The Naming Game

Luc Steels
SONY Computer Science Laboratory - Paris.
AI Laboratory Vrije Universiteit Brussel

November 17, 2004

Abstract

This paper defines a model for the Naming Game, which is seen as the simplest model for studying the self-organisation of communication. Agents directly name uniquely identified objects. Despite its simplicity, the semiotic dynamics of the game when played by a population is already complex, and understanding this dynamics is fundamental for understanding more complex language games, including those involving syntax.

[Paper in progress. Draft. Comments welcome.]

1 Introduction

This paper is part of efforts to develop theoretical insights in how a population of autonomous distributed agents can self-organise communication systems. It explores the paradigm of semiotic dynamics, in which agents expand or adapt their communication conventions and conceptual inventories as part of every game.

The Naming Game is the simplest possible Adaptive Language Game because agents directly name uniquely identified objects. They start without a shared set of conventions and have to self-organise a working communication system from scratch and culturally transmit it to incoming members. The system should be robust in the face of an influx and outflux of members and changes in the set of possible objects to be named.

The model of the Naming Game formalised in this paper was first presented in [5] and used in several computer simulations and robotic experiments thereafter [9]. From these experiments, we know a great detail about the semiotic dynamics generated by the Naming Game. But theoretical analysis is still lacking. This paper defines simple models of the game so that such analysis becomes possible.

The next section of this paper first defines the game. Then a first basic model is presented, which assumes that the probability that two agents use the same name for an object is zero (section 3). The behavior of this model is briefly examined in section 4. Next we examine (section 5) what happens when the probability of equal names is non-zero and so homonyms may occur. It is shown that the use of damping becomes important. Finally we examine some other aspects of the model, such as the acquisition of an existing inventory by a new agent, the effect of introducing an influx of new agents or new objects, and the effect of context. Some theoretical challenges are proposed in section 7 and some comparisons to other language models in section 8.

2 The Naming Game

We assume that games are played by agents which are members of a population $\mathcal{A} = \{a[1], ...a[n]\}$. The state of an agent a at time t is defined as $a_t = \langle \mathcal{I}_{a,t} \rangle$ where $\mathcal{I}_{a,t}$ is the language inventory known by the agent at time t. $\mathcal{S}_{a,t}$ is the set of all possible (single word) sentences which can be constructed using the agent's language inventory $\mathcal{I}_{a,t}$. T is the set of all time points and we assume that every language game takes a single unit of time.

Assume also a set of objects in the domain of discourse $\mathcal{O} = \{o_1, ..., o_n\}$.

The Coding and Decoding Game

A *Coding Game* COG_{NG} played by an agent a at time t is defined as:

$$COG_{a,t} = \langle o, \mathcal{C}, \sigma \rangle$$

where $\mathcal{C} \subseteq \mathcal{O}$ is the context, $o \in \mathcal{C}$ is an object, and $\sigma \in \mathcal{S}_{a,t}$ is a (single word) sentence in the case of success, or $\emptyset$ (the empty sentence) in the case of failure. The state of the agent may change as a side effect of the game: $\mathcal{I}_{a,t}$ is the language inventory available to the agent at the start of the game, and $\mathcal{I}_{a,t+1}$ the language inventory of a at time $t + 1$.

The score obtained in a coding game $COG_{a,t}$ is 1 (success), iff $\sigma \neq \emptyset$, otherwise it is 0.

We define a function $code : \mathcal{A} \times T \times \mathcal{O} \times \mathcal{P}(\mathcal{O}) \rightarrow \mathcal{S}$ such that $code_{a,t}(o, \mathcal{C}) = \sigma$ iff $COG_{a,t} = \langle o, \mathcal{C}, \sigma \rangle$

A *Decoding Game* DEG_{NG} played by an agent a at time t is defined as:

$$DEG_{a,t} = \langle \sigma, o, \mathcal{C} \rangle$$

where $\sigma \in \mathcal{S}_{a,t}$ is a sentence constructable from the agent's language inventory, $o \in \mathcal{C}$ in the case of success or $\emptyset$ (the empty object) in the case of failure, $\mathcal{C} \subseteq \mathcal{O}$ is the context. The state of the agent may change as a side effect of the game: $\mathcal{I}_{a,t}$ is the language inventory available to the agent at the start of the game, and $\mathcal{I}_{a,t+1}$ the language inventory of a at time $t + 1$.

