SIMULATING THE EVOLUTIONARY EMERGENCE OF LANGUAGE: A RESEARCH AGENDA

DOMENICO PARISI

Institute of Cognitive Sciences and Technologies
National Research Council,
via S. Martino della Battaglia 44, 00185 Rome
domenico.parisi@istc.cnr.it

1. Why simulations can be useful

If one is interested in studying the evolutionary emergence of human language, one is confronted with two formidable but well recognized problems. First, compared with animal communication systems, human language is a much more complex system for communicating with others, and therefore its evolutionary emergence must necessarily be correspondingly more complex. Second, we don't have much direct empirical evidence concerning how and when human language has emerged in the course of the evolution of the species, and therefore we are restricted to hypotheses and theories that must remain to a large extent speculative. With respect to these two problems there is not much that we can do. However, there is a third, less well recognized problem, and with respect to this problem there is something that we can, and should, do. Hypotheses and theories about the evolutionary emergence of language abound but it is notoriously very difficult to reach a consensus on which one to accept or reject. This is not only the inevitable consequence of a restricted empirical basis. Most hypotheses and theories about language's evolutionary emergence are only verbally formulated and, therefore, they tend to be insufficiently defined, precise, and articulated, and it is hard to derive from them uncontroversial empirical predictions. Hence, not only we do not have much direct empirical evidence on language's evolutionary emergence but we do not address the little, mostly indirect, evidence that we do have with well defined predictions uncontroversially derived from well defined hypotheses and theories.

This is where computer simulations can be of help. Computer simulations are hypotheses and theories which are expressed not in words or mathematical symbols, as is traditional in science, but as computer programs. When the program runs in the computer the results of the simulation are the empirical

predictions that are derived from the hypothesis or theory which is incorporated in the program. Why can this be of help? If a theory is expressed in the form of a computer program, i.e., as a simulation, the theory must necessarily be well defined, precise, and articulated because, otherwise, the program cannot be written or it will not run in the computer. Furthermore, since a simulation's results are the empirical predictions derived from the theory incorporated in the program, a theory expressed as a simulation necessarily (mechanically) and uncontroversially generates many detailed empirical predictions that can be confronted with the empirical data. Furthermore, simulations can be tools for thinking in that they can suggest new hypotheses and new types of empirical evidence.

Researchers in language evolution who do not use simulations but use the more traditional and well established methods of linguistics, psychology, anthropology, ethology, archaeology, and palaeontology, tend to be suspicious of computer simulations. Simulations simplify with respect to reality and to actual empirical phenomena, and these researchers are too well aware of the extreme complexity and diversity of empirical reality to become interested in such simplified models. But all scientific theories simplify with respect to reality, and they make us better understand reality only because they simplify, allowing us to capture the mechanisms and processes that underlie the variety and complexity of empirical phenomena and explain them. Since simulations are theories, they should simplify. The real problem is not that they simplify but that they should make the appropriate simplifications, including the aspects of reality which are critical in order to explain the particular phenomena of interest and leaving other aspects out. Therefore, computer simulations should not be criticized a priori and in general but they should be examined and evaluated case by case.

2. What types of simulations?

If one wants to reproduce in an artificial system - a computer simulation or a community of physical robots - the evolutionary emergence of language, one has to make a number of choices. A simulation of the emergence of language must necessarily assume the existence of a set of agents that represent the human population within which language has evolved. One first set of choices concern the nature of the agents. I believe that, to maintain a continuity between the biological and the cognitive sciences, these agents should possess three properties: (1) their behaviour should be controlled by a neural network, i.e., a

control system that simulates the physical structure and the physical way of functioning of the nervous system (Rumelhart and McClelland, 1986); (2) the agents should be embodied, i.e., the neural network should be part of a simulated physical body, with a given size, shape, and given sensory and motor organs (Nolfi and Floreano, 2000; Pfeifer and Scheier, 2001); (3) they should live in, and interact with, a simulated physical environment that includes conspecifics, other animals, objects, and possibly technological artefacts (Parisi, Cecconi, and Nolfi, 1990).

A second set of choices concern the nature of the process, or processes, of acquisition through which language emerges. If neural networks are used as the control system for the agents, it becomes a necessity to simulate a process of acquisition of whatever ability those agents will eventually possess because neural networks cannot be directly programmed by the researcher. The researcher defines the scenario within which some ability is acquired but then the simulation must show if, how, and in what conditions, the ability is actually acquired. Therefore, one has to decide which algorithm to use to "train" the networks, i.e., to go from networks that do not possess language to networks that do possess language. The problem is that underlying language and its evolutionary emergence there are many distinct acquisition processes and one must be able to simulate all of them, and how they interact with each other.

