IEEE Transactions in Evolution Computation, 5, 93-101, 2001

Evolution of Communication and Language Using Signals, Symbols, and Words

Angelo Cangelosi

Abstract

This paper describes different types of models for the evolution of communication and language. It uses the distinction between signals, symbols, and words for the analysis of evolutionary models of language. In particular, it show how evolutionary computation techniques, such as Artificial Life, can be used to study the emergence of syntax and symbols from simple communication signals. Initially, a computational model that evolves repertoires of isolated signals is presented. This study has simulated the emergence of signals for naming foods in a population of foragers. This type of model studies communication systems based on simple signal-object associations. Subsequently, models that study the emergence of grounded symbols are discussed in general, including a detailed description of a work on the evolution of simple syntactic rules. This model focuses on the emergence of symbol-symbol relationships in evolved languages. Finally, computational models of syntax acquisition and evolution are discussed. These different types of computational models provide an operational definition of the signal/symbol/word distinction. The simulation and analysis of these types of models will help understanding the role of symbols and symbol acquisition in the origin of language.

Index Terms: Evolution of language, Artificial Life, Symbol grounding, Neural networks

I. INTRODUCTION

A. *Icons, Indices, and Symbols*

Analyses of linguistic and communication systems are mainly based on the semiotic distinction between icons, indices, and symbols. These distinctions, originally introduced by Peirce [22], have been reproposed and slightly revised in recent language origin works (e.g., [12], [9]). Briefly, Peirce's original distinction between icons, indices, and symbols is based on the fact that an "icon" has physical resemblance with the object it refers to, an "index" is associated in time/space with an objects, and a "symbol" is based on a social convention or implicit agreement.

Harnad [11,12] distinguishes between three types of representations that are used by natural (and artificial) cognitive systems to build mental representations and classifications of the external environment. He hypothesises that symbols (words) originated as the names of perceptual categories that are based on iconic and categorical representations. Initially, each cognitive system builds an iconic representation of the perceived object. It corresponds to the sensory representation of an object, such as its projection in the retina. The retinal image of a horse is an iconic representation. Subsequently, this representation is processed and used to build a categorical representation. The object is represented by some essential (indexical) features that define its membership in a category. The category of horses is an example of categorical representation. These categorical representations [12] are useful to sort out the extensive perceptual variability of objects in the real world. Indeed, the ability of humans and animals to create categories, e.g. through categorical perception, constitutes the "groundwork" of cognition [11]. Upon this basis, it is possible to build more complex cognitive skills, such as language. The next level of representation is called symbolic. Symbolic representation makes it possible for us to name and describe our environment, in terms of objects' categories, their memberships, and their invariant features. The word "horse" is such a type of symbolic representation. Symbolic representations can be combined together to describe new entities and relations. For example, the word "horse" and "stripe" can be used together to describe the concept of "zebra." Symbols constitute the basis of language, especially in human languages.

Deacon [8,9] uses a hierarchy of referencing systems based on the three levels of iconic, indexical, and symbolic relationships. Icons are associated with entities in the world because of stimulus generalisation and conventional similarity. Indices are associated to world entities by spatio-temporal correlation or part-whole contiguity. They are typical of conditional learning based on simple stimulus association. Indexical references are used in common animal communication systems. Symbols have referential relationships to indexical relationships, and also to other symbols[1]. Human languages are based on the use of such symbols.

Recently, Deacon [9] proposed an explanation of the origin of language that is based on this hierarchical referencing system. His theory relies on the main distinction between communication with and without the use of symbolic representations to explain the evolutionary gap between

Angelo Cangelosi, Centre for Neural and Adaptive Systems and Plymouth Institute of Neuroscience, University of Plymouth, Plymouth PL4 8AA (UK). acangelosi@plymouth.ac.uk

[1] For a critique of Deacon's use of the term "reference" for describing relationships between symbols see [15]

animal and human communication systems. In fact, a variety of animal communication systems exist and have been studied in detail [14]. There is no apparent continuity between animal communication systems and complex human languages. No animal “simple languages” have been discovered, i.e., communication systems using some elementary forms of word combinations or syntax. The lack of simple languages helps explain the gap between animal and human communication. Deacon [8,9] ascribes this to the symbol acquisition problem. Indeed, the main difference between animal and human communication pertains to symbolic references. There is a significant difference between the animal indexical referencing system of simple object-signal associations and that of humans’ symbolic associations. In animals, simple associations between world entities and signals (e.g., monkeys’ calls) are mostly innate and can be explained by mere mechanisms of rote learning and conditional learning. A Vervet monkey always uses a call in association with a specific predator [5]. Instead, symbolic associations have double references, one between the symbol and the object, and the second between the symbol itself and other symbols. When a complex set of logical and syntactical relationships exist between symbols, we can call them words and distinguish grammatical classes of words. A language-speaking human knows that a word refers to an object and also that the same word has grammatical relationships with other words. Due to the possible combinatorial interrelationships between words, there can be an exponential growth of reference with each newly added word. Syntax allows the combination of more words to express new meanings. Therefore, each new word of the lexicon can be used to exponentially increase the overall number of meanings that the language can express.

B. Signals, Symbols, and Words in Language Evolution

In the last decade, computational modelling has been applied to the study of the evolution of language and communication. These models deal with different types of communication systems. Some rely on the use of simple signals, while others use symbolic communication systems or complex syntactical structures. Amongst the different types of computational approaches, evolutionary computation techniques, such as the synthetic approach of artificial life [20, 25], can be used to study the emergence of communication. This approach permits the study of the different stages of semiotic complexity, from simple associations between signals and objects to symbolic representations, and then to complex syntactic relationships between words.