The score obtained in a decoding game $DEG_{a,t}$ is 1 (success), iff $o \neq \emptyset$, otherwise it is 0.

We define a function $decode : \mathcal{A} \times T \times \mathcal{S} \times \mathcal{P}(\mathcal{O}) \rightarrow \mathcal{O}$ such that $decode_{a,t}(\sigma, \mathcal{C}) = o$ iff $DEG_{a,t} = \langle \sigma, o, \mathcal{C} \rangle$

The language game

A Naming Game NG is played by two agents $s \in \mathcal{A}$ and $h \in \mathcal{A}$ at time t, with $s \neq h$, and defined as

$$NG_{s,h,t} = \langle \mathcal{C}, t_s, \sigma, t_h \rangle$$

where $\mathcal{C} \subseteq \mathcal{O}$ is the shared context $t_s \in \mathcal{C}$ is the topic chosen by the speaker, σ is the sentence, $t_h \in \mathcal{C}$ is the topic guessed by the hearer. The relations between t_s, t_h, and σ is defined by a a coding and decoding game: $COG_{s,t} = \langle t_s, \mathcal{C}, \sigma \rangle$ and $DEG_{h,t} = \langle \sigma, t_h, \mathcal{C} \rangle$.

The score obtained in a game $score(NG_{s,h,t})$ is 1 iff $t_s = t_h$, otherwise it is 0. The score determines the communicative success of the game.

The language inventories of speaker and hearer change due to the coding and decoding game to become $\mathcal{I}_{s,t'}$ and $\mathcal{I}_{h,t'}$ and then get further changed based on the outcome of the game:

$$\mathcal{I}_{s,t+1} = SpeakerUpdateI(\mathcal{I}_{s,t'}, t_s, \sigma, score(NG_{s,h,t})) \text{ and}$$
$$\mathcal{I}_{h,t+1} = HearerUpdateI(\mathcal{I}_{h,t'}, t_h, \sigma, score(NG_{s,h,t}))$$

The Naming Game can also be defined in terms of the application of the coding and decoding functions defined earlier: $\sigma = code_{s,t}(t_s, \mathcal{C})$ and $t_h = decode_{h,t}(\sigma, \mathcal{C})$ iff $NG_{s,h,t} = \langle \mathcal{C}, t_s, \sigma, t_h \rangle$.

The goal of the agents is to maximise their cumulative score in the language game, and the details of their behavior should be designed to reach this goal. Initially the agents have no lexicon at all: $\forall a, \mathcal{I}_a = \emptyset$ and inventories must self-organise as a side effect of the game to maximise communicative success.

We are also interested to find out whether the system is adaptive: If a new agent enters the population, it should acquire the inventory already in place. If a new object is added to the domain, a new name should become adopted by the population.

Scenarios

A scenario for the Naming Game is a series of steps to be taken by both agents in order to play a game. Such scenarios should be realistic in the sense that (1) there should be no global agency dictating the agents what to do or intervening in any other way, (2) there should be no way for one agent to inspect or control the internal states of other agents, and (3) feedback should be realistic, i.e. through external behaviors only. The following diagram sketches a scenario for a successful naming game (i.e. $t_s = t_h$). The left column contains the actions of the speaker, the middle one are actions in the environment, and the right column contains the actions of the hearer. The context is initially set by the world and agents are assumed to be sensitive to the context by an attention system. The 'success' and pointing signals are non-verbal.

