Philosophical Transactions of the Royal Society London – A (2003)

The Role of Social and Cognitive Factors in the Emergence of Communication: Experiments in Evolutionary Robotics

Davide Marocco[1,3]
[1] University of Calabria,
Centro Interdipartimentale della Comunicazione
Arcavacata di Rende, 87036
Cosenza, Italy
davidem@ip.rm.cnr.it

Angelo Cangelosi[2]
[2]University of Plymouth
Institute of Neuroscience
and School of Computing
Drake Circus, PL4 8AA
Plymouth, UK
acangelosi@plymouth.ac.uk

Stefano Nolfi[3]
[3]National Research Council
Institute of Cognitive Science
and Technologies,
Viale Marx 15, 00137
Rome, Italy
nolfi@ip.rm.cnr.it

Abstract

Evolutionary robotics is a biologically inspired approach to robotics that is advantageous to studying the evolution of language. A new model for the evolution of language is presented. This model is used to investigate the interrelationships between communication abilities, namely linguistic production and comprehension, and other behavioral skills. For example, the model supports the hypothesis that the ability to form categories from direct interaction with an environment constitutes the ground for subsequent evolution of communication and language. A variety of experiments, based on the role of social and evolutionary variables in the emergence of communication, are described.

1. Introduction

The communication between autonomous agents, be they robots or simulated virtual agents, has recently attracted the interest of researchers from different fields. In engineering, the design and evaluation of communication systems is interesting due to its practical applications for agent-agent interaction and also for human-agent and human-robot communication (e.g. Lauria *et al*., 2002). For cognitive scientists, the development of computational models for the evolution of language permits the investigation of the role of sensorimotor, cognitive, neural and social factors in the emergence and establishment of communication and language (Cangelosi & Parisi, 2002).

Studies on the emergence of communication are often based on synthetic methodologies such as adaptive behavior and artificial life (Steels, 1997; Kirby, in press). A group of autonomous agents interact via language games to exchange information about the external environment. Their coordinated communication system is not externally imposed by the researcher, but emerges from the interaction between agents. In such models, the levels of detail of the representation of the agents and of their environment can vary significantly. This constitutes a continuum between abstract point models, at one end, and situated, embodied robots at the other. At one extreme, only the essential communicative properties of the agents and the environment are simulated. For example, the environment can consist of a list of abstract "meanings", and the agent consists of a function, or rule set, that maps these meanings to signals (e.g. Kirby, 2001; Oliphant, 1999). This approach is useful when one wants to study the dynamics of the auto-organization of lexicons and syntax and its dependence on single, pre-identified factors. An intermediate approach to language evolution is based on grounded simulation models (Harnad, 1990). The agents' environment is modeled with a high degree of detail upon which emergent meanings can be directly grounded. Each simulated agent has a set of sensorimotor, cognitive and social abilities that allow it to build, through interaction, a functional representation of the environment and use it to communicate (e.g. Cangelosi, 2001; Cangelosi & Harnad, 2000; Hazlehurst & Hutchins, 1998). This type of models supports the investigation of the interaction amongst various abilities of the agents for the emergence of language and the grounding of communication symbols in the environment and the agent's behavior.

At the other end of the continuum, the communicative behavior of embodied and situated robots results from the dynamical interaction between its physical body, the nervous and cognitive system and the external physical and social environment (Beer, 1995). For example, robots can interact and communicate among themselves (e.g. Steels & Vogt, 1997; Quinn, 2001), with virtual Internet agents (Steels, 1999) and with humans (Steels & Kaplan, 2000). Such an approach permits the study of the interaction between the different levels of a behavioral system, that is from sensorimotor coordination to high-level cognition and social interaction.

Amongst the robotic approaches to studying adaptive behavior, evolutionary robotics (Nolfi & Floreano, 2002) offers a series of advantages. Through evolutionary experiments, artificial organisms autonomously develop their behavior in close interaction with their environment. The main advantages of this approach are: (a) it involves systems that are embodied and situated (Brooks, 1991; Pfeifer and Scheier, 1999), and (b) it is an ideal framework for synthesizing robots whose behavior emerge from a large number of interactions among their constituent parts. This can be explained by considering that, in evolutionary experiments, robots are synthesized through a self-organization process based on random variation and selective reproduction where the selection process is based on the behaviors that emerge from the interactions among the robot's constituent elements and between these elements and the environment. This allows the evolutionary process to freely exploit interactions without the need to understand in advance the relation between interactions and emerging properties as it is necessarily required in other approaches that rely more on explicit design.

For these reasons the evolutionary robotics approach has been successfully applied to the synthesis of robots able to exploit sensorimotor coordination (Nolfi, 2002); on-line adaptation (Nolfi and Floreano, 1999); body and brain co-evolution (Lipson and Pollack, 2000); competing and cooperative collective behaviors; (Nolfi and Floreano, 1998, Martinoli, 1999; Baldassarre, Nolfi, and Parisi, 2002).

