents end up pushing the revolving door on the correct arms. When located in E^{01} and in E^{10} , the agent that experiences grey never stops emitting sound. The group remains in *S-state*, and both agents end up pushing the revolving door on the correct arms. When located in E^{00} , the group fails since it lacks the conditions to switch from the *S-state* to the *NS-state*. The group remains in *S-state* for the entire length of the trial, and both agents end up pushing the revolving door on the wrong arms. Group g_4 behaves in a very similar way to what observed in successful groups. However, in contrast to them, in group g_4 , the perception of black ground triggers long lasting signalling behaviour. When located in E^{11} , the group remains in *S-state*, and both agents end up pushing the revolving door on the wrong arms.

5.2 An evolutionary analysis

As shown in the previous section, to successfully discriminate E^{anti} from E^{clock} , the agents of successful groups g_1 and g_2 (i) complement each other's perceptual experience using communicative strategies based on their simple acoustic system, (ii) develop cooperative pushing strategies through mechanisms that set and 'parse' the categorisation outputs and (iii) exploit a certain amount of neural plasticity to continuously track environmental changes without being

perturbed by previous experience. Unsuccessful groups g_3 and g_4 fail in one environment because in that circumstance they are not capable of choosing, through acoustic interactions, the required pushing strategy.

The aim of this section is to provide empirical evidence that illustrates how and why the fitness function illustrated in Sect. 4.3 facilitates the emergence of the neural apparatus required by the agents to solve this cooperative task. In order to do so, we post-evaluated, using test P illustrated at the beginning of Sect. 5, each best group from generation 1 to generation 6,000 of all the 10 evolutionary runs. This analysis produced a huge amount of data, with which we managed to observe what happened during each evolutionary run. In particular, by comparing all the evolutionary runs on a series of performance measures obtained from tests P, we managed to isolate a distinctive pattern that can be considered the evolutionary milestone in the emergence of successful collective strategies. In the remainder of the section, we illustrate this pattern as observed between generation 100 and generation 200 of run₂.

As shown in Fig. 4a, during evolution, the best group at generation 200 shows higher final fitness than the best group at generation 100. However, if post-evaluated with test P, all the best groups between generation 100 and 200 do not differ in terms of percentage of success per environment. All of them are extremely successful in rotating the revolving door without touching the wrong arms in E^{10} and E^{01} , and very unsuccessful in performing the same task in the other two environments (see Fig. 4b). Post-evaluation tests showed also that failure in E^{00} and E^{11} of these groups are due to a high percentage of W^1 error (data not shown). These results tell us that all the groups between generation 100 and 200 are equally good, as far as, it concerns the behaviours rewarded by the fitness components F^1 and F^2 . That is, they are equally good in navigating the arena, in avoiding penalties such as those due to collisions with the arena walls, and in rotating the revolving door. Thus, the fitness increase observed in Fig. 4a can only be due to the emergence of behavioural traits rewarded by the fitness component F^3 . Indeed, our analysis revealed that this fitness increase is determined by the fitness component F^3 through its effects on the emergence of the mechanisms that regulate the emission of sound and the setting of the categorisation output.

In order to capture these effects, we plotted the proportion of time with sound in the arena and the average categorisation output of the group during the *approaching phase* (see Fig. 4c, dashed lines and continuous lines, respectively). The graphs indicate that less fit groups continuously emit sound and set their categorisation output to the lower bound in all environments. Fitter groups vary their signalling behaviour and categorisation output. That is, they always signal and have the categorisation output set to the upper bound in E^{10} , E^{01} and in E^{11} , and they never signal and have the categor-