Evolutionary computational models also offer the advantage to deal with the symbol grounding problem [12]. Computational cognitive models require an intrinsic link between the symbols used in the model, such as words, and their semantic referent in the external environment. In classical rule-based cognitive models, and in some evolutionary models of language, the intrinsic link between symbols and a world's entities does not exist. Organisms' internal symbols need sensorimotor grounding, otherwise the role of these models for understanding the evolution of cognition would be diminished. Instead, evolutionary and artificial life methodologies overcome the symbol grounding problem. For example, neural networks can be used for modelling organisms' neural and cognitive systems to build iconic and categorical representations. Subsequently, these representations can be used for high-level symbolic representations. In fact, simulated organisms can use symbols whose semantic referents are made up of categorical representations, such as the internal representations of a neural network. Organisms' iconic and categorical representations can be activated by the actual presence of their referents in the organism's world, therefore by directly grounded symbols in the external world.

The difference between different types of associations (e.g., simple indexical relationships versus symbolic associations) and their relation to the computational models of language evolution is represented graphically in Figure 1. The objects and symbols used in this example are taken from Savage-Rumbaugh & Rumbaugh’s [24] experimental stimulus set in ape language research. In this figure, the upper level always refers to the linguistic representations of signals/symbols, whilst the lower level refers to the objects present in the environment. Note that the relationship between symbols and objects, which constitutes the grounding of symbols into entities of the real world, is not a direct link between mental symbols and real objects. Instead, it is a link between mental entities (the symbols or words) and other mental entities (such as concepts) that constitute the semantic reference. Therefore, the objects in the lower level refer to a semantic categorisation that organisms can make of these objects. These representations are mediated by the organisms' sensorimotor and cognitive abilities. Solid foods, such as a banana and an orange, are represented with a link between themselves, and the two drinks are also linked. In fact, apes group food together because they require similar sensorimotor behaviour. In Rumbaugh's experiments, apes obtain food using a vending machine that gives solid food and pours drinks. Therefore animals learn that foods are "given," while drinks are "poured."

Figure 1a represents a communication system based on grounded signals. Communication relies on simple indexical associations between objects and signals. This situation refers to the models of the evolution of language that only focus on lexicon emergence (e.g., [4], [27]). Communication signals, that are directly grounded in the organisms' environment, do not have any symbolic or syntactic properties.

Figure 1b shows a system based on grounded symbols (e.g., [26], [2]). In the top layer there are links between symbols, and references between symbols and objects. Following the previous definition of symbol by Deacon [9] and [13], we can categorise this as a symbolic communication system due to the relationships between symbols.

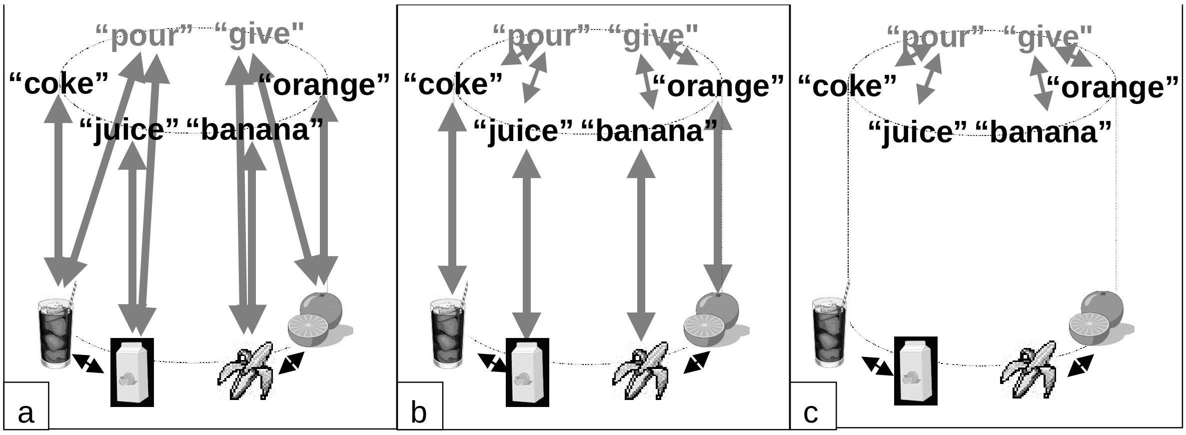

Figure 1 - Visualisation of the different types of associations in computational models of language evolution. 1a: Language based on simple indexical relationships between objects and signals, 1b: Language with grounded symbolic associations, 1c: Language with non-grounded symbolic associations. Object and word are inspired by Savage-Rumbaugh & Rumbaugh's (1978) experiments on ape language. See text for explanation.

A symbolic communication system based on words is one that simulates the evolution of grounded symbols, where the relationships between symbols are syntactic. For example, when the use of communication symbols is governed by a set of grammatical rules, these symbols can be classified into word classes, such as verbs, nouns, prepositions.

It is important to note that all these types of grounded symbolic communication systems can be simulated easily through evolutionary and artificial life methodologies. Other computational modelling techniques can simulate symbol systems but are limited in how they deal with the problem of symbol grounding. Figure 1c refers to models of the origin of symbols and words with no direct symbol grounding. At the top, only symbol-symbol associations exist. This type of system is not based on grounded symbolic associations as the links between symbols and objects are missing. Symbols are self-referencing and lack a direct grounding in the organisms' environment. These models are normally used to simulate the origin of syntax and words (e.g., [17])

The following two sections present models for the evolution of communication systems based on signals (section II) and symbols (section III). Section IV briefly discusses the use of computational models for the evolution of syntax and words.