These advantageous aspects of evolutionary robotics are of particular importance for modeling the evolution of language and communication. Sensorimotor coordination, social interaction, evolutionary dynamics and the use of neural systems can all have a potential impact in the emergence of coordinated communication. In this paper, new experiments are presented that study the emergence of communication in evolutionary robotics models. They are based on recent work by Nolfi and Marocco (2002) for the emergence of sensorimotor categorization. Nolfi and Marocco evolved the control system of artificial agents that are asked to categorize objects with different shapes on the basis of tactile information. Each agents uses proprioceptive information to actively explore objects using a three-segment arm. In addition, the agent uses the activation of one output node of its neural network controller as input. Agents are selected only for their performance in discriminating (categorizing) the objects using this unit, not for their ability to explore them. This results in the emergence of an active tactile exploration strategy that differentiate between objects of different shapes. Nolfi and Marocco's model is an example of explicit self-categorization.

In this new model, the robotic agents share the explicit categorization of objects. That is, the activation of the output nodes is the signal ("name") sent to another agent to instruct it on what to do with the object. Agents will be selected on their ability to manipulate objects correctly, not on their (linguistic) ability to name them correctly. A variety of experiments will test the role of different social and evolutionary variables. These will be used to analyze the role of sensorimotor, social and cognitive factors in the emergence of communication. The direct relations between behavioral and communication abilities, such us language production and comprehension, will also be discussed.

2. Method

The behavior of each agent consists of exploration within the environment, on the basis of tactile information, and the communication, about the type of objects that are in it. The environment consists of an open three-dimensional space in which one of two different objects is present in each epoch (Figure 1). The two objects used in this simulation are a sphere and a cube.

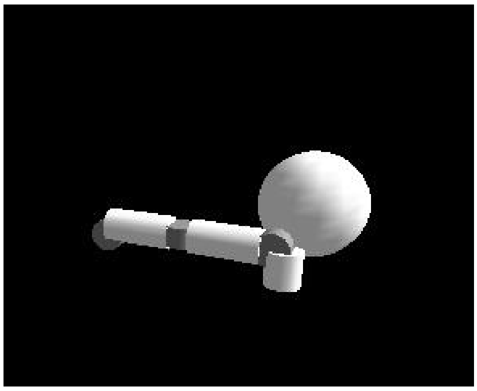

Figure 1 – The arm and a spherical object.

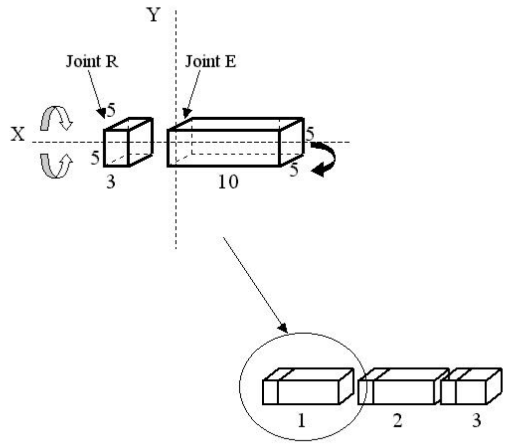

Figure 2 – A schematic representation of the arm.

Agents are provided with a 3-segments arm with 6 degrees of freedom (DOF) and extremely course touch sensors (see Figure 2). Each segment consists in a basic structure of two cylindrical bodies and two joints. This is replicated for three times. The basic structure consists of a shorter body of radius 2.5 and length 3 and a longer body of the same radius and length 10 for the first two segments. The length of the third segment is 5. This shorter segment represents a fingerless hand. The two bodies of each segment are connected by means of a joint (i.e. the *Joint E* in the Figure) that allows only one DOF on axis *Y*, while the shorter body is connected at the floor, or at the longer body, by means of a joint (i.e. the *Joint R*) that provides one DOF on axis *X*. In practice, the *Joint E* allows to elevate and lower the connected segments and the *Joint R* allows to rotate them in both direction. Notice that *Joint E* is free to moves only in a range between 0 and $\pi/2$, just like a human arm that can bend the elbow solely in a direction. The range of *Joint R* is [–π/2, +π/2]. Gravity is {0, –1, 0}. Each actuator is provided with a corresponding motor that can apply a maximum force of 50. Therefore, to reach every position in the environment the control system has to appropriately control several joints and to deal with the constraints due to gravity.

The sensory system consists of a simple contact sensor placed on each longer body that detects when this body collides with another, and two proprioceptive sensors that provide the current position of each joint.