The first process is biological evolution. Human language is learned during the first years of life but it is learned only because there are inherited predispositions that make the learning possible, as shown by the fact that non-human animals do not acquire a language even if they are exposed to human language. Hence, one thing that one must be able to simulate is the process of biological evolutionary emergence of these inherited predispositions for language learning, which can be either specific for learning language or can be more general predispositions which are species-specific for Homo sapiens but not specific for language. To simulate the evolutionary emergence of biologically inherited abilities or predispositions a procedure called genetic algorithm is used (Holland, 1975). A population of agents with individually different genotypes live in an environment and reproduce by generating offspring that inherit their genotypes. Genotypes encode specifications for the agents' neural networks. Reproduction is selective, with some individuals having more offspring than other individuals, and the offspring's genotypes are somewhat different from the genotypes of their parents because of random genetic mutations and sexual recombination. The two processes of selective reproduction and of constant addition of new variability due to mutations and

sexual recombination cause evolutionary changes in inherited genotypes and the emergence of initially absent abilities that make survival and reproduction more likely in the given environment.

The second process of acquisition which is involved in language is learning. A human being is not born with language but he or she learns the particular language which is spoken in his/her environment, although this is only possible because of the biologically inherited predispositions to learn any human language that I have already referred to. Language learning actually is language development, where by development I mean a process of acquisition which needs, and is modulated by, environmental input but goes through stages that are specified in the inherited genotype. This implies that the genotype of the agents must encode a developmental program, not just a set of already existing abilities or a set of predispositions to learn. A developmental program is a genetically inherited schedule for acquiring various behaviours in a sequence of temporal stages. What must evolve, therefore, is not only the content of these stages but also the sequence of stages itself.

Finally, human language involves a third type of acquisition process: cultural or, more precisely, linguistic evolution (language change). Language is learned from others, i.e., culturally. Hence, one has to simulate agents that not only learn in general terms but learn by imitating others. Learning from others creates a second form of evolution, cultural evolution. Behaviours are transmitted (imitated) from one generation to the next, and they are transmitted selectively, with some behaviours more imitated than other behaviours. Furthermore, behaviours can change either because of random errors in transmission, analogous to random mutations, because imitators recombine in new ways different aspects of the behaviour of different models, analogous to sexual recombination, and because of inventions and internal re-organization of the behavioural repertoire. As in biological evolution, this results in cultural evolution, with the emergence of new forms of culturally transmitted behaviour, in our case, new forms of language (language change).

Furthermore, groups of interacting agents tend to learn from each other and therefore to develop similar behaviours, including linguistic behaviours. Shared linguistic behaviours constitute historical languages. If a group of agents, for some reason, splits into separate sub-groups with little interactions between the sub-groups, distinct dialects or even different languages can emerge from a single initially shared language because culturally transmitted behaviours change all the time and the changes that occur in one sub-group tend

to diverge from those in other sub-groups. This is similar to biological phenomena of genetic divergence and the emergence of new species.

Some simulations based on the above assumptions and that address some of the phenomena of language evolution that I have briefly described have already been realized. However, most of the work in this area is still to be done, so what I am talking about is mainly a research agenda. (For a collection of articles on simulating the emergence of language, see Briscoe, 2002; Cangelosi and Parisi, 2002.)

3. Genetically inherited predispositions for language learning

Consider the species-specific biologically inherited pre-dispositions for language learning. As I have said, these predispositions can be either general or specific for language learning. Among the predispositions that are general and not specific for language learning there might be a species-specific tendency to learn by imitating others. Learning by imitating others may presuppose an ability to learn to predict the perceived effects of one's own actions (movements). This ability to predict the effects of one's actions appears to underlie learning by imitating any kind of behaviour, not only linguistic behaviour. Another algorithm, the backpropagation procedure (Rumelhart and McClelland, 1986), can be used to simulate both learning to predict the effects of one's own actions and learning to imitate the actions of others (Jordan and Rumelhart, 1992). In the first case, the learner compares the predicted effects of its own actions with their actual effects and, using the results of this comparison, adjusts the connection weights of its neural network in such a way that, in a series of learning experiences, any initial discrepancy between predicted and actual effects gradually disappears. In the second case the learner compares the predicted effects of its own actions with the perceived effects of the actions of another individual and adjusts its connection weights in such a way that it becomes eventually able to produce the same effects that are produced by the actions of others and, therefore, presumably, the same behaviours.

