Conventionalization of Linguistic Knowledge Under Communicative Constraints

Tao Gong
Andrea Puglisi
Physics Department
"La Sapienza" University
Rome, Italy
gtojty@gmail.com
Andrea.Puglis@roma1.infn.it

Vittorio Loreto
Physics Department
"La Sapienza" University
Rome, Italy
& ISI Foundation
Complex Systems Lagrange LAB
Turin, Italy
Vittorio.Loreto@roma1.infn.it

William S.-Y. Wang
Department of Electronic Engineering
The Chinese University of Hong Kong
China
wsywang@ee.cuhk.edu.hk

Abstract
The language game approach has recently been adopted to explore the conventionalization of linguistic knowledge in a social environment. Most contemporary studies focus on the dynamics of language games in random or predefined social networks, but neglect the reverse roles of communicative constraints in language evolution and social structures. This article, based on two forms of language games (the naming game and the category game), examines whether a simple, distance-based communicative constraint can affect the conventionalization of linguistic knowledge. The study bridges the gap between random networks and complex social structures, and illustrates that the internal properties of language games can influence the effects of communicative constraints and social structures.

Keywords
category game, communicative constraints, conventionalization, naming game

January 31, 2008; revised and accepted April 30, 2008

Biological Theory 3(2) 2008, 154–163.

Language evolves primarily via social contact among a limited number of individuals. A prerequisite for successful language use is conventional linguistic knowledge (Lewis 1969; Barr 2004). Recently, multi-agent simulations based on theoretical frameworks have been applied to the question of how these conventions become established without a central coordinator (see Wagner et al. 2003; Loreto and Steels 2007). Many proposed models in this line of research have adopted a language game approach (Loreto and Steels 2007) to explore the conventionalization (a process of social agreement in which one conforms one's language to that of others or of the community [Burling 2005]) of different types of linguistic knowledge. This approach views language as a complex adaptive system (Steels 2000; Wang 2006) and usually postulates a population of agents and an interaction protocol (a language game; Steels 2002) among them. Using the language game, agents carry out some communicative tasks. By inventing new forms of conceptualization and/or expression and adjusting available knowledge based on its utility, frequency, or social prestige, agents can gradually build up a communication system from scratch. Through iterated language games among agents, a set of linguistic knowledge will gradually become conventional in the population, and related statistical analyses can examine the dynamics of the game and provide quantitative data on human language and its evolution.

Many forms of language games have been proposed, including the naming game (Baronchelli et al. 2006), which studies the emergence of coherence (in the form of shared vocabulary), and the category game (Puglisi et al. 2008), which simulates the emergence of a set of conventional linguistic categories. These two games are briefly reviewed in the next section.

Although such language games reveal successful social strategies, statistical physicists (e.g., Dall'Asta et al. 2006; Kalampokis et al. 2007) seem more interested in exploring the dynamics of language games in predefined social structures, such as fully connected graphs, 1D/2D lattices, and scale-free (Barabási and Albert 1999) or small-world (Watts 1999) networks. Most of these latter explorations have neglected the reverse role of language games in social structure because in approaches such as these, a successful or failed language game cannot affect the predefined social connections among individuals.

As a social phenomenon, linguistic interactions can affect both the participants' knowledge and their social connections. An accumulation of failed or successful interactions can weaken or strengthen social connections, while local communicative constraints during linguistic interactions, such as geographical or social distance, can alter the probability that agents will interact with each other, thus affecting individual or group similarities on a global scale (Axelrod 1997). Local communicative constraints, in existence long before the emergence of complex social structures, can continue to cast their influence on the formation of mutual understanding and social structures (Nettle 1999). Therefore, it is premature to proceed from random networks directly to complex structures before conducting a careful study of these constraints and their roles in language evolution and social structures.

With all this in mind, we present here a simulation study, based on the naming game and the category game, whose aim is to explore the role of a simple, distance-based communicative constraint on the conventionalization of shared vocabulary and linguistic categories. The results and related analyses reveal some essential differences in these language games, which can cause different performances under the same communicative constraint. These findings contribute not only to the discussion of mutual influence between linguistic communications and social structures, but also to the exploration of language game dynamics in complex networks.

In what follows, we first briefly review the naming game and the category game to give readers some concrete ideas about the language game approach. We then introduce the distance-based communicative constraint. Based on the simulation results using these language games, we illustrate the effect of this constraint on the conventionalization of linguistic knowledge, point out some different internal properties of these language games, and analyze their distinct performances under the same communicative constraint, ending with our conclusions.

The Language Games

The Naming Game

The naming game studies the conventionalization of lexical knowledge among idiolects. It simulates how, during linguistic interactions, individuals exchange words (names) to arrive at what to call a particular object in their environment. Through some local strategies and iterated naming games among agents, a collective agreement on a shared word-meaning mapping can emerge in a population of agents.

In the naming game model (Baronchelli et al. 2006), N homogeneous agents use invented or learned words to describe an object (meaning). Each agent has an unlimited, initially empty inventory (memory) in which to store the exchanged words. In one version of the game, two agents (a speaker and a hearer) are randomly chosen from the population. The speaker first utters a word to the hearer. If its inventory is empty, the speaker will randomly invent a word; otherwise, it will utter an available word randomly selected from its inventory. If the hearer has the same word in its inventory, the game is successful, and both agents delete words other than the uttered one from their inventories; otherwise, the game fails, and the hearer learns the uttered word and adds it to its inventory.

