Chapter 15
Mathematical Modeling of Language Games

Vittorio Loreto, Andrea Baronchelli,
and Andrea Puglisi

Abstract In this chapter we explore several language games of increasing complexity. We first consider the so-called *Naming Game*, possibly the simplest example of the complex processes leading progressively to the establishment of human-like languages. In this framework, a globally shared vocabulary emerges as a result of local adjustments of individual word-meaning association. The emergence of a common vocabulary only represents a first stage while it is interesting to investigate the emergence of higher forms of agreement, e.g., compositionality, categories, syntactic or grammatical structures. As an example in this direction we consider the so-called *Category Game*. Here one focuses on the process by which a population of individuals manages to categorize a single perceptually continuous channel. The problem of the emergence of a discrete shared set of categories out of a continuous perceptual channel is a notoriously difficult problem relevant for color categorization, vowels formation, etc. The central result here is the emergence of a hierarchical category structure made of two distinct levels: a basic layer, responsible for fine discrimination of the environment, and a shared linguistic layer that groups together perceptions to guarantee communicative success.

1 Introduction

Language emergence and evolution has recently gained growing attention through multi-agent models and mathematical frameworks to study their behavior. Lan-

---

V. Loreto (✉) · A. Puglisi
Dipartimento di Fisica, “Sapienza” Università di Roma, Piazzale Aldo Moro 5,
00185 Rome, Italy
e-mail: loreto@roma1.infn.it

V. Loreto
ISI Foundation, Viale Settimio Severo 65, 10133 Villa Gualino, Turin, Italy

A. Baronchelli
Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya,
Campus Nord B4, 08034 Barcelona, Spain
e-mail: andrea.baronchelli@upc.edu

A. Puglisi
CNR (INFM-SMC and ISC), Piazzale Aldo Moro 5, 00185 Rome, Italy
e-mail: andrea.puglisi@roma1.infn.it

S. Nolfi, M. Mirolli (eds.), *Evolution of Communication and Language in Embodied Agents*, 263
DOI 10.1007/978-3-642-01250-1_15, © Springer-Verlag Berlin Heidelberg 2010

guage is based on a set of cultural conventions socially shared by a group. But how are these conventions established without a central coordinator and without telepathy? The problem has been addressed by several disciplines, but it is only in the last decade that there has been a growing effort to tackle it scientifically using multi-agent models and mathematical approaches (cf. Nowak and Krakauer 1999; Steels 2005; Castellano et al. 2009 for a review). Initially these models focused on the emergence of a shared vocabulary, but increasingly attempts are made to tackle grammar (Steels 1998; Nowak and Krakauer 1999; Nowak et al. 2000; Brighton et al. 2005; Puglisi et al. 2008).

In this chapter we explore several modalities of interactions among the individuals of a population corresponding to language games of increasing complexity. We first consider the so-called *Naming Game* (Sect. 2), which possibly represents the simplest example of the complex processes leading progressively to the establishment of human-like languages. It was expressly conceived to explore the role of self-organization in the evolution of language (Steels 1995, 1996) and it has acquired, since then, a paradigmatic role.

The next step we consider is the so-called *Category Game* (Sect. 3). In this case one focuses on the process by which a population of individuals manages to categorize a single perceptually continuous channel, each stimulus being represented as a real-valued number ranging in the interval [0, 1]. The problem of the emergence of a discrete shared set of categories out of a continuous perceptual channel is a notoriously difficult problem relevant for color categorization, vowels formation, etc. The central result here is the emergence of a hierarchical category structure made of two distinct levels: a basic layer, responsible for fine discrimination of the environment, and a shared linguistic layer that groups together perceptions to guarantee communicative success. Remarkably, the number of linguistic categories turns out to be finite and small, as observed in natural languages.

2 The Naming Game

The *Naming Game* (NG) possibly represents the simplest example of the complex processes leading progressively to the establishment of complex human-like languages. It was expressly conceived to explore the role of self-organization in the evolution of language (Steels 1995, 1996) and it has acquired, since then, a paradigmatic role in the whole field of semiotic dynamics. The original paper (Steels 1995), focuses mainly on the formation of vocabularies, i.e., a set of mappings between words and meanings (for instance, physical objects). In this context, each agent develops its own vocabulary in a random private fashion. But agents are forced to align their vocabularies, through successive conversations, in order to obtain the benefit of cooperating through communication. Thus, a globally shared vocabulary emerges, or should emerge, as a result of local adjustments of individual word-meaning association. The communication evolves through successive conversations, i.e., events that involve a certain number of agents (two, in practical implementations) and meanings. It is worth remarking that here conversations are particular cases of language games, which, as already pointed out in Wittgenstein (1953a, 1953b), can be used to describe linguistic behavior, even if they can include also non-linguistic behavior, such as pointing.

This original seminal idea triggered a series of contributions along the same lines and many variants have been proposed over the years. It is particularly interesting to mention the work proposed in Ke et al. (2002), that focuses on an imitation model which simulates how a common vocabulary is formed by agents imitating each other, either using a mere random strategy or a strategy in which imitation follows the majority (which implies non-local information for the agents). A further contribution of this paper is the introduction of an interaction model which uses a probabilistic representation of the vocabulary. The probabilistic scheme is formally similar to the framework of evolutionary game theory (Hurford 1989; Oliphant and Batali 1996; Oliphant 1997; Nowak et al. 1999a; Nowak 2006), since to each agent a *production* and a *comprehension* matrices are associated. Differently from the approach of Evolutionary Language Game (Nowak et al. 1999a), here the matrices are dynamically transformed according to the social learning process and the cultural transmission rule. A similar approach has been proposed in Lenaerts et al. (2005).

In the next section we shall present a *minimal* version of the NG which results in a drastic simplification of the model definition, while keeping the same overall phenomenology. This version of the NG is suitable for massive numerical simulations and analytical approaches. Moreover, the extreme simplicity allows for a direct comparison with other models introduced in other frameworks of statistical physics as well as in other disciplines.

The Minimal Naming Game The simplest version of the NG (Baronchelli et al. 2006b) is played by a population of N agents, on a fully connected network, trying to bootstrap a common name for a given object. Each player is characterized by an inventory of word-object associations he/she knows. All agents have empty inventories at time $t = 0$. At each time step ($t = 1, 2, \ldots$), two players are picked at random and one of them plays as speaker and the other as hearer. Their interaction obeys the rules described in Fig. 1.

