# How producer biases can favor the evolution of communication: An analysis of evolutionary dynamics

Marco Mirolli, Domenico Parisi

Institute of Cognitive Sciences and Technologies, CNR, 44 Via San Martino della Battaglia, 00185 Rome, Italy

**Abstract**

As any other biological trait, communication can be studied under at least four perspectives: mechanistic, ontogenetic, functional, and phylogenetic (Tinbergen, 1963). Here we focus on the following phylogenetic question: how can communication emerge given that both signal-producing and signal-responding abilities seem to be adaptively neutral until the complementary ability is present in the population? We explore the problem of co-evolution of speakers and hearers with artificial life simulations: a population of artificial neural networks evolving a food call system. The core of the paper is devoted to the careful analysis of the complex evolutionary dynamics demonstrated by our simple simulation. Our analyzes reveal an important factor which might solve the phylogenetic problem: the spontaneous production of good (meaningful) signals by speakers due to the need for organisms to categorize their experience in adaptively relevant ways. We discuss our results with respect both to previous simulative work and to the biological literature on the evolution of communication.

**Keywords:** evolution of communication, producer bias, receiver bias, evolutionary dynamics

## 1  Introduction

### 1.1  The Phylogenetic Problem of Communication

The Nobel prize Nikolas Tinbergen, one of the founders of modern ethology, classified ethological questions into four main categories (Tinbergen, 1963). According to Tinbergen, animal behavior must be understood from four perspectives:

1. mechanistic: what are the mechanisms (e.g., neural or psychological) which determine the behavior?

2. ontogenetic: what are the genetic and environmental factors that contribute to the behavior's development?

3. functional: what are the effects of a given behavior on survival and reproduction?

4. phylogenetic: how did the behavior evolve during the history of the species?

As Tinbergen suggested, a full explanation of any given behavior (or biological trait) requires not only detailed answers to all these questions, but also an analysis of the relationships between the different but clearly related factors (e.g., how the mechanisms underlying a given behavior arise during ontogenetic development or arose during phylogenetic evolution). Nonetheless, as an heuristic rule, it seems appropriate to start studying each kind of problem separately, letting the relationships between them manifest themselves during this study.

In the present paper, the behavior that we are dealing with is communication, and the question we will focus on is the phylogenetic one. Furthermore, as it is true for most, even if by no means all, investigations in the simulation of adaptive behavior and artificial life communities, our approach is quite general and theoretically oriented. In other words, we will not address the evolution of any particular communication system of any given species. On the contrary, we are interested in the general question of how communication (any kind of communication) can emerge phylogenetically during the evolutionary history of a species. The rationale for doing this is that communication seems to pose a very general phylogenetic problem, a problem of co-evolution between signallers (speakers) and receivers (hearers).

The emergence of communication, in fact, requires the co-evolution of both good speakers and good hearers. Good speakers are individuals that produce signals that systematically co-vary with situations that are relevant for survival and reproduction. Good hearers are individuals that react to signals in ways that are appropriate to the situations with which signals systematically co-vary. Good speakers without good hearers are *useless* because signals are not understood and reacted to appropriately. Good hearers without good speakers are *meaningless* because one cannot react appropriately to signals that do not co-vary with some relevant feature of the world. As Maynard-Smith (1997) puts it: "It's no good making a signal unless it is understood, and a signal will not be understood the first time it is made" (p. 208). This simple fact renders the emergence of communication tricky. How can communication evolve if the traits which are necessary for its emergence – namely, good speaking and good hearing capabilities – taken in isolation are adaptively neutral in that they do not, *per se*, increase the reproductive chances of the individuals that possess them?

This phylogenetic problem of communication must not be confounded with an adaptive problem, which also stems from the complementarity between speakers and hearers. How can communication systems which do not benefit equally emitters and receivers emerge? Why should speakers send useful signals if they gain no advantage in doing so (as it seems to be the case for many alarm and food calls)? And why should receivers appropriately respond to signals if they are just being manipulated by signallers to the interests of the latter (as it seems to be the case for many aggressive signals or mating displays)? This adaptive problem is clearly related to the co-evolutionary one in that both depend on the complementarity between signallers and receivers. But they are clearly different problems: the problem of co-evolution has to do with the phylogenetic question of Tinbergen, while the adaptive problem has to do with the functional question. Furthermore, the phylogenetic problem seems to be more general: while communication poses the adaptive problem only in those communication systems in which the interests of hearers and speakers conflict or at least diverge (Searcy & Nowicki, 2005), the problem of initiating co-evolution between signallers and receivers is always present.

The rest of the paper is structured as follows. The next two subsections are devoted to briefly review previous related work in traditional evolutionary biology (1.2) and in the adaptive behavior/artificial life fields (1.3). In section 2 we describe our model. In section 3 we show the surprising results of our simulations and propose a possible explanation for them, which represents an interesting solution to the phylogenetic problem. In section 4 we present a detailed analysis of a single evolutionary run which makes it possible to fully understand the complex evolutionary dynamics demonstrated by our simple model. This analysis confirms our hypothesis on the phylogenesis of communication, which we further test, in section 5, by comparing the standard, base-line, simulation with two control simulations. Having reached a full understanding of the forces that are at play in our model, in section 6 we discuss the relevance of our work with respect to both the adaptive behavior community and to the general (biological) theory on the evolution of communication. Finally, section 7 concludes the paper with a summary and some general comments.

## 1.2    The Biological View: Receiver Biases in the Emergence of Communication

The recent biological literature on communication has focused mostly on manipulative communication, due to the controversial but deeply influential work of Dawkins & Krebs (1978). [1] These authors gave a definition of communication which *identified* it with manipulative behavior: according to Dawkins and Krebs communication occurs "when an animal, the actor, does something which appears to be the result of selection to influence the sense organs of another animal, the reactor, so that the reactor's behavior changes to the advantage of the actor". This theoretical framework was put forward as a reaction to the general view of communication of early ethologists (e.g. Tinbergen, 1952) according to which the function of communicative signals was to ensure the survival of the group or the species. This species-centered view of natural selection was challenged by the gene-centered view according to which the basic unit of selection must be the smallest unit of reproduction, that is the gene, and any advantage that a trait can produce for the species, the group or even the individual must be understood in terms of the selective advantage of 'egoistic genes' (Dawkins, 1976). Hence, the definition of Dawkins and Krebs: a conception of communication as manipulation was in fact more in line with the new gene-centered view than the old, cooperative conception of early ethologists.

The historical importance of Dawkins and Krebs' framework is probably due to the the fact that it has ever been the only comprehensive theoretical framework on the topic, and one which has produced a number of interesting empirical predictions (Hauser, 1996). Nonetheless, the identification of communication with manipulation seems to be unjustified both theoretically and empirically. The theoretical point is that the adaptive problem posed by communication is symmetrical: pure adaptive reasoning implies that just as the signaller must get an advantage for sending information, the receiver must get an advantage in order to respond to signals. Empirically, just as there are communication systems in which the result of the communicative interaction seems to benefit only (principally) the signaller, such as the aggressive displays produced when competing for some resources or the begging calls produced by newborns for receiving extra feeding from parents, there are also communication systems which seem to provide benefit to both the actors of communication, such as mating calls or signals emitted during cooperative endeavors like group hunting, and communication systems which seem to benefit only (principally) the receiver, such as the food or alarm calls produced by many species. We say "*seem* to benefit" because we want to stress that the functional, adaptive role of what we consider communicative behavior is something that must discover, rather than presuppose. In other words, it is an empirical question who and to what extent is benefited by (what we perceive as) a communicative behavior, not something that can be decided *a priori*, by including it in the definition of communication itself.

While the communication-as-manipulation framework is well-suited for dealing with the first kind of communication systems, it appears less useful when dealing with cooperative communication, and completely misleading when dealing with communication systems which seem to benefit principally the receiver. Indeed, the definition of communication of Dawkins and Krebs was criticized by several authors (i.e. Maynard-Smith & Harper, 1995; VanRhijn & Vodegel, 1980) but the reply of Krebs & Dawkins (1984) was to accept some of the criticisms as just little amendments to the basic framework, and their way of framing the problem of the evolution of communication as the evolution of manipulation has been informing most of both the theoretical and empirical biological literature (again, see Hauser, 1996).

The deep influence of this manipulative view of communication is clear in the standard treatment which biologists have been giving to the phylogenetic problem of communication. In fact, a possible solution to this problem has been suggested in the biological literature through the notion of the exploitation of the receiver (or sensory) bias (Guilford & Dawkins, 1991; Maynard-Smith & Harper, 2003). The basic idea is that in order to manipulate the receiver, the speaker could exploit some behavioral bias of the receiver: if some environmental feature triggers a specific behavior in some organism that has an adaptive value for another one, the latter can produce signals that resemble that environmental feature in order to manipulate the behavior of the former for its own interests. This idea is certainly sound and can

---

[1]The very brief historical sketch of the biological thought regarding the evolution of communication made in the present and next paragraphs follows the discussion of Hauser (1996), especially section 2.2.

probably explain the evolution of several manipulative communication systems. But it seems much less adequate for explaining the emergence of communication systems which seem to advantage the hearer, like alarm and food calls which are known to be present in many species, including monkeys, chickens, squirrels, and dogs (for a recent review, see Searcy & Nowicki, 2005). It seems more reasonable to suppose that this kind of communication systems could emerge from the opposite mechanism, which we can call the *exploitation of a producer bias*. If organisms have some bias to produce behaviors that systematically co-vary with features of the environment which have an adaptive value, those behaviors can be used by other organisms as cues for their own purposes. In this way, the biased behaviors become signals, the organisms that exploit them become hearers and the individuals that produce them become speakers. But is the idea of a producer bias reasonable? Where could a producer bias come from? Some recent modeling work (reviewed below) has provided a suggestion: producers might be biased towards good signaling by the need for organisms to internally categorize experiences in adaptive ways. With the work presented in this paper we show that this idea of a producer bias is in fact sound, and might constitute a fundamental way to solve the phylogenetic problem of communication, in both artificial and natural organisms.

## 1.3    Related Modeling Work

The study of the evolution of communication and language through computational models is a very lively field, but most of the research done in this area is specifically devoted to language (i.e. Cangelosi, 2001; Hurford, 1989; Kirby, 2002; Parisi, 1997; Steels, 2006; for a general overview, see Cangelosi & Parisi, 2002), which is not the focus of the present paper. There are also a number of interesting models which explicitly address the evolution of (animal) communication (for a recent review, see Wagner, Reggia, Uriagereka & Wilkinson, 2003). However, most of these models take a functional perspective rather than a phylogenetic one (according to Tinbergen's classification). In other words, most of the adaptive behavior research on the evolution of communication has been focusing on which are the conditions which can favor the emergence of a communication system, and, in particular, on the adaptive problem described above (i.e. Ackley & Littman, 1994; Cangelosi & Parisi, 1998; Mirolli & Parisi, 2005a; Noble, 1999; Oliphant, 1996). This is odd, since adaptive behavior/artificial life simulations could give a fundamental contribution especially to the study of phylogenetic problems. In fact, empirical data on the evolutionary history of species are rather scanty, and necessarily indirect. You simply cannot reload the tape of evolution and directly observe what happened, or what would have happened in different scenarios. But you can run artificial evolutionary experiments as many times as you want, and by comparing the results of simulations with different set-ups you can test the influence of any kind of factor on the (artificial) evolutionary process. In this way, you can use your virtual experiments to guide speculations, the gathering of empirical data, and the interpretation of the available empirical evidence (Bedau, 1998; Dennett, 1994; Di Paolo, Noble & Bullock, 2000).