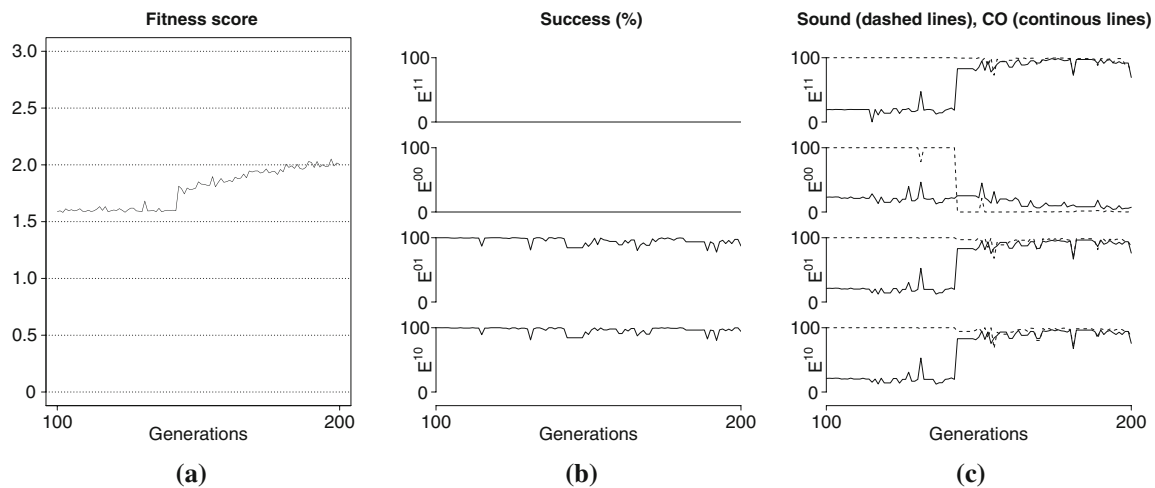


Fig. 4 Graphs showing various performance measures, of the best groups from generation 100 to 200, of run₂. **a** Final fitness (FF) of the best groups at each generation, as observed during evolution. **b** and **c** show data gathered by post-evaluating the above-mentioned best groups with test P illustrated at the beginning of Sect. 5. In particu-

lar, the graphs refer to: **b** the percentage of success in each environment, **c** the proportion of time with sound in the arena during the *approaching phase* (dashed lines) and the average categorisation output during the *approaching phase* (continuous lines), in each environment

isation output set to the lower bound in E^{00} . This distinctive differentiation, observed only in fitter groups, between what the agents do when located in E^{10} , E^{01} and in E^{11} and what they do when located in E^{00} , can be explained by the emergence of mechanisms that regulate the sound and categorisation output in response to environmental stimuli. In particular, at generation 200, the perception of black ground triggers the emission of sound, and the perception of grey ground interrupts signalling behaviour. The categorisation output varies in response to the perception of sound. With no sound in the arena, the agents set their categorisation output to the lower bound; with sound in the arena, the agents set their categorisation output to the upper bound.⁴ Due to these causal links, during the second phase of the task, one of the agents emit sound when the group is located in E^{10} , E^{01} and both of them signal when the group is located in E^{11} . None of the agents signals when the group is located in E^{00} . Note that, at this evolutionary stage, variations of the categorisation output do not induce any change in the pushing strategy. Regardless of their state, the groups always push the revolving door in an anticlockwise direction.

It is worth mentioning that the simultaneous appearance of the causal relationships between the perception of the painted

zones and signalling, and between signalling and the values of the categorisation output, is the phenomenon that enables the emergence of successful groups. This is because the emergence of single causal links, frequently observed in unsuccessful evolutionary runs, proved not capable of bootstrapping a virtuous evolutionary process. For example, in some unsuccessful evolutionary runs, evolution produced groups in which the categorisation output varies randomly and signalling behaviour varies in an unstructured way due to the emergence of mechanisms that causally link the emission of sound to the perception of black or grey ground. Since signalling is not directly rewarded by the fitness function, these groups are as fit as groups in which sound signalling behaviour varies randomly. Thus, the phenotypic traits of the former groups have the same chance of being transmitted to the following generations than the phenotypic traits of the latter groups. We noticed that, in these circumstances, the mechanisms that causally link the emission of sound to the colour of the painted zone tend to disappear from the population before further evolutionary transitions causally link the perception of sound to the setting of the categorisation output.