- The speaker selects an object from the current context.
- The speaker retrieves a word from its inventory associated with the chosen object, or, if its inventory is empty, invents a new word.
- The speaker transmits the selected word to the hearer.
- If the hearer has the word named by the speaker in its inventory and that word is associated to the object chosen by the speaker, the interaction is a success and both players maintain in their inventories only the winning word, deleting **all** the others.
- If the hearer does not have the word named by the speaker in its inventory, or the word is associated to a different object, the interaction is a failure and the hearer updates its inventory by adding an association between the new word and the object.

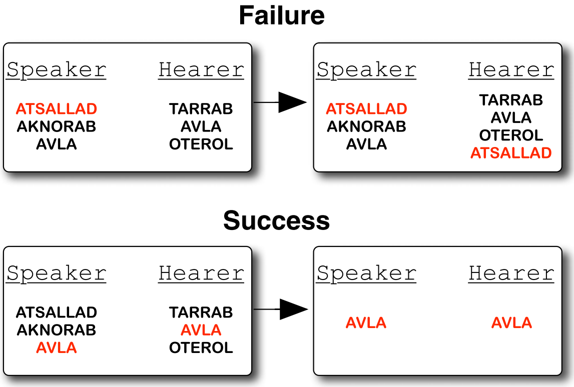


Fig. 1 Naming Game. Examples of the dynamics of the inventories in a failed (*top*) and a successful (*bottom*) game. The speaker selects the word highlighted. If the hearer does not possess that word he includes it in his inventory (*top*). Otherwise both agents erase their inventories only keeping the winning word (*bottom*)

One assumes that the number of possible words is so huge that the probability that two players invent the same word at two different times for two different objects is practically negligible (this means that homonymy is not taken into account here, though the extension is trivially possible) and so the choice dynamics among the possible words associated with a specific object are completely independent. As a consequence, without loss of generality, one can reduce the environment to one single object ($M = 1$).

In this perspective it is interesting noting that in Komarova and Niyogi (2004), it was formally proven, adopting an evolutionary game theoretic approach, that languages with homonymy are evolutionary unstable. On the other hand, it is commonly observed that human languages contain several homonyms, while true synonyms are extremely rare. In Komarova and Niyogi (2004) this apparent paradox is resolved remarking that if we think of "words in a context," homonymy does indeed disappear from human languages, while synonymy becomes much more relevant. In the framework of the NG, homonymy is not always an unstable feature (see Sect. 3 and Puglisi et al. 2008 for an example) and its survival depends in general on the size of the meaning and signal spaces (Gosti 2007).

These observations match perfectly also with the assumption of the NG, according to which speaker and hearer are able to establish whether the game was successful by subsequent action performed in a common environment. For example, the speaker may refer to an object in the environment he wants to obtain and the hearer then hands the right object. If the game is a failure, the speaker may point to (non-verbal communication) or get the object himself, so that it is clear to the hearer which object was intended.

Macroscopic Analysis The first property of interest is the time evolution of the total number of words owned by the population $N_w(t)$, of the number of different words $N_d(t)$, and of the success rate $S(t)$ (Fig. 2).

We can distinguish three phases in the behavior of the system. Very early, pairs of agents play almost uncorrelated games and the number of words hence increases over time as $N_w(t) = 2t$, while the number of different words increases as $N_d(t) = t$. In the second phase the success probability is still very small and

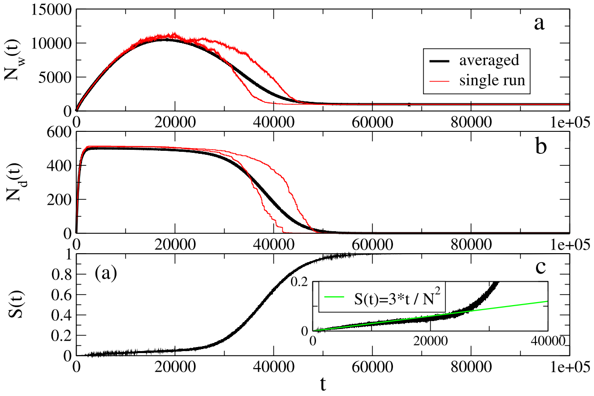


Fig. 2 Naming Game. (**a**) Total number of words present in the system, $N_w(t)$; (**b**) Number of different words, $N_d(t)$; (**c**) Success rate $S(t)$, i.e., probability of observing a successful interaction at time t. The *inset* shows the linear behavior of $S(t)$ at small times. The system reaches the final absorbing state, described by $N_w(t) = N$, $N_d(t) = 1$ and $S(t) = 1$, in which a global agreement has been reached. From Baronchelli et al. (2006b)

agents' inventories start getting correlated, the $N_w(t)$ curve presenting a well identified peak. The process evolves with an abrupt increase in the number of successes and a further reduction in the numbers of both total and different words. Finally, the dynamics ends when all agents have the same unique word and the system is in the attractive convergence state. It is worth noting that the developed communication system is not only *effective* (each agent understands all the others), but also *efficient* (no memory is wasted in the final state).

After a transient period, the system undergoes spontaneously a disorder/order transition to an asymptotic state where global coherence emerges, i.e., every agent has the same word for the same object. It is remarkable that this happens starting from completely empty inventories for each agent. The asymptotic state is one where a word invented during the time evolution took over with respect to the other competing words and imposed itself as the leading word. In this sense the system spontaneously selects one of the many possible coherent asymptotic states and the transition can thus be seen as a symmetry breaking process. The dynamics of the NG is characterized by the following scaling behavior for the convergence time t_{conv}, the time and the height of the peak of $N_w(t)$, namely $t_{\max}$ and $N_w^{\max} = N_w(t_{\max})$. It turns out that all these quantities follow power law behaviors: $t_{\max} \sim N^{\alpha}$, $t_{\text{conv}} \sim N^{\beta}$, $N_{\max} \sim N^{\gamma}$, and $t_{\text{diff}} = (t_{\text{conv}} - t_{\max}) \sim N^{\delta}$, with exponents $\alpha = \beta = \gamma = \delta = 1.5$ (with a subtle feature around the disorder-order transition where an additional timescale emerges). The values of those exponents can be understood through simple scaling arguments (Baronchelli et al. 2006b).[1]

---

[1] Here the time is the number of binary interactions.