This notwithstanding, there are only very few models in the adaptive behavior literature which directly deal with the phylogenesis of communication. The first one is the seminal work of Werner & Dyer (1992). In their work, Werner and Dyer evolved communication in a population of artificial organisms which were divided in females and males. Females could see approaching males and emit signals but could not move. On the other hand, males could move and hear the signals emitted by females but could not send signals and could not see females. Hence, females had to emit signals in order to guide males to reach for them so that they could mate. The population evolved under what Werner and Dyer called a 'XGA', that is, a genetic algorithm in which fitness is not explicitly calculated: rather, organisms reproduce as soon as a male and a female meet. Both recurrent neural networks and look-up tables were used as control mechanisms for artificial organisms, and in both cases the population succeeded in evolving a communication system. Form the point of view of the present paper, the interest of this work lies in the fact that the authors try to describe the evolutionary history through which communication emerges. In summary, they divide this history in 5 stages, which can be summarized as follows: (1) at the beginning, both males and females behave randomly; (2) males who do not move become extinct; (3) males who only go straight (ignoring signals) take over the population; (4) females evolve signals that

can help males to find them, and, consequently, males evolve the ability to appropriately respond to these signals; (5) finally, there is a further improvement of the communication system which help optimize mate-finding. Although Werner and Dyer's paper is quite interesting in their trying to give an account of communication phylogeny, their description is somewhat vague: for example, regarding the transition between stage (3) and (4), that is, the crucial transition between non-communicating and communicating organisms, they just say: "Males appeared that turn when in the same row or column as a female. At this point, females had evolved that produce a signal telling the males how to find them" (p. 667). From this passage it appears as if the ability to emit useful signals evolves first, and the ability to respond appropriately to those signals comes later. This would be in line with our hypothesis of communication emerging through a producer bias, rather than a receiver one, as proposed by biologists. But Werner and Dyer are not very clear on this point, and they do not report a detailed analysis of whether and why this is the case. Hence, from Werner and Dyer's work it is not possible to extract general principles about the phylogenetic problem which we are dealing with here.

Another very interesting study of the phylogeny of communication is that by Quinn (2001), which explicitly investigates the problem of the co-evolution of speakers and hearers. Using a genetic algorithm, Quinn evolved a team of two simulated robots which had to move in the environment while remaining close to one another. Robots had only proximity sensors and so had to negotiate how to solve the task by some form of (communicative) interaction, and could not rely on dedicated communication channels. Evolved robots solve their task in a way that can be summarized as follows: (1) both agents rotate anti-clockwise; (2) the agent (agent A) which firstly faces the other one (agent B) moves towards it; (3) agent A stays close to agent B by slightly moving back and forth; (4) as agent B faces agent A it starts moving backward and is followed by agent A. The back and forth movements of agent A seem to function as signals for agent B meaning something like "I am here, waiting for you to start moving." In fact, if an agent perceives the signal while still rotating, it will start moving backward, taking the leading role. On the contrary, if an agent perceives the other agent not moving back and forth (that is, not signaling), it will start signaling and subsequently it adopts the follower role. Quinn presents an analysis of the evolutionary emergence of such a simple form of communicative interaction. This analysis demonstrates that the back and forth behavior which functions as a signal evolved by adapting previously existing behaviors which had other functions to serve a communicative function. In particular, before assuming its communicative function the back and forth movement served as an obstacle avoidance behavior which evolved in order to avoid dangerous collisions between two robots moving one towards the other. This work is particularly important for our purposes in that it demonstrates one way to solve the phylogenetic problem of emitter/receiver co-evolution: communication might evolve through re-adaptation of non-communicative behaviors to serve a new, communicative function.

Another study which explicitly addresses the evolutionary origins of communication in artificial agents is that of (Marocco & Nolfi, 2006). In their work Marocco and Nolfi evolved a group of four simulated robots for the ability to solve a collective navigation problem: the four robots had to equally distribute themselves between two target areas, two robots in each area. Robots had proximity (infrared) sensors, a ground sensor for detecting target areas, a sensor detecting their own produced signals, and four communication sensors, detecting signals emitted by other robots from four orthogonal directions (front, rear, left, right). Signals were emitted through a dedicated communication output, with continuous activation, and could be perceived by other robots only from a small distance. Robots evolved using a standard genetic algorithm. The best simulation run produced groups of robots which were able to efficiently solve their task by using 5 different signals: signal A is emitted by a robot that is looking for a target area and does not detect other signals; signal B is produced by a robot looking for a target area and detecting another robot sending either signal A or B; signal C is produced by a robot located in a target area and not perceiving any other signal; signal D is a 'null' signal, that is, a zero intensity signal which is produced by a robot located outside a target area which detects a signal C; signal E is produced by two robots which are together inside the same target area. Marocco and Nolfi analyze the evolutionary emergence of the evolved signals, stressing the fact that communication abilities co-evolve with individual (non-communicative) ones. In particular, they focus on how the meanings of

signals are grounded not only on robots' simple sensory-motor interactions, but rather on already evolved behavioral capacities. For example, the presence of an ability (behavior) to remain in a target area acts as a prerequisite for the emergence of signal C, which communicates the location of the target areas to other robots. Marocco and Nolfi do not discuss explicitly the phylogenetic problem we are dealing with here, but from their analyzes it seems that signaling abilities tend to systematically precede responding abilities. If true, this is particularly interesting here, because it is strictly related to the hypothesis we are advancing in the present paper, namely, that communication can evolve due to the exploitation, by receivers, of already existing producer biases towards good signaling.

Finally, three previous studies which focused on the adaptive problem of communication (the problem of altruism) are relevant for the present work, in that they represent its immediate inspiration. In a model very similar to the one presented here, Cangelosi & Parisi (1998) were able to evolve a communication system which benefits only hearers in a population of ecological neural networks. Discussing their surprising results, Cangelosi and Parisi suggested that even though in their model there was no advantage for speakers to produce useful signals, communication could evolve thanks to the link between communicative behavior and other 'cognitive' abilities. The argument runs as follows. In order to behave adaptively, organisms need to internally categorize experiences in ways that are relevant for their survival. As a consequence, if communicative behavior depends on the internal representation of experience, which is true in Cangelosi and Parisi's simulations and is assumed to be true also in real organisms, then produced signals will tend to be correlated to adaptively useful categories. As a result, the evolution of good signallers might be explained as a by-product of their independently evolving ability to categorize the environment.

The same idea has been re-stated by Marocco, Cangelosi & Nolfi (2003), who tried to evolve a communication system in a population of neural networks controlling a robot arm whose task was to categorize objects by physically interacting with them. But in that work, notwithstanding the supposed 'cognitive' pressure toward good signal production, the evolution of such a communication system proved to be very difficult. In fact, a good communication system evolved only in 7 out of 10 replications of the simulation, even if (a) the networks were cognitively pre-adapted to solve the task before the introduction of communication, and (b) the speaker was always the parent of the hearer, so that there was a strong selective pressure for the evolution of communication due to kin-selection.

In a recent paper (Mirolli & Parisi, 2005a), which is the immediate precursor of the present one, we have shown that a communication system which provides benefits only for hearers is not evolutionary stable, in a model very similar to the one used in Cangelosi & Parisi (1998), unless other factors prevent altruistic signallers to be exploited by cheaters. Though we believe that Cangelosi and Parisi's idea of a cognitive pressure on the production of good signals is sound, our previous results, together with those of Marocco and co-workers, cast doubt on it. Indeed, it seems that the supposed cognitive pressure on the production of good signals cannot be sufficient to explain the evolutionary emergence of an altruistic communication system, for at least two reasons: first, the cognitive pressure can produce an *individual* tendency to produce good signals, but cannot explain why signals are *shared* in the population; second, the *emergence* of a good communication system does not explain its *evolutionary stability*, that is why the communicating population is not invaded by mutant individuals that cheat conspecifics by producing misleading signals. The simulations and analyzes presented in this paper are intended to make clarity on this topic. In particular, we will present detailed analyzes of the complex evolutionary dynamics exhibited by our simulations which will serve to answer the following questions: Does the need for categorizing experience in adaptive ways really constitute a drive towards good individual signalling behavior? What else is needed for good communication to emerge at the level of the population? And can these factors, by themselves, render the communication system evolutionary stable? To relate these questions to the general problems posed by communication which we discussed above: Are producer biases towards good signaling possible solutions to the phylogenetic problem of communication? Can they also represent solution to the adaptive problem posed by altruistic communication systems?
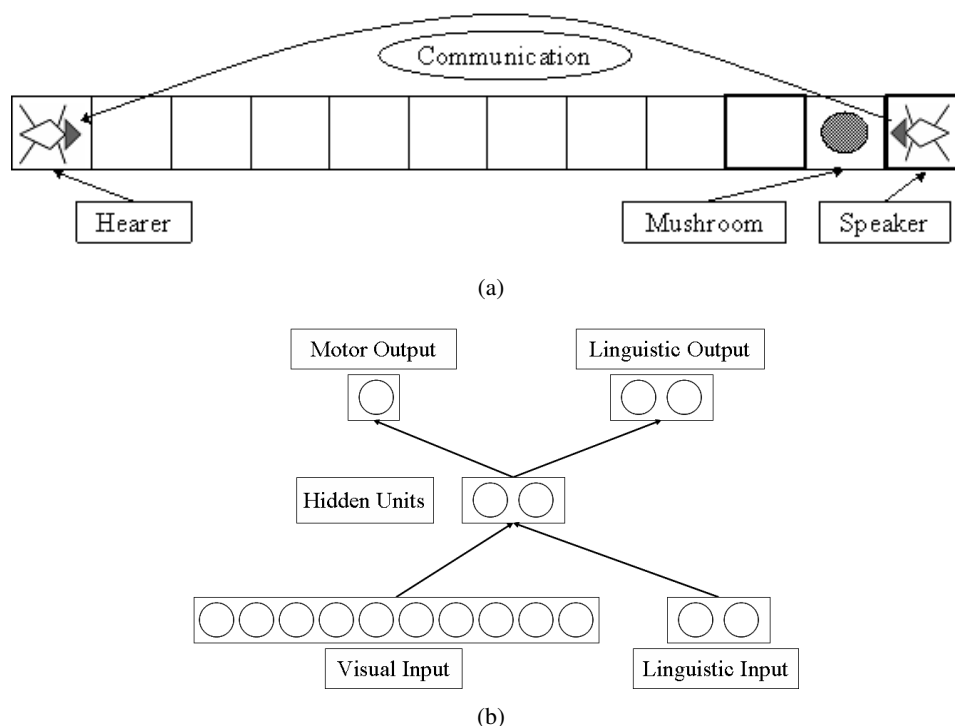
(a)



(b)

Figure 1: (a) Schematic representation of the environment. (b) The neural network (arrows represent connections from all the neurons of the sending group to all the neurons of the receiving group).

## 2 The Model

Our model consists of a population of 100 organisms living in a one-dimensional environment (a corridor, shown in figure 1a). An individual's behavior is controlled by a neural network and the individual's fitness is a function of the number of movements the individual makes and the number and type of mushrooms it eats during its life. There are 420 possible mushrooms, each different from all the others in its perceivable properties, encoded as strings of 10 bits in the visual input units of an organism's neural network. Half of the mushrooms are edible, in that they increase the fitness of the individual that eats them, while the other half are poisonous, in that they decrease the individual's fitness. The architecture of the neural network is shown in figure 1b: it includes 10 visual input units, 2 communicative input units, 2 hidden units, 1 motor output unit and 2 communicative output units. The network's connection weights are encoded as real values in the range [-4, 4] in the genome of the organisms. The visual range of our organisms is limited in that the visual input units encode a mushroom's properties only if the organism is in the cell immediately preceding the mushroom cell, otherwise the organism sees nothing, i.e., the activation of all visual input units is set to 0. The motor output unit of the network has a step activation function which determines whether the organism moves one step forward (if the activation is 1) or stays still (if the activation is 0). In our model signals consist in the two-dimensional vectors of the activations (in [-1, 1]) of the communicative output units of an organism (the speaker), which are copied in the communicative input units of another organism (the hearer).