The controller of each individual consists of an artificial neural networks with 11 sensory neurons connected to 3 hidden neurons. These connect with 8 output neurons. The first 9 sensory neurons encode the angular position (normalized between 0.0 and 1.0) of the 6 DOF of the joints and the state of the three contact sensors located in the three corresponding segments of the arm. The other 2 sensory neurons receive their input from the other agents. The first 6 motor neurons control the actuators of the corresponding joints. The output of the neurons is normalized between [0, +π/2] and [–π/2, +π/2] in the case of elevation or rotational joints respectively and is used to encode the desired position of the corresponding joint. The motor is activated so to apply a force (up to 50) proportional to the difference between the current and the desired position of the joint. The last 2 output neurons encode the signal to be communicated to the other agents. This works as a small winner-takes-all cluster, where the neuron with the highest activation is set to 1 and the other to 0.

The activation state of internal neurons was updated accordingly to the following equations (output neurons were updated according to the logistic function):

$$A_j = t_j + \sum w_{ij} O_i$$
$$O_j = \tau_j O_j^{(t-1)} + (1-\tau_j)(1+e^{-A_j})^{-1} \qquad \textbf{(1)}$$
$$0 \le \tau_j \le 1$$

With *Aj* being the activity of the *j*th neuron (or the state of the corresponding sensor in the case of sensory neurons), *tj* the bias of the *j*th neuron, *Wij* the weight from the *i*th to the *j*th neuron, *Oi* the output of the *i*th neuron. *Oj* is the output of the *j*th neuron, *τj* the time constant of the *j*th neuron.

Each individual was tested for 36 epochs, each epoch consisting of 150 sensorimotor cycles. At the beginning of each epoch the arm is fully extended. A spherical or a cubic object is placed in a random selected position in front of the arm. The position of the object is randomly selected between the following intervals: 15.0 <= X <= 25.0; Y = 7.5; –5.0 <= Z <= 5.0). The object is a sphere (15 units in diameter) during even epochs and a cube (15 units in side) during odd epochs so that each individual has to discriminate the same number of spherical and cubic objects during its lifetime.

In addition to the proprioceptive information, agents also receive in input a 2-bit signal produced by some other agent in the population, such as the parent or any agent from the population (linguistic comprehension task). The protocol of interaction and communication between agents was systematically varied and is analyzed in section 3.

Before they act as speaker, agents undergo a linguistic production task. That is, each agent is put in the environment and asked to interact with the object. The value of the two output neurons in the last cycle of the epoch is saved and used as the signal produced to "name" the object.

A genetic algorithm is used to evolve the behavior of agents. The genotype of each agent consists of 81 parameters that include 67 weights, 11 biases, and 3 time constants. Each parameter is encoded with 8 bits. Weights and biases are normalized between –5.0 and 5.0, time constants are normalized between 0.0 and 1.0.

The fitness rewards the behavior of the agent with the current object in the environment. Good communication behavior does not produce any fitness gain for the speaker. Following the behaviors evolved in Nolfi & Marocco's (2002) simulation, the agent has to touch and stay in contact with one object (the sphere) and has to avoid as much as possible to touch the other object (cube). The fitness of individuals is computed by summing the number of cycles in which the agent touches the sphere or does not touch the cube. Fitness scores decrease for each cycle the agent touches the cube or when it does not touch the sphere.

A population of 80 agents is used in each simulation. During selection, the 20 agents with the highest fitness (i.e. behavioral performance) reproduce and each make 4 offspring. The genotype of each offspring is then subject to mutation with an overall probability of 2%. That is, each bit has a 2% probability of being mutated, by generating a random binary value. There is generational overlap between the population of parents and that of new offspring. The first will only act as speakers and cannot reproduce anymore. The population of new offspring will be subject to the fitness test and will reproduce at the end of their lifetime.

Evolutionary simulation of embodied robotic agents can be time consuming and computationally expensive. To reduce the time necessary to test individual behaviors and to model the real physical dynamics as accurately as possible, the rigid body dynamics simulation SDK of Vortex*TM* was used[1]. This was linked to the EvoRobot simulator (Nolfi, 2000).

3. Results

The simulation model was used to run a series of experiments on the role of various social and evolutionary variables in the emergence of shared communication. The first independent variable refers to the selection of speakers (SPEAKER) with two levels: Parent or All. In the first case, each agent receives communication signals only from its own parent. In the second level of the variable, each agent

[1] http://www.cm-labs.com/products/vortex/

can receive signals from any individual of the previous population. This factor is aimed at investigating the role of different social groups of speakers in facilitating shared communication.

The second independent variable manipulated during experiments consists in the time in which communication is allowed (COMMUNICATION) with two levels: From_0 and From_50. In the first case, agents were allowed to communicate from the initial random generation. In the second level of the variable, agents start to communicate between themselves only at generation 50, i.e. after they have evolved a good ability to touch/avoid the two objects. Through this variable it will be possible to investigate the initial behavioral and cognitive abilities necessary to evolve communication.

For each of the 4 conditions (2 SPEAKER × 2 COMMUNICATION), 10 replications were executed, by changing the initial random population. Fifty generations were necessary to pre-evolve an optimal behavior of object manipulation to be used in the From_50 conditions. Table 1 reports the communication success in each condition in terms of good populations and percentage of good speaker in the population. The criterion for deciding whether a population has successfully evolved communication depends on the fact that, at the last generation, at least 50% of agents produce two signals that differentiate the two objects.