2.1 Symmetry Breaking: A Controlled Case

Consider now a simpler case in which there are only two words at the beginning of the process, say A and B, so that the population can be divided into three classes: the fraction of agents with only A, n_A, the fraction of those with only the word B, n_B, and finally the fraction of agents with both words, n_{AB}. Describing the time evolution of the three species is straightforward:

$$\begin{aligned} \dot{n}_A &= -n_A n_B + n_{AB}^2 + n_A n_{AB}, \\ \dot{n}_B &= -n_A n_B + n_{AB}^2 + n_B n_{AB}, \\ \dot{n}_{AB} &= +2n_A n_B - 2n_{AB}^2 - (n_A + n_B) n_{AB}. \end{aligned} \tag{1}$$

The system of differential equations (1) is deterministic. It presents three fixed points in which the system can collapse depending on the initial conditions. If $n_A(t=0) > n_B(t=0)$ $[n_B(t=0) > n_A(t=0)]$, at the end of the evolution we will have the stable fixed point $n_A = 1$ $[n_B = 1]$ and, consequently $n_B = n_{AB} = 0$ $[n_A = n_{AB} = 0]$. If, on the other hand, we start from $n_A(t=0) = n_B(t=0)$, the equations lead to $n_A = n_B = 2n_{AB} = 0.4$. The latter situation is clearly unstable, since any external perturbation would make the system fall in one of the two stable fixed points.

Equations (1), however, are not only a useful example to clarify the nature of the symmetry breaking process. In fact, they also describe the interaction among two different populations that converged separately on two distinct conventions. In this perspective, (1) predict that the larger population will impose its conventions. In the absence of fluctuations, this is true even if the difference is very small: B will dominate if $n_B(t=0) = 0.5 + \epsilon$ and $n_A(t=0) = 0.5 - \epsilon$, for any $0 < \epsilon \leq 0.5$ and $n_{AB}(t=0) = 0$. Data from simulations show that the probability of success of the convention of the minority group n_A decreases as the system size increases, going to zero in the thermodynamic limit ($N \to \infty$). A similar approach has been proposed to model the competition between two languages in the seminal paper Abrams and Strogatz (2003). Here it is worth remarking the formal similarities between modeling the competition between synonyms in a NG framework and the competition between languages: in both cases a synonym or a language are represented by a single feature, e.g. the characters A or B, for instance, in (1). The similarity has been made more evident by the subsequent variants of the model introduced in Abrams and Strogatz (2003) to include explicitly the possibility of bilingual individuals. In particular in Wang and Minett (2005); Minett and Wang (2008) deterministic models for the competition of two languages have been proposed, which include bilingual individuals. In Castelló et al. (2006) a modified version of the voter model including bilingual individuals has been proposed, the so-called AB-model. In a fully connected network and in the limit of infinite population size, the AB-model can be described by coupled differential equations for the fractions of individuals speaking language A, B or AB, that are, up to a constant normalization factor in the timescale, identical to (1).

2.2 The Role of the Interaction Topology

Social networks play an important role in determining the dynamics and outcome of language change. The first investigation of the role of topology was proposed in 2004, at the 5th Conference on Language evolution, Leipzig (Ke et al. 2008). Since then many approaches focused on adapting known models on topologies of increasing complexity: regular lattices, random graphs, scale-free graphs, etc.

The NG, as described above, is not unambiguously defined on general networks. As already observed in Chap. 5, when the degree distribution is heterogeneous, the order in which an agent and one of its neighbors are selected does matter, because high-degree nodes are more easily chosen as neighbors than low-degree nodes. Several variants of the NG on generic networks can be defined. In the *direct NG* (*reverse NG*) a randomly chosen speaker (hearer) selects (again randomly) a hearer (speaker) among its neighbors. In a *neutral* strategy one selects an edge and assigns the role of speaker and hearer with equal probability to the two nodes (Dall'Asta et al. 2006b).

On low-dimensional lattices consensus is reached through a coarsening phenomenon (Baronchelli et al. 2006a) with a competition among the homogeneous clusters corresponding to different conventions, driven by the curvature of the interfaces (Bray 1994). A scaling of the convergence time as $\mathcal{O}(N^{1+1/d})$ has been conjectured, where $d \leq 4$ is the lattice dimension. Low-dimensional lattices require more time to reach consensus compared to a fully connected graph, but a lower use of memory. A similar analysis has been performed for the AB-model (Castelló et al. 2006). The effect of a small-world topology has been investigated in Dall'Asta et al. (2006a) in the framework of the NG and in Castelló et al. (2006) for the AB-model. Two different regimes are observed. For times shorter than a crossover time, $t_{\mathrm{cross}} = \mathcal{O}(N/p^2)$, one observes the usual coarsening phenomena as long as the clusters are one-dimensional, i.e., as long as the typical cluster size is smaller than $1/p$. For times much larger than t_{cross}, the dynamics is dominated by the existence of shortcuts and enters a mean field like behavior. The convergence time is thus expected to scale as $N^{3/2}$ and not as N^3 (as in $d = 1$). Small-world topology allows to combine advantages from both finite-dimensional lattices and fully connected networks: on the one hand, only a finite memory per node is needed, unlike the $\mathcal{O}(N^{1/2})$ in fully connected graphs; on the other hand the convergence time is expected to be much shorter than in finite dimensions. In Castelló et al. (2006), the dynamics of the AB-model on a two-dimensional small world network, has been studied. Also in this case a dynamic stage of coarsening is observed, followed by a fast decay to the A or B absorbing states caused by a finite size fluctuation (Fig. 3).

The NG has been studied on complex networks as well. Here the convergence time t_{conv} scales as N^{β}, with $\beta \simeq 1.4 \pm 0.1$, for both Erdös–Renyi (ER) (Erdös and Rényi 1959, 1960) and Barabási-Albert (BA) (Barabási and Albert 1999) networks. The scaling laws observed for the convergence time are general robust features not affected by further topological details (Dall'Asta et al. 2006b; Dall'Asta and Baronchelli 2006). Finally, it is worth mentioning how the NGs with

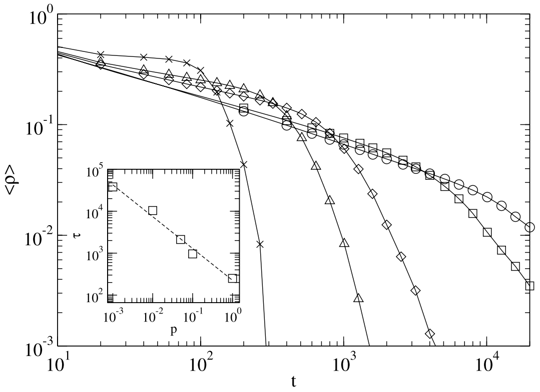


Fig. 3 AB-model. Time evolution of the average density $\langle\rho\rangle$ of bilingual individuals in small-world networks for different values of the rewiring parameter p. From *left* to *right*: $p = 1.0$, 0.1, 0.05, 0.01, 0.0. The *inset* shows the dependence of the characteristic lifetime τ on the rewiring parameter p. The *dashed line* corresponds to a power law fit $\tau \sim p^{-0.76}$. From Castelló et al. (2006)

local broadcasts on random geometric graphs have been investigated in Lu et al. (2008) as a model for agreement dynamics in large-scale, autonomously operating wireless sensor networks.