These two learning processes may underlie the successive stages of prelinguistic and then linguistic behaviour in the child from birth to 1 year and on. Stage 1: the production of all kinds of sounds in the very first months of life, when the child is learning to predict the acoustic consequences of his or her phono-articulatory movements. Stage 2: the emergence of babbling at 4-6 months, when the child is learning to imitate his or her own sounds (self-

imitation). Stage 3: the tendency of the sounds produced by the child in the second semester of life to resemble the sounds of the particular language spoken in the child's environment, when the child is learning to imitate the sounds produced by others. Stage 4: the first emergence of true language at around 1 year of age, when the child is learning to produce the same sounds produced by another individual in response to the same objects and actions. The separate stages of this developmental process have already been simulated (Floreano and Parisi, 1992) but the challenge is to evolve a developmental genotype that encodes the entire developmental sequence.

Stage 4 explains the typically referential nature of human language in contrast to the mostly non-referential nature of animal communication. Linguistic signals acquire their referential meaning because particular sounds are systematically paired, in the learner's experience, with particular objects and actions. Signals that are paired with objects, independently of the specific action which is done with respect to the object, become nouns, while signals paired with actions, independently of the specific object on which the action is done, become verbs. One can also simulate the successive emergence of other parts of speech (Parisi, Cangelosi and Falcetta) and the emergence of complex signals made up of simpler signals (Cangelosi, 1999).

Learning to predict the effects of one's actions and learning by imitating others can be based on species-specific biologically inherited predispositions of Homo sapiens that underlie language learning but are not limited to language. This is consistent with the idea that these same predispositions may underlie other typically human traits such as constructing and using all sorts of technological artefacts by predicting the effects of one's actions on the artefact and the effects of the artefact on the environment. Language learning, however, can also be based on biologically inherited predispositions that are specific for language such as a particular sophistication of the sensory-motor sub-network, or module, that maps heard sounds into phono-articulatory movements or the ability to parse and construct linguistic signals that are combinations of smaller signals (syntax). However, in evaluating the se claims of linguistic specificity one must consider that deaf children are able to learn non-acoustic sign languages (which may have preceded the evolutionary emergence of acoustic languages) and that linguistic syntax might emerge from already existing and more general abilities to analyze and to generate complex actions as combinations of simpler actions. Simulations can be of help here by allowing us to test alternative hypotheses. For example, would a population of agents that has developed a language in the acoustic/phono-articulatory mode

be able to easily switch to a language in the visual/motor mode? Or, can we demonstrate a general ability to parse and construct complex actions as combinations of simpler actions even in simulated organisms with no language?

4.   Adaptive functions of language

Other interesting research issues concerning language evolution that can be addressed with simulations of the type I have described are the adaptive role of linguistic behaviour and the function of language as a tool for communicating with oneself (thinking), and not only with others, which appears to be another property that distinguishes human language from animal communication systems.

The present framework assumes that language has evolved because it was adaptive but one can create and study different evolutionary scenarios in which one contrasts various adaptive roles for language. For example, language is a social behaviour and it involves at least two agents, the speaker and the hearer. Therefore, one can ask: Is language adaptive for the speaker or for the hearer? Imagine an agent which is looking for food but cannot recognize if an encountered but distant mushroom is edible or poisonous. The only option for the agent is to get near to the mushroom and eat the mushroom if it is edible and go away if it is poisonous. In these circumstances the agent would benefit from the linguistic behaviour of another agent which is nearer to the mushroom and tells the first agent if the mushroom is edible or poisonous, saving the first agent's time and energy if the mushroom is poisonous. However the behaviour of the speaker is advantageous for the hearer but is not advantageous for the speaker or can even be disadvantageous if the two individuals compete for survival. So why should speakers evolve in this scenario? And if there are no speakers, there is no language. In fact, if one makes a simulation the results of the simulation show that in this scenario language will not evolve (Mirolli and Parisi, 2005). The behaviour of the speaker is altruistic in that it increases the survival and reproductive chances of the hearer but decreases its own. Therefore, the genes underlying speaking behaviour tend to disappear from the population and language does not emerge. Language emerges only if the speaker and the hearer share the same genes, as predicted by kin selection theory. By benefiting the hearer, the altruistic behaviour of the speaker increases the survival and reproductive chances of an individual which, by sharing the same genes of the speaker, also is a good speaker. Therefore, the genes underlying speaking behaviour remain and diffuse in the population.

