Baronchelli, A., Felici, M., Caglioti, E., Loreto, V. & Steels, L. (2005). Self-organizing communication in language games. Paper presented at the First European Conference on Complex Systems ECCS'05, 14-18 Nov 2005, Paris, France.

Self-Organizing Communication In Language Games

A. Baronchelli[†], M. Felici[†], E. Caglioti[‡], V. Loreto[†] and L. Steels[♯]

† Physics Dept. and INFM-SMC *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
‡ Mathematics Dept. *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
♯ SONY CLS, 6 rue Amiot, 75005 Paris, FRANCE

Keywords:
Language games, Communication, Self-organization, Negotiation dynamics, Shared conventions, Social bookmarkings.

Abstract:
From the point of view of semiotic dynamics language is an evolving complex dynamical system. In this perspective, unrevealing the mechanisms that allow for the birth of shared conventions is a major issue. Here we describe a very simple model in which agents negotiate conventions and reach a global agreement without any intervention from the outside. In particular we focus on the possibility of predicting on which of the several competing conventions the agreement is reached. We find from simulations that early created conventions are favored in the competition process and this advantage can be quantified. Beyond the specific results presented here, we think that this paper provides an example of a new way of investigating language features where simple models allow for the investigation of precise problems and, possibly, for analytical approaches.

Corresponding Author:
Andrea Baronchelli
Physics Dept. and INFM-SMC *La Sapienza* University, P.le A. Moro 2, 00185 Rome, ITALY
Tel: +39 06 4991 3450, e-mail: Andrea.Baronchelli@roma1.infn.it

Recently the study of the self organization and evolution of language and meaning has led to the idea that language can be seen as a complex dynamical system [1]. In this perspective, the theoretical tools developed in complex systems science acquire a central role for the study of the self generating structures of language systems.

Language is a semiotic system that relates signs (words, grammatical constructions) to the world through the intermediary of conceptualization. Such a semiotic system can then be used for communication in the sense, for example, that signs can be used to draw attention to objects of the physical world. Once relaxed the hypothesis of staticity of a language, a natural and very interesting question is how new conventions, developed in local interactions among few individuals, can become stable in the whole population. Said in different words, the problem is to determine the behavior of a system of many components, endowed with an individual structure, that interact with each others. The analogy with problems of n-body systems of interacting particles, well known in physics, is then obvious and stimulating.

Here we discuss an extremely simple multi-agent model, the Naming Game [Baronchelli A, Felici M, Caglioti E, Loreto V and Steels L: unpublished 2005], in which agents play pairwise games in order to negotiate conventions, i.e. associations between forms and meanings. In this way the attention is focused only on cultural spreading, without resorting to any evolutionary issues [2, 3].

In particular we deal with a population of N agents whose aim is to agree on the name to give to a certain object (the presence of a single object corresponds to the denial of homonymy). Each agent is characterized by its *inventory*, i.e. a list of words that can be dynamically updated. Agents have empty inventories at time $t = 0$ and at each time step ($t = 1, 2, ..$) two players are picked at random to play an interaction: one of them plays as *speaker* and the other as *hearer*. Their interaction obeys the following rules:

- The speaker randomly extracts a word from its inventory, or, if its inventory is empty, invents a new word.
- If the hearer has the word selected by the speaker in its inventory, the interaction is a <u>success</u> and both players maintain in their inventories only the winning word, deleting all the others.
- If the hearer does not have the word selected by the speaker in its inventory, the interaction is a <u>failure</u> and the hearer updates its inventory adding the new word.

The model is able to describe the emergence of a communication system where a unique form (or name) is assigned by all the individuals to the same meaning (here the object). This is clearly showed in Figure 1, where the evolution of total number of words present in the system is shown. At