Similarly, some evolutionary runs produced groups in which signalling behaviour varies randomly and the categorisation output varies in a structured way due to the emergence of mechanisms that causally link the setting of the categorisation output to the perception of black or grey ground. Due to the fitness component F^3 , these groups tend to be comparatively fitter than groups in which the categorisation output varies randomly. Therefore, they invade the population. However, a population made only of these groups turns out to be an

⁴ Note that, graphs in Fig. 4c exclude the existence of alternative causal relationships between the colour of the painted zone and the categorisation output. For example, if the agents set the categorisation output to the upper bound in response to the perception of black ground and to the lower bound in response to the perception of grey ground, then the average CO values recorded in tests P in E^{10} , and in E^{01} would be close to 0.5 rather than close to 1, due to the fact that the agents experience different colours.

evolutionary deadlock. In order to make any further progress, evolution needs to favour groups whose neural mechanisms underpin completely different behaviours. Moreover, these mechanisms should appear all at once to avoid to pass through less fit or equally fit behavioural alternatives. For example, it would be hard for evolution to favour the appearance of further behavioural traits (e.g. any causal links between sound output and the perception of painted zones) that, although necessary to built successful strategies, may not be immediately adaptive (i.e. associated to a fitness increase).

The simultaneous appearance of the causal relationships between the perception of the painted zones and signalling, and between signalling and the values of the categorisation output observed in both successful runs has also been found in run₃ and run₄. In run₃ and run₄, these transitions appear much later than in run₁ and run₂, around generation 5,000 for run₃, and around generation 2,000 for run₄. We wondered whether continuing these simulations longer than 6,000 generations would have produced a successful group. However, further tests in which we let these simulations run for up to 8,000 generations did not produce any successful group. Clearly, the evolutionary transitions illustrated in Fig. 4, it represent a significant step further towards the emergence of successful collective strategies. However, at the first appearance of the causal relationships between the perception of the painted zones and signalling, and between signalling and the categorisation output, groups are still far away from successfully accomplishing their task in all the environments. In order to get to this final point, best groups go through various behavioural changes, in which the causal relationships between the perception of the painted zones and signalling behaviour change rather frequently, and in different ways in the two successful evolutionary runs.

By looking at the fitness curves shown in Fig. 3, we notice that, in all the four runs, there is a very short evolutionary time, referred to as phase I, in which the agents develop the mechanisms to perform phototaxis without colliding with the arena walls. In run₂, phase I is characterised by the emergence of the mechanisms illustrated in Fig. 4. Very soon, all the runs reach phase II, characterised by a plateau with $FF \simeq 2$ (see Fig. 3). At this stage, the groups are capable of accomplishing the task only in E^{anti} , and they systematically commit W^1 error when located in E^{clock} . The fitness components approximately contribute to the groups' final fitness in the following: $F^1 = 1$, $F^2 \simeq 0.5$ and F^3 oscillates between 0 and 0.8. The oscillations of the final fitness recorded almost for the entire phase in run₁, run₃, and run₄, are due to F^3 . This is because, for the agents of these best groups, the value of CO does not seem to be causally determined by any of their perceptual experiences, such as the perception of sound, or of the painted zone. In the last part of the phase in run₁ and run₂, there are only small oscillations as the agents set the categorisation output, according to the dynamics illustrated

in Fig. 4. In phase II, we observe the emergence of important mechanisms for simple social behaviour. For example, in all the four runs, there are groups for which the perception of sound triggers signalling behaviour in agents that are not signalling yet. This type of social interaction allows the agents to engage themselves in simple turn-taking. Most of these turn-taking appear and disappear throughout phase II, since they are adaptively neutral. Those that remain into the populations are the one responsible for the transition from phase II to phase III.

In phase III, characterised by a plateau with $FF \simeq 2.5$, the groups are capable of accomplishing the task in three environments out of four. The main difference between phase II and phase III groups is that in phase III, the agents possess neural mechanisms that causally link the setting of the categorisation output to the state of the group, and they vary signalling behaviour in response to the perception of the painted zones. Moreover, these groups possess neural mechanisms by which different values of the categorisation output trigger different behavioural responses. Thus, agents begin approaching the revolving door by keeping the central light on their left or on their right side according to the values of the categorisation output generated by the module M^C . However, run₁'s agents in E^{11} , run₃'s agents in E^{00} and run₄'s agents in E^{11} , systematically commit W^1 error, since in these environments, they are not capable of appropriately defining the state of the group to trigger the correct pushing strategy. For groups in run₁, run₃ and run₄, the fitness components approximately contribute to the groups' final fitness in the following: $F^1 = 1$, $F^2 \simeq 0.75$ and $F^3 \simeq 0.75$, with very small variability amongst groups of different generations. In run₂, the rather large variability amongst groups of different generations is due to the fact that, for specific sequencing of environments, the groups are capable of accomplishing the task in all environments. In particular, to be successful in E^{00} , the agents need to approach the revolving door in NS -state. However, this happens only if the agents encounter this environment before anyone else or after having experienced E^{11} . This is because, only the perception of a black zone by both agents inhibits signalling behaviour previously triggered by an experience in a E^{anti} type of environment.