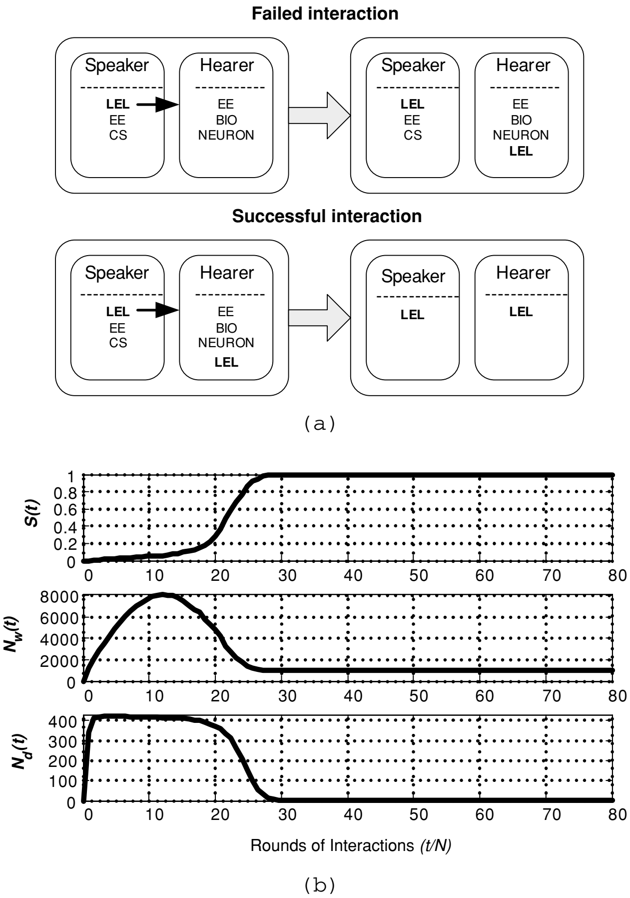

Figure 1.
The naming game and its dynamics. (a) Two examples of the naming game (adapted from Baronchelli et al. 2006). (b) The dynamics of the naming game in a fully connected graph with $N = 100$ agents and $t = 8 \times 10^3$ games. The top panel traces $S(t)$, the middle $N_w(t)$, and the bottom $N_d(t)$.

These strategies are exemplified in the two examples of the naming game shown in Figure 1(a). In the first (top) game, the speaker utters "LEL," a randomly selected word. Since the hearer's inventory does not contain "LEL," this game fails, and the hearer adds "LEL" to its inventory. In the second game, the speaker utters "LEL." Since the hearer already has this word in its inventory, the game is successful, and both agents delete other words in their inventories and preserve only "LEL."

The dynamics of the naming game can be traced by several indices, such as $N_w(t)$, the total number of words in the population; $N_d(t)$, the number of different words in the population; and $S(t)$, the average successful rate of naming games among all pairs of agents. Based on these indices, Figure 1(b) summarizes the dynamics of the naming game in a fully connected graph. A detailed analysis of the dynamics can be found in Baronchelli et al. (2008).

The Category Game

The category game extends the naming game to consider the conventionalization of semantic categories and their common word labels among idiolects. In the category game model (Puglisi et al. 2008), N individuals describe stimuli from a single analogical perceptual space. These stimuli are represented by real-valued numbers ranging from 0.0 to 1.0. Agents can perceive these stimuli and establish perceptual categories to discriminate them. Each perceptual category corresponds to a discrete subinterval within [0.0, 1.0]. Agents can assign a list of words to describe the stimuli from this subinterval. If one of these words was successfully used in previous category games, it will become the preferred word of that perceptual category. Agents have unlimited inventories to store both perceptual categories and their word labels. If some perceptual categories share the same preferred word, a linguistic category that unifies the subintervals of these categories will emerge. If most agents in the population develop a number of common linguistic categories (their unified subintervals have similar boundaries and their preferred words are identical), a "categorical language" represented by common categorical knowledge is said to emerge in the population. In linguistics, an example of such categorical knowledge is the color terms; in a language, a number of perceptually different color stimuli can be called by an identical color term, such as red, green, or blue. In addition to color categories and color terms, the category game can simulate the emergence and conventionalization of other categorical knowledge in a population of agents.

In a category game, two agents (a speaker and a hearer) are randomly selected, and M (≥ 2) stimuli are presented to them. The speaker discriminates these stimuli and uses a word to describe one stimulus (topic). The hearer tries to guess the topic from other contextual stimuli based on both the speaker's word and the hearer's own inventory of perceptual categories. In this model, an agent's discrimination ability is restricted by its perceptual power, denoted by $d_{\min}$, the minimal numerical distance required for distinguishing two stimuli. In an individual category game, the minimal distance between any presented stimuli is $d_{\min}$, although, in general, stimuli can take any numerical value, constrained only by the numerical precision used in the experiment.