2.3 *Variants of the Naming Game*

A variant of the NG has been introduced with the aim of mimicking the mechanisms leading to opinion and convention formation in a population of individuals (Baronchelli et al. 2007). In particular, a new parameter, β, has been added mimicking an *irresolute attitude* of the agents in making decisions ($\beta = 1$ corresponds to the NG). The parameter β is simply the probability that, in a successful interaction, both the speaker and the hearer update their memories erasing all opinions except the one involved in the interaction (see Fig. 1). This negotiation process displays a non-equilibrium phase transition from an absorbing state in which all agents reach a consensus to an active (not frozen as in the Axelrod 1997 model) stationary state characterized either by polarization or fragmentation in clusters of agents with different opinions. Very interestingly, the model displays the non-equilibrium phase transition also on heterogeneous networks, in contrast with other opinion-dynamics models, like for instance the Axelrod model (Klemm et al. 2003), for which the transition disappears for heterogeneous networks in the thermodynamic limit.

A hybrid approach, combining vertical and horizontal transmission of cultural traits, has been proposed in Ke et al. (2002) while an evolutionary version of the Naming Game has been introduced in Lipowski and Lipowska (2008).

3 The Category Game

Categories are fundamental to recognize, differentiate and understand the environment. According to Aristotle, categories are entities characterized by a set of properties which are shared by their members (Barnes 1995). A recent wave in cognitive science, on the other hand, has operated a shift in viewpoint from the object of categorization to the categorizing subjects (Lakoff 1987; Gardner 1987): categories are culture-dependent conventions shared by a given group. In this perspective, a crucial question is how they come to be accepted at a global level without any central coordination (Steels 1998; Kirby and Christiansen 2005; Steels and Belpaeme 2005; Belpaeme and Bleys 2005; Baronchelli et al. 2006b; Komarova et al. 2007). The answer has to be found in communication, that is the ground on which culture exerts its pressure. An established breakthrough in language evolution (Maynard-Smith and Szathmary 1997; Hurford et al. 1998; Nowak et al. 2002; Kirby and Christiansen 2005) is the appearance of linguistic categories, i.e. a shared repertoire of form-meaning associations in a given environment (Labovin et al. 1973; Lakoff 1987; Gardner 1987; Garrod and Anderson 1987; Taylor 1995; Steels 1998; Coehn and Lefebvre 2005). Different individuals may in principle perceive, and even conceptualize, the world in very different ways, but they need to align their linguistic ontologies in order to understand each other.

In the past there have been many computational and mathematical studies addressing the learning procedures for form-meaning associations (Hurford 1989; Briscoe 2002). From the point of view of methodology, the evolutionary scheme, based on the maximization of some fitness functions, has been extensively applied (Nowak and Krakauer 1999; Nowak et al. 1999b). Recent years, however, have shown that also the orthogonal approach of self-organization can be fruitfully exploited in multi-agent models for the emergence of language (Steels and Belpaeme 2005; Belpaeme and Bleys 2005; Baronchelli et al. 2006b). In this context, a community of language users is viewed as a complex dynamical system which has to develop a shared communication system (Steels 2000; Komarova 2006). In this debate, a still open problem concerns the emergence of a small number of forms out of a diverging number of meanings. For example, the few "basic color terms," present in natural languages, coarse-grain an almost infinite number of perceivable different colors (Berlin and Kay 1969; Saunders and Brakel 1997; Lindsey and Brown 2006).

Following this recent line of research, our work shows that an assembly of individuals with basic communication rules and without any external supervision may evolve an initially empty set of categories, achieving a non-trivial communication system characterized by a few linguistic categories. To probe the hypothesis that cultural exchange is sufficient to this extent, individuals in our model are never replaced (unlike in evolutionary schemes Nowak and Krakauer 1999; Nowak et al. 1999b), the only evolution occurring in their internal form-meaning association tables, i.e., their "mind." The individuals play elementary language games (Wittgenstein 1953a; Steels 1996) whose rules constitute the only knowledge initially shared by the population. They are also capable of perceiving analogical stimuli and communicating with each others (Steels and Belpaeme 2005; Belpaeme and Bleys 2005).

3.1 The Category Game Model

Our model (Puglisi et al. 2008) involves a population of N individuals (or players), committed in the categorization of a single analogical perceptual channel, each stimulus being represented as a real-valued number ranging in the interval [0, 1].

3.1.1 Modeling Categories

Here we identify categorization as a partition of the interval [0, 1] in discrete subintervals, from now onwards denoted as perceptual categories. This approach can also be extended to categories with prototypes and fuzzy boundaries, for instance adding a weight structure upon it. Typical proposals in literature, such as prototypes with a weight function equal to the inverse of the distance from the prototype (Belpaeme and Bleys 2005), are exactly equivalent to our "rigid boundaries" categories. Moreover, all the results of our experiment can be easily generalized to multi-dimensional perceptual channels, provided an appropriate definition of category domains is given. It should be kept in mind that the goal of our work is to investigate why the continuum of perceivable meanings in the world is organized, in language, in a finite and small number of subsets with different names, with a no immediate (objective) cause for a given partition with respect to other infinite possibilities. Apart from the evident example of the partition of the continuous light spectrum in a small number of "basic color terms," this phenomenon is widespread in language: one can ask, for example, what objective differences allow to distinguish a cup from a glass; one can present a multi-dimensional continuum of objects able to "contain a liquid" (including also objects given as a prize), but a natural discontinuity between cups and glasses does not appear; our model, even reducing the phenomenon to the case of a 1-dimensional continuum, unveils a mechanism which can be easily extended to any kind of space, once it has been provided with a topology. The mechanism we propose for the discrete partition in linguistic subsets (categories) does not depend on the exact nature of this topology, which is of course a fundamental, yet different, matter of investigation.