Phase IV, characterised by a plateau with $FF \simeq 3$, is reached only by run₁ and run₂ and it refers to groups that are almost 100% successful in all the environments (see Fig. 3a, b). The transition from phase III to phase IV is due to mutations that 'tunes' the causal relationships between the perception of the painted zone, signalling behaviour and the setting of the categorisation output in a way that, the agents, whether located either in E^{anti} or in E^{clock} , by interacting through sound signals, generate different values of the categorisation output. Consequently, the agents perform different pushing strategies, which bring them to push the correct arms of the revolving door.

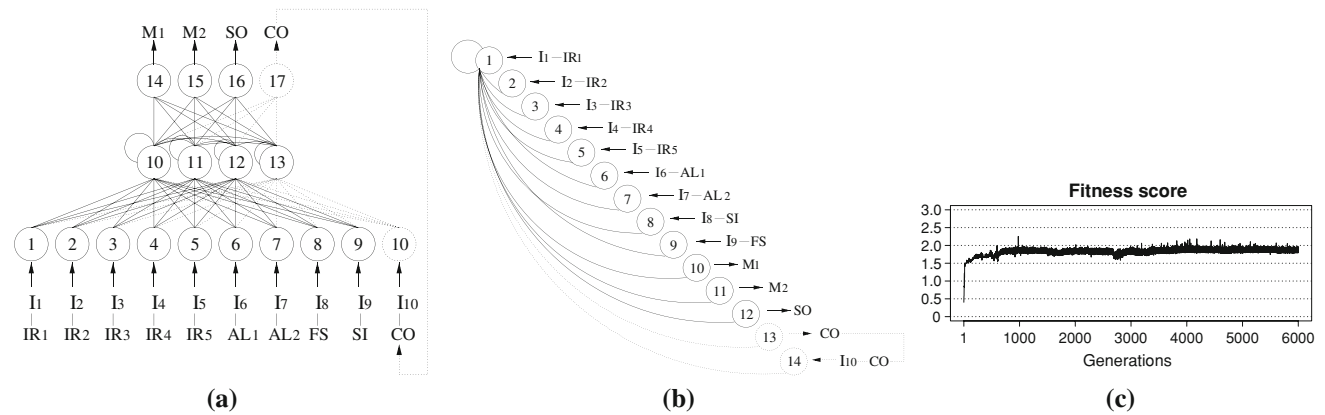


Fig. 5 Non-modular **a** three layers control architectures with a fully recurrent hidden layer (neurons 10–13), and **b** fully recurrent control architectures. **c** Graph showing the fitness of the best groups at each generation of an unsuccessful evolutionary run. In this simulation, the agents are controlled by a non-modular three layers control architec-

tures with categorisation output. In **a** and **b**, dotted lines refers to those structures added to the network to take into account the categorisation output. In both architectures, the categorisation output produced at time t is sensed by the network at time $t + 1$. In **b**, only the efferent connections for one neuron are drawn

6 Discussion

The analysis, detailed in Sect. 5.2, is an important contribution of this study which shows the most significant evolutionary transitions that paved the way for the emergence of the mechanisms underpinning reciprocal communication, of categorisation capabilities and of successful cooperative strategies. In this section, we further comment the outcome of this analysis with observations concerning the architecture of the agents' controller, and the significance of the fitness component F^3 .