The individuals' perceptual categories and their words can be adjusted during category games using some local strategies. Such strategies are exemplified in the two examples of the category game shown in Figure 2(a). In the upper game, since the two stimuli fall into the same perceptual category, the speaker has to discriminate the topic ("a") by creating a new boundary in its rightmost category at the position (a + b)/2, thus separating that category into two new perceptual categories. These new categories inherit the word inventory ("green" and "olive") of their parent category. The speaker then browses the list of words in the category that contains "a." If there was a previous successful game with this category, its preferred word is sent to the hearer; otherwise, the newly created word ("brown") becomes the preferred word and is sent to the listener. Since the

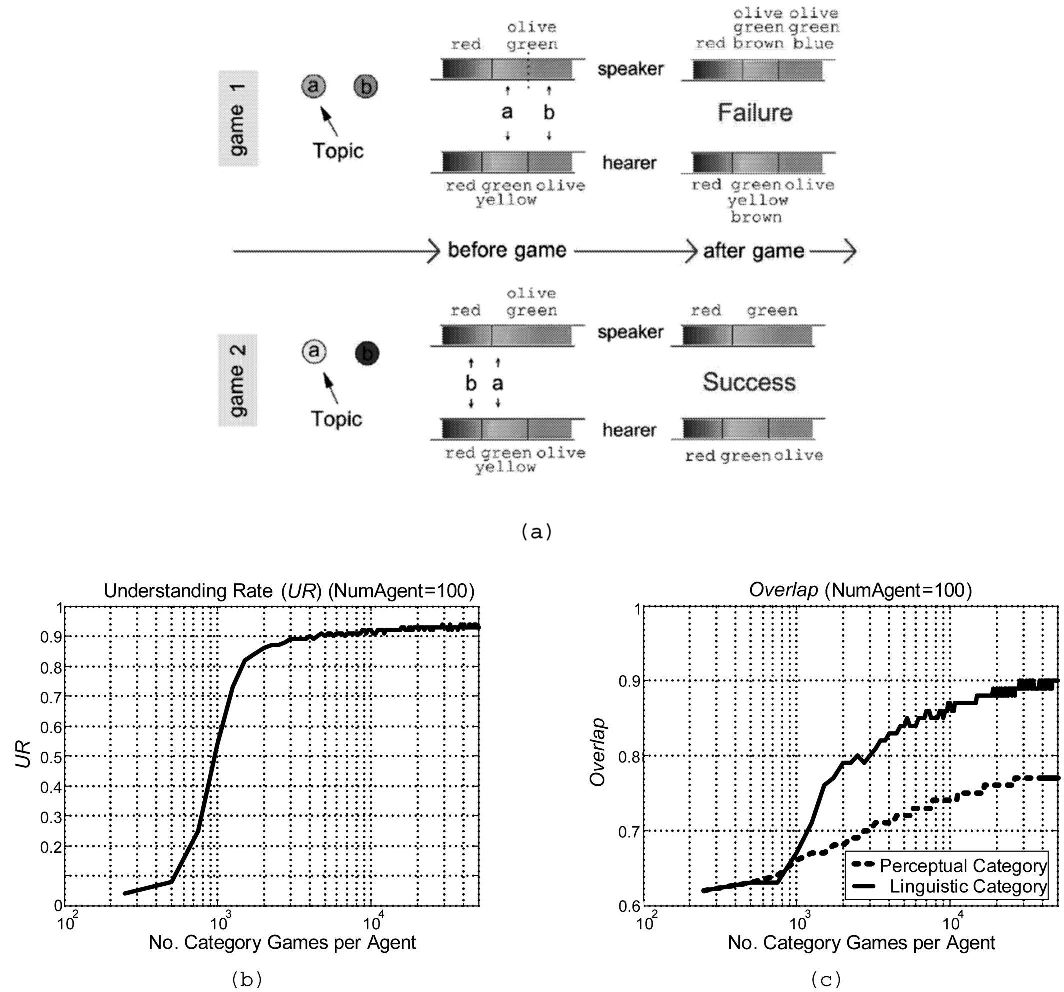

Figure 2.
The category game and its dynamics. (a) Two examples of the category game (adapted from Puglisi et al. 2008). The round objects are stimuli presented in these games, among which the topics are pointed to with arrows. Each agent has a colorful banner to represent the perceptual space, which is partitioned by bars into perceptual categories, whose word inventories are listed above or below. (b) and (c) The dynamics of the category game in a fully connected graph with $N = 100$ agents, $d_{\min} = 0.01$, and $t = 5 \times 10^4$ games, traced by *UR* (b) and *Overlap* (c).

hearer does not have "brown" in its inventory, this game fails. Then, the speaker points at "a," and the hearer discriminates the topic, adding "brown" to the inventory of its corresponding category. In the second game, the speaker chooses the topic "a," which can already be discriminated by a perceptual category whose preferred word is "green." The speaker then sends "green" to the hearer. The hearer knows "green" and points at the topic contained in its corresponding perceptual category. This game is successful, and both agents eliminate all the competing words in their used perceptual categories, leaving only "green." If ambiguity occurs, in the sense that the speaker's word is associated with many categories containing the topic, the hearer chooses randomly among the categories.