Each individual of each generation lives for 420 trials, one for each mushroom. In each trial the individual is put in the start cell of the corridor, one of the mushrooms is put at the end of the corridor, and another individual is chosen randomly from the population to act as speaker. The speaker is placed near the mushroom and emits a signal through its communicative output units which is received by the tested organism (the hearer) through its communicative input units. The trial ends either if the hearer reaches the mushroom and eats it or after 11 input-output cycles. An individual's fitness increases with the number of edible mushrooms eaten and decreases with the number of poisonous mushrooms eaten and the number of movements made. More precisely, fitness is calculated according to the following

formula:

$$f(x) = \frac{an_e - bn_p - cn_m}{norm} \qquad (1)$$

where $n_e$ and $n_p$ are, respectively, the number of edible and poisonous mushrooms eaten by x, $n_m$ is the number of movements made by x, $a$, $b$ and $c$ are three constants (set to 30, 5, and 1, respectively), and $norm$ is the normalization factor (maximum possible fitness) which equals 4200.

After all the individuals of one generation have lived their lives, the next generation is created by selecting individuals with a probability proportional to their fitness (using the roulette wheel method with windowing), [2] making them reproduce sexually (with double-point crossover), and mutating the genes (connection weights) of the offspring with a probability of 0,1%.

Simulations run for 2000 generations, during which we calculate various statistics, the most important ones being average fitness and what we call 'communication system quality'. For a communication system to be good, three conditions must apply: (a) objects belonging to the same category must be signalled in similar ways, (b) objects belonging to different categories must be signalled in different ways, and (c) all individuals must signal in the same way, which implies that the communication system is shared in the population. In our model, signals are two-dimensional vectors, which can be represented as points in a two-dimensional space whose coordinates are the activations of the two communicative output units of speakers. Hence, the communication system of a population consists of two clouds of points: the points that represent the signals emitted by all the organisms in presence of all edible mushrooms and the points corresponding to the signals emitted in presence of all poisonous mushrooms. Condition (a) above is satisfied if the two clouds are small (meaning that mushrooms of the same category are signaled in similar ways). Condition (b) is satisfied if the two clouds are distant from one another (meaning that mushrooms belonging to different categories are signaled in different ways). Condition (c), which regards the sharedness of the communication system, is implicitly satisfied by the other two conditions, since the clouds represent the communication system of the entire population. Our communication system quality is calculated as the average between the two measures that make a communication system good: (1) the mean distance of the points of each cloud from the cloud's geometrical center, and (2) the distance between the geometrical centres of the two clouds.

## 3 Results, with a Possible Explanation

In the scenario of our simulations the emergence of a good communication system would provide an advantage only to hearers. A hearer who receives good signals could in fact use the received signals for deciding to approach mushrooms signaled as 'edible' or to stay still in the case the received signal stands for 'poisonous'. This would allow the hearer to avoid wasting its energy by approaching poisonous mushrooms and thus would increase its fitness. For speakers, however, there is no benefit whatsoever in producing good signals. On the contrary, since individuals compete for reproduction, producing good signals is a purely altruistic trait in that, by giving an advantage to competitors, it indirectly decreases the reproductive chances of the good signaler. In other words, our simulative scenario poses both the phylogenetic problem and the adaptive problem due to the altruistic character of this kind of food calls.

The results of 10 out of 50 replications of the simulation, which are representative of all the others, are shown in figure 2.

The results are quite surprising: average fitness and communication system quality fluctuate cyclically between very low and very high values. This means that communication continually emerges but it is also continually disrupted. Since this very unstable pattern is present in all the replications of the simulation, the mechanisms that generate it must be very strong and reliable. What might these mechanisms be?

---

[2] 'Windowing' consists in reducing the fitness of each subject by the fitness of the worst subject, so that the fitness of all individuals are >= 0. For an explanation of the details of the genetic algorithm, see Mitchell, 1996 and Mirolli & Parisi, 2005a. The results reported in the next section seem quite robust with respect to the details of the genetic algorithm: for example, runs with slightly different mutation rates or with asexual reproduction (i.e. without cross-over) provided qualitatively similar results.
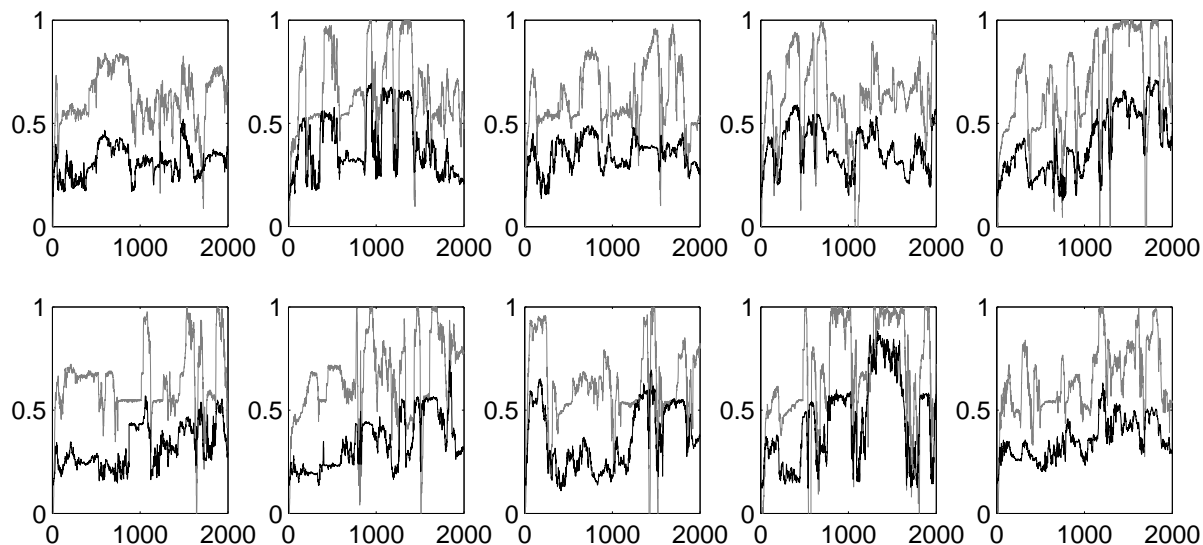
Figure 2: Average fitness (gray line) and communication system quality (black line) of 10 replications of the simulation.

The sudden drops in the communication system quality can be explained by referring to the above mentioned adaptive problem: that is, as a consequence of the altruistic character of communication. Once a good communication system has emerged in the population bad speakers are selected against good speakers because they act as cheaters: they can take advantage of the good signals they receive while misleading other organisms by producing bad signals. The result is that bad signalers (cheaters) rapidly invade the population with the consequence of a sudden decrease in the quality of the communication system. But why a good communication system should emerge in the first place or re-emerge after being disrupted? Clearly, in our simulation something quite powerful must be at work which is able to solve the phylogenetic problem, pushing the communication system towards good quality. We hypothesize that the tendency of communication to emerge is due to two factors.

The first factor is the 'cognitive' pressure suggested by Cangelosi & Parisi (1998). In order to survive and reproduce organisms must categorize perceived mushrooms according to their quality, moving forward in the corridor when they perceive edible mushrooms and staying still when they perceive poisonous mushrooms. For producing this kind of behavior, mushrooms must be appropriately categorized in the neural network's hidden units. In other words, in order for an organism to behave efficiently, its network's connection weights must be such that mushrooms belonging to the same category elicit similar activation patterns in the hidden units (similar internal representations), while mushrooms belonging to different categories elicit different internal representations. But, given the architecture of our organisms' neural networks, the signal emitted by an organism in the presence of a mushroom directly depends on the internal representation elicited in the hidden units by the perception of the mushroom. This means that the adaptive pressure to categorize mushrooms appropriately causes, indirectly, a pressure to produce good signals: mushrooms belonging to the same category elicit similar internal representations, which, in turn, tend to elicit similar signals; vice versa, mushrooms belonging to different categories elicit different internal representations which tend to elicit different signals.

However, this cognitive explanation of the spontaneous emergence of a good communication system cannot be enough. The cognitive pressure can in fact explain only the fact that each individual tends to produce good signals, that is, signals that systematically co-vary with the category of the mushrooms perceived by the individual. It cannot explain why the communication system is shared in the population, that is, why different individuals tend to produce similar signals, a condition which is necessary for good communication and which is considered in our measure of the communication system quality.
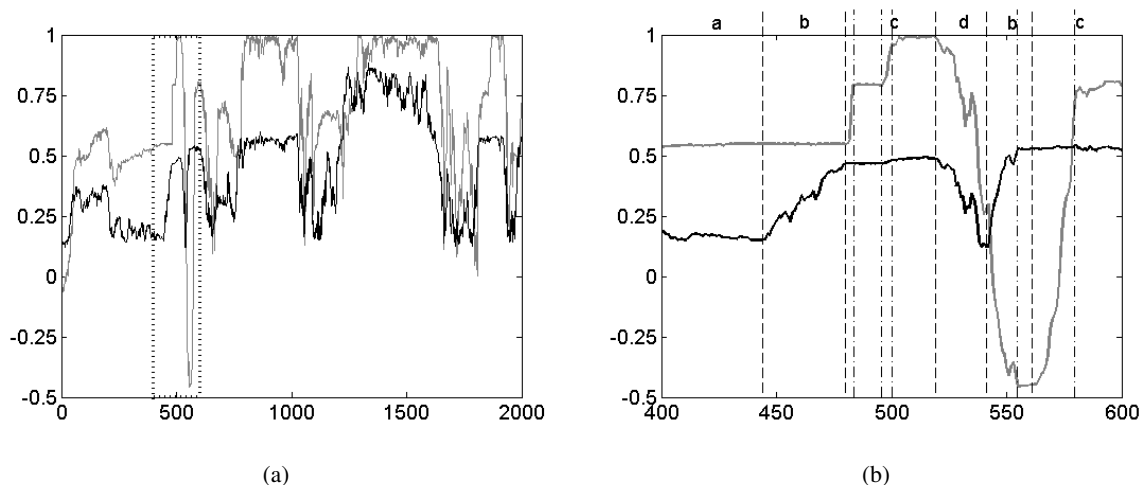
Figure 3: (a) Average fitness (gray line) and communication system quality (black line) of one replication of the simulation. (b) Zooming in on the 200 generations from generation 400 to generation 600 enclosed in the dotted lines of graph (a). The dashed lines of graph (b) separate macro-evolutionary phases, while the dot-dashed lines subdivide this evolutionary period in the 11 sub-phases described in table 1. See text for details

(Remember that the measure is taken at the populational level, considering the signals emitted by all individuals). In order to explain the sharedness of the communication system we need to take into consideration another factor, namely genetic convergence. In fact, if organisms have similar genes (in particular for the weights connecting the visual input units to the hidden units and the hidden units to the communicative output units), they will have similar (categorical) internal representations and will tend to produce similar signals. Genetic similarity can in turn be explained by genetic convergence due to selection: only a few individuals, the fittest, reproduce in each generation and, as a consequence, their genes will tend to spread in the population reducing genetic diversity.

This explanation of our results, if correct, might provide a possible powerful solution to the very general phylogenetic problem of the evolution of communication. But even if the idea of possible cognitive pressure towards good signaling has been proposed by several researchers (Cangelosi & Parisi, 1998; Marocco et al., 2003; Mirolli & Parisi, 2005a), to our knowledge nobody has ever really demonstrated its presence in any model of the evolution of communication, let alone discussed its possible significance for the evolution of real communication systems. The rest of the paper is devoted to do just that: test our explanation of the peculiar results of our simulation (next two sections), and discuss its relevance for biological theory on the evolution of communication.