The modular structure described in Sect. 4.1 and the fitness component F^3 described in Sect. 4.3 have been included into the model after having explored a variety of different alternatives, with no modules and no categorisation output, and in which the agents were uniquely scored for their capability to perform phototaxis and to rotate the revolving door. In particular, we tested two different architectures: a three layers network shown in Fig. 5a, in which the state of the sensory and motor neurons were updated in the same way as the corresponding neurons in module M^M , and the state of internal neurons were updated in the same way as neuron 3–6 in module M^C ; and a fully recurrent network shown in Fig. 5b, in which the state of sensory and non-sensory neurons were updated in the same way as the corresponding neurons in module M^C (see Sect. 4.1 for details). With both networks, we carried out several simulations varying in the three layers network, the number of internal neurons, and in the fully recurrent network, the number of neurons that did not receive external input and were not used to set the agents' actuators.

Repeated failure brought us first to the introduction of the categorisation output. Dotted lines in Fig. 5a and 5b refer to the structural changes added to networks to take into account

the categorisation output. This output was introduced to directly score the agents, through the fitness component F^3 , for their ability to categorise the environments with respect to the combination of colour of the painted zones. Note that without F^3 , both communication and the capability to carry out cooperative strategies, if taken in isolation, are adaptively neutral, since they do not, per se, increase the reproductive chance of the groups that possess them. In other words, without F^3 , a good communication system without the capability to perform different cooperative strategies is useless because there is no fitness gain in having good communication if it does not contribute to carry out the cooperative task. The capability to perform different cooperative strategies without an effective communication system is also useless as there is no fitness gain in being potentially capable of performing actions without being capable of knowing under which circumstances to perform them.

F^3 guides evolution towards the emergence of the mechanisms for communication and categorisation without directly rewarding specific types of communication. In early evolutionary stages, communication systems are generally too simple and primitive to trigger the required collective responses. Thus, if performances are evaluated only with fitness component F^1 and F^2 , variability in signalling behaviour does not correspond to a concrete fitness difference amongst the groups. With F^3 , early generations groups in which any form of signalling interactions favour the discrimination between E^{anti} from E^{clock} through the generation of the categorisation output, have a selective advantage over groups in which signalling interactions do not produce the same effect. Thus, the former groups take over the population. In this way, we open an alternative evolutionary path to the origin of signalling. However, for all the simulations with no modular architecture, with or without categorisation output, the fitness

curve was always following the same trend, with evolution reaching very quickly and permanently a plateau around a fitness score corresponding to the performance of a group capable of accomplishing the task either in E^{anti} or in E^{clock} . In other words, these runs did not progress further than phase II described in Sect. 5.2. Figure 5c shows the fitness curve of an unsuccessful evolutionary run, in which the agents were controlled by a three layers network with the categorisation output. Only with the use of the modularised controller, categorisation output and fitness component F^3 produced the desired results. What did modularisation add to the previous unsuccessful models?

The results of our study suggest that structural properties of the agent's controller facilitate the emergence of the behavioural variability that is necessary for selecting those variants that several generations later give birth to successful groups. While in evolutionary runs with non-modularised controllers all the groups, since the earlier generations, were either signalling all the time or never; in evolutionary runs with modularised controllers, we observed the coexistence in the same population of some variability in signalling behaviour. As shown in Sect. 5.2, the fitness function F^3 favoured the emergence of successful groups by rewarding those signalling behaviours that induced systematic variation in the generation of the categorisation output. We wish to remark that, in spite of the hand-coded modularisation of the neuro-controller, communicative behaviour stays firmly and intentionally outside the bag of assumptions made by the experimenter, and entirely left to evolution.

In order to conclude, we would like to draw the attention of the reader to an important issue not mentioned before, which concerns the memory structures required by the agents to solve this cooperative task. Memory is needed by the agents because the action of pushing the revolving door happens several time steps later with respect to the perception of the environmental cues that represent the sources of 'information' for deciding the pushing strategy. Results of post-evaluation tests indicate that, quite unexpectedly, in both best evolved strategies, memory structures and means for communication are rather intertwined.⁵ In particular, we saw that, during the *approaching phase*, the agents proved to be extremely sensitive to small disruptions on their acoustic system. That is, their capability to successfully accomplish the task is largely reduced if, even for short intervals (i.e., 1 s), sound is present while no agent is emitting or if the agents are made deaf to their own signals. Thus, it seems that sound is employed by the agents not only as a means to acoustically interact, as illustrated above, but also as a kind of external digital memory which, together with the categorisation output, reminds them, during the *approaching phase*, what is the pushing strategy

they are currently employing. Sound as an external memory seems to be an efficient solution, in particular, if compared to the alternative of having individual neural structures dedicated to memory. This is because, on the one hand, sound as an external memory reduces the cognitive load imposed by the task. The agent do not need to remember what pushing strategy to employ during the *approaching phase*. They simply act in response to the perception of sound. On the other hand, the external memory is shared by both agents, reducing the risks of failure due to forgetting, and improving the robustness of the collective strategies.