3.1.2 Negotiation Dynamics

Each individual has a dynamical inventory of form-meaning associations linking perceptual categories (meanings) to words (forms), representing their linguistic counterpart. Perceptual categories and words associated to them co-evolve dynamically through a sequence of elementary communication interactions, simply referred as games. All players are initialized with only the trivial perceptual category [0, 1], with no name associated to it. At each time step, a pair of individuals (one playing as speaker and the other as hearer) is selected and presented with a new "scene," i.e., a set of $M \geq 2$ objects (stimuli), denoted as $o_i \in [0, 1]$ with $i \in [1, M]$. The speaker discriminates the scene, adding new category boundaries to isolate the topic, then

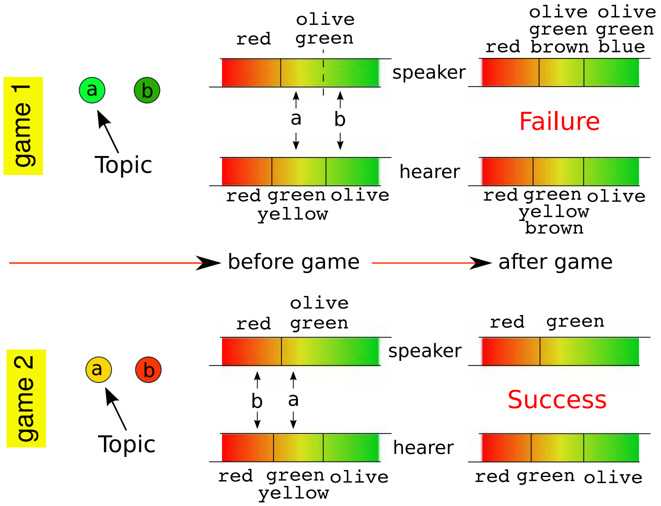


Fig. 4 Rules of the game. A pair of examples representing a failure (game 1) and a success (game 2), respectively. In a game, two players are randomly selected from the population. Two objects are presented to both players. The speaker selects the topic. In game 1 the speaker has to discriminate the chosen topic ("a" in this case) by creating a new boundary in his rightmost perceptual category at the position $(a+b)/2$. The two new categories inherit the words-inventory of the parent perceptual category (here the words "green" and "olive") along with a different brand new word each ("brown" and "blue"). Then the speaker browses the list of words associated to the perceptual category containing the topic. There are two possibilities: if a previous successful communication has occurred with this category, the last winning word is chosen; otherwise the last created word is selected. In the present example the speaker chooses the word "brown," and transmits it to the hearer. The outcome of the game is a failure since the hearer does not have the word "brown" in his inventory. The speaker unveils the topic, in a non-linguistic way (e.g., pointing at it), and the hearer adds the new word to the word inventory of the corresponding category. In game 2 the speaker chooses the topic "a", finds the topic already discriminated and verbalizes it using the word "green" (which, for example, may be the winning word in the last successful communication concerning that category). The hearer knows this word and therefore points correctly to the topic. This is a successful game: both the speaker and the hearer eliminate all competing words for the perceptual category containing the topic, leaving "green" only. In general when ambiguities are present (e.g., the hearer finds the verbalized word associated to more than one category containing an object), these are solved making an unbiased random choice

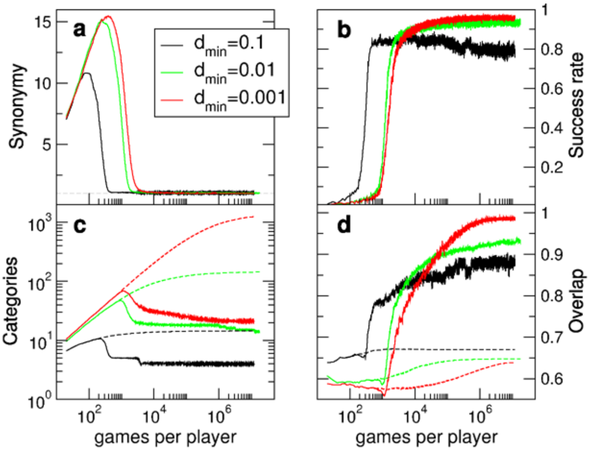
he names one object and the hearer tries to guess it. The word to name the object is chosen by the speaker among those associated to the category containing the object, with a preference for the one which has been successfully used in the most recent game involving that category. A correct guess makes the game successful. Based on game's outcomes, individuals may update their category boundaries and the inventory of the associated words: in a successful game both players erase competing words in the category containing the topic, keeping only the word used in that game; in failed games, the speaker points out the topic and the hearer proceeds to discriminate it, if necessary, and then adds the spoken word to its inventory for that category. Detailed example of the game are given in Fig. 4.

The perceptive resolution power of the individuals limits their ability to distinguish objects/stimuli that are too close to each other in the perceptual space: in order to take this into account, we define a threshold $d_{\min}$, inversely proportional to their resolution power.[2] In a given scene, the M stimuli are chosen to be at a distance larger than this threshold, i.e., $|o_i - o_j| > d_{\min}$ for every pair (i, j). Nevertheless, objects presented in different games may be closer than $d_{\min}$. The way stimuli are randomly chosen characterizes the kind of simulated environment: simulations will be presented both with a homogeneous environment (uniform distribution in [0, 1]) and more natural environments (e.g., without loss of generality, the distributions of the hue sampled from pictures portraying natural landscapes).