II. Evolution of Communication Using Signals

Evolutionary computation has recently been applied to studying the emergence and auto-organisation of communication lexicons. Some models have been used for the simulation of the emergence of simple lexicons in populations of simulated organisms (e.g., [23], [4], [19]), in small communities of robots [27], or in on line Internet agents [26]. In these studies, organisms evolve shared lexicons for describing entities and relations of the environment. These models, that focus on lexicon emergence, do not make any explicit and direct reference to the role of syntax in language origin. Their aim is to model the early stages of the evolution of animal-like communication.

A recent study by Cangelosi and Parisi [4] has simulated the evolution of signals for naming foods in a population of foragers. This type of model studies communication systems based on simple signal-object associations. Organisms learn and evolve simple stimulus associations between objects in the environment and signals. Communication signals only have referential relationships with the world's entities.

A. *Evolution of Signals: Model Setup*

The simulation scenario is inspired by the use of communication signals observed in small groups of animals. Animals have evolved the use of signals to communicate about predators [5] or to communicate information about the location and quality of food [14]. The simulation scenario uses the exchange of communicative signals between pairs of organisms concerning the quality of food. More specifically, individual organisms signal each other if encountered "mushrooms" are edible or poisonous. The organisms live in an environment that contains two types of mushrooms: edible and poisonous. Organisms reproduce on the basis of their ability to eat the edible mushrooms and avoid the poisonous ones. They must first categorise an encountered mushroom as either edible or poisonous, and then they must respond by approaching and eating edible mushrooms and by going away from poisonous ones.

The simulation methodology is based on Econet models [21]. There is a population of 100 organisms. Each individual lives in an environment of 20x20 cells that contains 20 randomly distributed mushrooms (e.g., Figure

2). Ten mushrooms are edible and the other 10 are poisonous. At the beginning of its life, an individual organism is placed in a randomly selected cell with a randomly selected orientation. When an organism happens to step on a cell containing a mushroom, the organism eats the mushroom and it disappears.

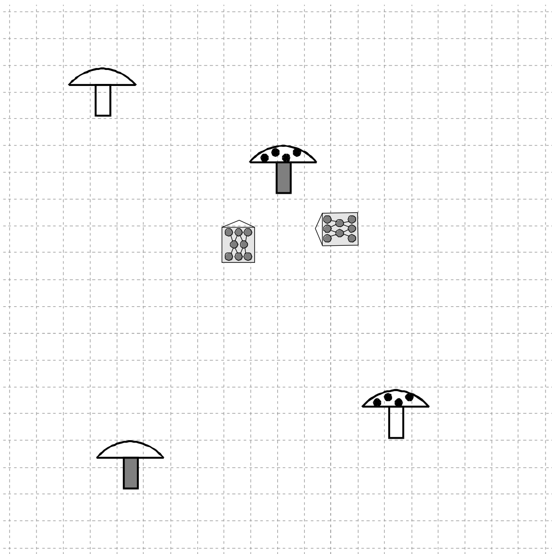

Figure 2 - Environment for the foraging task with edible and poisonous mushrooms. Each organism is controlled by a neural network.

The behaviour of each organism is controlled by a feedforward neural network with 14 input units, 5 output units, and 5 hidden units (Figure 3). One input unit encodes the location of the single nearest mushroom as the mushroom's angle measured clockwise from the organism's current facing direction. This angle is mapped in the interval from 0 to 1. Ten input units encode the mushroom's perceptual properties. The 10 edible mushrooms are encoded as 10 patterns of 10 bit, with each pattern obtained by changing a single bit, randomly chosen, in the prototypical pattern 1111100000. Similarly, the 10 poisonous mushrooms are encoded as 10 single-bit deviations from the prototype 0000011111. The 3 remaining input units (signal-encoding input units) encode one of 8 possible perceived signals: 111, 110, 100, etc. Two of the 5 output units encode a movement of the organism in the environment. The organism can either proceed one step forward (11), turn 90 degrees to the left (10) or to the right (01), or just do nothing (00). The remaining 3 output units (signal-encoding output units) encode one of 8 possible emitted signals in the same way as the signal encoding input units.

A starting population of 100 neural networks with the same architecture and randomly assigned connection weights is generated initially in each simulation. The individual's energy (fitness) is increased every time the organism eats an edible mushroom and it is decreased if the organism eats a poisonous mushroom. At the end of life, the organisms are ranked in terms of their energy and the 20 individuals with the most energy are allowed to reproduce by generating 5 offspring each. An offspring has the same connection weights of its single parent with the exception of some genetic mutations. Ten percent of the weights are modified by adding or subtracting a small random number. The process is repeated for 1000 generations. The selective reproduction of the individuals with most energy and the constant addition of variation to the genetic pool of connection weights through the genetic mutations results in an increase in average energy across the 1000 generations and the evolutionary emergence of the behaviour of approaching and eating the edible mushrooms and avoiding the poisonous ones.

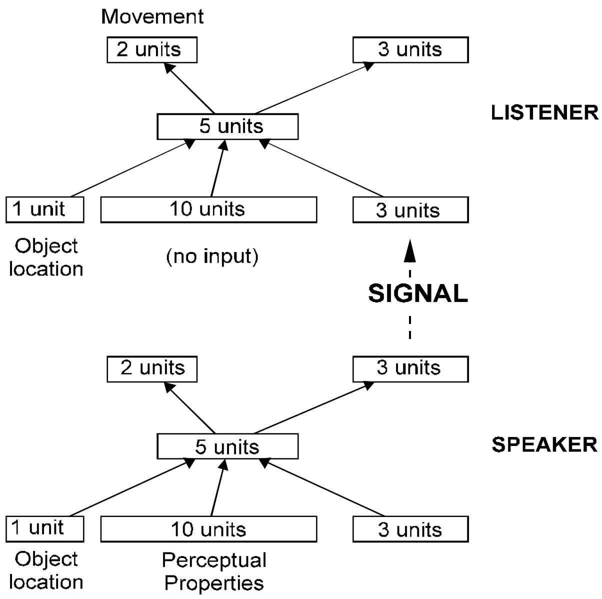

Figure 3 - Neural network architecture for the listener and speaker organisms. Note the exchange of the communication signal between the three linguistic units.