## 4 Evolutionary Dynamics of Communication

### 4.1 Zooming in on Evolution

In the previous section we have given an explanation of the repeated rise and fall of communication in our model based on four factors: (1) the cognitive pressure toward spontaneous individual good signalling, (2) the genetic pressure toward sharedness of the communication system due to the convergence of the population, (3) the direct benefits brought by responding appropriately to informative and shared signals, and (4) the pressure against good (shared) signalling produced by the altruistic character of communication. In order to test this explanation of the complex evolutionary dynamics demonstrated by our simulations, we take a single run of the simulation, select a portion of it (from generation 400 to generation 600), and try to analyze this portion in detail. Figure 3 shows the process of zooming in on the 200 generations we selected.

| phase | macro-phase | generations | description |
|---|---|---|---|
| 1 | a | 400-444 | communication is bad and ignored |
| 2 | b | 445-480 | communication improves gradually |
| 3 | c | 481-485 | fitness improves rapidly |
| 4 | c | 486-495 | communication is good and partially exploited |
| 5 | c | 496-500 | fitness improves rapidly |
| 6 | c | 501-519 | communication is fully exploited |
| 7 | d | 520-541 | communication system quality and fitness decrease |
| 8 | b1 | 541-555 | fitness decreases but communication improves |
| 9 | b1 | 556-561 | communication is good but not understood |
| 10 | c1 | 562-580 | fitness improves rapidly |
| 11 | c1 | 581-600 | communication is good and partially exploited |

Table 1: Evolutionary phases of the analyzed simulation.

We choose this particular portion of this run for analysis because, looking at the dynamics of average fitness and communication system quality, it seems to clearly exemplify the typical cycle of rise and fall of communication which repeats itself continually in all the runs. As can be seen from the picture, we have divided the 200 generations into 11 sub-parts by identifying transitions consisting in significant changes in average fitness and/or communication system quality. These are summarized in table 1.

For the sake of clarity, we have grouped some of these sub-parts together by identifying 4 macro-evolutionary phases which can be described as follows:

**Phase** (a) (generations 400-444): Communication system quality is low and communicative input is ignored, resulting in an average fitness of about 0.55, which is the maximum value that can be reached without the aid of communication;

**Phase** (b) (generations 445-480): Communication system quality improves because of the cognitive and genetic pressures towards good communication;

**Phase** (c) (generations 481-519): Good signalling starts to be exploited and fitness increases until it is near to the maximum value;

**Phase** (d) (generations 520-541): Bad speakers are selected against good ones because they cheat others and invade the population: the result is a sudden decrease in communication system quality and, consequently, fitness.

Once the communication system quality has reached its minimum it starts to grow again (phase (b1): generations 542-561) until it reaches a very high value, communication re-starts to be exploited, and fitness increases again (phase (c1): generations 562-600).

## 4.2 Analyzes

In order to understand the evolutionary dynamics of these 200 generations of our simulation we have collected a number of statistics and plotted them together in figure 4. An evolving population can be considered as a cloud of points in a multi-dimensional genotype space. In our case, each dimension corresponds to one of the 32 connection weights of an individual's neural network (encoded as real values in the individual's genome). The points correspond to the individuals of the population and each point's coordinates are the values of the individual's genes (connection weights). In order to understand how the population moves in its genetic space we consider the centroid of the population, that is, the geometrical center of the cloud of points representing the individuals of each generation. In dynamical systems terms (Beer, 2000), the centroid is a dynamical system moving in discrete time (time-steps correspond to generations), in a 32-dimensional continuous space (with each dimension bounded in [-4;4]), with a very complex evolution operator (which is implicitly defined by the rules which govern

individual life and the selection, reproduction, and mutation operators) and highly stochastic: there is stochasticity in the individuals' interactions and in all the genetic algorithm's operators.

Since the genetic convergence of the population is necessary, according to our hypothesis, for explaining the sharedness of the spontaneously emerging communication system, an interesting statistic to take is the genetic variance of our population. But since our explanation postulates different genetic pressures on different sets of genes, we wanted to determine whether those different genetic pressures would result in different genetic variances. Consequently, we divided our system (the genome of the evolving population) into three sub-systems corresponding to the three distinct functional roles that different sets of genes play with respect to communication: the non-communicative genes, the communicative-input genes, and the communicative-output genes (see figure 5a). Genetic variance for each gene (connection weight) $x$ ($V_x$) is calculated, standardly, according to the following formula:

$$V_x = \frac{\sum_{i=1}^{N}(\bar{w}_x - w_x^i)^2}{N} \tag{2}$$

where $N$ is the number of individuals in the population (i.e. 100), $\bar{w}_x$ is the mean value for weight $x$, and $w_x^i$ is the weight $x$ of the $i_{th}$ individual. Consequently, the variances ($V^k$) for the three categories ($k \in K$) of weights (non-communicative, communicative-input, and communicative-output) are calculated as follows:

$$V^k = \frac{\sum_{x=1}^{N^k} V_x}{N^k norm} \tag{3}$$

where $N^k$ is the number of weights of the $k_{th}$ category (i.e. 28, 4 and 4 for non-communicative, communicative-input, and communicative-output weights, respectively), $V_x$ is the variance on weight $x \in k$ and $norm$ is the normalization factor, so that $V^k$ is in [0;1] (since weights are bound in [-4, 4] the theoretical maximum variance for each gene is 16, obtained when half of the population has a value of -4 and the other half a value of 4). Figure 4a shows these variances as they change during the 200 generations under analysis.

Figure 4b shows the displacement of the centroids of the three functional sub-sets of genes (non-communicative, communicative-input and communicative-output) between generation x and generation x-1. Formally, we plotted the following function:

$$y(k) = dist(C_x^k, C_{x-1}^k) \tag{4}$$

where $k$ is the plotted category (non-communicative, communicative-input and communicative-output), $C_i^k$ is the centroid of category $k$ at generation $i$, and $dist(a, b)$ is the euclidean distance between points $a$ and $b$.

Figure 4c represents the evolution of the populational communication system itself. As discussed above, as signals are represented by points in the two-dimensional space of the communicative-output units of organisms, the communication system can be represented as the two populational centroids of the points that represents signals emitted in response to edible mushrooms and signals emitted in response to poisonous mushrooms, respectively. In figure 4c we have plotted the two coordinates of the two centroids of the signals used for edible and poisonous mushrooms.

Figure 4d shows the comparison between four values: the quality of the populational communication system, the average quality of the individuals' communication systems, the quality of the populational representation system, and the average quality of the individuals' representation systems. As discussed above, the communication system quality that we have discussed so far is calculated on the centroids of signals emitted by all the individuals of one generation for all the edible and poisonous mushrooms. But we can also calculate the quality of the individuals' communication systems: for each individual we calculate its communication system quality by considering the centroids of the signals emitted by that particular individual. By averaging those quantities over all individuals in a population we can look at the quality of individual communication systems regardless of the sharedness of communication in the population. Just as the signal emitted by an individual for a given mushroom is the two-dimensional point whose coordinates are the activations of the two communicative output units of that individual for

that mushroom, so the internal representation of that individual for that mushroom is the point whose coordinates are the activations of the individual's hidden units when it sees the mushroom. As a consequence, we can calculate both populational and average individual representation quality in the same way as we calculate communication system quality, but considering the points represented by the individuals' internal representations of mushrooms in their hidden units instead of those represented by the signals emitted through the communicative output units.

The last two images of figure 4 show the evolution of the genes of the best individual of each generation (figure 4e) and of the populational centroid (figure 4f). The two images are bitmaps, where each point represents in a gray-scale the value of one of the 32 genes of a given generation (with white corresponding to very low, negative, values and black to very high, positive, values). As usual, genes are grouped according to their functional role (non-communicative, communicative-input, and communicative-output).

## 4.3 Explaining Evolutionary Dynamics

In this section, we will try to understand the mechanisms that underlie the evolutionary dynamics of our simulation with the help of the various statistics plotted in figure 4. For reasons of space, we will not describe in detail all that happens during the 11 evolutionary phases that are summarized in table 1, even though this is something that can easily be done. Rather, we will focus just on the macro-phases described above, with a particular emphasis on those which involve important changes in the communication system quality of the population, that is, phases (b), (d), and (b1) (see figure 3b). In particular, we will try to answer the following three questions: (1) How and why does communication improve during phase (b)? How and why is communication disrupted during phase (d)? How and why does communication improve again during phase (b1)?

**Phase (a):** Even though phase (a) is an evolutionary stable phase, with respect to both average fitness and communication system quality, we can see that during this phase the populational centroid moves slightly in the non-communicative genetic sub-space, while it does not move significantly in the other two sub-spaces (figure 4b). This continuous displacement of the non-communicative centroid is due to the competition between a few genotypes which have different non-communicative genes but are phenotypically quite similar in that they provide comparable fitness. This is shown by the fact that during this period some of the non-communicative genes of the best individuals change continually (specifically genes number 2, 14, 17, 19, 20, 21 and 22, see figure 4e), resulting in corresponding changes in the populational centroid (figure 4f). The simultaneous presence of competing genotypes during this phase is demonstrated also by the existence of some genetic variance in the non-communicative genes (figure 4a). [3]

**Phase (b):** Phase (b) is characterized by the slow increase in communication system quality at the level of the population. How and why does this happen? During this phase, the activity in the non-communicative genetic space due to the competition between genotypes with equivalent fitness continues, as demonstrated by the graphs of figure 4a, b, e and f. However, during this competition the best individual of the population starts to have a different communicative-output gene (the first one, see figure 4e). As a consequence, the communicative-output populational centroid starts to slowly displace itself (figure 4b) until the population has converged on the new value of the mutated gene (figure 4f). During this phase signals are ignored because during phase (a) they used to be uninformative (indeed, as shown