3.2 Hierarchical Coordination

A resume of the main results of our experiments is given in Fig. 5. The evolution of the population presents two main stages: (1) a phase where players do not understand each other, followed by (2) a phase where communication has reached an averagely high success thanks to the emergence of a common language, still with evolving perceptual categories and a finite fraction of failures due to slightly unaligned categories and ambiguities. The first phase is marked by the growth and decline of synonymy; see Fig. 5a. Synonymy, in the context of the "naming game" (an individual object to be named), has been already studied (Baronchelli et al. 2006b), and a similar evolution was observed and explained. All individuals, when necessary, create new words with zero probability of repetition: this leads to an initial growth of the vocabulary associated to each perceptual category. New words are

Fig. 5 **Simulations results** with $N = 100$ and different values of $d_{\min}$: (**a**) Synonymy, i.e. average number of words per category; (**b**) Success rate measured as the fraction of successful games in a sliding time windows games long; (**c**) Average number of perceptual (*dashed lines*) and linguistic (*solid lines*) categories per individual; (**d**) Averaged overlap, i.e., alignment among players, for perceptual (*dashed curves*) and linguistic (*solid curves*) categories



---

[2] In psychology, $d_{\min}$ is equivalent to the so-called Just Noticeable Difference ("JND") or Difference Limen ("DL").

spread through the population in later games and, whenever a word is understood by both players, other competing words for the same category are forgotten.[3] This eventually leads to only one word per category. During the growth of the dictionary the success rate, see Fig. 5b, is very small. The subsequent reduction of the dictionary corresponds to a growing success rate which reaches its maximum value after synonymy has disappeared. In all our numerical experiments the final success rate overcomes 80% and in most of them goes above 90%, weakly increasing with the final number of perceptual categories. Success is reached in a number of games per player of the order of 5×10^2, logarithmically depending on N, and it remains constant hereafter.

The set of perceptual categories of each individual follows a somewhat different evolution (see dashed lines in Fig. 5c). The first step of each game is, in fact, the discrimination stage where the speaker (possibly followed by the hearer) may refine his category inventory in order to distinguish the topic from the other objects. The growth of the number of perceptual categories n_{perc} of each individual is limited by the resolution power: in a game two objects cannot appear at a distance smaller than d_{min} and therefore $n_{\text{perc}} < 2/d_{\text{min}}$. The minimal distance also imposes a minimum number of categories $1/d_{\text{min}}$ that an individual must create before his discrimination process may stop. The average number of perceptual categories per individual, having passed $1/d_{\text{min}}$, grows sub-logarithmically and for many practical purposes it can be considered constant.

The success rate is expected to depend on the alignment of the category inventory among different individuals. The degree of alignment of category boundaries is measured by an *overlap function* O

$$O = 2 \sum_{i<j} \frac{o_{ij}}{N(N-1)} \quad \text{with } o_{ij} = \frac{2 \sum_{c_i^j} (l_{c_i^j})^2}{\sum_{c_i} (l_{c_i})^2 + \sum_{c_j} (l_{c_j})^2}, \tag{2}$$

where l_c is the width of category c, c_i is one of the categories of the ith player, and c_i^j is the generic category of the "intersection" set obtained considering all the boundaries of both players i and j. The function returns a value proportional to the degree of alignment of the category inventories, reaching its maximum unitary value when they exactly coincide. Its study, see dashed curves of Fig. 5d, shows that alignment grows with time and saturates to a value which is, typically, in between 60%–70%, i.e., quite smaller than the communicative success. This observation immediately poses a question: given such a strong misalignment among individuals, why is communication so effective?

The answer has to be found in the analysis of polysemy, i.e. the existence of two or more perceptual categories identified by the same unique word. Misalignment, in fact, induces a "word contagion" phenomenon. With a small but non zero probability, two individuals with similar, but not exactly equal, category boundaries, may

---

[3] Extensions of this model can be devised in order to account for cases where words are not always erased, but instead they can became more "specialized," eventually yielding to the emergence of a hierarchy of category names.

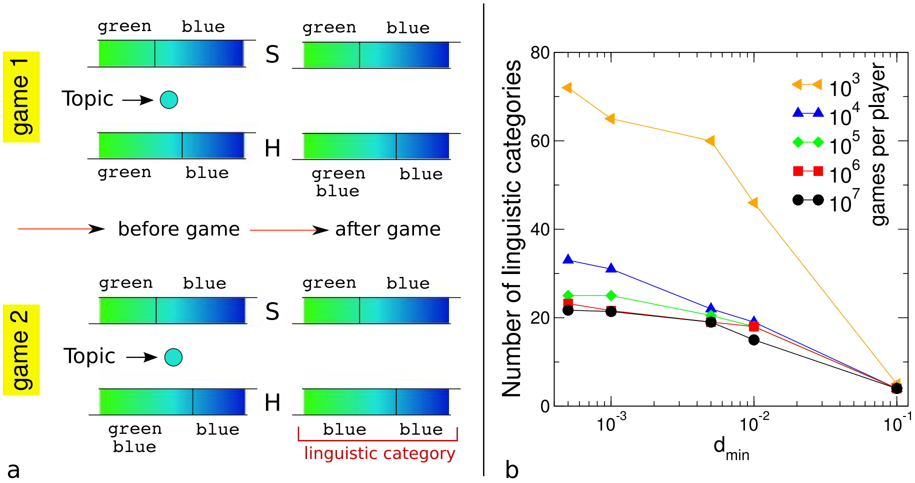


Fig. 6 **Saturation in the number of linguistic categories**: (**a**) A "word contagion" phenomenon occurs whenever the topic falls in a gap between two misaligned categories of two playing individuals. In the shown examples two individuals play two successive games. In game 1 the speaker (S) says "blue" and the hearer (H), unable to understand, adds "blue" as a possible word for his leftmost category; successively (game 2) the speaker repeats "blue" and the hearer learns this word as the definitive name for that perceptual category; both left and right perceptual categories of the hearer are now identified by the same name "blue" and they can be considered (for the purpose of communication) as a single linguistic category; (**b**) Final number of linguistic categories as a function of d_{min} at different times, with $N = 100$. As the time increases the number of linguistic categories saturates. At large times, for small d_{min}, the number of linguistic categories becomes independent of d_{min} itself. Concerning size dependence, only a weak (logarithmic) dependence on N, not shown, is observed

play a game with a topic falling in a misalignment gap, as represented in Fig. 6a. In this way a word is copied to an adjacent perceptual category and, through a second occurrence of a similar event, may become the unique name of that category. Interfering events may occur in between: it is always possible, in fact, that a game is played with a topic object falling in the bulk of the category, where both players agree on its old name, therefore canceling the contagion. With respect to this cancelling probability, some gaps are too small and act as almost perfectly aligned boundaries, drastically reducing the probability of any further contagion. Thus, polysemy needs a two-step process to emerge, and a global self-organized agreement to become stable. On the other hand, polysemy guarantees communicative success: perceptual categories that are not perfectly aligned tend to have the same name, forming true linguistic categories, much better aligned among different individuals. The topmost curve of Fig. 5d, displays the overlap function measured considering only boundaries between categories with different names:[4] it is shown to reach a much higher value, even larger than 90%.