	context $\mathcal{C}$	
1. Speaker chooses $t_s \in \mathcal{C}$		
2. Speaker codes t_s as σ		
	$\longrightarrow \quad \sigma \quad \longrightarrow$	
		1. Hearer decodes σ as t_h
		2. Hearer points to t_h
	$\leftarrow$ pointing $\leftarrow$	
3. Speaker compares to t_s		
	$\rightarrow$ success $\rightarrow$	
4. Speaker updates $\mathcal{I}_{s,t+1}$		3. Hearer updates $\mathcal{I}_{h,t+1}$

Here is a scenario in the case of failure. There is an option for the speaker to point to the topic in order to give the opportunity to the hearer to learn a new association faster:

	context $\mathcal{C}$	
1. Speaker chooses $t_s \in \mathcal{C}$		
2. Speaker codes t_s as σ		
	$\longrightarrow \quad \sigma \quad \longrightarrow$	
		1. Hearer decodes σ as t_h
		2. Hearer points to t_h
	$\leftarrow$ pointing $\leftarrow$	
3. Speaker compares to t_s		
	$\rightarrow$ failure $\rightarrow$	
4. Speaker points to t_s		
	$\rightarrow$ pointing $\rightarrow$	
5. Speaker updates $\mathcal{I}_{s,t+1}$		3. Hearer updates $\mathcal{I}_{h,t+1}$

A Remark on Objects

The present formalisation makes a short-cut with respect to the objects in the domain. It would be more accurate to make a distinction between the set of objects in reality $\mathcal{O}_r$ and the set of objects as perceived by the speaker $\mathcal{O}_s$ or the hearer $\mathcal{O}_h$. It is well known in real world robotics that object or event recognition is extraordinarily difficult and may not yield the same results due to different spatial positions of speaker and hearer, differences in light conditions, different expectations, etc. Moreover feedback on success in the game (whether $t_s = t_h$) is not so obvious either and usually has to go through real world actions. For example, the speaker may name an object in order to obtain it and the hearer then gives the desired object or not. Often feedback is of course much less direct and success or failure may only become apparent much later. These problems are not considered here, but are of extreme importance for embodied agents playing grounded language games. This paper assumes that $\mathcal{O}_s = \mathcal{O}_h$ and that agents are able to signal extra-linguistically that $t_s = t_h$.

The same issue arises also for the context. There is no guarantee that two agents exactly share the same context, in fact that would be rare. At best there is enough overlap so that at least the topic is part of the joined context. In this paper we assume however that the context is completely shared.

Requirements

Any concrete model for the Naming Game must satisfy a number of requirements:

1. The first requirement is obviously that communicative success should increase to reach total success, particularly in a stationary population and for a stationary domain.

2. We also want efficiency of coding. Optimal coding occurs if there is a unique mapping from names to objects and vice-versa in all agents. Efficiency determines many other aspects, such as storage space required, time needed for coding or decoding, as well as viability in cultural transmission. Clearly if the system of conventions is not optimal it is more difficult to acquire.

3. The model should not only lead a population to achieve an inventory which is sufficiently shared to allow success in communication, but it

should be resilient within certain parameter boundaries to the influx and outflux of members of the population, or to changes to the domain of possible objects.

4. The model should also be to a reasonable degree noise-tolerant. Noise can occur in the signal being transmitted, the feedback received to determine success in the game, or the utilisation of internal knowledge.

3 The Basic Model

A model for the Naming Game must define the structure of the language inventory $\mathcal{I}_a$, as well as the methods by which agents use it in coding and decoding. The model must also define how agents expand or adapt the language inventory. In a first phase, we will assume that the probability that agents use (indepently) the same name for the same object is zero and so homonyms cannot arise. We also assume that the context plays no role, i.e. $\mathcal{O} = \mathcal{C}$.