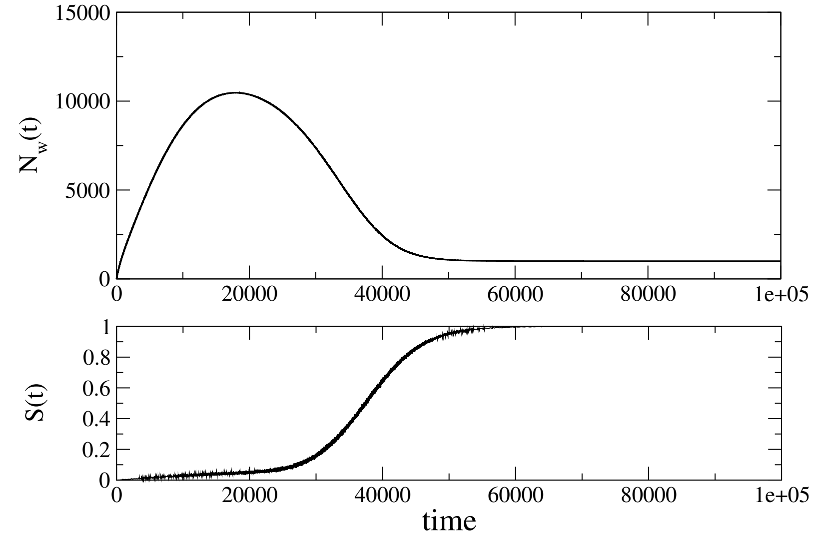

Figure 1: **System evolution:** We show here the evolution in time of the total number of words $N_w(t)$ and the success rate $S(t)$ for a population of $N = 1000$ agents. As it is clear the population is able to build up an effective communication system in which $N_w = N$, i.e. each agent has only one word. Moreover the success rate in this state is equal to 1 indicating that all interactions are successful.

the beginning the curve grows due to the invention process. It follows a further period of growth in which agents perform unsuccessful interactions. While increasing in size, however, inventories correlate with each other, so that at a certain point the number of words, having reached a maximum, starts decreasing. This is due to successful communications between agents, that increase in their number till the system reaches a final convergence state in which all the agents have the same unique word, thus being able to perform only successful interactions. In Figure 1 it is reported the success rate curve too. Obtained averaging over several different runs of the process, this curves shows the probability that an interaction at a given time is successful. As just discussed, the system evolves to a situation of convergence through an intermediate state in which words are eliminated while the success probability increases. It is important noting that the developed communication system is not only effective (every agent understands all the others) but also efficient (no memory or computational resources are wasted for comprehension in the final state).

Given that the population does, at the end, agree on a convention, an interesting question is whether we can predict on *which* convention the agreement takes place. In fact, the same process in which agents negotiate with each other can be seen as a process in which different conventions compete to survive. According to the rules of our model all different words are equiv-

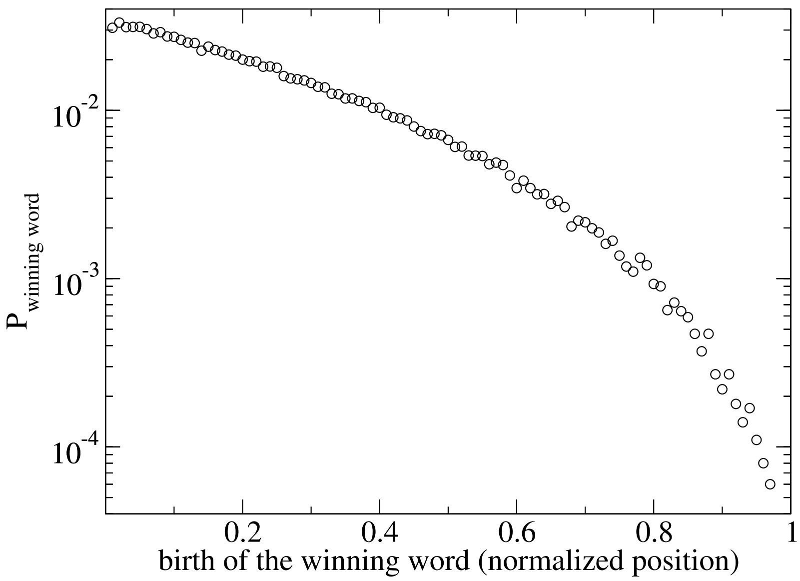

Figure 2: **Winning word probability distribution (order):** We investigate how the probability for a word to dominate is affected by its invention order. We label the first invented word with 1, the second invented word with 2 and so on till the last invented one. Then labels are normalized with the tag of the last invented word. Early invented words are more likely to become dominant since they have more chances to propagate at the beginning. Data are obtained for a population of $N = 1000$ agents by averaging results from 10^5 runs.