Table 1 – Data on the emergence of communication in each experimental condition. The first line contains the number of populations (out of 10) where communication emerged. The second line contains the average percentage of good speakers for the 10 replications and the average for the best performing population (value between brackets).

SPEAKER		COMMUNICATION From 0	COMMUNICATION From 50
Parent	# good pops	5	7
	% speakers (best pop)	*27% (75%)*	*63% (100%)*
All	# good pops	0	0
	% speakers (best pop)	*7% (20%)*	*5% (27%)*

The results of the number of populations that evolve shared communication clearly show that it is only when the parents act as the speakers there is a selective pressure for the emergence and preservation of a shared communication system. In particular, 7 populations out of 10 reach a stable communication system when language is introduced after agents have learned to use both objects. Figure 3 shows the fitness curves and the proportion of good speaker in the best seed of the condition From_50 - Parent speaker.

When communication is introduced directly from the initial random population, the probability of evolving a good language, together with a good behavior, is lower (5 populations out of 10). This advantage for evolving languages after the basic behavioral skills have evolved is similar to that observed by Cangelosi & Parisi (2001) in a grounded simulation model on the emergence of verbs and nouns.

When agents listen to all individuals of the previous generation, no stable communication exists in the last generations. In fact, during evolution good lexicons sometimes emerge for a short time, but they are not maintained or further developed by the whole population. A temporary good lexicon is defined as the case in which at least 20% of agents use two different signals to name the two objects. In 8 of the 10 From_50 - All speaker populations, such temporary appearances of good signal production is observed. Figure 4 shows the best population in the From_50 - All speaker conditions. Here the longest period of good production only lasts for 17 generations, with a maximum peak of best language at 41%.

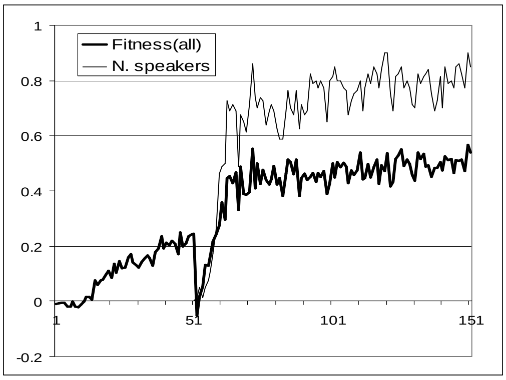

Figure 3 – Data for the best population of the condition Parent speaker - From_50.

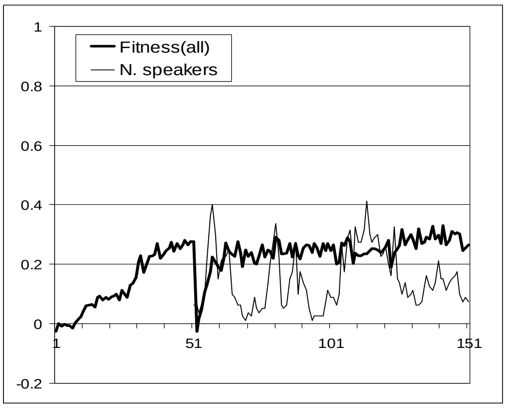

Figure 4 – Data for the best population in condition All speakers - From_50.

The lexicon produced by agents in successful replications has been tested to investigate whether individuals actually use this language in a meaningful way, i.e. avoid the cube when the signal produced in response to the cube is used, and touch the sphere when the other signal is used. Figure 5 shows the behavior of an agent that interacts with the cube with or without language. This tests the linguistic comprehension ability of agents. The pictures on the left column (Figure 5 - left) show the behavior of the agent when no input signal is used. The agent needs to touch the cube, at least once (in cycle 95), to identify it as a cube and then retract from it. The pictures on the right (Figure 5 - right) show the behavior of the agent when the signal "10" is used as additional input. This signal is produced by the parent organism at the end of the interaction with a cube. During this scene, the agent does not need to touch the cube at all because the signal "10" identifies it as a cube. The meaning of "10" can be interpreted as "cube"[2], because the listener treats the object as a cube, and the speaker produces it after its interaction with a cube. When the signal "01" is used, the agent touches the object regardless of its shape. In this case, "01" has the meaning of "sphere".