The Language Inventory

The language inventory consists of a set of 3-tuples, called associations, relating an object to a name with a strength γ. The strength reflects the experience of the agent in using that particular association. Formally, $\mathcal{I}_{a,t} = \mathcal{Q} \times \mathcal{N} \times [0.0, 1.0]$ where $\mathcal{Q} \subseteq \mathcal{O}$ is the set of objects nameable by the agent, $\mathcal{N}$ is the set of names, and $0.0 \leq \gamma \leq 1.0$ is the strength of an association. An inventory can be displayed as a graph as in figure 1, which shows the inventories:
$\mathcal{I}_{a[1],t} = \{\langle o_1, n_1, 0.5\rangle, \langle o_1, n_2, 0.3\rangle, \langle o_2, n_2, 0.6\rangle\}$ (left) and
$\mathcal{I}_{a[2],t} = \{\langle o_1, n_1, 0.7\rangle, \langle o_2, n_1, 0.2\rangle, \langle o_2, n_2, 0.8\rangle\}$ (right).

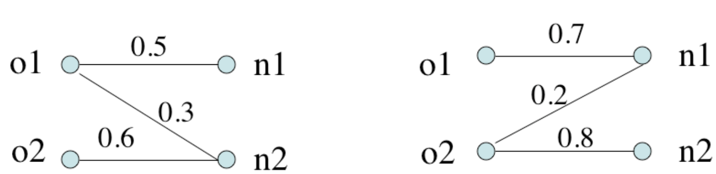

Figure 1: Example inventories for two agents, $a[1]$ (left) and $a[2]$ (right).

Decision strategies for coding and decoding

The language inventory of speaker and hearer are next transformed into a payoff matrix so that a game-theoretic framework becomes applicable for defining and analysing a game.

Given a speaker s and a hearer h then the payoff matrix $M = |\mathcal{I}_{s,t}| \times |\mathcal{I}_{h,t}|$ is constructed with a row for each $r_i \in \mathcal{I}_{s,t}$ and a column for each relation $r_j \in \mathcal{I}_{h,t}$. $M(i,j) = 1$ iff $r_i = \langle o_i, n_i, \gamma_i \rangle$, $r_j = \langle o_j, n_j, \gamma_j \rangle$ with $n_i = n_j$ and $o_i = o_j$, otherwise it is 0.

Assuming that the players want to maximise the payoff, then

- $code_{s,t}(o_k, \mathcal{C}) = n_k$ iff $r_k = \langle o_k, n_k, \gamma_k \rangle \in \mathcal{I}_{s,t}$ and $\forall r_j \in \mathcal{I}_{s,t}, r_j = \langle o_j, n_j, \gamma_j \rangle$ and $n_j \neq n_k \rightarrow \gamma_j < \gamma_k$.
- $decode_{h,t}(n_k, \mathcal{C}) = o_k$ iff $r_k = \langle o_k, n_k, \gamma_k \rangle \in \mathcal{I}_{h,t}$, $\gamma_k > 0$, and $\forall r_j \in \mathcal{I}_{h,t}, r_j = \langle o_k, n_k, \gamma_k \rangle$ and $o_j \neq o_k \rightarrow \gamma_j < \gamma_k$.

In other words, both speaker and hearer pick the 'best' strategy for playing the game, which is the strategy with the highest γ. When two associations have equal strength the choice is random. When the strength γ of an association is zero, the hearer does not consider it. u_s and u_h are the associations chosen by the speaker and hearer respectively.

An example of a pay-off matrix for the inventories
$\mathcal{I}_{a[1],t} = \{\langle o_1, n_1, 0.5\rangle, \langle o_1, n_2, 0.3\rangle, \langle o_2, n_2, 0.6\rangle\}$ and
$\mathcal{I}_{a[2],t} = \{\langle o_1, n_1, 0.7\rangle, \langle o_2, n_1, 0.2\rangle, \langle o_2, n_2, 0.8\rangle\}$ is equal to

speaker/hearer	$\langle o_1, n_1, 0.7\rangle$	$\langle o_2, n_1, 0.2\rangle$	$\langle o_2, n_2, 0.8\rangle$
$\langle o_1, n_1, 0.5\rangle$	1	0	0
$\langle o_1, n_2, 0.3\rangle$	0	0	0
$\langle o_2, n_2, 0.6\rangle$	0	0	1

then $code_{s,t}(o_1, \mathcal{C}) = n_1$ and $decode_{h,t}(n_1, \mathcal{C}) = o_1$ and so the game gets a payoff of 1 (communicative success).