The results of this simulation may suggest the hypothesis that language has first emerged in small groups of kin related individuals. However, this implies that different kin groups will speak different languages whereas language is particularly useful when it can be used more widely. How did a language which is spoken and understood by larger groups of non-kin individuals emerge? One possibility is that language has emerged in a single, small group of kin related individuals, and then it has been culturally inherited in a progressively larger group of more distantly related individuals that were descendants of the original group. Another possibility, of course, is that language was used in situations in which it was useful to both speakers and hearers, such as the speaker asking for something from the hearer and the hearer being interested in knowing what is asked.

Still another possibility is that language was used from its very beginning not only for communicating with others but also for communicating with oneself. In fact, language will emerge if we modify the simulation scenario that I have described and require that the hearer, when it hears a signal from the speaker specifying that the mushroom is edible, has to repeat the signal to itself in order to keep it in memory while approaching the edible mushroom. This is using language for communicating with oneself, not with others. The results of the simulation show that if language is used for communicating with oneself, language emerges even if the speaker has not the same genes of the hearer (Mirolli and Parisi, 2005). The reason is that the hearer must also be a good speaker in order to speak appropriately with itself and be able to increase its reproductive chances. This simulation may suggest the hypothesis that using language to communicate with oneself, i.e., to think, may have been an adaptive pressure for the emergence of language from its earliest evolutionary stages, instead of supposing that the use of language for thinking is a recent discovery which requires an already well developed language.

5.   Language change and differentiation

Finally, to demonstrate the scope of simulations I want to briefly mention another, very different, simulation that addresses language change and the historical process of linguistic differentiation which creates patterns of similarities and differences among genetically (historically) related languages. In this simulation what is simulated is the process of diffusion of farming in Europe which originated in Anatolia nine thousand years ago and reached the whole of Europe in 3-4 millennia. In the simulation the entire European

territory is divided into relatively small cells with properties that specify the suitability of each cell for farming, with cells more favourable for farming and cells less favourable (sea, mountains, deserts). Farmers (or, more probably, the technology of farming, at least at increasing distances from the point of origin in Anatolia) follow particular paths in their diffusion in Europe which depend on the appropriateness of the particular territories for farming, with some paths dividing into diverging paths at specific points. It turns out that the resulting tree of paths followed by the simulated farmers in their diffusion in Europe has some similarities with the tree of genetic relatedness of the languages spoken in Europe (Parisi, Antinucci, Cecconi and Natale, in press). This gives some support to theories on the origin of Indo-European languages in Anatolia nine thousand years ago with respect to other theories that hypothesize a more recent origin in the region to the north of the Black Sea and Caspian Sea.

Acknowledgements

References

Briscoe, E. J. (2002). *Linguistic evolution through language acquisition: Formal and computational models.* Cambridge: Cambridge University Press.

Cangelosi A. (1999). Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In D. Floreano, J. Nicoud, and F. Mondada (eds.), *Advances in Artificial Life*, New York, Springer, 1999, pp. 654-663.

Cangelosi, A., & Parisi, D. (eds.) (2002). *Simulating the Evolution of Language.* New York: Springer.

Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, (MIT Press, 1992).

Jordan, M.I., & Rumelhart, D.E. (1992). Forward models: supervised learning with a distal teacher. *Cognitive Science*, 16, 307-354.

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

Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics.* Cambridge, MA: MIT Press

Parisi, D., Antinucci, F., Cecconi, F., & Natale, F. (in press). Simulating the expansion of farming and the differentiation of European languages. In B. Laks, and D. Simeoni (eds.) *Origins and Evolution of Language*. New York: Oxford University Press

Parisi, D., Cangelosi, A., & Falcetta, I. (2002). Verbs, nouns, and simulated language games. *Italian Journal of Linguistics*, 14, 99-114.

Parisi, D., Cecconi, F., & Nolfi, S. (1990). Econets: neural networks that learn in an environment. *Network*, 1, 149-168.

Parisi, D., & Floreano, D. (1992). Prediction and imitation of linguistic sounds by neural networks. In A. Paoloni (ed.), *Proceedings of the 1st Workshop on Neural Networks and Speech Processing*. Rome, Fondazione Bordoni, 50-61.

Pfeifer, R., & Scheier, C. (2001). *Understanding intelligence*. Cambridge, MA: MIT Press.

Rumelhart. D.E. and McClelland, J.L. *Parallel Distributed Processing. Explorations in the Microstructure of Cognition.* Cambridge, Mass., MIT Press, 1986.