In a fully connected graph, category game dynamics can be evaluated by two indices, Understanding Rate (*UR*) and *Overlap*. To measure *UR*, all pairs of agents play "virtual" category games without updating their perceptual categories and the word lists associated with those categories, and the percentage of successful games in all these virtual games is then calculated. *Overlap* evaluates the alignment of subinterval boundaries of the perceptual or linguistic categories. It is calculated by the overlap function (Puglisi et al. 2008), shown in equation (1):

$$O = 2\sum_{i<j} o_{ij}/N(N-1),$$

$$o_{ij} = \frac{2\sum_{c_{ij}} (lc_{ij})^2}{\sum_{c_i} (lc_i)^2 + \sum_{c_j} (lc_j)^2}, \tag{1}$$

where lc is the width of the subinterval of category c, c_i is a category of player i, and c_{ij} is the category intersection set obtained based on the category boundaries of both players i and j. o_{ij} indicates the degree of alignment between categories in players i and j, which will reach 1.0 if the subinterval boundaries of these two sets of categories are identical. The dynamics of the category game in a fully connected graph is shown in Figure 2(b)–(c), and a detailed discussion of the dynamics can be found in Puglisi et al. 2008.

Both of these language games have been proposed as a way to abstract the self-organizing process of conventionalization of linguistic knowledge. In a fully connected graph, the emergence of shared lexical knowledge using the naming game (indicated by *S*) and of common categorical knowledge using the category game (indicated by *UR* and *Overlap*) is similar, both starting from a disordered state, then undergoing a sharp transition from no knowledge to conventional linguistic knowledge, and ending up at a saturation state. The similarity in dynamics reflects the similarities in the strategies adopted in these language games. For instance, each agent in these games has a dynamical inventory in which to store names and perceptual categories and their word labels. If necessary, agents can randomly create the name to call a particular object or a perceptual category to discriminate a particular topic from other contextual stimuli. If a game is successful, both participants eliminate all the competing words, leaving only the word uttered in that game; otherwise, the speaker clarifies the topic by pointing and the hearer adds the speaker's word to its inventory.

The Distance-based Communicative Constraint

In this section we adopt the naming game and the category game to examine whether a distance-based communicative constraint on these language games can affect the conventionalization of linguistic knowledge. To introduce the constraint, as shown in Figure 3, we put all agents in a 2D torus, X^2 (X is the side length of this torus). Agents can take discrete positions in X^2 and their positions are not overlapped. After playing a language game, an agent can move to one of its eight unoccupied, adjacent discrete locations. Compared with other 2D structures, there is no physical boundary in a torus; if agents at the rightmost (leftmost, top, or bottom) positions move further right (left, up, or down), they will reappear in the leftmost (rightmost, bottom, or top) positions. This structure can represent either a physical world or a virtual one, such as the distributions of opinions or social status.

The distance-based communicative constraint is defined as follows: in an X^2, language games only take place between agents whose coordinates are within a limited block distance (D_x and D_y), as shown in equations (2) and (3):

$$\left|x_i - x_j\right| \leq D_x \quad \text{or} \quad \left|x_i - x_j - 0.5X\right| \leq D_x, \qquad (2)$$

$$\left|y_i - y_j\right| \leq D_y \quad \text{or} \quad \left|y_i - y_j - 0.5X\right| \leq D_y, \qquad (3)$$

where x_i, y_i are agent i's coordinates in X^2. The second part of each equation concerns the special feature of torus.

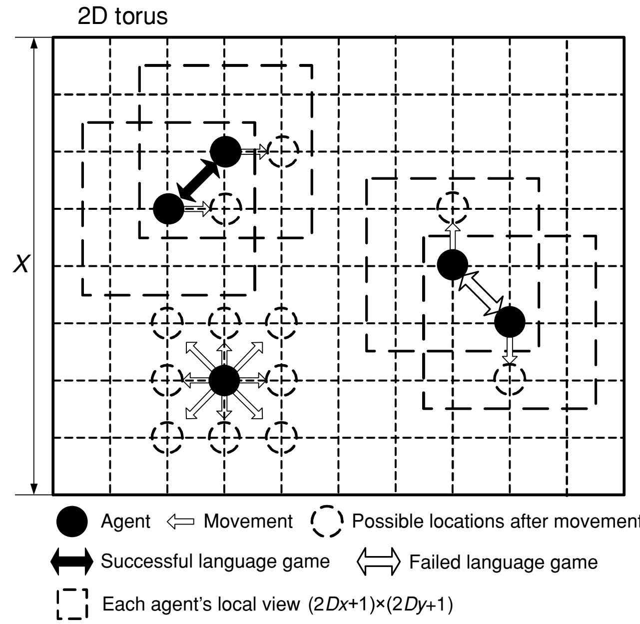

Figure 3.
The 2D torus with moving agents.

D_x and D_y can represent either a geographical constraint such as the city-county distance or a social constraint such as dissident opinions. It acts as an individual's local view, under which an agent can contact at most $(2D_x + 1) \times (2D_y + 1) - 1$ (itself) other agents if all of them are in its nearby locations. If D_x and D_y are 1, each agent can only interact with those lying in its eight adjacent locations. This distance constraint also provides a bias in agent movement. If a language game between two agents is successful, they tend to bind together by moving jointly to maintain their block distance within D_x and D_y (i.e., if either of them moves, it tends to move to one of its unoccupied, adjacent locations [if any] so that their block distance after the movement is still within D_x and D_y). If the game fails, this binding may break down (i.e., if either of them moves, it moves randomly to one of its unoccupied, adjacent locations [if any]). This constraint is local and much simpler than those defined by global complex networks. It may trigger clusterization, i.e., the emergence of social clusters, each containing some agents who share similar linguistic knowledge but may not necessarily interact directly with each other. These clusters and their shared knowledge provide the prototypes of complex social structures and their communal languages.