Assuming a slightly different language inventory for $a[2]$:
$\mathcal{I}_{a[1],t} = \{\langle o_1, n_1, 0.5\rangle, \langle o_1, n_2, 0.3\rangle, \langle o_2, n_2, 0.6\rangle\}$ and
$I_{a[2],t} = \{\langle o_1, n_1, 0.3\rangle, \langle o_2, n_1, 0.8\rangle, \langle o_2, n_2, 0.8\rangle\}$ then the payoff matrix is

speaker/hearer	$\langle o_1, n_1, 0.3\rangle$	$\langle o_2, n_1, 0.8\rangle$	$\langle o_2, n_2, 0.8\rangle$
$\langle o_1, n_1, 0.5\rangle$	1	0	0
$\langle o_1, n_2, 0.3\rangle$	0	0	0
$\langle o_2, n_2, 0.6\rangle$	0	0	1

The players now choose: $code_{s,t}(o_1) = n_1$ and $decode_{h,t}(n_1) = o_2$ and payoff for both is 0 (communicative failure).

Updating the Language Inventory

The defining characteristic of Adaptive Language Game models is that the inventories of both speaker and hearer change as a side effect of every game. The method by which they do this is defined as follows:

There are four cases:

1. $\nexists r_i = \langle t_s, n_i, \gamma_i \rangle \in \mathcal{I}_{s,t}$ (The speaker has no name for the topic.)
In that case the speaker adds a new relation between the topic and a new name $\tilde{n}$ to his inventory, with an initial strength γ_{init} (INVENTION):
$\mathcal{I}_{s,t+1} = \mathcal{I}_{s,t} \cup \{\langle t_s, \tilde{n}, \gamma_{init} \rangle\}$.

2. $\nexists r_i = \langle t_i, \sigma, \gamma_i \rangle \in \mathcal{I}_{h,t}$ (The hearer does not know σ, the name used by the speaker.)
In that case, the hearer infers the topic t_h based on pointing by the speaker and adds a new relation to his inventory between t_h and σ, with an initial strength γ_{init}(ADOPTION):
$\mathcal{I}_{s,t+1} = \mathcal{I}_{s,t} \cup \{\langle t_h, \sigma, \gamma_{init} \rangle\}$.
The game is considered a communicative failure.

3. $t_s = t_h$ (The hearer guessed the same topic as chosen by the speaker.)
Speaker and hearer increase the strength of the association that was used with Δ_{inc} (ENFORCEMENT), and decrease the strength of competitors (lateral INHIBITION). Competitors are relations that either use another name for the same object, in which case they are decreased with Δ_{n-inh}, or that have associated another object with the same name, in which they are decreased with Δ_{o-inh}
$SpeakerUpdateI(\mathcal{I}_{s,t}, t_s, \sigma, score(NG_{s,h,t}))$ is defined as
$\mathcal{I}_{s,t+1} = \{r_i | r_i = \langle o_i, n_i, \gamma_i \rangle \in \mathcal{I}_{s,t}$ with $o_i \neq t_s$ and $n_i \neq \sigma\} \;\cup$

$$\{\langle t_s, \sigma, \gamma_i + \Delta_{inc} \rangle \text{ for } u_s = \langle t_s, \sigma, \gamma_i \rangle \in \mathcal{I}_{s,t}\} \;\cup$$

$$\{r_j | r_j = \langle t_s, n_j, \gamma_j + \Delta_{n-inh} \rangle \text{ with } r_j \neq u_s \quad \cup$$

$\{r_j | r_j = \langle o_j, \sigma, \gamma_j + \Delta_{o-inh} \rangle$ with $r_j \neq u_s$