alent. The only feature that could differentiate them is their invention moment. In Figure 2 we investigate the role of the invention *order*. It emerges that the probability for a word to become the one the agents will agree upon strongly depends on the moment of its creation, indeed. To investigate the role of creation order we label the first invented word with cardinal number 1, the second invented word with 2, and so on till the last invented word. Then we divide the label of each word by the label of the last invented one so that results from different simulation runs become comparable. Performing several runs and taking memory of the label of the winning words, we are then able to estimate the probability that the winning word has a given label. From Figure 2, it clearly emerges that early invented words have bigger chances to dominate. This can be explained considering that the sooner a word is invented the higher are its opportunities to propagate. Moreover at the beginning of the game agents have small inventories, and this reduces the probability that an existing word is not played by a speaker who holds it. Finally, in Figure 3 we analyze the role of invention *time*. Obviously, the advantage of early invented words is found again, but interestingly in the domain of time this advantage can be quantified. Indeed, in Figure 3 it is

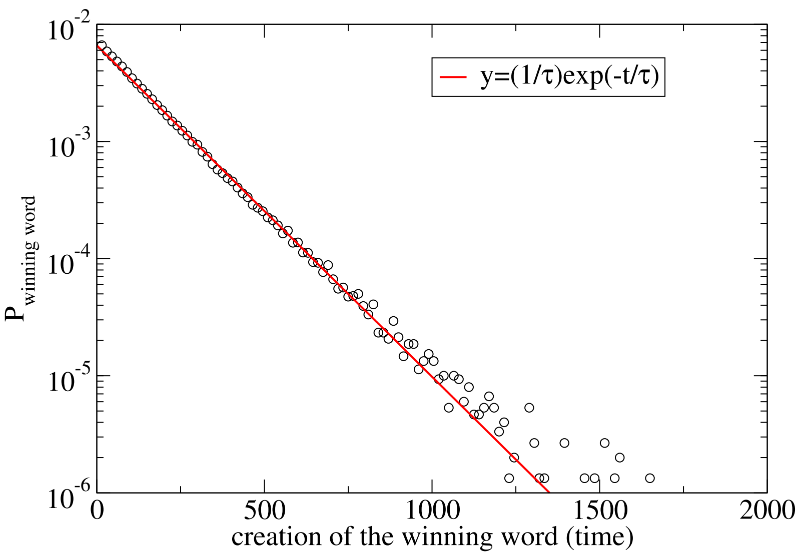

Figure 3: **Winning word probability distribution (time):** We plot here the probability that a word created at a given time becomes the dominating one. Data, relative to a population of $N = 1000$ agents, are well fitted by an exponential distribution (here $\tau \simeq 150$).

shown that data from simulations are well fitted by an exponential distribution. In Figure 2 this behavior was not found due to the presence of a cutoff (the last invented word) that is absent here.

In conclusion, we have seen how the view of language as an evolving systems casts several interesting issues that can be addressed in the framework of complex science methods. A profitable approach consists, in our opinion, in the definition and study of simple models that allow for precise investigations of specific problems. Then, in this perspective, the definition of a reasonable model is in itself an important goal. Here we have presented an interesting model in which agents negotiate conventions according to elementary rules and manage to reach a global agreement. We have also shown that, even though all words (or conventions) are equivalent in the model, the moment of their invention affects (exponentially) their probability of becoming dominant.

1 Acknowledgements

The authors thank L. Steels and C. Cattuto for many stimulating discussions. A. Baronchelli and V. L. have been partly supported by the ECAgents project funded by the Future and Emerging Technologies program (IST-FET) of the European Commission under the EU RD contract IST-1940. The information provided is the sole responsibility of the authors and does not reflect the Commission's opinion. The Commission is not responsible for any use that may be made of data appearing in this publication. A. Barrat and L.D. are partially supported by the EU under contract 001907 (DELIS).

References

[1] Steels L: Language as a complex adaptive system. In Shoenauer M (ed): Proceedings of PPNS VI. Lecture Notes in Computer Science, Berlin, Germany, September 2000. Springer-Verlag.

[2] Hurford J: Biological evolution of the Saussurean sign as a component of the language acquisition device, Lingua 1989; 77:187-222.

[3] Nowak MA, Plotkin JB and Krakauer D: The evolutionary language game, J. Theor. Biol. 1999; 200:147-162.