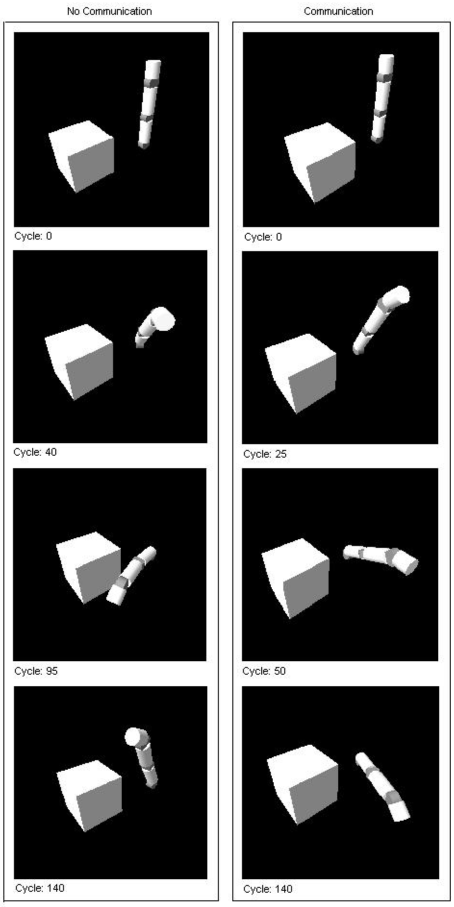

Figure 5 – Agent's interaction with the cube and test of linguistic understanding ability. Left column: Only the proprioceptive input is given to the agent. Right column: An additional communication signal is given as input. This is produced by another agent at the end of its interaction with a cube. Figures from the best individual of a From_50 - Parent speaker population.

Fitness data shows that final scores in the 4 experimental conditions reflect the pattern of results on the emergence of successful communication. The two conditions with Parent speakers reach the highest fitness scores, with a significant advantage for the From_50 populations (e.g. average fitness of best individuals = 0.55; fitness peak in best population = 0.72) versus the From_0 population (average = 0.45, peak = 0.66). The baseline for the behavior without communication is the fitness at generation 50 of the From_50 simulation, before agents start to communicate (average = 0.44, peak = 0.52). Consider that the maximum hypothetical fitness score is 1. This can never be reached because, for example, at the beginning of each epoch some negative fitness cycles are always necessary for agents to reach the spherical object and start gaining fitness.

4. Discussion

There are several issues that can be discussed regarding these results, and what we can learn from the model. A series of questions will be used to analyze the results.

Question 1: Is there any benefit to be in a population where good communication has emerged?

Question 2: Is there any direct advantage to evolving a good linguistic comprehension ability?

To answer the first question, it is possible to compare the fitness results in the simulations where no shared communication emerged, and those where good communication systems evolved. The condition in which communication emerged more frequently (From_50, Parent speaker) will be used as example. In this condition, 7 populations evolved good languages, whilst 3 did not. Figure 6 shows the average fitness of the good communication populations (thick lines) and that of the no communication populations (thin lines). The chart clearly shows that agents who use communication reach fitness values that are higher that those not communicating. This is true both for the fitness of the best individual and for that of the whole population. For example, at the final generation the average fitness of the 7 successful communication replications is 0.35, while it is 0.21 for the 3 unsuccessful populations. Moreover, the fitness in these 3 populations remains relatively constant during the simulation. In the first 50 generations after communication is permitted (i.e. from 50 to 100), there is no increase and the average fitness at generation 100 is very similar to that at generation 50. In the remaining generations, the agents gain some extra fitness points, which are due to the continuation of the evolutionary algorithm search.

The extra fitness gain in populations that evolve communication is easily explained by the direct benefits for the behavior (i.e. fitness) of using two different signals: one for the cube, and one for the sphere. As already shown in Figure 5, during the interaction with a cube the input of its "name" produces significant improvements to behavioral performance. Agents do not need to touch the object to recognize it, and therefore do not lose fitness due to such exploratory behavior. In addition, they gain fitness in every cycle. There is also some benefit for the use of the signal for the sphere. If an agent initially is told that there is a spherical object in the environment, it can go directly

[2] This signal can also be interpreted as the verb *avoid*, instead of as the noun *cube*. In fact, in this model it is not possible to distinguish between syntactic word classes (cf. Cangelosi & Parisi 2001 and Cangelosi 2001 for a discussion)

towards the object and touch it, without having to use some interaction cycles for recognizing the object as a non cube.

The previous explanations also answer the second question, since they identify a direct adaptive advantage for evolving a good comprehension ability.

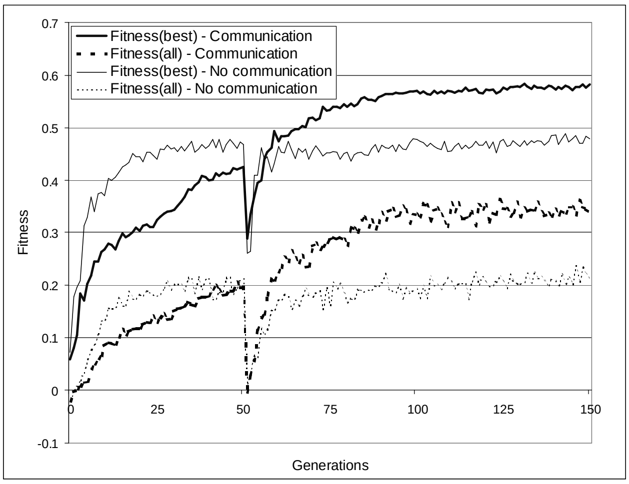

Figure 6 – Average fitnesses of the conditions From_50 - Parent speaker. Thick lines refer to the average fitness of the 7 replications where good communication emerged (continuous line for the best agent and dotted line for the average of all agents). Thin lines refer to the average fitness of the 3 replications where no shared communication emerged.