---

[3]The reason of the relatively high average genotypic variance in the communicative-output genes during the entire period we are discussing lies in the simultaneous presence of two opposite alleles in the last gene. In fact, high and low values for that gene alternate in the best individual (figure 4e), while the populational centroid's gene remains completely stable during all the 200 generations (figure 4f). Evidently, two alleles of the last gene co-exist in the population throughout this period, due to the fact that during this period this gene has no effect on fitness nor on communication system quality. This co-existence is demonstrated by the fact that the populational centroid value for that gene is neither of the two values present in the best individual, but rather a mean between the two (figure 4f). This means that about half of the population possess one of the two alleles while the other half possesses the other.
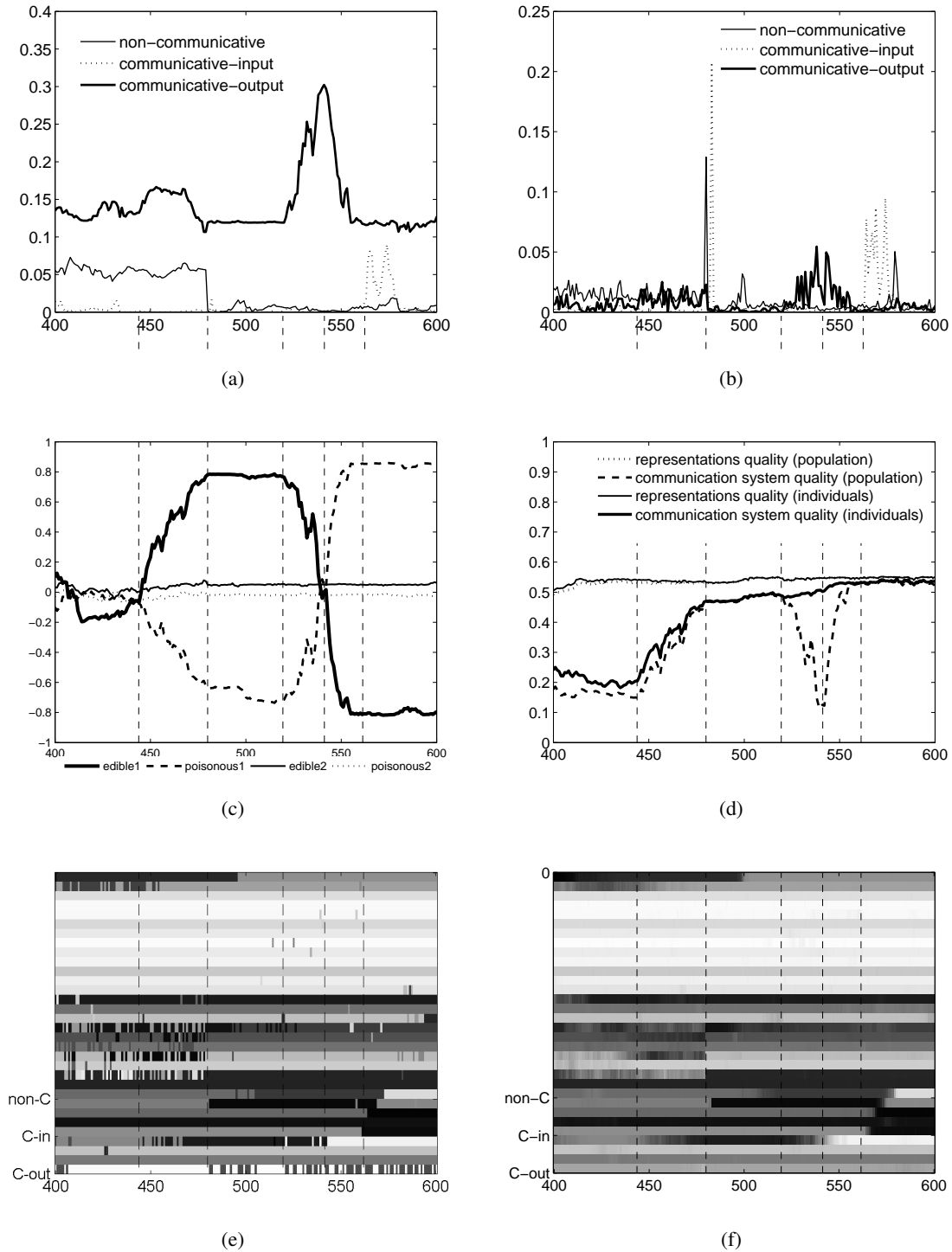
Figure 4: Analyzes of evolutionary dynamics during 200 generations (from generation 400 to generation 600) of one replication of the simulation. (a) Average variances of the non-communicative, communicative-input, and communicative output genes. (b) Displacement of the non-communicative, communicative-input, and communicative-output populational centroids. (c) Coordinates of the two centroids of the signals used for edible and poisonous mushrooms. (d) Quality of the communication and representation systems of the population and average quality of individuals' systems. (e-f) Gray-scale bitmap of the genes of (e) the best individual of each generation and (f) of the populational centroid; genes are grouped according to their functional role: non-communicative (non-C), communicative-input (C-in), and communicative-output (C-out). The lines on or under the graphs divide the 200 generations under analysis into six evolutionary macro-phases. See text (section 4.2) for details.

in figure 4c, during phase (a) there was almost no signalling at all, in that the centroids of the signals for both edible and poisonous mushrooms have a value of about 0). As a result, there cannot be any selective pressure whatsoever regarding communicative-output genes. And in fact, during this phase average fitness does not change at all (figure 3). Hence, the new communicative-output gene spreads in the population just for genetic drift. But how can simple genetic drift in a single gene produce such an increase in the quality of the communication system of the population? As we can see in figure 4d, notwithstanding the slight variability in the non-communicative genes, the population has already converged at the level of internal representations: the populational representation quality is in fact very good, almost as good as average individual categorization quality (figure 4d). It is just this convergence on categorized and shared representations of mushrooms that makes it possible for the spreading in the population of the new communicative-output gene to bring about a simultaneous increase in both individual and populational communication system quality. In fact, the new communicative-output gene just makes manifest in the communicative-output units (in particular in the first unit, see figure 4c) the categorization of mushrooms already present in the hidden units.

**Phase (c):**   During phase (c) the good signals produced spontaneously by speakers get exploited by receivers thanks to the appearance of favorable mutations in both the non-communicative and the communicative-input genes. Thanks to the fact that they permit the exploitation of good signals, these mutations generate two sudden increases in average fitness. The first fitness increase happens at the beginning of phase (c), and is due to two sudden jumps of the population in the non-communicative and communicative-input genetic sub-spaces. These jumps are evident in the sudden stabilization of both the best and the centroid non-communicative genes (which produces a sudden decrease in non-communicative genes variance (figure 4a), and in the sudden change in the first linguistic-input gene (figures 4e and f). Those sudden changes are also evident in the high peaks of the non-communicative and communicative-input centroids displacement (figure 4b). The second increase in average fitness is due to a further displacement of the non-communicative genes centroid (figure 4b). In turn, this depends on a favorable mutation in the first gene of the best individual, which rapidly spreads to the whole population (figures 4e and f).

**Phase (d):**   At the end of phase (c) all organisms both produce similar, useful signals, and correctly interpret the signals they receive, resulting in very good average fitness (figure 3b). How and why is communication disrupted during phase (d)? This is clearly due to the invasion of the population by cheaters. At the very beginning of phase (d) a new mutation occurs in the first communicative-output genes which makes the mutant individual produce signals which are different from those of all the other organisms in the population. As is clear from figure 4e, the best individual happens to be one with a new allele for that gene (figure 4e). This new allele starts to spread in the population, even though slowly, because individuals that possess this allele cheat the others. When they act as speakers, they produce the signal that is interpreted as meaning 'edible' when there is a poisonous mushroom and the signal interpreted as meaning 'poisonous' when there is an edible mushroom. As a consequence, hearers are mislead and obtain lower fitness. The fact that the new allele makes individuals invert their signals, and does not simply make individuals produce signals that do not distinguish between edible and poisonous mushroom is demonstrated by the fact that while during phase (d) populational communication system quality decreases, average individuals communication system quality slightly increases (figure 4d), meaning that organisms with the new alleles in the first communicative-output gene produce, individually, even better signals than those with the old alleles, but the new signals have opposite meaning with respect to old ones. When the new allele has spread in half of the population, the communication system quality of the population has reached its minimum, because half of the organisms signal 'edible' with a high value in the first communicative-output unit and 'poisonous' with a low value in the same unit, as in the preceding generations, while the other half, with the new allele, do the opposite, signalling 'edible' with a low value in the first communicative-output unit and 'poisonous' with a low value. This explains why at generation 540 the mean value of the first communicative-output unit is 0 for signals emitted for both edible and poisonous mushrooms (figure 4c), while individual communication system quality is quite high

(figure 4d). The presence of two opposite alleles in the population of the first communicative-output gene explains also the peak in the genetic variance of the communicative-output genes (figure 4a): now there is high variance not only in the last non-communicative gene but also in the first one, that produces a variance which is about twice as great as that of the preceding generations. Finally, the simultaneous presence of the two alleles is also manifest in the fact that while the first communicative-output gene of the populational centroid (figure 4f) changes between a very low value (dark gray) to the mean value (0, represented by medium gray) to a very high value (bright gray), the best individuals have always either the dark allele or the bright one (figure 4e). This means that the change in the populational centroid is due to a change in the proportion of the two competing alleles in the population.

**Phase (b1):**   How and why does communication quality increases again during phase (b1)? Phase b1 is just a continuation of phase (d). The population keeps on moving in the communicative-output space (figure 4b) towards convergence on the new allele for the first communicative-output gene (figure 4a) because this gene continues to provide a relative selective advantage due to cheating. As a result, populational communication system quality starts to increase again (figure 4d), as the population converge to the new signalling behavior: low values in the first communicative-output unit for 'edible' and high values for 'poisonous' (figure 4c). The reason why average fitness keeps on decreasing during this phase while communication system quality increases (figure 3b) is the same that explains the convergence of the population to the new allele for the first communicative-output gene: even though the new communicative-output gene makes speakers produce good signals, those signals are misinterpreted by hearers, which have communicative-input genes adapted to the previous communication system, which was just opposite to the new one. As a result, average fitness remains at its minimum value for a few generations.

**Phase (c1):**   Finally, at the beginning of phase (c1) a series of mutations occur that allow the population to re-adapt to the new communication system. Chronologically, changes in the fourth, second, and first communicative input genes and in the last non-communicative gene of the best individuals (figure 4e) are rapidly followed by analogous changes in the populational centroid (figure 4f). As a consequence, the population rapidly displaces itself in both the communicative-input and non-communicative subspaces (figure 4b). At generation 580 the population has converged in both the communicative-input and non-communicative genetic spaces (figure 4a), good signals are partially exploited, and nothing relevant happens any more.

In conclusion, we can say that the analysis of the dynamics of the population during the 200 generations we have taken into consideration has clearly confirmed the overall interpretation of the continuous rise and fall in both communication system quality and average fitness that we have provided in section 3. But this is not enough. Even if our analysis has been demonstrated to be correct for the evolutionary period we have studied, we cannot still be sure that the same is true for all the periods of all the replications of the same simulation. In other words, the analysis of a single run is not enough for the extrapolation of general principles. In order to test whether the tendency towards the emergence of communication due to cognitive and genetic pressures, is general rather than relative to a single run of the simulation, we compare the average results of 50 runs of the simulation discussed so far (which from now on we will call the *base-line* simulation) with those of two new control conditions, which we call the *no-cognitive-pressure* and the *no-communication* simulations. This is done in the next section.

# 5   Control Experiments

## 5.1   The No-Communication and No-Cognitive-Pressure Simulations

The no-cognitive-pressure simulation runs exactly as the base-line simulation but organisms have neural networks with a different architecture, shown in figure 5b.
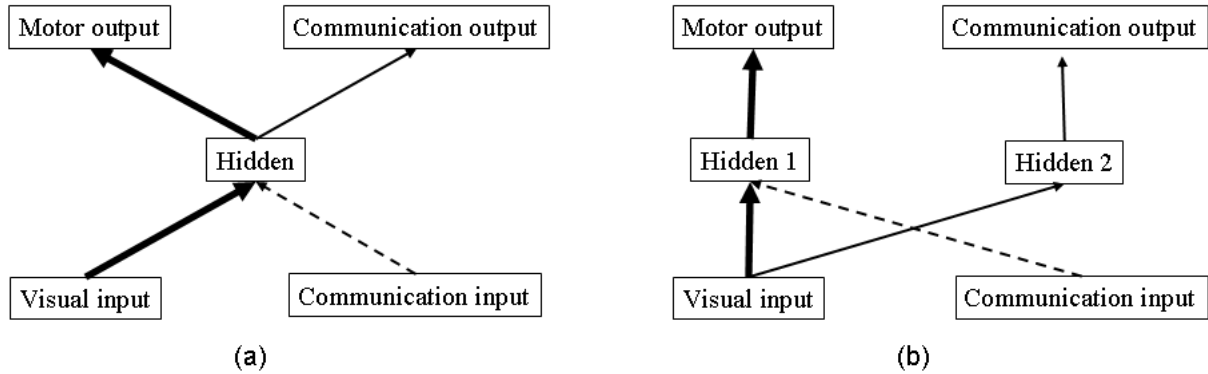
Figure 5: Comparison between the neural network used in the base-line and no-communication simulations (a) with that used in the no-cognitive pressure simulation (b). Connection weights are grouped according to their roles in communication: non-communicative (thick arrows), communicative-output (thin arrows), and communicative-input (broken arrows).