This communicative constraint can affect interactions among agents in two ways: (1) through the local view, in which an agent can only interact with others within its local view; (2) through the movement bias, by which agents can contact others previously outside its local view and form clusters.

To evaluate the effects of this constraint on the formation of conventional linguistic knowledge, we carry out two

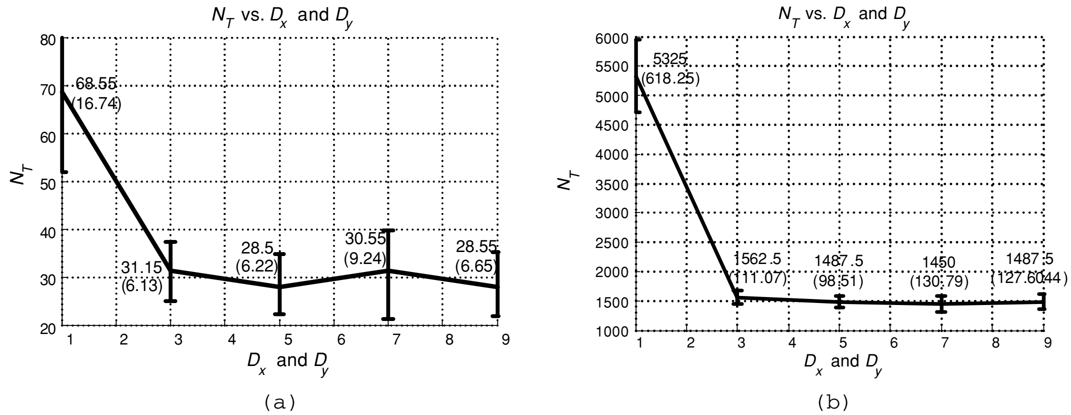

Figure 4.
The dynamics of the naming game (a) and that of the category game (b) in Experiment 1 based on N_T (numbers outside brackets are the average values, and those inside brackets are the standard deviations values).

experiments. In Experiment 1, 100 agents are located in a 10^2 torus (each location in the torus is occupied by an agent) and D_x and D_y vary. In this situation, the local view is the only factor to affect interactions. In Experiment 2, 100 agents are put into tori having different side lengths, but D_x and D_y remain constant. In this situation, the movement bias also affects interactions.

In a unit time step of these experiments, all agents are selected one by one following a random sequence. Once an agent is chosen, it will play a language game with one of the others lying within its distance constraint (if any). After that, it will move (only in Experiment 2), based on the game result (successful or failed), to one of its unoccupied, adjacent positions (if any). Noticing the different timescales in which conventionalization is achieved in a fully connected graph, we set the total number of time steps in the naming game to 100 (the maximum number of possible interactions is $100 \times 100 = 10^4$) and in the category game to 5×10^4 (the maximum number of possible interactions is 5×10^6). The actual interactions depend on X, D_x and D_y, and on agent displacement in the torus.

In each condition, the results of 20 simulations are collected. In each simulation, we set up 200 sampling points in the total timescale to calculate several indices. For the naming game, these indices include S, N_d, and N_T (the number of time steps required to reach the highest S); for the category game, these indices include *UR*, *Overlap* of linguistic categories, and N_T (the number of time steps required to reach the highest *UR*). If all agents gradually share a common word-meaning mapping or a set of common linguistic categories, S will increase to 1.0 and N_d will decline to 1, and *UR* and *Overlap* will increase to 1.0, too. If clusterization occurs, S and N_d will not reach 1.0 and neither will *UR* and *Overlap*. In this situation, a small number of N_d in the naming game can reflect the number of isolated clusters (although a big N_d may indicate that agents are still at the misunderstanding stage) and N_T in both games can indicate the efficiency of conventionalization in the whole population or in different clusters.

The Simulation Results

Experiment 1: Under Different Distance Constraints (D_x and D_y)

In this experiment, all agents lie in a 10^2 torus, and D_x and D_y change from 1 to 9 with a step of 2. The conventionalization of linguistic knowledge is achieved in all conditions using both forms of language games. Figure 4 illustrates the dynamics of these games in this experiment using N_T.

As shown in Figure 4, when D_x and D_y are so small that not every two agents in the whole group can directly interact with each other, the conventionalization of linguistic knowledge is still achieved via intermediate agents, and all agents form one cluster. With the increase in D_x and D_y, the conventionalization process is accelerated, since agents with bigger local views can contact more individuals in the group to share their names or align the subintervals of their perceptual categories and acquire common preferred words in these categories. But when D_x and D_y become quite large, the acceleration effect on conventionalization becomes less explicit (N_T will not further decrease) since agents can already interact with all others in the group.

Experiment 2: Under Different Torus Sizes (X)

In this experiment, 100 agents are randomly located in tori with various sizes but the local views of these agents remain at 5. In the naming game, the torus size changes from 10 to 50 with a step of 10. In the category game, it changes from 10 to 90 with a step of 20. Figure 5 shows the dynamics of these language games in this experiment using S, N_T, and N_d for the naming game, and *UR*, *Overlap* of linguistic categories, and N_T for the category game.