During its lifetime, an organism wanders in its environment. In each cycle, one particular mushroom happens to be the mushroom closest to the organism. If the mushroom is sufficiently near to the organism, i.e., it is located in one of the 8 cells adjacent to the organism's cell, the organism perceives both the location of the mushroom (its angle with respect to the organism's facing direction) and its perceptual properties (the pattern of 10 bits). However, if the mushroom is more distant, the organism can perceive the mushroom's location but not its perceptual properties. The 10 input units encoding the mushroom's perceptual properties all have 0 activation value.

The evolution across 1000 generations of three different populations is compared. One population has no language. When an individual encounters a mushroom that is not located in one of the 8 cells adjacent to the individual's cell, the organism can perceive the direction in which the mushroom lies but not the mushroom's perceptual properties. In a second type of population the language is provided externally by the researcher and it does not evolve. When an individual belonging to this population encounters a mushroom, the three input units of its neural network that encode perceived signals have an activation pattern of '100' if the encountered mushroom is edible and an activation pattern of '010' if it is poisonous. The signals produced by the organisms are ignored.

In the third type of population, language is not externally provided but it instead evolves autonomously. The scenario, which has been inspired by [16], is the following. Like the organisms of the other two populations, an individual can perceive the nearest mushroom's perceptual properties only if the mushroom is close enough. However, in each cycle, a second individual is selected randomly from the population and is placed next to the first individual. This way it is exposed to the same perceptual input as the first individual with the only difference that the second individual has access to the perceptual properties of the mushroom whatever the distance of the mushroom. The only task for the second individual is to label the mushroom for the first individual. The binary output of its signal-encoding output units in response to the perceptual properties of the mushroom is used as input to the signal-encoding input units of the first individual. Therefore, in this last population, when an individual encounters a mushroom, it has always access to a linguistic signal produced by a conspecific. However, in this population, unlike the previous population, the quality of the signals provided by conspecifics is not guaranteed. Whatever signal is generated by the conspecific's neural network, the signal is input to the neural network of the individual that must decide whether to approach or go away from the mushroom. Hence, the language can be useful to these organisms only if it evolves appropriately. The fitness of the listening organism only depends on its own ability to reach/avoid mushroom, and not on its linguistic ability. The speaking organism does not receive any fitness payoff.

B. *Evolution of Signals: Results*

We ran 10 replications and analysed the final fitness at generation 1000 and the structure of evolved lexicons in populations with auto-organisation of communication. The comparison of the fitness values in the three populations shows that language is a useful addition to the evolutionary adaptation of these organisms. The organisms with no language have an average energy of a little more than 150 units at the end of evolution while the two populations with language have an average energy of more than 250 units. On the other hand, the two populations with language do not differ very much from each other. Although, predictably, the population with externally provided language has a more regular increase in average energy than the population with evolved language, the two populations reach an equivalent level of energy at the end of evolution.

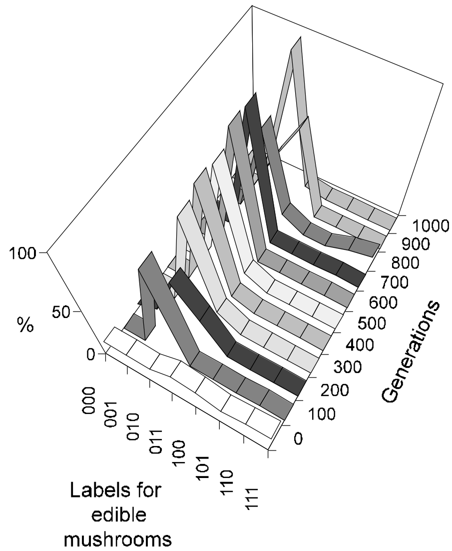

Figure 4 - Frequency distribution of the 8 possible signals for the edible mushrooms. Although there are some oscillations, the population evolves a language that tends to consistently use the pattern '010' to label edible mushrooms.

It is interesting to examine what linguistic signals evolve in the third population. Figure 4 shows the frequency distribution of the 8 possible signals for the edible mushrooms in a sample simulation. Although there are some oscillations, the population evolves a language that tends to consistently use the pattern '010' to label edible mushrooms (the pattern '110' is evolved to label poisonous mushrooms). In other populations, a similar tendency to evolve the use of only two signals was found. However, no consistent pattern was observed in evolved signals. A population can be said to possess an efficient language if (a) functionally distinct categories (in our case, edible and poisonous mushrooms) are labelled with distinct signals, (b) a single signal tends to be used to label all the instances within a category, (c) all the individuals in the population tend to use the same signal to label the same category. (Clark [7] has argued that principles similar to these govern the child's acquisition of the lexicon.) According to these criteria, the language evolved by our population appears to be rather efficient. Similar results were obtained in the other replications of the simulation although different pairs of signals emerged for the two categories of mushrooms.

This simulation shows that a population of simple artificial organisms living in a simple environment can evolve an efficient language with an informative function to help the individuals interact with their environment. Due to sensory limitations, an individual can perceive the location but not the perceptual properties of a distant mushroom. This represents a serious handicap because an individual can adopt an informed decision on whether to approach or go away from an encountered mushroom only if the mushroom is very close. In these circumstances, the population evolves a simple language in the sense that individuals tend to generate distinctive signals for edible and for poisonous mushrooms and these signals are used by other individuals to decide whether to approach or avoid a mushroom. This type of language is based on signal use, rather than on symbolic communication, because there are independent pairs of signal-object associations between the elements of the lexicon and the mushrooms existing in the organisms' environment. There are no relationships between the two signals indicating poisonous and edible mushrooms.