7 Conclusions and future work

Recent technological innovations provided scientists with the means to develop computational models of evolving embodied agents. Amongst the various applications, these technologies are used to investigate biological/ethological issues concerning the evolution of individual and social behaviours and their underlying causal mechanisms (see Harvey et al. 2004). Following a theoretical line of investigation focused on the evolutionary origin of communication, the model illustrated in this paper provides further evidence of the evolutionary circumstances which facilitate the emergence of reciprocal communication in a population of evolving agents. This model is characterised by the fact that communication is needed by a group of two homogeneous agents to choose, on the basis of perceptual cues that are only partially available to each single agent, the required cooperative strategy between two possible alternatives. Communication allows the agents to reciprocally complement their partial perceptual experience, to correctly and repeatedly categorise the environment in which they behave, and to develop common and effective cooperative strategies. The scientific contribution of this research work is in showing how reciprocal communication can originate from scratch in a context in which it contributes to perform a categorisation task. We demonstrated that artificial evolution managed to successfully assemble the neural mechanisms required by the agents for (i) developing a simple form of reciprocal communication, (ii) distinguishing two different types of environment and (iii) performing a simple cooperative task. Acoustic communication fulfils its function thanks to its effect on the categorisation output, which, in turn, determines the agents' pushing strategy.

Two elements were responsible for the evolution of successful strategy: the modular structure of the agents' controller, and the selective pressure which made signalling behaviour an adaptive trait even if not yet directly helpful to allow the agents to open the revolving door. The modular architecture helped in generating the phenotypic variability in signalling behaviour required by artificial selection to reward traits that proved to be more adaptive. In early stages

⁵ A detailed description of these post-evaluation tests can be found at <http://laral.istc.cnr.it/elio.tuci/suppPagn/J13/suppMatJ13.html>.

of successful evolutionary runs, the fitness component F^3 contributed to keep into the evolving populations those signalling behaviours that, although extremely simple and not particularly helpful to correctly categorise the environments, were relevant to the categorisation task.

Concerning future work, we will first look at the operational principles of the neural mechanisms underlying successful as well as unsuccessful cooperative strategies. The work described in this paper will be complemented by further analysis on the structural properties of the evolved neuro-controllers. There are several aspects of this model that, being arbitrarily chosen amongst a variety of alternatives, should be further questioned. For example, alternative solutions concerning the structural properties of the agents' controller as well as the parameters of the evolutionary algorithms should be further investigated. The modular approach described in Sect. 4.1 and the evolutionary regime described in Sect. 4.2, although more effective than other choices, generated adequate variability in signalling behaviour only in 2 out of 10 randomly seeded evolutionary runs. The other eight evolutionary runs quickly converged on populations, in which the groups conformed to single strategies. That is, both agents were either signalling all the time or never. Further investigations will focus on the suitability of functional (e.g. the activation function of the neurons) and structural (e.g. the sound signalling system of the agent and the neural network connectivity) alternatives that facilitate the emergence of phenotypic variability with respect to sound signalling behaviour. More effective models will allow us to look at the origins of communication in more complex simulated scenarios.