---

[4] We define name of a perceptual category the word that an individual would choose, according to the rules of the model, to communicate about an object discriminated by that category (i.e., the last-winning word or the last created word). Of course, if there is a unique word associated with

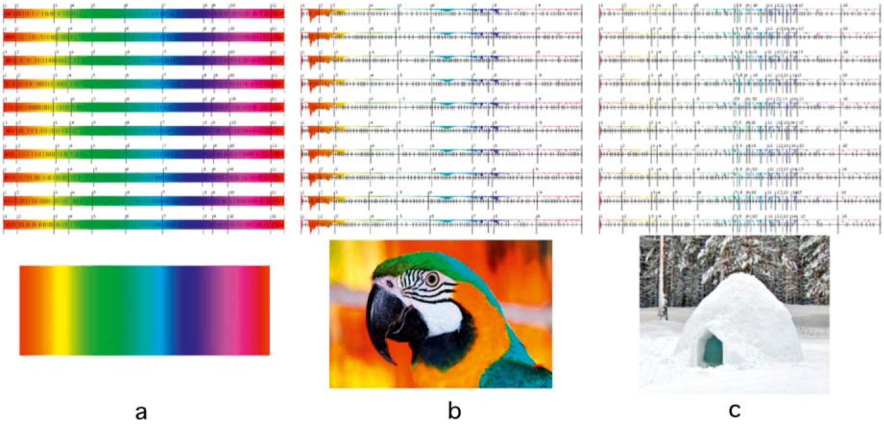


Fig. 7 Categories and the pressure of environment. Inventories of ten individuals randomly picked up in a population of $N = 100$ players, with $d_{\min} = 0.01$, after 10^7 games. For each player the configuration of perceptual (*small vertical lines*) and linguistic (*long vertical lines*) category boundaries is superimposed to a colored histogram indicating the relative frequency of stimuli. The labels indicate the unique word associated to all perceptual categories forming each linguistic category. Three cases are presented: one with uniformly distributed stimuli (**a**) and two with stimuli randomly extracted from the hue distribution of natural pictures (**b** and **c**). One can appreciate the perfect agreement of category names, as well as the good alignment of linguistic category boundaries. Moreover, linguistic categories tend to be more refined in regions where stimuli are more frequent: an example of how the environment may influence the categorization process

The appearance of linguistic categories is the evidence of a coordination of the population on a higher hierarchical level: a superior linguistic structure on top of the individual-dependent, finer, discrimination layer. The linguistic level emerges as totally self-organized and is the product of the (cultural) negotiation process among the individuals. The average number of linguistic categories per individual, n_{ling}, Fig. 5c (solid curves), grows together with n_{cat} during the first stage (where communicative success is still lacking), then decreases and stabilizes to a much lower value. Some configurations of both category layers, at a time such that the success rate has overcome 95%, are presented in Fig. 7, using different sets of external stimuli.

The analysis, resumed in Fig. 6b, of the dependence of n_{ling} on $d_{\min}$ for different times, makes our findings robust and, to our knowledge, unprecedented. As the resolution power is increased, i.e., as $d_{\min}$ is diminished, the asymptotic number of linguistic categories becomes less and less dependent upon $d_{\min}$ itself. Most importantly, even if any state with $n_{\text{ling}} > 1$ is not stable, we have the clear evidence of a saturation with time, in close resemblance with metastability in glassy systems (Mezard et al. 1987; Debenedetti and Stillinger 2001). This observation allows to give a solution to the long-standing problem of explaining the finite (and small)

---

a category (which is most often the case after homonymy has almost disappeared), the definition above identifies that word as the name of the category.

number of linguistic categories n_{ling}. In previous pioneering approaches (Steels and Belpaeme 2005; Belpaeme and Bleys 2005) the number of linguistic categories n_{ling} was trivially constrained (with a small range of variability) by d_{min}, with a relation of the kind $n_{\text{ling}} \propto 1/d_{\text{min}}$, implying a divergence of n_{ling} with the resolution power. In our model we have a clear indication of a finite n_{ling} even in the continuum limit, i.e., $d_{\text{min}} \to 0$, corresponding to an infinite resolution power.

4 Conclusions

Language dynamics is a promising field which encompasses a broader range of applications with respect to what described in this chapter (Loreto and Steels 2007). In many biological, technological and social systems, a crucial problem is that of the communication among the different components, i.e., the elementary units of the systems. The agents interact among themselves and with the environment in a sensorial and non-symbolic way, their communication system not being predetermined nor fixed from a global entity. The communication system emerges spontaneously as a result of the interactions of the agents and it could change continuously due to the mutations occurring in the agents, in their objectives as well as in the environment. An important question concerns how conventions are established, how communication arises, what kind of communication systems are possible and what are the prerequisites for such an emergence to occur.

In the framework of the so-called Category Game we have shown that a simple negotiation scheme, based on memory and feedback, is sufficient to guarantee the emergence of linguistic categories in a population of individuals endowed with the ability of forming perceptual categories. The Category Game reproduces a typical feature of natural languages: despite a very high resolution power, the number of linguistic categories is very small. For instance, in many human languages, the number of "basic color terms" used to categorize colors usually amounts to about ten (Berlin and Kay 1969; Saunders and Brakel 1997; Lindsey and Brown 2006), in European languages it fluctuates between 6 and 12, depending on gender, level of education, and social class, while the light spectrum resolution power of our eyes is evidently much higher. Finally we believe that these results could be important both from the point of view of language evolution theories, possibly leading to a quantitative comparison with real data[5] (Selten and Warglien 2007) and suggesting new experiments (e.g., different populations sizes and ages), and from the point view of applications (e.g., emergence of new communication systems in biological, social and technological contexts; Steels 2006; Cattuto et al. 2007).

In this perspective, the emergence of a common vocabulary only represents a first stage while it is interesting to investigate the emergence of higher forms of

---

[5] A collection of available experimental data can be found in *The World Color Survey*, http://www.icsi.berkeley.edu/wcs.

agreement, e.g., compositionality, categories, syntactic or grammatical structures. It is clear how important it would be to cast a theoretical framework where all these problems could be defined, formalized and solved. That would be a major input for the comprehension of many social phenomena as well as for devising new technological instruments.