III. Evolution of Symbolic Communication

Some language evolution models have focused on symbolisation and the use of communication systems based on symbolic representations. The use of symbolic representations implies some form of symbol combination, since symbols have double references: one with external world's entities, another with some of the existing symbols. This type of model can be seen as a first approach to the study of the evolution of syntax. It also permits a systematic analysis of the problem of symbol acquisition. For example, comparisons can be made between symbol acquisition in animal models (e.g., chimpanzees) and computational models (e.g., artificial neural networks). An example of such models is Cangelosi's [2] work on the emergence of simple syntactic rules based on symbol combination.

A. Evolution of Symbols: Model Setup

In this study, an evolutionary computation methodology is used which is similar to that used in the previous model. The model's behavioural task is influenced directly by Savage-Rumbaugh & Rumbaugh's [24] ape language experiments. A population of 80 organisms must perform a foraging task by collecting edible mushrooms and avoiding toadstools. There are six categories of mushrooms: three edible mushrooms (big, medium, and small) and three toadstools (big, medium, and small). Once an edible mushroom is approached, organisms must identify its size category in order to gain fitness. As toadstools must be avoided, no further classification of their type is required. These foraging stimuli resemble those in Savage-Rumbaugh & Rumbaugh's [24] study on ape language. As we already mention in Section I, chimpanzees were fed through a vending machine that could "give" solid foods and "pour" drinks. Therefore animals had to learn not only the names (lexigrams) for the single foods/drinks but also a lexigram for the different types of solid foods to be "given" and another lexigram for different types of liquid to be "poured."

Organisms live in a 2D environment measuring 100 by 100 cells. At the beginning of each epoch there are 1200 randomly distributed mushrooms, 200 per category. A mushroom is characterised by a binary string of 18 perceptual features. These features will be used by the organism's neural network to identify the mushroom type and appropriate action. A set of 3 binary features always set to 1 identifies the mushroom category whilst the remaining bits are either 0 or 1. Therefore, the 200 mushrooms of each category share a common binary prototype. When this 18-bit string is input to the organism's neural network, the mushroom should be classified into one of the six categories.

Each time an organism collects an edible mushroom, its fitness is increased by one point if the correct category of mushroom is identified. Identification is based upon the level of activation of one output unit. When a toadstool is collected, the fitness decreases by one point. At the end of their lifetime, the fittest 20 organisms are selected and reproduce 4 offspring each. The organism's genotype is made up of the neural network's connection weights. Ten percent of each offspring's connection weights are mutated randomly.

A 3-layer feedforward neural network controls the behaviour of the organism (Figure 5). In the input layer, 3 units encode the location of the closest mushroom and 18 units encode their binary features. Eight input units are used for the 8 communication symbols. The network has 5 hidden units. In the output layer, 3 units control the organism's behaviour (movement and identification of mushroom category), and 8 units are used to encode the mushroom names. These symbolic output units are organised in two clusters of competitive winner-takes-all units (one cluster of 2 units, the other of 6 units). Since only one unit per cluster can be active, each mushroom will be named using two symbols.

During the first 300 generations, organisms evolve the ability to differentiate between the 6 types of mushrooms. Organisms do not communicate and do not use the 8 symbolic input and output units. Only the closest mushroom's location and the 18-bit feature string are available as input. From generation 301 to 400, organisms can communicate by using the 8 linguistic input/output units. During these generations, the new 80 organisms live together with their 20 parent organisms. Only the 80 offspring will forage and reproduce. The parents serve as speakers and teachers for naming the mushroom categories. This interaction constitutes the process of cultural transmission between the children's and the parents' generations. During each action, the parent network receives the 18-bit feature as input and produces two output symbols describing the mushroom. These symbols are used as input to the child's neural network. Ten percent of the time the child also receives the 18-bit string as input. This facilitates the evolution of good languages as the availability of the mushroom features is rare. Therefore, the parent's linguistic description becomes an important source for discriminating between mushroom categories. The child's network first uses the parent's symbols to decide what action to perform. Then it uses the parent's symbols for a naming task and an imitation task. The error backpropagation algorithm is used in both learning tasks, and the parent's two-symbol string is used as teaching input. Some noise (a random number between $\pm .5$) is added to the error between the child's output symbols and the parent's output symbols to introduce variability in the process of