The difference is that in this condition there are two groups of hidden units (composed by 2 neurons each): one receives connections from both the visual and the communicative input units and sends connections to the motor output units, the other receives connections from the visual input units and send connections to the communicative output units. In this way we artificially avoid any impact of categorization on communicative behaviour: since the communicative output does not depend on the same neurons on which the motor output depends, with the new architecture there cannot be any cognitive pressure towards good signal production. [4]

Organisms of the no-communication simulation have the same neural network as those of the base-line simulation (figure 5a) but in this case communication is not allowed. In other words, in the no-communication simulation organisms cannot send signals and, consequently, tested individuals never receive any communicative input. Under such conditions, the only way for an individual to behave correctly is to move towards the mushroom until the mushroom is close enough to be clearly perceived and then to decide whether to eat the mushroom or refrain from doing so. As a consequence, the maximum fitness that can be obtained by the individuals of the no-communication simulation is 0.55. The reason is that these organisms must always pay the cost of moving towards mushrooms: a cost that can be avoided, when the mushroom happens to be poisonous, by organisms that receive reliable signals (in the simulations with communication). Note that the communicative weights (those that link the communicative input and output units with the hidden units) are never used in this simulation, so they are adaptively neutral and subject to genetic drift. Nonetheless, we can still test what signals are produced by the networks located near to a mushroom and, consequently, measure the communication system quality of the no-communication simulation even if signals are never used.

## 5.2 Results

Figure 6 shows the results of the 3 simulations (base-line, no-communication, and no-cognitive-pressure) in terms of communication system quality (a) and average fitness (b).

The results (averaged over 50 replications of each simulation) clearly support our hypothesis regarding the cognitive pressure towards the emergence of a good communication system. The communication system quality in the no-communication simulation is quite high (about 0.5, see figure 6a), being not much lower than the value reached in a 'kin-selection' simulation (about 0.6) which was good enough to allow the population to reach almost maximum fitness (the kin-selection simulation and its results are

---

[4]We also ran a no-cognitive pressure simulation without the second group of hidden units, connecting directly the visual input group to the communication output group. This simulation gave qualitatively identical results as those of the reported simulation.
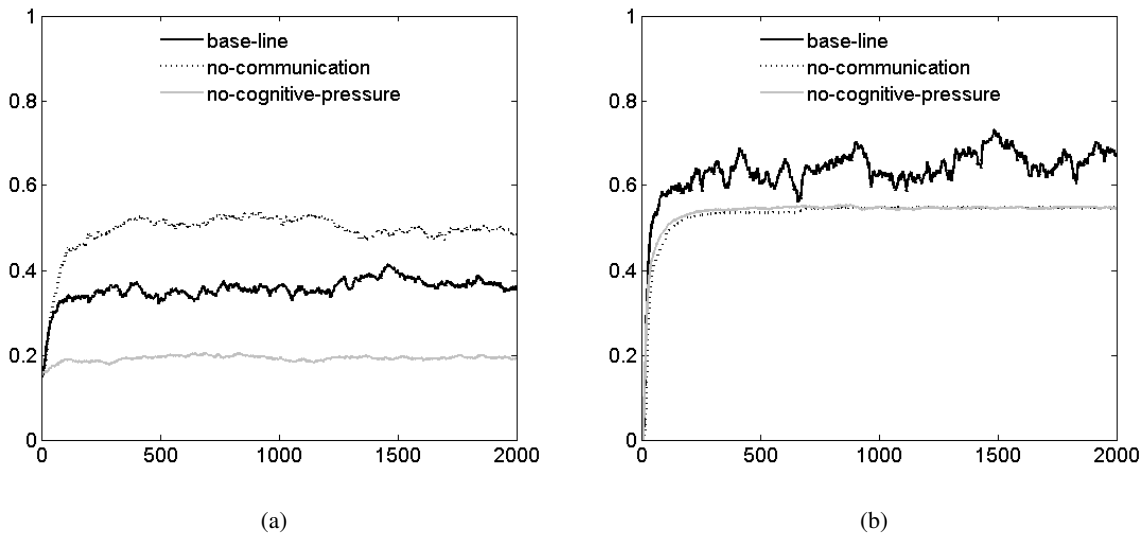
Figure 6: Communication system quality (a) and average fitness (b) of the base-line, no-communication, and no-cognitive-pressure simulations. Average results of 50 replications.

reported in Mirolli & Parisi, 2005b). This means that the need to categorize mushrooms and the genetic convergence of the population are sufficient to produce a good communication system even in populations in which communication is not allowed and the communicative-output weights are subject to genetic drift. Furthermore, the fact that the quality of the communication system in the no-communication condition is considerably higher than in the base-line condition (about 0.35) means that, if signals are used, there is a direct selective pressure *against* good speaking, due to the altruistic character of communication.

The results of the no-cognitive-pressure simulation confirm our analysis. If we prevent any influence of cognition on signal production by manipulating the architecture of the neural network, a good communication system never emerges, resulting in an average communication system quality of about 0.2, which is considerably lower than that of the base-line condition.

The results in terms of average fitness (figure 6b) are quite clear. In both the no-communication and the no-cognitive-pressure simulations fitness reaches the maximum value achievable without communication: in the first case because communication is not allowed, in the second case because signals are useless and cannot be exploited. Since the organisms of the base-line simulation can sometimes take advantage of communication, their average fitness is a little (but not much) higher.

In figure 7a we have plotted the variances of non-communicative and communicative-output connection weights of the base-line simulation (as usual, we show the average results of 50 replications).

There is indeed a considerable genetic convergence since both variances are very low, which explains the sharedness of the spontaneously emerging communication system, but the communicative-output weights have a much higher variance (about 0.1) than the non-communicative weights (about 0.01). The explanation is to be found in the selective pressure against good signal production due to the altruistic character of communication. As we have seen in the previous section, the cheaters which invade the population when the communication system is good are those organisms that, due to mutations, have communicative-output weights that are different from those on which the population has converged and which, consequently, let them produce misleading signals. In other words, the variance of the communicative-output weights is higher than the variance of the non-communicative weights because each time the population has genetically converged and the quality of the communication is high, there is a strong and reliable pressure to possess communicative-output genes that differ from those of the population, in order to cheat competitors.

This analysis is further confirmed by the comparison between the variances on the communicative-

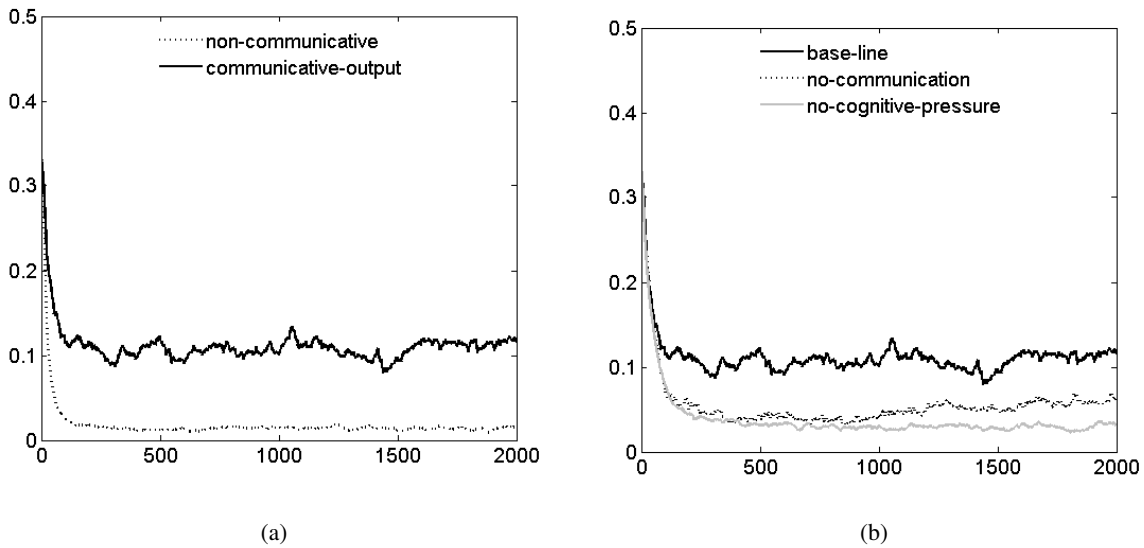(a)                                    (b)

Figure 7: Average variances on the non-communicative and communicative-output weights of the base-line simulations (a) and average variances on the communicative-output weights of the base-line, no-communication, and no-cognitive-pressure simulations (b). Average results of 50 replications of the simulations.

output genes in the three conditions: base-line, no-communication, and no-cognitive-pressures (figure 7b). The three different values reflect the different pressures on those genes in the three conditions. In the no-cognitive-pressure simulation, since there is no spontaneous tendency, in individuals, to produce a good communication system, the low variance of communicative-output weights (about 0.03) is explained by the genetic convergence of the population. The fact that the variance of communicative-output weights in the no-communication condition is higher (about 0.06) and tends to grow after reaching its minimum is due to the fact that in that simulation there is no pressure whatsoever on the communicative-output genes, so they are subject to genetic drift. (While in the no-cognitive-pressure simulation there is some mild pressure to generate bad signals). Nonetheless, the variance of the communicative-output genes is still higher in the base-line simulation than in the no-communication simulation. This confirms our hypothesis that when communication is used there is a reliable selective pressure against good speakers, resulting in a pressure against the genetic convergence on the communicative-output weights: mutations on those genes that generate cheaters are rewarded by selection and spread in the population.

## 6   Discussion

In this section, we discuss our work with the respect to both computational models of the emergence of communication (6.1) and the biological literature on the evolution of communication (6.2).

### 6.1   Synthetic Modeling of the Emergence of Communication

The work presented in this paper clearly demonstrates the soundness of the hypothesis that the emergence of communication can be favored by a cognitive pressure towards spontaneous good signalling. Indeed, we have shown that the need to categorize stimuli in adaptive ways can act as a powerful cognitive pressure to individually produce good signals, while the sharedness of the communication system between different individuals of the same population can be favored by the genetic convergence of the population itself. These two factors, taken together, generate a 'producer bias' which can represent an important factor for the solution of the phylogenetic problem of communication. But can this producer bias, by itself, solve also the adaptive problem of communication? In our simulations the emerged communication

system is continually disrupted by the invasion of the population by cheaters, but in the simulations reported by Cangelosi & Parisi (1998) the cognitive and genetic pressures were enough for communication to both evolve and remain stable. What is the explanation for this difference?

From our analysis, it seems clear that the cognitive pressure cannot be sufficient, by itself, for the evolutionary stability of a communication system which gives benefits only to one of the two actors of a communicative interaction (in our case the hearer). The reason is that a population of altruistically communicating individuals is always subject to be invaded by cheating individuals which benefit from the received signals but do not reciprocate (or *lie*, thus misleading competing individuals), *unless some other factor does not prevent this from happening*. In Mirolli & Parisi (2005a) we have shown three possible factors which can block the invasion of cheaters in a communicating population: kin selection (Hamilton, 1964), the social learning of the communication system through docility (Simon, 1990) and the use of signals not only for communication but also for individual purposes (see also Mirolli & Parisi, 2006) such as as memory aids. Other possible factors are sexual selection through the handicap principle (Zahavi, 1975; Grafen, 1990; Bullock, 1998) and reciprocal altruism (Trivers, 1971; Axelrod, 1981). Why, then, does the communication system in the simulations of Cangelosi and Parisi turn out to be evolutionary stable even though none of the factors discussed above are present?