Finally it is important to mention that in the last few years a potentially very interesting experimental platform appeared: the World Wide Web. Though only a few years old, the growth of the Web and its effect on the society have been astonishing, spreading from the research in high-energy physics into other scientific disciplines, academe in general, commerce, entertainment, politics and almost anywhere where communication serves a purpose. Innovation has widened the possibilities for communication. Social media like blogs, wikis, and social bookmarking tools allow the immediacy of conversation, with unprecedented levels of communication speed and community size. In this perspective the web is acquiring the status of a platform for *social computing*, able to coordinate and exploit the cognitive abilities of the users for a given task. In this sense, it is likely that the new social platforms appearing on the web, could rapidly become a very interesting laboratory for social sciences in general, and for studies on language emergence and evolution in particular.

Acknowledgements A.B. aknowledges support of Spanish MCI through the Juan de la Cierva program and from the Spanish MEC (FEDER), under project No. FIS2007-66485-C02-01, as well as additional support through ICREA Academia, funded by the Generalitat de Catalunya.

References

Abrams, D., & Strogatz, S. (2003). Modelling the dynamics of language death. *Nature*, *424*, 900.

Axelrod, R. (1997). The dissemination of culture: a model with local convergence and global polarization. *The Journal of Conflict Resolution*, *41*(2), 203–226.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509.

Barnes, J. B. (Ed.) (1995). *The complete works of Aristotle: the revised Oxford translation. Bollingen series* (Vol. LXXI). Princeton: Princeton University Press.

Baronchelli, A., Dall'Asta, L., Barrat, A., & Loreto, V. (2006a). Topology induced coarsening in language games. *Physical Revue E*, *73*, 015102.

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006b). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics*, P06014.

Baronchelli, A., Dall'Asta, L., Barrat, A., & Loreto, V. (2007). Non-equilibrium phase transition in negotiation dynamics. *Physical Revue E*, *76*, 051102.

Belpaeme, T., & Bleys, J. (2005). Explaining universal color categories through a constrained acquisition process. *Adaptive Behavior*, *13*, 293–310.

Berlin, B., & Kay, P. (1969). *Basic color terms: their universality and evolution*. California: University of California Press. (Reprinted 1991).

Bray, A. (1994). Theory of phase-ordering kinetics. *Advances in Physics*, *43*(3), 357–459.

Brighton, H., Smith, K., & Kirby, S. (2005). Language as an evolutionary system. *Physics of Life Reviews*, *2*, 177–226.

Briscoe, T. (Ed.) (2002). *Linguistic evolution through language acquisition*. Cambridge: Cambridge University Press.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Revues of Modern Physics*.

Castelló, X., Eguíluz, V., & San Miguel, M. (2006). Ordering dynamics with two non-excluding options: bilingualism in language competition. *New Journal of Physics*, *8*(12), 308.

Cattuto, C., Loreto, V., & Pietronero, L. (2007). Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences of the United State of America*, *104*, 1461–1464.

Coehn, H., & Lefebvre, C. L. (Eds.) (2005). *Handbook of categorization in cognitive science*. New York: Elsevier.

Dall'Asta, L., & Baronchelli, A. (2006). Microscopic activity patterns in the naming game. *Journal of Physics A*, *39*, 14851–14867.

Dall'Asta, L., Baronchelli, A., Barrat, A., & Loreto, V. (2006b). Nonequilibrium dynamics of language games on complex networks. *Physical Revue E*, *74*(3), 036105.

Dall'Asta, L., Baronchelli, A., Barrat, A., & Loreto, V. (2006a). Agreement dynamics on small-world networks. *Europhysics Letters*, *73*(6), 969–975.

Debenedetti, P. G., & Stillinger, F. H. (2001). Supercooled liquids and the glass transition. *Nature*, *410*, 259.

Erdös, P., & Rényi, A. (1959). On random graphs I. *Publicationes Mathematicae (Debrecen)*, *6*, 290.

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, *7*, 17.

Gardner, H. (1987). *The mind's new science: a history of the cognitive revolution*. New York: Basic Books.

Garrod, G., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, *27*, 181.

Gosti, G. (2007). *Role of the homonymy in the naming game*. Undergraduate thesis, "Sapienza" Univ. of Rome.

Hurford, J. (1989). Biological evolution of the Saussurean sign as a component of the language acquisition device in linguistic evolution. *Lingua*, *77*, 187.

Hurford, J. R., Studdert-Kennedy, M., & Knight, C. (Eds.) (1998). *Approaches to the evolution of language: social and cognitive bases*. Cambridge: Cambridge University Press.

Ke, J., Minett, J., Au, C.-P., & Wang, W. (2002). Self-organization and selection in the emergence of vocabulary. *Complexity*, *7*(3), 41–54.

Ke, J., Gong, T., & Wang, W.-Y. (2008). Language change and social networks. *Computer Physics Communications*, *3*(4), 935–949. Originally presented at the 5th conference on language evolution, Leipzig, Germany, March 2004.

Kirby, S., & Christiansen, M. (Eds.) (2005). *Language evolution*. Oxford: Oxford University Press.

Klemm, K., Eguíluz, V. M., Toral, R., & San Miguel, M. (2003). Nonequilibrium transitions in complex networks: a model of social interaction. *Physical Revue E*, *67*(2), 026120.

Komarova, N. L. (2006). Population dynamics of human language: a complex system. In *Frontiers of engineering: reports on leading-edge engineering from the 2005 symposium* (pp. 89–98).

Komarova, N. L., & Niyogi, P. (2004). Optimizing the mutual intelligibility of linguistic agents in a shared world. *Artificial Intelligence*, *154*(1–2), 1–42.

Komarova, N. L., Jameson, K. A., & Narens, N. (2007). Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology*, *51*, 359–382.

Labovin, W., Bailey, C.-J., & Shuy, R. (1973). *New ways of analyzing variation in English*. Washington: Georgetown University Press.

Lakoff, G. (1987). *Women, fire and dangerous things*. Chicago: Chicago University Press.

Lenaerts, T., Jansen, B., Tuyls, K., & de Vylder, B. (2005). The evolutionary language game: an orthogonal approach. *Journal of Theoretical Biology*, *235*(4), 566–582.

Lindsey, D. T., & Brown, A. M. (2006). Universality of color names. *Proceedings of the National Academy of Sciences of the United State of America*, *103*, 16608.

Lipowski, A., & Lipowska, D. (2008). *Computational approach to the emergence and evolution of language—evolutionary naming game model*.