There are other implementation details that bear upon the nature of the evolved cooperative strategies and of the communication protocols. The genetic relatedness of the individuals of a group is one of these details. We employed groups in which the agents share the same genetic material. The genetic homogeneity of the agents helps to avoid the effects of combinatorial affinity. Combinatorial affinity refers to a cooperative scenario in which genetically non-related individuals are effective only in combination with a subset of all possible partners. Combinatorial affinity may emerge in an evolutionary scenario, in which individuals are evaluated only with a subset of all possible partners. Due to biases in the evaluation procedure, the agents' fitness may result over or under-estimated, leading to the emergence of fragile strategies strongly dependent on the affinity amongst the partners. However, in spite of this problem, whose effects can be minimised by carefully adjusting the evaluation procedure, the study of our scenario in heterogeneous groups may generate different evolutionary dynamics and alternative solutions for what concerns collective strategies and provide further

interesting insight about the origins of reciprocal communication. The sound signalling system is another implementation detail on which to concentrate future work. As shown in Sect. 4.1, we employed a signalling system as simple as possible. The agents emit the same single tone signal, and they do not have any built-in mechanism to distinguish self- and non-self-signal. There are various ways in which the signalling system can be modified and made more complex. The consequences produced by any possible modification of the signalling system on the evolution of collective strategies cannot be easily predicted. Thus, they have to be systematically investigated with further simulations.

We believe that the main contribution of ER models is to provide proof-of-concept demonstrations of facts that cannot be directly observed and/or scientifically investigated, such as those concerning the evolutionary history of the behaviour of living organisms. In this spirit, ER models should keep on generating simple ecological scenarios which, by working on the characteristics of the agents sensory and motor capabilities as well as on the properties of the environment, generate the neural mechanisms underlying incrementally complex individual and social skills. Concerning communication, our intention is to invest more energy into the study of the origin of social interactions that can be used to communicate about states of affairs that are distant both in space and in time (i.e. referential communication). As far as we know, these types of interaction are not very common in biological organisms, with the waggle dance of honeybees being the most studied example (Von Frisch 1967). Nevertheless, they are absolutely ordinary in human language. Thus, ER models focused on the evolutionary origin of referential communication may shed light on the origin of language and provide further evidence concerning its biological roots. Moreover, the interest on referential communication is also due to the fact that these types of social interaction supposedly require cognitive functions (e.g. the capability to form mental representations of perceptual experience and to store them in memory) whose structural and functional identities are highly debated in cognitive science. ER models on the evolution of referential communication may be effective tools to generate helpful insights that can be further tested in a more neuro-scientific experimental setup. The work described in Williams et al. (2008), although in a very simple scenario, has already begun to look at these issues. The model described in this paper requires only minimal changes to contemplate referential communication. In particular, it suffices to limit the agents' acoustic perceptual field, so that sound is heard only if agents are at a distance significantly shorter than the arena length. This would be an interesting direction for future work along a line of research which

is meant to strengthen the significance of ER models to the above aforementioned biological and cognitive science debates.

Acknowledgements The author would like to thank his former colleagues at IRIDIA (Université Libre de Bruxelles, Belgium), in particular, Christos Ampatzis and Marco Montes De Oca for his valuable support during the development of this research work, and Carlo Pinciroli for his support with ODE. The author also thanks his current colleagues at LARAL (ISTC-CNR, Italy), in particular, Marco Mirolli and Vieri Santucci, for stimulating discussions and feedback during the preparation of this paper, and Prof. Justus Piater for the logistic support provided during his three months visit at Université de Liege, Belgium. Special thanks to Cécile and Leonardo. This research study was supported by the *ITALK* project (EU, ICT, Cognitive Systems and Robotics Integrating Project, Grant No. 214668).