Question 3: Is there any "direct" advantage to evolving good linguistic production abilities?

This question is more difficult to answer. In fact, there seems to be no direct fitness advantage to the agents to speaking well. Individuals only update their fitness when they hear others speaking. When agents act as speakers, some have already reproduced, whilst the others have not been selected at all. In the condition Parent speaker, agents only speak to their own children. Therefore, the kinship relationship can partially explain this apparent altruistic behavior and the indirect fitness gain for the common genes shared by the parent and its offspring (e.g. Ackley & Littman, 1994). The benefits of kin selection can also explain the successful evolution of communication in the Parent speaker versus the All speaker conditions. However, there is another important phenomenon to be considered. In the Parent speaker conditions, the linguistic input to each listener is constant, since its parent will always use the same signal for the same object. In addition, when the parent is a good speaker (i.e. it uses two different signals to refer to the two objects), its signals are more reliable. The child can then try to use them to improve its fitness performance. In the All speaker conditions, the high variability of the linguistic input coming from all agents of previous generation can be too unreliable, and agents will tend to ignore it.

In the All speaker conditions, some communication abilities also emerge, although the number of good speakers never reaches the critical amount needed to allow it to remain stable until the end of the simulation (cf. Figure 4). In addition, in the Parent speaker conditions, there are three cases when shared communication does not evolve. According to the altruistic, kin selection explanation, all Parent speaker populations should evolve communication because of it indirect adaptive advantage. The fact that this does not always happen raises the issues of understanding the relation between linguistic comprehension/production abilities and other behavioral/cognitive abilities (question 4), and the identification of factors that cause and favor the emergence of shared communication (question 5). First, the data in Table 1 indicates that it is easier to evolve good communication when language is introduced after the pre-evolution of good behavioral capacities (7 out of 10 populations) than when agents are allowed to communicate from the initial generation (5 out of 10 seeds). In addition, the onset of effective communication (i.e. when at least 20% of agents speak well) is much earlier in the From_50 populations (on average after 16 generations) that in the From_0 simulations (on average after 41 generations). This data is consistent with Cangelosi and Parisi's (2001) model on the evolution of syntactic languages. This research showed that agents learn languages more efficiently when communication is introduced after the pre-evolution of good behavioral skills. Effectively, the pre-evolution of good behavior "prepares" a cognitive ground upon which good linguistic abilities can start to develop. Analyses of the categorical perception effects in language learning models have shown that language uses and modifies the space of similarities between members of different perceptual and linguistic categories (Cangelosi & Harnad, 2000).

Question 4: What is the relation between comprehension, production and behavioral abilities?

Question 5: What are the underlying factors that cause and favor the emergence of communication?

To understand better the relations between communication abilities and behavioral skills, the correlations between fitness scores and a measure of the quality of produced language have been computed. Figure 7 and 8 present the averages of the fitness curves, the proportions of good speakers (i.e. language index), the fitness/language correlation r_{all} for the whole population, and the fitness/language correlation r_{best} for the best 20 agents. Figure 7 refers to the 7 successful populations of the From_50 - Parent speaker condition. Figure 8 refers to data from the remaining 3 populations without communication. For the computation of the language index based on the proportion of good speakers, an agent is classified as good speaker when it produces two opposite signals respectively for the two objects in at least half of the 36 epochs. The Pearson *r* correlations index was used.

Overall, the two figures show that the correlation between the fitness of all agents and their language production index is positive and quite high ($r_{all} \approx 0.5$) <u>after</u> good communication emerges. This can explain the maintenance of good communication, since it reflects a link between good speaking abilities and good comprehension (i.e. behavioral fitness).