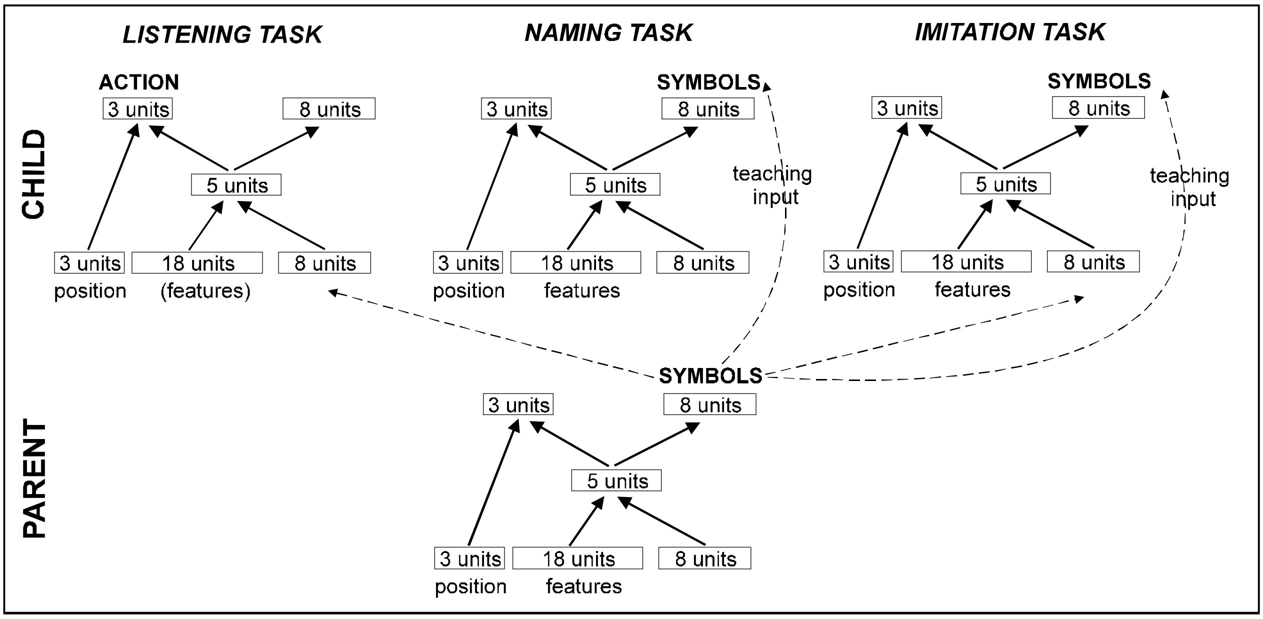

Figure 5 - Neural network architecture and learning tasks during the parent-child communicative interaction.

cultural transmission. It is important to notice that in the next generation, the new offspring will only inherit their parents' pre-learning connection weights. Any weight changes resulting from the backpropagation algorithm will not be transmitted to the next generation.

B. *Evolution of Symbols: Model Results*

The simulation from generation 1 to 300 was repeated 10 times, using different initial random populations (i.e., neural networks with different random weights). At generation 300, the foraging task fitness in 9 out of 10 populations reached an optimal level. Indeed, analysis of the behaviour of the best organisms shows that all toadstools were avoided and all edible mushrooms were approached and correctly identified.

The 9 successful populations were used in the second stage of simulation from generation 301 to 400. In this stage, communication was permitted and organisms could learn how to name mushrooms from their parents. Eighteen different simulations were performed (9 populations * 2 initial random lexicons). The percentages of the different types of evolved languages are shown in Table 1. Single-signal languages are characterised by the use of only one linguistic output cluster that differentiates between the semantic categories of mushrooms. In the first cluster, one unit is used for each category, whilst in the other cluster the same unit is active for all categories. Signal-combination languages are those characterised by the use of both clusters to communicate differences between semantic categories. However, in these languages there is no compositionality, i.e. no parallelism exists between the topology of the semantics (i.e. hierarchical structure of the mushroom categories) and that of the units in the two clusters. The verb-noun languages are a special case of signal-combination. These languages are also compositional because there is a clear parallelism between the semantic structure and the linguistic clusters. In the first cluster two different units are used. One unit is constantly associated with all mushroom categories to be avoided, the other with all mushrooms to be approached. The units in the second cluster are associated with the subcategories referring to mushroom size (big, medium, small.)

In 11 of the 18 runs, populations evolved good languages, i.e. the use of at least four signal/signal-combinations to distinguish the four emerged behavioural categories (the whole group of toadstools, and the three categories of edible mushrooms). These emergent semantic categories did not distinguish between the subcategories of toadstools. These shared languages emerged through a process of auto-organisation of the lexicon, due to the interaction between organisms and the process of cultural transmission. In the remaining 7 populations the emerged language was poor. That is, some mushroom types were labelled incorrectly due to the lack of some signal/signal-combination. Therefore the fitness remained very low since some mushrooms were incorrectly described and collected.

	Single-signal	Signal-comb.	Verb-noun
Good lang.	9%	27%	64%
Imperfect lang.	14%	29%	57%

Table 1: Percentages of the different types of languages in 18 replications (generation 400).

We are interested in identifying the different types of languages that have emerged. In particular, we want to focus on signal-combination rules that resemble known syntactical structures, such as the verb-noun rule.

Considering the populations in which good communication evolved, 10 of the 11 languages use combinations of signals. Out of these 10 populations, 3 populations use various combinations of two signals, and the remaining 7 use a verb-noun rule. In fact, in the two-unit cluster, each linguistic unit is systematically associated to only one of the high-order categories edible/poisonous. One "verb" symbol is always used for all toadstools ("avoid") and the other for all edible mushrooms ("approach"). The units in the 6-word cluster are used for distinguishing single "nouns" (mushroom size subcategories) with which the two verbs systematically couple. This is used to identify the verb-noun rule. An example of such a language at the generation 400 of a sample simulation is shown in Figure 6.

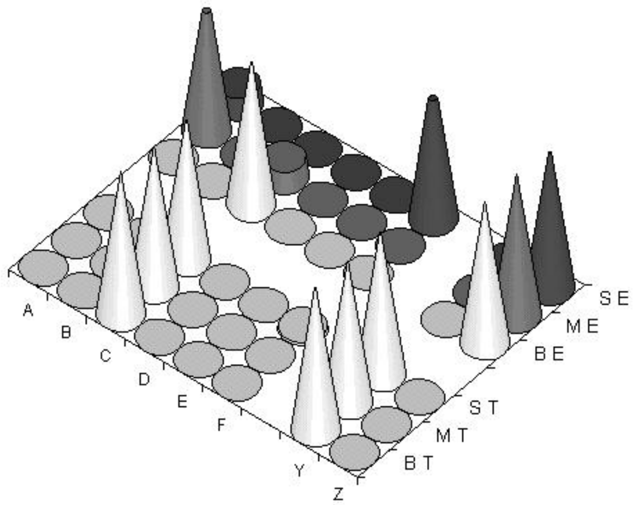

Figure 6: Final language in a sample population. Note that the language has a perfect "verb-noun" structure since all toadstools (ST, MT, BT, respectively for Small, Medium and Big Toadstool) use the output linguistic unit "Y" (= "avoid"), and all edible mushrooms (SE, ME, BE, respectively for Small Edible, Medium Edible, and Big Edible mushrooms) use the linguistic unit "Z" (= "approach"). The name of each mushroom category is indicated by the output linguistic units A-F (for differentiating mushroom subcategories by their size).