References

- Ampatzis C, Tuci E, Trianni V, Dorigo M (2008) Evolution of signaling in a multi-robot system: categorization and communication. *Adapt Behav* 16(1):5–26
- Balch T, Arkin RC (1994) Communication in reactive multiagent robotic systems. *Auton Robot* 1(1):27–52
- Baldassarre G, Nolfi S, Parisi D (2003) Evolving mobile robots able to display collective behaviour. *Artif Life* 9:255–267
- Beer RD, Gallagher JC (1992) Evolving dynamic neural networks for adaptive behavior. *Adapt Behav* 1(1):91–122
- Cangelosi A, Parisi D (1998) The emergence of a “language” in an evolving population of neural networks. *Connect Sci* 10(2):83–97
- Cangelosi A, Parisi D (eds) (2002) *Simulating the evolution of language*. Springer, London, UK
- Cao YU, Fukunaga AS, Kahng A (1997) Cooperative mobile robotics: antecedents and directions. *Auton Robot* 4(1):7–27
- Dawkins R, Krebs JR (1978) Animal signals: information or manipulation? In: Krebs J, Davis N (eds) *Behavioural ecology: an evolutionary approach*. Blackwell, Oxford, pp 282–309
- Di Paolo EA (2000) Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adapt Behav* 8(1):27–48
- Floreano D, Mitri S, Magnenat S, Keller L (2007) Evolutionary conditions for the emergence of communication in robots. *Curr Biol* 17:514–519
- Fong T, Nourbakhsh I, Dautenhahn K (2002) A survey of socially interactive robots. *Robot Auton Syst* 42(3–4):143–166
- Goldberg DE (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, MA
- Harnad S (ed) (1987) *Categorical perception: the groundwork of cognition*. Cambridge University Press, New York
- Harvey I, Di Paolo EA, Wood R, Quinn M, Tuci E (2004) Evolutionary robotics: a new scientific tool for studying cognition. *Artif Life* 11(1–2):79–98
- Hauser MD (1996) *The evolution of communication*. MIT Press, Cambridge, MA
- Marocco D, Nolfi S (2007) Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connect Sci* 19(1):53–74
- Marocco D, Cangelosi A, Nolfi S (2003) The emergence of communication in evolutionary robots. *Phil Trans R Soc London Ser A* 361:2397–2421
- Maynard-Smith J, Harper DG (1995) *Animal signals: models and terminology*. *J Theor Biol* 177:305–311
- Maynard-Smith J, Price GR (1973) The logic of animal conflict. *Nature* 246:15–18
- Mirolli M, Parisi D (2008) How producer biases can favor the evolution of communication: an analysis of evolutionary dynamics. *Adapt Behav* 16(1):27–52
- Mondada F, Pettinaro GC, Guignard A, Kwee IV, Floreano D, Deneubourg JL, Nolfi S, Gambardella LM, Dorigo M (2004) SWARM-BOT: a new distributed robotic concept. *Auton Robot* 17(2–3):193–221
- Nolfi S, Floreano D (2000) *Evolutionary robotics: the biology, intelligence, and technology of self-organizing machines*. MIT Press, Cambridge, MA
- Pfeifer R, Scheier C (2001) *Understanding intelligence*. MIT Press, Cambridge, MA
- Quinn M (2001) Evolving communication without dedicated communication channels. In: Kelemen J, Sosik P (eds) *Advances in artificial life: 6th European conference on artificial life*. Springer, Berlin, Germany, pp 357–366
- Quinn M, Smith L, Mayley G, Husbands P (2003) Evolving controllers for a homogeneous system of physical robots: structured cooperation with minimal sensors. *Phil Trans R Soc Lond Ser A* 361:2321–2344
- Steels L (2003) Evolving grounded communication for robots. *Trends Cognit Sci* 7(7):308–312
- Støy K (2001) Using situated communication in distributed autonomous mobile robots. In: *Proceedings of the 7th Scandinavian conference on artificial intelligence*. IOS Press, Amsterdam, The Netherlands, pp 44–52
- Trianni V, Dorigo M (2006) Self-organisation and communication in groups of simulated and physical robots. *Biol Cybern* 95:213–231
- Tuci E, Ampatzis C, Vicentini F, Dorigo M (2008) Evolving homogeneous neuro-controllers for a group of heterogeneous robots: coordinated motion, cooperation, and acoustic communication. *Artif Life* 14(2):211–230
- Vicentini F, Tuci E (2006) *Swarmod: a 2d s-bot’s simulator*. Technical Report TR/IRIDIA/2006-005, IRIDIA, CoDE, Université Libre de Bruxelles. Available at <http://iridia.ulb.ac.be/IridiaTrSeries.html>
- Von Frisch K (1967) *The dance language and orientation of bees*. Harvard University Press, Cambridge, MA
- Williams PL, Beer RD, Gasser M (2008) Evolving referential communication in embodied dynamical agents. In: Bullock S, Noble J, Watson R, Bedau M (eds) *Proceedings of the 11th international conference on the simulation and synthesis of living systems, ALife XI*. MIT Press, Cambridge, MA, pp 702–709