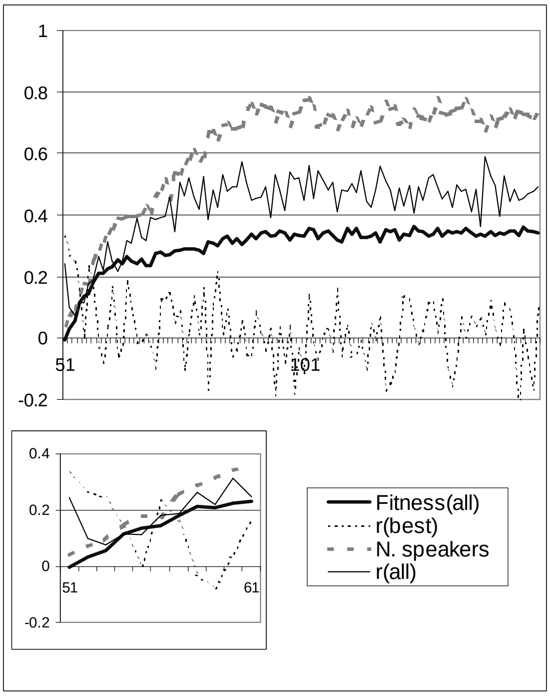

Figure 7 – Fitness curve of the whole population, number of good speakers, fitness/language correlation r_{all} for the whole population and fitness/language correlation r_{best} for the best 20 agents (i.e. future parents and speakers). Average curves over the 7 successful From_50-Parent speaker populations. Only the data for generations 51-150 are shown. (see text for discussion)

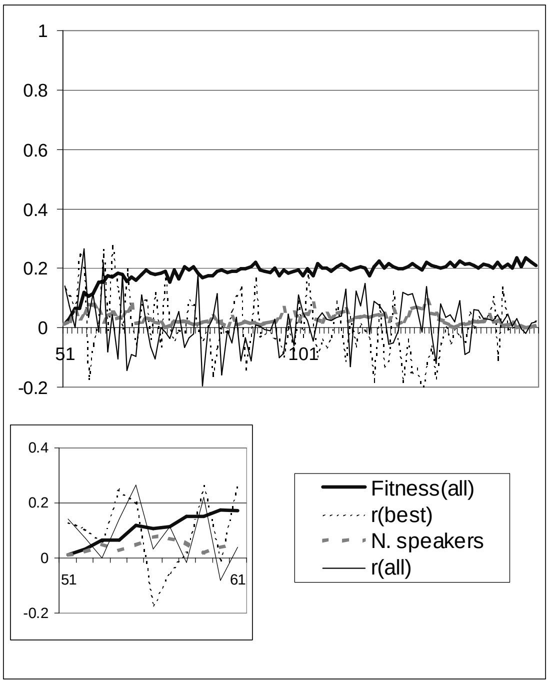

Figure 8 – Data for the 3 unsuccessful From_50-Parent speaker populations. (see text for discussion)

There is no correlation between the fitness of the 20 best performing agents and their speaking ability. The r_{best} stays around 0 (with peaks of ±0.2) in both groups, for the majority of generations. However, this correlation differs significantly in the initial generations of the two groups of populations with and without communication. At generation 51, there is a high positive fitness/language correlation for the 20 best organisms (r_{best}=0.33) of the populations that succeed. The correlation is much lower in the populations that do not succeed (r_{best}=0.13). This indicates that, to evolve communication, it is necessary to be in a population where there is an initial positive correlation between the fitness and language of the best performing individuals. This initial correlation could be due to the role of hidden units, where the linguistic production ability and the comprehension/fitness abilities interact. As it has been shown previously (Cangelosi & Parisi, 1998), the ability to recognize and categorize the two objects can produce quite distinct activation patterns in the hidden units. These will, in turn, increase the possibility of initially producing different signals for the two categories of objects.

The initial difference in the correlation r_{best} between the successful and unsuccessful populations quickly disappears. The correlation index becomes 0 after approximately 5 generations (cf. smaller chart of Figures 7 and 8, which zoom in the first 10 generations). However, the initial advantage of this high r_{best} correlation has the effect of supporting the fitness/language correlation r_{all}. In the unsuccessful populations, this correlation goes down and stays around 0 for all subsequent generations. Instead, in the successful population, r_{all} never reaches 0, and it starts to grow since generation 4. The initial strengthening of the link between production and fitness in the whole population will subsequently help the establishing and maintenance of a shared lexicon.

The analyses of such correlations explain the fine interrelationships between language production, language comprehension, and fitness. In addition, it highlights the role of cognitive factors in supporting and favoring the emergence of communication.

Conclusions

To summarize, the simulation of this evolutionary robotics model of the evolution of communication shows that: (a) the emergence of language brings direct benefits to the agents and the population, in terms of increased fitness and comprehension ability; (b) there is a benefit in communicating with your kin-related agents (e.g. between parents and children), since this improves the possibilities of successfully evolving shared lexicons also by maintaining stable and reliable signals; (c) good sensorimotor and cognitive abilities permit the establishment of a link between production and comprehension/behavioral abilities; (d) the kinship relation between speaking parents and listening offspring does not fully explain the emergence of communication, since the r_{best} stays around 0 for most of the generations – instead, this is important in the early stages of communication because it exploits the cognitive benefits of positive production/fitness correlations.

Most of these results have important implications for the theories and hypotheses on the origins of language. For example, this simulation highlights and explains the role of cognitive factors in the emergence of communication (Burling, 1993). In particular, the model supports the hypothesis that the ability to form categories constitutes the grounding for the subsequent evolution of words and language (Harnad, 1996; Cangelosi & Harnad, 2000). In addition, future developments of this model could also have an impact on computational investigations of the mirror neuron hypothesis for the origins of language (Arbib, 2002).