The language evolution model under discussion is supposed to be based on symbolic referencing, rather than simple object-signal associations. It is necessary to test the symbolic value of such evolved languages. Tests for symbol acquisition have been developed in animal language studies. For example, in the study by Savage-Rumbaugh & Rumbaugh [24] the test consisted of experiments where chimpanzees learned initially to associate "pour" with the name of two drinks (coke and juice) and "give" with the name of two solid foods (banana and orange). Subsequently, animals were taught new names of drinks and foods and then the researchers checked if apes where able to generalise the correct type of verb. The results showed that under certain language training conditions animals are able to learn real symbolic associations and make correct rule generalisations. Chimpanzees correctly associated the new drinks' name with the verb "pour," and the new foods' name with "give."

A similar symbol acquisition test was developed for the foraging task of the model presented in this section. Organisms are first taught to associate the verb "avoid" with the names of two toadstool categories and "approach" with the names of two edible mushroom categories. Subsequently, they are taught the name of a new toadstool and the name of a new edible mushroom. No direct feedback for the verb association is given during the learning of these new names. Finally, neural networks are tested to establish whether they learned to use the "verb-noun" rule to associate the correct verb with the new names. Ten replications of this test were executed. The results show that in 70% of populations (7 out of 10) the learned language is actually based on symbolic associations between the mushrooms' names and the two verbs [2]. It indicates that in this model most of the organisms' neural networks use a symbolic strategy when learning linguistic symbols and the syntactic rules for combining them. However, further and more systematic analysis of the acquisition of predicate-argument rules in this type of neural networks will be needed to fully assess the syntactic structures handled by these networks.

IV. COMPUTATIONAL APPROACHES TO THE EVOLUTION OF WORDS AND SYNTAX

The previous type of model simulates the emergence of simple forms of syntax, such as two-signal symbol combination and verb-noun compositionality. The evolved communication system is based on the use of symbols, since organisms are able to generalise the use of the verb-noun rule to each new entry in the lexicon. It uses symbols that are also directly grounded in the organisms' environment. However, the complexity of the evolved syntax is too simple and quite distant from the level of complexity of human languages.

Language and syntax have often been modelled using different computational techniques. Among these, neural networks have been used extensively for language simulation [6]. For example, recurrent neural network architectures have been often used for word prediction tasks. Networks were trained using complex grammatical sentences, with many levels of recursion. The study showed that neural networks were able to abstract the grammatical structure hidden in the sentences. They were also able to represent, in the layer of hidden units, the different classes of words, such as nouns and verbs.

Some recent computational studies have used evolutionary techniques to study the emergence of syntax and words. They simulate complex languages in which it is possible to identify "words", i.e. symbols that belong to specific grammatical classes, such as verbs, nouns, prepositions, etc... For example, Kirby [17] studied the emergence of compositionality and recursive grammars. It showed how compositionality can emerge without natural selection, using a simple mechanism of cultural transmission of language. In [1], populations of recurrent neural networks learn context-free grammars. They are used to understand the role of critical periods in language acquisition.

Both neural network and evolutionary computation approaches have produced interesting results. For example, neural network models have been useful in understanding the mechanisms of neural processing of language and syntax and the processes of language acquisition. Evolutionary models have supported the integration of the process of language acquisition with that of evolution in the study of the role of cultural transmission and phylogenesis in language origin.

Current models of syntax acquisition and evolution have some limitations. Among these, one particular shortcoming is the fact that these models solely focus on syntax, rather than on the evolution and acquisition of both syntax and lexicon. In fact, most of them are based on the referential system described in Figure 1c. They only simulate the top level of a language system, that of words and word-word relationships. The links between words and their semantic referents in the external environment are not simulated directly. Sometimes, an abstract system of semantic grounding is used, with the use of other "symbols" to represent semantics (e.g. when a modeller uses a list of words to denote semantic categories). Using this approach, the associations that organisms learn are mainly self-referential symbol-symbol relationships. The importance of providing a model with a mechanism for grounding symbols, and the way it can affect the type of results that the model produces, can be shown by analysis of the simulation presented in section III. The initial setup of the model provided a two-level hierarchy of six low-level mushroom categories and two high-level categories. These categories were not explicitly given in input to the organisms, i.e. the organisms did not have the list of mushroom names from which to select a specific meaning during communication. The making and selection of meanings to communicate depended on the organisms' evolving behavioural skills. In fact, the final results showed that the emerged lexicon was organised around four grounded semantic categories (the whole category of poisonous mushrooms, and the three subcategories of edible mushrooms). This shows that when the system is allowed to ground its own semantics, unexpected sets of semantic categories can emerge. If organisms had bee provided with a non-grounded, ready-made symbolic representation of meanings, results could have been different because of the bias provided by the modeller's fixed semantics. Moreover, a non-grounded approach would have significant limitations in models that wanted to study the possible existence (or not) of sequential stages of syntax complexity in the evolution of language. A researcher using a non grounded approach would have to define an *a priori* series of stages of semantic complexity upon which syntax would be biased to gradually develop. In a symbol grounded approach other autonomous factors (such as the emergence of different stages of behavioural complexity during organism's adaptation), would be free to affect (or not) the evolution of different stages of syntax complexity.