$HearerUpdateI(\mathcal{I}_{h,t}, t_h, \sigma, score(NG_{s,h,t}))$ is defined as

$\mathcal{I}_{h,t+1} = \{r_i | r_i = \langle o_i, n_i, \gamma_i \rangle \in \mathcal{I}_{s,t}$ with $o_i \neq t_h\} \;\cup$

$\{\langle t_h, \sigma, \gamma_i + \Delta_{inc} \rangle$ for $u_h = \langle t_h, \sigma, \gamma_i \rangle \in \mathcal{I}_{h,t}\} \;\cup$

$\{r_j | r_j = \langle t_h, n_j, \gamma_j + \Delta_{n-inh} \rangle$ with $r_j \neq u_h \;\cup$

$\{r_j | r_j = \langle o_j, \sigma, \gamma_j + \Delta_{o-inh} \rangle$ with $r_j \neq u_h$

4. $t_s \neq t_h$. (The hearer guessed a different object as chosen by the speaker.)
Two things happen: Speaker and hearer decrease the strength of the association that was used with Δ_{dec} (DAMPING) and the hearer guesses an alternative meaning for σ based on guessing the meaning of the topic pointed out through non-verbal means and adds a new relation to his inventory (as in step 2 of the algorithm) unless this relation already exists. Damping leads to the following updates:

$SpeakerUpdateI(\mathcal{I}_{s,t}, t_s, \sigma, score(NG_{s,h,t}))$ is defined as
$\mathcal{I}_{s,t+1} = \{r_i | r_i = \langle o_i, n_i, \gamma_i \rangle \in \mathcal{I}_{s,t}$ with $o_i \neq t_s$ and $n_i \neq \sigma\} \;\cup$
$\{\langle t_s, \sigma, \gamma_i + \Delta_{dec} \rangle$ with $\langle t_s, \sigma, \gamma_i \rangle \in \mathcal{I}_{s,t}$

$HearerUpdateI(\mathcal{I}_{h,t}, t_h, \sigma, score(NG_{s,h,t}))$ is defined as
$\mathcal{I}_{h,t+1} = \{r_i | r_i = \langle o_i, n_i, \gamma_i \rangle \in \mathcal{I}_{h,t}$ with $o_i \neq t_h$ and $n_i \neq \sigma\} \;\cup$
$\{\langle t_h, \sigma, \gamma_i + \Delta_{dec} \rangle$ with $\langle t_h, \sigma, \gamma_i \rangle \in \mathcal{I}_{h,t}\}$

Consequently the main parameters of this Naming Game model are defined by a 5-tuple: $\langle \gamma_{init}, \Delta_{inc}, \Delta_{n-inh}, \Delta_{o-inh}, \Delta_{dec} \rangle$.

Measures

Before studying the behavior of the model it is useful to define the following measures:

The cumulative success $\mathcal{CS}_{j,n}$ for the last series of $n-j$ games at game j is defined as:

$$\mathcal{CS}_{j,n} = \frac{1}{n} \sum_{i=j,j+n} score(NG_i) \tag{1}$$

.

The non-zero strength inventory of an agent $\mathcal{I}^o_{a,j}$ contains the set of all associations with non-zero strength: $\mathcal{I}^o_{a,j} = \{r_i | r_i = \langle o_i, n_i, \gamma_i \rangle \in \mathcal{I}_{a,j}$ and $\gamma_i \neq 0\}$.

The cumulative non-zero strength inventory size $IS^a_{j,n}$ of an agent $a_t = \langle \mathcal{I}^o_{a,t} \rangle$ at game j for the last series of $n-j$ games is defined as:

$$IS^a_j = \frac{1}{n} \sum_{i=j,j+n} |\mathcal{I}^o_{a,j}| \tag{2}$$

.