Using the naming game, the lexical conventionalization in Experiment 2 follows two regimes. When X is smaller than 30, agents can encounter others through movement and get their words converged, and all agents can form one cluster and share

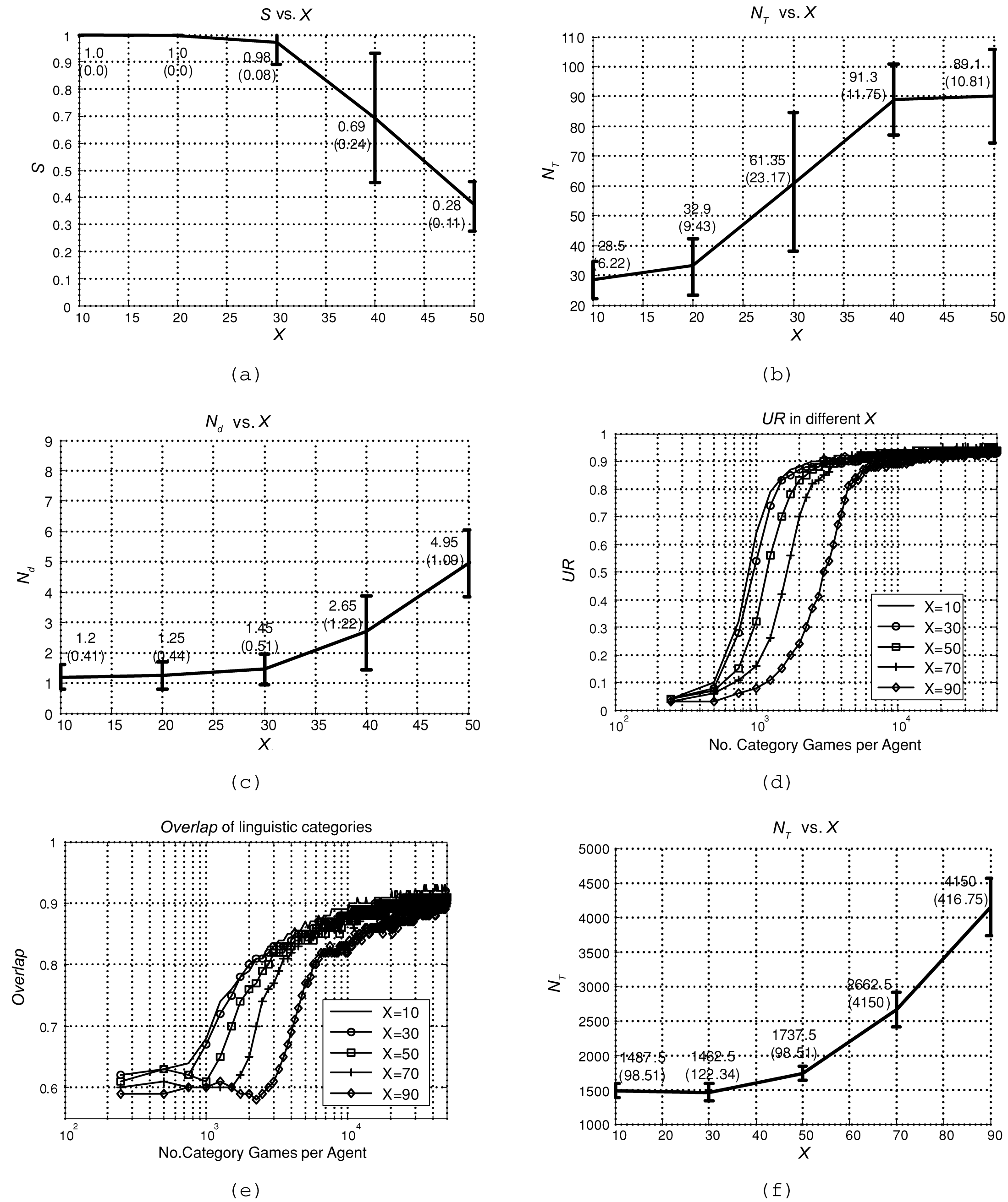

Figure 5.

The dynamics of the language games in Experiment 2. As for the naming game, the average (numbers outside brackets) and standard deviation values (numbers within brackets) of S, N_T, and N_d are shown in (a), (b), and (c). As for the category game, conventionalization is achieved in all conditions. Therefore, the *UR* and *Overlap* of linguistic categories in some particular simulations are shown in (d) and (e), and the average (numbers outside brackets) and standard deviation values (numbers within brackets) of N_T in all conditions are shown in (f).

a common word. Once X exceeds 30, this one-step movement does not allow for agents to meet many others. S begins to drop, and both N_T and N_d increase. In this situation, clusterization occurs: some isolated clusters each sharing a common word gradually emerge, as shown by the slight increase in N_d. The low S of the whole group also indicates that these clusters share different words. Moreover, these clusters and their shared words are temporally stable, as illustrated by the values of S and N_d, which remain stable for a long time in particular simulations. This occurs because once these clusters are formed, the movement bias can restrain their members from moving freely, and reduce the chance of contacts between members of different clusters. Clusterization in Experiment 1 illustrates a "local convergence, global polarization" phenomenon (Axelrod 1997): agents within clusters can clearly understand each other via a shared word, but those between clusters cannot, as they employ different words. This phenomenon partially reflects the coexistence of many languages in the world, and it is largely caused by communicative constraint on local interactions.