Future models of the evolution of syntax should include the essential grounding of words into the organisms' environment. This will release the researcher from the task of deciding which meanings to input to the system. The simulation approach proposed here, and other approaches such as the robotic modelling of the evolution of language [26, 27], are clear examples of how symbols can directly and autonomously ground their meaning in the organisms' environment. However, all words in a lexicon need to be directly grounded. In fact, once a basic set of grounded words has emerged, additional words can acquire grounded meanings by mechanisms of grounding transfer [3].

V. Conclusion

Computational modelling, and in particular evolutionary computation, helped to renew the interest for a scientific approach to the studies on language origin and evolution. In the last decade, many models have been developed for the simulation of the emergence of language and communication in evolving population of interacting organisms. Some models studied the emergence of lexicons and signal-object associations, others have simulated the evolution of symbols and syntax directly.

The recent language evolution theories of Harnad [13] and Deacon [9] have focused on the role of symbols and symbol acquisition for the understanding of the origins of language. The ability of cognitive systems, in particular in humans, to build symbolic representations from simple indexical associations constitutes the main basis for the further development of linguistic abilities, and in particular for the evolution of syntax. Simple indexical associations are the basis for animal communication systems through the use of signals. Symbolic representations are the basis for communication using symbols, and in particular for the use of words and syntax in human languages.

This paper has proposed the use of the distinction between signals, symbols, and words for the analysis of language evolution models. Moreover, it has stressed the need for simulating the grounding of symbols and words. Evolutionary computation permits the design of such type of language evolution models. The simulation and analysis of these models will help to understand the role of symbols and symbol acquisition in language origin.

VI. Acknowledgement

This research was partially supported by an Award for Newly Appointed Lecturers of The Nuffield Foundation (NUF-NAL # SCI/180/97/116) and a PhD studentship of the University of Genoa.

VII. REFERENCES

[1] Batali J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (eds), *Artificial Life IV*, Cambridge, MA: MIT Press, 160-171.

[2] Cangelosi A. (1999). Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life*, Berlin: Springer-Verlag, 654-663

[3] Cangelosi A., Greco A., & Harnad S. (2000). From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories. *Connection Science*, 12(2), 143-162

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

[5] Cheney D.L. & Seyfarth R.M. (1990). *How monkeys see the world: Inside the mind of another specie*. Chicago, IL: Chicago University Press.

[6] Christiansen M.H. & Chater N. (in press). Connectionist natural language processing: The state of the art. *Cognitive Science*

[7] Clark E. (1993). *The lexicon in acquisition*. Cambridge, MA: Cambridge University Press.

[8] Deacon T.W. (1996). Prefrontal cortex and symbol learning: Why a brain capable of language evolved only once. In B.M. Velichkovsky & D.M. Rumbaugh (eds), *Communicating meaning: The evolution and development of language*, Mahwah NJ: LEA Publishers, 103-138.

[9] Deacon T.W. (1997). *The Symbolic Species: The coevolution of language and human brain*, London: Penguin.

[10] Greenfield P.M. & Savage-Rumbaugh S. (1990). Grammatical combination in *Pan paniscus*: Process of learning and invention in the evolution and development of language. In S.T. Parker & K.R. Gibson (eds), *Language and intelligence in monkeys and apes*, Cambridge University Press, 540-579.

[11] Harnad S. (Ed.) (1987). *Categorical Perception: The groundwork of cognition*. New York: Cambridge University Press.

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

[13] Harnad S. (1996) The origin of words: A psychophysical hypothesis. In B.M. Velichkovsky & D.M. Rumbaugh (eds), *Communicating meaning: The evolution and development of language*, Mahwah NJ: LEA Publishers

[14] Hauser M.D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

[15] Hurford J. (1998). Review of Terrence Deacon, 1997 The Symbolic Species: The co-evolution of language and the human brain. *The Times Literary Supplement*, October 23rd, 1998, 34

[16] Hutchins E. & Hazelhurst B. (1995). How to invent a lexicon. The development of shared symbols in interaction, In N. Gilbert e R. Conte (Eds.) *Artificial societies: The computer simulation of social life*, London: UCL Press.

[17] Kirby S. (in press). Syntax without Natural Selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert-Kennedy & J. Hurford (Eds.) *Approaches to the Evolution of Language*, Cambridge University Press.

[18] Kirby S. (1999). Syntax out of learning: The cultural evolution of structured communication in a population of induction algorithms. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life*, Berlin: Springer-Verlag, 694-703.

[19] Oliphant M. & Batali J. (1997). Learning and the emergence of coordinated communication. *Centre for Research in Language Newsletter*, 11(1).

[20] Parisi D. (1997). An Artificial Life approach to language. *Mind and Language*, 59, 121-146.

[21] Parisi D., Cecconi F. & Nolfi S. (1990). ECONETS: Neural networks that learn in an environment. *Network*, 1: 149-168.

[22] Peirce C.S. (1978). *Collected papers. Vol. II: Element of logic*, C. Hartshorne & P. Weiss (Eds.), Cambridge, MA: Belknap.

[23] Saunders, G.M., & Pollack, J.B. (1996). The evolution of communication schemes over continuous channels. Proceedings of the SAB'96 Conference on the Simulation of Adaptive Behavior. Cambridge, MA: MIT Press.

[24] Savage-Rumbaugh S. & Rumbaugh D.M. (1978). Symbolization, language, and chimpanzees: A theoretical reevaluation on initial language acquisition processes in four young Pan Troglodytes. *Brain and Language*, 6: 265-300.

[25] Steels L. (1997) The synthetic modeling of language origins. *Evolution of Communication*, 1(1): 1-37.

[26] Steels L. & Kaplan F. (1999). Collective learning and semiotic dynamics. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life*, Berlin: Springer-Verlag, 679-688.

[27] 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, 474-482