A possible answer lies in the genetic algorithm used by Cangelosi and Parisi, in particular in their peculiar mutation operator. As in the present work, in Cangelosi & Parisi (1998) the neural networks' connection weights are codified in the genome of organisms as real values. But in their simulations weights are initialized in the range [-1; 1] and each weight has a probability of 10% of being mutated *by adding to the current value a random number in [-1; 1]*. As a result, weights have no limits. This peculiar way of implementing the mutation operator might prevent the invasion of cheaters for the following reason. As demonstrated by our analysis of evolutionary dynamics, during the periods in which the communication system quality is low, received signals are ignored. Consequently, communication-output connection weights are under genetic drift in that they do not affect the fitness of any organism: neither that of the speakers, which is never influenced by its own produced signals, nor that of the hearers, since hearers ignore incoming signals. But a genetic drift without any limits on the connection weights is likely to produce weights of very high absolute values. And a mutation which changes a weight's value by adding a random number in the interval [-1; 1] will tend to be behaviorally uninfluential for weights with very high absolute values. This is especially true for the signalling behavior of the organisms simulated in Cangelosi & Parisi (1998), since the transmission function of 'linguistic' neurons in that simulation is the step function, whose value can be changed only by changing the sign of the input received by the neurons. The consequence is that once a good communication system has emerged as a result of cognitive and genetic pressures, cheaters cannot emerge because mutations on the communication-output connection weights are completely neutral in that they cannot affect signaling behavior.

On the contrary, in our simulations mutations can change radically the value of connection weights (the mutated value is substituted by a random value always in the same range), so cheaters can emerge when a mutation to one of the communicative-output connection weights changes the signalling behavior of an organism. The same is true for the robotic experiment of Marocco et al. (2003), in which connection weights are coded as strings of 8 bits and mutations act on each bit. In this way, all the bits coding for a single weight can change simultaneously and, in general, a mutation in one single bit (for example in the bit coding for the weight's sign) can have a radical effect on the value of the weight and, consequently, on communicative behavior. And in fact, the cognitive pressure towards good signalling is not sufficient for the emergence of communication in their robots.

The mutation operator used by Cangelosi and Parisi has two fundamental problems, one theoretical and one practical. Theoretically, it seems not to be biologically plausible. In real organisms, mutations can happen in many different ways (see Maki, 2002) and they can always have any kind of effect on the mutating gene: from no effect to the substitution of the codified protein. Hence, there seems to be no biological analogous to the mechanism, present in the simulations of Cangelosi and Parisi, which prevents a gene which has been subject to genetic drift from having non-neutral mutations. This may sound as a poor argument since, in a sense, *any* mutation method used in current genetic algorithms (including the

one used in the simulations of the present paper) can be considered to be biologically implausible. The reason is that the genotype-phenotype mappings are typically highly biologically unrealistic (as it is the case for the current work, in which the genotype is constituted by connection weights which are directly mapped to an individual's phenotype). But in addition to the general problem of working on unrealistic genotypes, the mutation operator used by Cangelosi and Parisi has the further specific problem of tending to prevent the occurrence of non-neutral mutations in genes which have been subject to genetic drift. This can have important consequences on a population's evolutionary dynamics which can in turn produce misleading results. Pragmatically, the reason for not adopting the kind of mutation operator used in Cangelosi & Parisi (1998) is that it is too easily subject to falling in local maxima. In general, during a genetic search, some parts of the genome tend to evolve early and other parts lately (either because some part of the genome codes for a trait which is a precondition for the evolution of the trait coded from other parts or simply because of chance). With a mutation system like the one we are discussing, the genetic drift of the genes which are not functional during a given evolutionary phase can prevent the subsequent adaptive evolution of those genes just because they have reached absolute values which are too high with respect to the range on which mutations act. Consequently, this kind of mutation operator can seriously compromise the evolutionary process.

Another take-home message for modelers which are interested in the emergence of communication in artificial agents (be they simulated or real robots) regards the importance of studying the evolution of communication together with the evolution of other, non-communicative behaviors. This is something which several other researchers have already been arguing for (i.e. Nolfi, 2005; Quinn, 2001; Werner & Dyer, 1992). The usual argument, to which we completely subscribe, stems from taking a functional perspective: communication is an adaptive trait which is shaped by its function. Hence, communication can only be understood if it is grounded in the environmental and social needs of communicating organisms. But our work suggests another, practical, reason for making communication emerge from the needs posed by other non-communicative tasks. The reason is that the need for internally categorize experiences in adaptive ways for purposes other than communication can play a major role in solving the phylogenetic problem of communication. As our no-cognitive pressure simulations demonstrate, if the evolution of producer biases towards communication is not allowed, it is much more difficult to evolve a communication system in artificial organisms. And this is true even for cases in which communication would be very useful and the adaptive problems posed by some communication system are not present or solved by other factors, such as kin selection. And if the only thing that agents have to do is to communicate, the presence of producer biases is prevented.

Furthermore, the favorable effect that having to represent one's environment in adaptive ways can have on the emergence of communication can show up only if the control systems of agents permit this to happen. This is the case for neural networks but not for other kinds of control system used in the adaptive behavior literature on the evolution of communication such as look-up tables (e.g. Di Paolo, 1998; MacLennan & Burghardt, 1993; Steels, 1996). Neural networks, as the control systems of real organisms, that is brains, process information in a parallel and distributed way, and the same processing units are involved when perceiving different situations and for producing different behaviors. It is the link between communicative and non-communicative behavior which tends to result in producer biases towards communication. If the control system of an artificial agent is a look-up table, in which the relationships between each possible perception and each possible action are completely independent the ones from the others, the emergence of a producer bias would be very unlikely, and only due to chance. On the contrary, neural networks make the interdependencies between different behaviors explicit, thus providing a systematic mechanism for the emergence of producer biases.

This implies two different views of 'modules' in control systems. In a neural network, and very likely in the brain itself, modules are purely functional but structurally they can include and use subsets of units (neurons) that are also included in and used by other modules with different functions. On the other hand, in non-neural computational models such as look-up tables modules are both functionally and structurally distinct entities. If cognitive pressure and producer bias really exist and have an important role in explaining the evolutionary emergence of communication, neural models appear to be more

appropriate than non-neural models to capture the evolutionary emergence of communication.

## 6.2 Evolution of Communication

As discussed in the introduction, the fact that communication requires two independent capabilities, speaking (producing useful signals) and hearing (understanding perceived signals), poses two related problems: an adaptive problem and a phylogenetic problem. The adaptive problem is a problem of altruism: why should speakers produce good signals if they gain no benefit in doing so? And, conversely, why should hearers respond adequately to signals if they don't receive any advantage from that? Notwithstanding the emphasis of the biological literature on manipulative communication, the situation seems to be completely symmetrical, both theoretically and empirically. And the existence of natural communication systems that seem not to benefit both speakers and hearers needs to be explained either by kin selection or by the handicap principle or by some other means.

But besides the adaptive problem there is the more general phylogenetic problem: what evolutionary paths can lead to the emergence of communication systems, since speaking and hearing abilities are complementary and hence adaptively neutral if they are not both present at the same time? The biological literature has provided a partial answer to this question with the concept of a receiver (or sensory) bias: if some stimuli trigger in some organism a behavioral response that advantages another organism, the latter can produce signals that resemble those stimuli in order to manipulate the behavior of the former. In this way, the problem of co-evolution is by-passed through the exploitation of behavioral biases in the receivers which are assumed to be present for reasons other than communication. There are several reasons for receiver biases to be present. In particular, they can be both the results of selective processes independent from communication (for example, detection of prey), and non-selected, incidental traits which are simply consequences of how the sensory systems and the brain of a species are formed (see Endler & Basolo, 1998; Guilford & Dawkins, 1991). Since its introduction in the biological literature on the evolution of communication the concept of a receiver bias has inspired a lot of both empirical (e.g. Basolo, 1995; Proctor, 1992; Ryan, Fox, Wilczynski & Rand, 1990) and theoretical/computational research (e.g. Arak & Enquist, 1995; Ryan, Phelps & Rand, 2001).

In the present paper we have argued for the complementary idea of the exploitation of a producer bias, which has been so far highly overlooked in the biological literature on the evolution of communication. In fact, though ideas similar to that of producer bias have appeared in the literature under the label of 'intention movements' (i.e. movements necessarily preceding an action; see, e.g., Krebs & Dawkins, 1984; Tinbergen, 1952), these ideas have not received enough theoretical attention and have not inspired any substantial empirical research. The reason for this is, we suppose, twofold: first, as discussed in the introduction, the idea of producer biases is fundamentally related to *informative* communication, while the influential conception of communication of Dawkins and Krebs has led the biological literature to focus on *manipulative* communication; second, ideas similar to the producer bias have mostly been discussed in relation to the adaptive problem of communication (in particular in relation to the handicap principle: see, for example, Lotem, Wagner & Balshine-Earn, 1999; Noble, 1998), while they have never been explicitly and clearly proposed as a possible general solution to the phylogenetic problem of communication.

We argue that just as the exploitation of receiver biases is an important mechanism which can trigger the evolution of manipulative communication systems, the exploitation of producer biases can play an analogous role in the historical emergence of communication systems which (principally) benefit hearers, like alarm and food calls. Furthermore, the simulations presented in this paper suggest also one possible mechanism which can lead to producer biases: namely, the cognitive pressure due to the need for organisms to internally represent stimuli according to adaptively relevant categories. In fact, as our simulations have shown, good categorization of stimuli can result in the spontaneous production of good signals which can be thus be exploited by hearers.

But is the hypothesis of a cognitive pressure towards good signalling behavior plausible? Does it hold only for our simulations and related artificial systems? Or is it also a possible candidate for explaining the evolutionary emergence of real communication systems? We think that the hypothesis is indeed

plausible and that it should be taken in consideration when trying to explain the evolutionary emergence of animal communication systems, in particular those which seem to benefit principally receivers. After all, the only assumption which needs to be made for our hypothesis to work is the link, in the organisms' brains, between communicative and non-communicative behaviors. In fact, if such a link exists, then the need for individuals to categorize the environment in adaptive ways will generate a bias towards the production of useful signals. And this assumption seems quite reasonable: what should signals reflect if not what an organism have in their mind/brain?

Note that notwithstanding our use of a representationalist jargon, the hypothesis of possible cognitive pressures towards good signaling does not depend on taking a representationalist stance in the representationalists vs. dynamicists debate. Even the most radical enthusiast of the dynamical approach to cognition and the most critic or skeptic on the use of the notion of representation for explaining adaptive behavior (Beer, 2003; Brooks, 1991; Harvey, 1996) acknowledges that an organism's behavior depends also on internal (neural) dynamics. Hence, the same single assumption is needed, from a dynamical systems point of view, for accepting the possibility of a possible cognitive pressure towards good signaling: namely, the assumption that signaling behavior is internally linked to the brain structures that govern also other non-communicative behaviors. In fact, for an organism's behavior to be adaptive, different adaptive conditions will be correlated with different internal dynamics, which in turn will tend to be reflected in different signaling behavior. But this is exactly what the hypothesis on a cognitive pressure towards good signalling states: that produced signals will tend to spontaneously reflect adaptively relevant situations due to the need for organisms to cope adaptively with their environment.

Our hypothesis is also empirically testable. One way to test it is the following. Take an animal capable of sending signals. Train it to send a signal in at least two situations which differ from the point of view of the animal's survival (in presence of edible vs. poisonous items, as in our simulation, or for asking two significantly different kinds of things, like food vs. a partner for mating). Our hypothesis of the cognitive pressure on good signalling predicts that the animal will tend to *spontaneously differentiate* between the signals emitted in the two different situations.