Using the category game, with the increase in X, the time steps required for a high degree of mutual understanding also increase, while increases in both *UR* and *Overlap* of linguistic categories are delayed. However, using the category game, clusterization does not occur even in a much bigger torus $X = 90$; in all tori, after a number of time steps, agents can achieve a set of conventional linguistic categories having a high *UR* and *Overlap*. These results suggest that using the category game, the movement bias in the communicative constraint does little to help the formation of clusters or to restrain agents from freely interacting with others.

Collectively, the simulation results in both experiments show the role of the communicative constraint in the conventionalization of linguistic knowledge. First, a big cluster-sharing common linguistic knowledge can be formed among individuals whose local views may not allow them to interact directly with each other. Second, there is a correlation between the local view and world size. Under a fixed world, the increase in the local view can accelerate the conventionalization process. Under a fixed local view, an increase in world size can delay this process in the whole population and it may trigger clusterization, as shown in the naming game. The condition of enlarging local views is reminiscent of today's growing mass media and the development of a "global village" over recent centuries, while a fixed local view with increasing world sizes reflects the fact that people still have a relatively limited view. The simulations in this article address a scenario with these two competing conditions; other activities, such as opinion formation (Rosvall and Sneppen 2007), may follow a similar scenario.

Despite these general tendencies, under the simulation settings used here the same communicative constraint fails to trigger clusterization during the category game, which inspires us to seek the underlying differences between these language games.

Discussion

These two types of language games display both similarities (listed earlier) and differences. In the naming game, the object is fixed; what differs are the names to call it. Based on local imitation strategies, two agents can easily make their inventories converge after only a few naming games. And once converged, future naming games between these agents remain successful. Even if the convergence breaks down due to interference from other agents, it can quickly be reestablished within a limited number of naming games.

In the category game, however, the topic and other contextual stimuli are randomly selected from a continuous perceptual space. Given the various contexts, to categorize the same topic, agents may partition the perceptual space differently by creating perceptual categories with different boundaries. Due to a mismatch of their boundaries, in these categories the preferred words to call a topic may not easily converge. The uncertainty in the stimulus selection will cause the boundary alignment of perceptual categories and the "word contagion" phenomenon (the spread of preferred words among adjacent perceptual categories to form linguistic categories; Puglisi et al. 2008) in individuals' inventories to continue for a long time, until they become slower and slower and, finally, imperceptible. Therefore, without external interference, the acquisition of categorical knowledge between two agents cannot be efficiently achieved within a limited number of category games, even though the agents are endowed with similar discrimination and learning mechanisms. Since it is utterly impossible to experience all stimuli in the continuous perceptual space, after a number of category games, a temporary solution to the problem of how to achieve a certain degree of mutual understanding emerges, in which a number of linguistic categories that roughly partition the whole space into some discrete subintervals—each using a different preferred word—are gradually shared among agents. This solution is largely determined by the repertoire of external stimuli and the properties of the communicative task involved in the category game (Puglisi et al. 2008).

Considering these different properties, a successful naming game can clearly indicate the convergence of two agents' lexical knowledge and the movement bias based on it is a good factor to form and maintain social clusters. However, the success of a single or even a few category games is insufficient to indicate the convergence of participants' categorical knowledge and the same movement bias cannot play a similar role in building up social clusters.

As shown in Figures 1(b) and 2(b), the number of games per agent (time steps) required to achieve conventionalization (T_{share}) based on the naming game is much smaller than that based on the category game. This timescale difference also helps explain the different success rates of the two language games under the same communicative constraint in Experiment 2. Instead of qualitative analyses, we can suggest a quantitative analysis on this point.

In Experiment 2, agents need to explore the torus to contact each other. If we assume that agents are randomly hopping on the torus, the mean square displacement of agent position should follow equation (4):

$$\langle (x(t) - x(0))^2 \rangle \sim 2Dt, \tag{4}$$

where D is a diffusion coefficient. In the dilute limit and absence of constraints, if each agent on average plays and moves once in a time step, D is of order 1. With a slightly reduced value of D, equation (4) is expected to hold for longer times, even in the presence of constrained interactions among agents (e.g., when displacement is correlated after a successful game).

When an agent explores the torus, following equation (4), a repertoire of category boundaries and category names also travels through the system together with this agent. Small pieces of these repertoires among agents are distributed by means of constrained interactions, and can in principle travel faster: constrained interactions allow information to hop on the torus with a diffusion coefficient not larger than D_x (D_y), both of which are kept constant in Experiment 2.

A combination of agent diffusion and constrained interaction drives the travel of categorical information through the system. The time steps required for this information to travel across the whole torus should follow equation (5):

$$T_{\text{diff}} \sim X^2/D, \tag{5}$$

where D is a constant not smaller than 1 and not larger than D_x.

While an agent travels, through category games, it learns the repertoires of surrounding agents. As shown in Experiment 2, in a time T_{share}, it manages to share a large part of its category and word label inventories with the population it has contacted during its lifetime.