The cumulative non-zero strength inventory size $\mathcal{IS}^{\mathcal{A}}_{j,n}$ for a population $\mathcal{A}$ of m agents for the last series of $n-j$ games is defined as:

$$\mathcal{I}S^{\mathcal{A}}_{j,n} = \frac{1}{|\mathcal{A}|} \sum IS^a_{j,n} \tag{3}$$

.

The non-zero strength inventory size of an agent or a population is more interesting than the number of words because the same words may be used in more than one way by the agent. Moreover if the coding is optimal, inventory size is equal to the number of objects as there is only one relation for each object.

4 Behavior of the Model

There are a number of phenomena that are observed in computer simulations of this model (see figure 2):

1. The population reaches communicative success even if $\Delta_{inc} = 0$, $\Delta_{inh} = 0$, $\Delta_{dec} = 0$. This is simply because agents invent and adopt words from each other (case 1 and 2), and so all words eventually propagate in the population. The strength of all associations stays at γ_{init}. This shows that simply taking communicative success as ultimate measure for the adequacy of a language game model is not enough. It is important to look at the efficiency of coding because this determines the time to reach success as well as viability of cultural transmission.

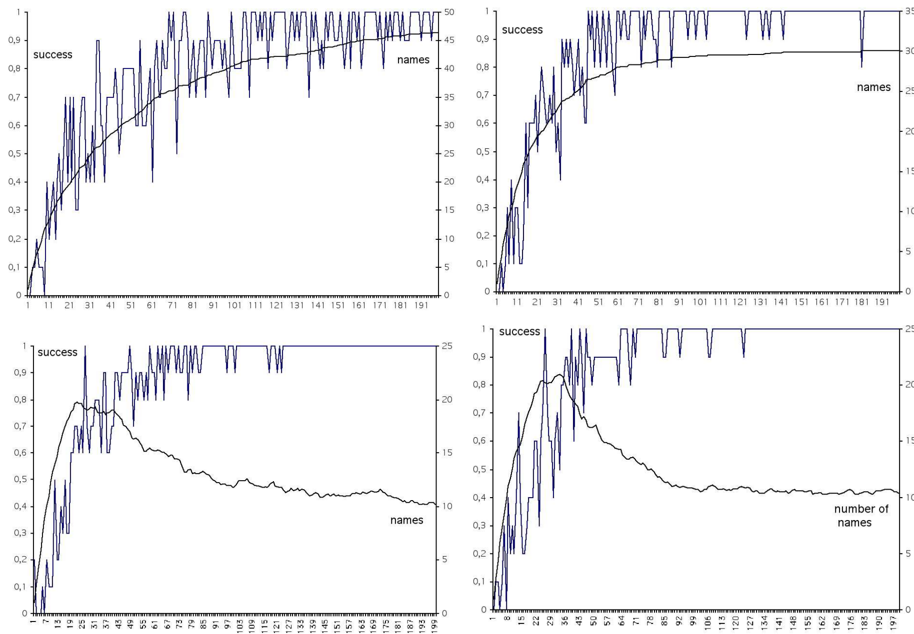

Figure 2: Behavior of the Naming Game model for different parameter settings. The size of the population $\mathcal{A}$ and the number of objects $\mathcal{O}$ is always equal to 10. $\mathcal{CS}_{j,10}$ and $\mathcal{IS}_{j,10}$ are shown for 2000 games with $\gamma_{init} = 0.5$ When this is equal to the number of objects to be expressed, the inventory is optimal. Top left shows $\Delta_{inc} = 0, \Delta_{inh} = 0, \Delta_{dec} = 0$. Top right (enforcement) $\Delta_{inc} = 0.1, \Delta_{inh} = 0, \Delta_{dec} = 0$. Bottom left (enforcement and lateral inhibition) $\Delta_{inc} = 0.1, \Delta_{inh} = 0.2, \Delta_{dec} = 0$. Bottom right (enforcement, lateral inhibition, damping) $\Delta_{inc} = 0.1, \Delta_{inh} = 0.2, \Delta_{dec} = 0.1$.