Further simulations will address in more detail the role of sensorimotor coordination, cognitive, neural and social factors in the emergence of complex communication systems, such as syntactic languages. For example, the authors plan to investigate (a) the factors that favor the emergence of syntactic lexicon within such an evolutionary robotics model, (b) whether listening to our own language might contribute to the development of a communication ability, and (c) whether language and communication might lead to the development of internal categories that, aside from communication, can be used by the robot to better fulfill its own goals.

Acknowledgements

The work of D. Marocco and A. Cangelosi has been partially supported by the UK EPSRC (Grant: GR/N01118).

References

Ackley D.H., Littman M.L. (1994). Altruism in the evolution of communication. In: Brooks R, Maes P (eds) *Proceedings of the Fourth Artificial Life Workshop*. MIT Press, Cambridge

Arbib M. (2002). Grounding the mirror system hypothesis for the evolution of the language-ready brain. In Cangelosi & Parisi (2002).

Baldassarre G., Nolfi S. and Parisi D. (2002). Evolving mobile robots able to display collective behaviours, *Technical Report*. Rome: Institute of Cognitive Science and Technologies, CNR.

Beer, R. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72, 173-215

Braitenberg, V. (1984). *Vehicles*. Cambridge, MA: MIT Press

Brooks R.A. (1991b) New approaches to Robotics. *Science*, 253:1227-1232.

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

Cangelosi A., Harnad S. (2000). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication*, 4, 117-142

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. (2001). How nouns and verbs differentially affect the behavior of artificial organisms. In J.D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, London: LEA, 170-175

Cangelosi A., Parisi D. (Eds.) (2002). *Simulating the Evolution of Language*. London: Springer.

Floreano D., Urzelai J. (in press). Neural morphogenesis, synaptic plasticity, and evolution. *Theory in Biosciences*.

Harnad S. (1990). The Symbol Grounding Problem. *Physica D*, 42: 335-346.

Harnad S. (1996). The origin of words: A psychophysical hypothesis. In Durham W., Velichkovsky B. (Eds.) *Communicating Meaning: Evolution and Development of Language*. NJ: Erlbaum.

Hazlehurst B., Hutchins E. (1998). The emergence of propositions from the co-ordination of talk and action in a shared world. *Language and Cognitive Processes*, 13: 373-424

Kirby S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation and Cognitive Science*, 5(2): 102-110

Kirby S. (in press). Natural language from artificial life. *Artificial Life*.

Lauria S., Bugmann G., Kyriacou T., Klein E. (2002). Mobile robot programming using natural language. *Robotics and Autonomous Systems*

Lipson H., Pollack J.B. (2000). Automatic design and manufacture of robotic lifeforms. *Nature*, 406: (6799) 974-978.

Martinoli A. (1999). *Swarm intelligence in autonomous Collective robotics: from tools to the analysis and synthesis of distributed control strategies*. PhD Thesis. Lausanne - EPFL

Nolfi S. (2000). Evorobot 1.1 User Manual. *Technical Report*, Institute of Psychology, Rome

Nolfi S. (2002). Power and Limits of Reactive Agents. *Neurocomputing*, 42:119-145.

Nolfi S., Floreano D. (1998). Co-evolving predator and prey robots: Do 'arm races' arise in artificial evolution? *Artificial Life*, 4(4), 311-335.

Nolfi S., Floreano D. (1999). Learning and Evolution, *Autonomous Robots, 7(1):* 89-113.

Nolfi S., Floreano D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA: MIT Press

Nolfi S., Floreano D. (2002). Synthesis of autonomous robots through artificial evolution, *Trends in Cognitive Sciences* 1, 31-37.

Nolfi S., Marocco D. (2002). Active perception: A sensorimotor account of object categorization. *From Animals to Animats 7. Proceedings of the 7th Intl. Conf. on Simulation of Adaptive Behavior.*

Oliphant M. (1999). The learning barrier: Moving from innate to learned systems of communication. *Adaptive Behavior*, 7(3/4).

Pfeifer R., Scheier C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press

Quinn M. (2001). Evolving communication without dedicated communication channels. In J. Kelemen and P. Sosík (eds.), *Proceedings of ECAL01*, Springer. pp. 357-366

Steels L. (1997) The synthetic modeling of language origins. *Evolution of Communication*, 1(1).

Steels L. (1999). *The Talking Heads Experiment (Volume 1: Words and Meanings)*. Laboratorium: Antwerpen.

Steels L., Kaplan F. (2000). AIBO's first words: The social learning of language and meaning. *Evolution of Communication*, 4(1)

Steels L., Vogt P. (1997). *Grounding adaptive language games in robotic agents*. In P. Husband & I. Harvey (eds). Proceedings of the Fourth European Conference on Artificial Life, London: MIT Press