The comparison of these two characteristic times, T_{diff} and T_{share}, determines the fate of the system:

(1) If $T_{\text{diff}} < T_{\text{share}}$ (i.e., when $X < \sqrt{T_{\text{share}} D}$), each agent visits the world in a shorter time than the time required to achieve shared categorical knowledge, then clusterization (i.e., different "category languages") is not possible in this regime;
(2) If $T_{\text{diff}} > T_{\text{share}}$ (i.e., when $X > \sqrt{T_{\text{share}} D}$), an intermediate "clusterized" regime may appear during the time window $T_{\text{share}} < T < T_{\text{diff}}$.

Note that T_{share} also depends of course on the agent density and $D_x(D_y)$. For example, it is impossible to simply increase X^2 (to increase T_{diff}) and keep a constant N, since T_{share} would also increase. The system volume should be increased at a constant density, which requires a proportional increase of N.

With our choices of parameters, the largest possible T_{share} in the category game is of the order of 10^4, under which the simulation results indicate that the system is in regime (1) and clusterization is impossible. We suspect that an increase of X^2 (and of N) by a factor of 10 should cause regime (2) to be visible. Of course, given times longer than T_{diff}, regime (1) will eventually come back. In the naming game, since conventionalization among agents is achieved on a much shorter T_{share}, it is much easier to observe clusterization. Asymptotically in time, regime (1) will also eventually govern the naming game anyway.

Besides the situation in Experiment 2, the internal properties of these language games may also influence the effect of the social structures that provide more complex topological constraints. For instance, since coherence can be efficiently established via a limited number of naming games, agents with many social connections (hubs) become important drivers of the convergence of the whole population. Meanwhile, based on a simple, distance-related communicative constraint and some naming-game-like interactions, similar complex networks can be triggered, as exemplified in the studies of lexical convergence and language change in complex networks (e.g., Kalampokis et al. 2007; Baronchelli et al. 2008; Ke et al. 2008). However, using the category game, the hub's role becomes inexplicit and it is difficult to predict the global structures under the same communicative constraint.

All these analyses suggest that, unless we clearly examine the internal properties of language games and their relations with local communicative constraints, we should not directly play various language games under the same predefined social structures, since under the same structure, different language games may have different dynamics, and different language games may trigger different global social connection patterns.

Conclusions

This article presents a simulation study whose aim is to explore the conventionalization of linguistic knowledge using two language games under a local, distance-based communicative constraint. The simulation results and related analyses suggest that different properties of the communicative tasks involved in language games can affect the role of communicative constraints and social structures. This finding, largely neglected by previous explorations, may guide the future study of social structure effects on language evolution. The framework proposed here, together with the statistical analyses in correlated conditions, also makes it possible to carry out a systematic study of the mutual influence of local interactions, conventionalization of linguistic knowledge, and global social structures.

Acknowledgments

This study is partly supported by the EU under RD contract IST-1940 (ECAgents). The authors are indebted to Dr. Andrea Baronchelli from Universitat Politècnica de Catalunya for valuable comments on this work.

References

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

Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286: 509–512.

Baronchelli A, Felici M, Loreto V, Caglioti C, Steels L (2006) Sharp transition towards shared vocabularies in multi-agent systems. Journal of Statistical Mechanics: Theory and Experiment: P06014.

Baronchelli A, Vittorio L, Steels L (2008) In-depth analysis of the naming game dynamics: The homogeneous mixing case. International Journal of Modern Physics C 19: 785–812.

Barr DJ (2004) Establishing conventional communication systems: Is common knowledge necessary? Cognitive Science 28: 937–962.

Burling R (2005) The Talking Ape: How Language Evolved. Oxford: Oxford University Press.

Dall'Asta L, Baronchelli A, Barrat A, Loreto L (2006) Nonequilibrium dynamics of language games on complex networks. Physics Review E 74: 036105.

Kalampokis A, Kosmidis K, Argyrakis P (2007) Evolution of vocabulary on scale-free and random networks. Physica A: Statistical Mechanics Applications 379: 665–671.

Ke JY, Gong T, Wang WSY (2008) Language change and social networks. Communications in Computational Physics 3: 935–949.

Lewis D (1969) Convention: A Philosophical Study. Cambridge, MA: Harvard University Press.

Loreto V, Steels L (2007) Social dynamics: Emergence of language. Nature Physics 3: 758–760.

Nettle D (1999) Using social impact theory to simulate language change. Lingua 108: 95–117.

Puglisi A, Baronchelli A, Loreto V (2008) Cultural route to the emergence of linguistic categories. Proceedings of the National Academy of Sciences of the USA 105: 7936–7940.

Rosvall M, Sneppen K (2007). Dynamics of opinions and social structures. arXiv:0708.0368v1 [physics.soc-ph].

Steels L, ed (2000) Language as a Complex Adaptive System. Proceedings of the 6th International Conference on Parallel Problem Solving from Nature. Lecture Notes in Computer Science 1917: 17–26.

Steels L (2002) Grounding symbols through evolutionary language games. In: Simulating the Evolution of Language (Cangelosi A, Parisi D, eds), 211–226. London: Springer.

Wagner K, Reggia JA, Uriagereka J, Wilkinson GS (2003) Progress in the simulation of emergent communication and language. Adaptive Behavior 11: 37–69.

Wang WSY (2006) Language is a complex adaptive system. Journal of Tsinghua University (Philosophy and Social Sciences) 21(6): 5–13.

Watts DJ (1999) Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton, NJ: Princeton University Press.