It is of the most importance that a first confirmation of this prediction can indeed already be found in recent empirical experiments on Japanese monkeys performed by Hihara, Yamada, Iriki & Okanoya (2003). In these experiments a monkey was trained to produce a generic call in two different conditions: (a) for receiving food and (b) for receiving a tool with which it could reach a distant food. Surprisingly, without any reward for doing so, the trained monkey started to spontaneously differentiate the sounds emitted in the two conditions. The authors of the experiments interpret their striking finding as the result of a change, in the monkey, from emotional to intentional vocalizations. But the spontaneous differentiation of monkey's calls can be explained, less speculatively, by the simple fact that in the two conditions the monkey's brain was in two different states: one corresponding to the need for a piece of food and the other corresponding to the need of a tool with which to reach the food. As predicted by our hypothesis, different internal states, be they 'conscious' (as Hihara et al. suggest), 'intentional' or 'emotional', tend to produce different (communicative) behaviors. To put it shortly: what you have in mind will influence what you say. The signals an organism produce will tend to reflect the way in which the organism categorizes its experience.

Finally, we would like to note that the hypothesis that communication might phylogenetically emerge thanks to the exploitation of producer biases is far more general than the hypothesis of the cognitive pressure towards spontaneous good signalling. In fact, though we think that the need to adaptively categorize the environment does constitute a strong mechanism for producer biases to appear, there is no reason for producer biases to have exclusively cognitive origins. On the contrary, there are good reasons to suppose that other, non-cognitive, mechanisms can equally generate producer biases. Indeed, it is reasonable to think that most emotional communication might have evolved thanks to producer biases: since emotional states are associated with specific internal bodily changes (e.g. changes in heart rate, blood pressure...), the visible consequences of these changes can be exploited by 'receivers' to infer the emotional state of the 'producer', thus resulting in the evolution of an emotional communication signal through a non-cognitive producer bias.

# 7   Summary and Conclusion

In this paper we have presented a simple model which was intended to shed some light on an important and very general problem posed by the evolution of communication, that is, the phylogenetic problem posed by the apparent necessity of the simultaneous co-evolution of both good signal production and good signal understanding, which seem to be two completely independent abilities. Our evolutionary model demonstrated a quite unusual behavior: a sort of limit cycle, present in all the replications of our simulations, in which both individuals' fitness and the quality of the communication system continually oscillate between very low and very high values. In order to understand the mechanisms that produce this striking result, we used two methods. First, we isolated a typical cycle of rise and fall of communication that takes place in one replication of the simulation and we analyzed the evolutionary dynamics during this period. Second, we compared the average results of the base-line simulation with two other simulations in which we disabled communication (no-communication simulation) or we changed the neural networks which control the behavior of our artificial organisms in such a way that cognition was entirely separated from communication (no-cognitive-pressure simulation). The results of our analyzes suggest a possible general solution to the phylogenetic problem of communication: that is, the possibility for communication systems to emerge thanks to the exploitation, by receivers, of producer biases which are due to the need for organisms to categorize perceived stimuli in adaptively relevant ways. We think that the idea of producer biases is both sound and largely unexplored in the biological literature on the evolution of communication, and might represent an important addition to the complementary very well know idea of receiver bias. Indeed, just as the idea of the exploitation of receiver biases have proven to be quite powerful in explaining the evolutionary emergence of manipulative communication systems (like mating signals sent by males to attract females), we think that the idea of the exploitation of producer biases is likely to play an analogous role in explaining the evolutionary emergence of informative signals, like the alarm and food calls used by many species.

## Acknowledgements

## References

Ackley, D. H. & Littman, M. L. (1994). Altruism in the evolution of communication. In Brooks, R. A. & Maes, P. (Eds.), *Artificial Life IV: Proceedings of the International Workshop on the Synthesis and Simulation of Living Systems*, (pp. 40–48)., Cambridge, MA. MIT Press.

Arak, A. & Enquist, M. (1995). Conflict, receiver bias and the evolution of signal form. *Philosophical Transactions of the Royal Society, Series B*, *349*, 337–344.

Axelrod, R. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Basolo, A. L. (1995). Phylogenetic evidence for the role of a pre-existing bias in sexual selection. *Proceedings of the Royal Society of London B*, *259*(1356), 307–11.

Bedau, M. A. (1998). Philosophical content and method of artificial life. In T. W. Bynam & J. H. Moor (Eds.), *The Digital Phoenix: How Computers are Changing Philosophy* (pp. 135–152). Oxford: Basil Blackwell.

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, *4*(3), 91–99.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*(4), 209–243.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence Journal*, *47*, 139–159.

Bullock, S. (1998). A continuous evolutionary simulation model of the attainability of honest signalling equilibria. In Adami, C., Belew, R., Kitano, H., & Taylor, C. (Eds.), *Proceedings of the Sixth International Conference on Artificial Life*, (pp. 339–348).

Cangelosi, A. (2001). Evolution of communication and language using signals, symbols and words. *IEEE Transactions on Evolutionary Computation*, *5*(2), 93–101.

Cangelosi, A. & Parisi, D. (1998). The emergence of a language in an evolving population of neural networks. *Connection Science*, *10*(2), 83–97.

Cangelosi, A. & Parisi, D. (Eds.). (2002). *Simulating the evolution of language*. London: Springer-Verlag.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. & Krebs, J. R. (1978). Animal signals: information or manipulation? In J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach* (pp. 282–309). Oxford: Blackwell Scientific Publications.

Dennett, D. C. (1994). Artificial life as philosophy. *Artificial Life*, *1*(1), 291–292.

Di Paolo, E. (1998). An investigation into the evolution of communication. *Adaptive Behavior*, *6*(2), 285–324.

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In Bedau, M., McCaskill, J., Packard, N., & Rasmussen, S. (Eds.), *Artificial Life VII*, (pp. 497–506)., Cambridge, MA. MIT Press.

Endler, J. A. & Basolo, A. L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in ecology & evolution*, *13*(10), 415–420.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*, 517–546.

Guilford, T. & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, *42*, 1–14.

Hamilton, W. D. (1964). Genetic evolution of social behavior. *Journal of Theoretical Biology*, *7*(1), 1–52.

Harvey, I. (1996). Untimed and misrepresented: connectionism and the computer metaphor. *AISB Quarterly*, *96*, 20–27.

Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.

Hihara, S., Yamada, H., Iriki, A., & Okanoya, K. (2003). Spontaneous vocal differentiation of coo-calls for tools and food in japanese monkeys. *Neuroscience Research*, *45*, 383–389.

Hurford, J. R. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, *77*(2), 187–222.

Kirby, S. (2002). Natural language from artificial life. *Artificial Life*, *8*(2), 185–215.

Krebs, J. R. & Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioral Ecology: An Evolutionary Approach* (second edition ed.). chapter 15, (pp. 380–402). Oxford: Blackwell Scientific Publications.

Lotem, A., Wagner, R., & Balshine-Earn, S. (1999). The overlooked signaling component in non-signaling behavior. *Behavioral Ecology*, *10*, 209–212.

MacLennan, B. & Burghardt, G. M. (1993). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, *2*(2), 161–187.

Maki, H. (2002). Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annual Review of Genetics*, *36*, 279–303.

Marocco, D., Cangelosi, A., & Nolfi, S. (2003). The emergence of communication in evolutionary robots. *Philosophical Transactions to the Royal Society A: Mathematical, Physical and Engineering Sciences*, *361*(1811), 2397–2421.

Marocco, D. & Nolfi, S. (2006). Origins of communication in evolving robots. In Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J., Marocco, D., Miglino, O., Meyer, J.-A., & Parisi, D. (Eds.), *From animals to animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behaviour*, LNAI, (pp. 789–803)., Berlin. Springer Verlag.

Maynard-Smith, J. (1997). *The theory of evolution*. Cambridge: Cambridge University Press.

Maynard-Smith, J. & Harper, D. G. (1995). Animal signals: models and terminology. *Journal of Theoretical Biology*, *177*, 305–311.

Maynard-Smith, J. & Harper, D. G. (2003). *Animal Signals*. Oxford University Press.

Mirolli, M. & Parisi, D. (2005a). How can we explain the emergence of a language which benefits the hearer but not the speaker? *Connection Science*, *17*(3-4), 325–341.

Mirolli, M. & Parisi, D. (2005b). Language as an aid to categorization: A neural network model of early language acquisition. In Cangelosi, A., Bugmann, G., & Borisyuk, R. (Eds.), *Modelling language, cognition and action: Proceedings of the 9th Neural Computation and Psychology Workshop*, (pp. 97–106)., Singapore. World Scientific.

Mirolli, M. & Parisi, D. (2006). Talking to oneself as a selective pressure for the emergence of language. In Cangelosi, A., Smith, A., & Smith, K. (Eds.), *The Evolution of Language: Proceedings of the 6th International Conference on the Evolution of Language*, (pp. 214–221). World Scientific Publishing.

Mitchell, M. (1996). *Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.

Noble, J. (1998). Tough guys don't dance: Intention movements and the evolution of signalling in animal. In Pfeifer, R., Blumberg, B., & Kobayashi, H. (Eds.), *Proceedings of the Fifth International Conference on the Simulation of Adaptive Behavior*, Cambridge, MA. MIT Press.

Noble, J. (1999). Cooperation, conflict and the evolution of communication. *Adaptive Behavior*, *7*(3/4), 349–370.

Nolfi, S. (2005). Emergence of communication in embodied agents: Co-adapting communicative and non-communicative behaviours. *Connection Science*, *17*(3-4), 231–248.

Oliphant, M. (1996). The dilemma of saussurean communication. *Biosystems*, *37*(1-2), 31–38.

Parisi, D. (1997). An artificial life approach to language. *Brain and Language*, *59*, 121–146.

Proctor, H. C. (1992). Sensory exploitation and the evolution of male mating behaviour: a cladistic test using water mites (acari: Parasitengona). *Animal Behaviour*, (44), 745–752.

Quinn, M. (2001). Evolving communication without dedicated communication channels. In Kelemen, J. & Sosik, P. (Eds.), *Advances in Artificial Life: Sixth European Conference on Artificial Life*, (pp. 357–366)., Prague, Czech Republic. Springer.

Ryan, M. J., Fox, J. H., Wilczynski, W., & Rand, A. S. (1990). Sexual selection for sensory exploitation in the frog physalaemus pustulosus. *Nature*, *343*(6253), 66–67.

Ryan, M. J., Phelps, S. M., & Rand, A. S. (2001). How evolutionary history shapes recognition mechanisms. *Trends in Cognitive Sciences*, *5*(4), 143–148.

Searcy, W. A. & Nowicki, S. (2005). *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems*. Princeton, NJ: Princeton University Press.

Simon, H. A. (1990). A mechanism for social selection and successful altruism. *Science*, *250*, 1665–1668.

Steels, L. (1996). Emergent adaptive lexicons. In Maes, P., Mataric, M., Meyer, J., Pollack, J., & Wilson, S. (Eds.), *From Animals to Animats 4: Proceedings of the Fourth International Conference On Simulation of Adaptive Behavior*, (pp. 562–567)., Cambridge, MA. The MIT Press.

Steels, L. (2006). Experiments on the emergence of human communication. *Trends in Cognitive Sciences*, *10*(8), 347–349.

Tinbergen, N. (1952). Derived activities: Their causation, biological significance, origin and emancipation during evolution. *Quarterly Review of Biology*, *27*, 1–32.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift fur Tierpsychologie*, *20*, 410–433.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

VanRhijn, J. & Vodegel, R. (1980). Being honest about one's intentions: An evolutionary stable strategy for animal conflicts. *Journal of Theoretical Biology*, *85*, 623–641.

Wagner, K., Reggia, J. A., Uriagereka, J., & Wilkinson, G. S. (2003). Progress in the simulation of emergent communication and language. *Adaptive Behavior*, *11*(1), 37–69.

Werner, G. M. & Dyer, M. G. (1992). Evolution of communication in artificial organisms. In C. Langton, C. Taylor, D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II* (pp. 659–687). Redwood City, CA: Addison-Wesley.

Zahavi, A. (1975). Mate selection. a selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214.