2. Making Δ_{inc} positive (ENFORCEMENT) decreases the inventories that are used by the population but the coding is still rather inefficient because the strength of several names (for the same object) tends to 1.0 and there is no force pulling them down later.

3. When Δ_{inh} becomes a factor, lateral INHIBITION starts to play a role and after an initial rise the inventories start to decrease to reach an optimal performance (the number of names is equal to the number of objects). So both enforcement and inhibition are needed to achieve an optimal repertoire.

4. When Δ_{dec} becomes a factor, DAMPING starts to work. However it has less of an effect than might be expected because the 'good' associations may get damped also. The effect is only relevant if an agent invents a new word that is unknown by another one and gets sanctioned for it. This factor therefore prevents some variation and so there is a slightly lower rate of inventory increase in the beginning. But then the negative impact starts to win over the positive one.

A summary of the evolution of the non-zero strength inventory for each parameter setting is given in figure 2.

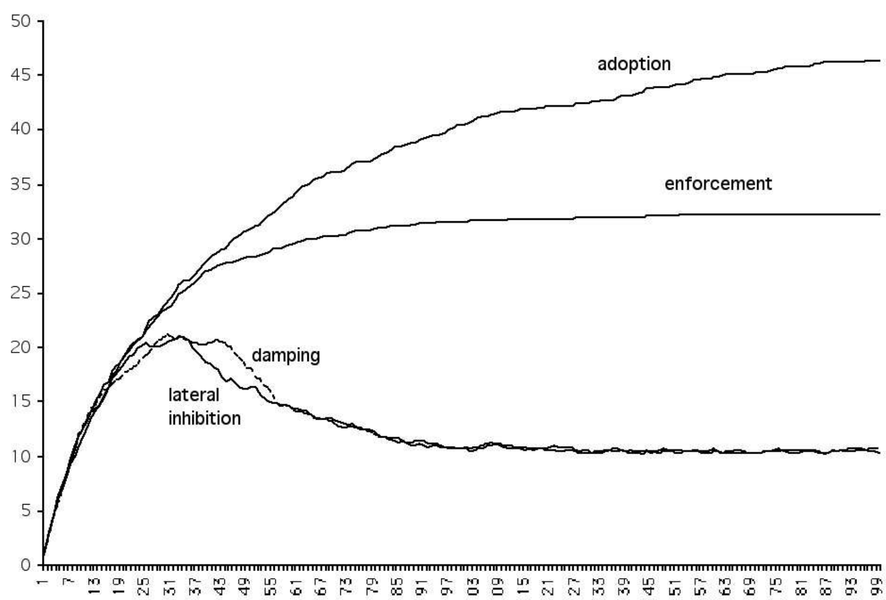

Figure 3: The evolution in average non-zero inventory size for the four parameter settings. Enforcement combined with lateral inhibition leads to the most efficient coding.

A typical example of an inventory obtained after 5000 games and 10 objects named by a population of 10 agents is given below. $\Delta_{inc} = 0.1, \Delta_{inh} = 0.2, \Delta_{dec} = 0$. An optimal coding has been reached.

```
o-1: GOPI (1.00) o-2: PURO (1.00)
o-3: NOWI (1.00) o-4: WOWE (1.00)
o-5: PIHE (1.00) o-6: MINI (1.00)
```

```
o-7: HEVE (1.00) o-8: FIFU (1.00)
o-9: DABE (1.00) o-10: KISO (1.00)
```

This mapping lists for each object the percentage of agents that uses a particular word.

Here is an example of an inventory obtained after 5000 games, 10 objects and 10 agents for $\Delta_{inc} = 0.1, \Delta_{inh} = 0.0, \Delta_{dec} = 0$. There is no pressure to settle on an optimal coding. There is still communicative success because all agents learn after a while all names.

```
o-1: VIHE (0.20) WABO (0.10) DAFA (0.27) WAKU (1.00)
o-2: WEHO (0.42) SOKU (1.00) TEWU (0.25) BUHO (1.00) LEWE (0.10)
o-3: VESU (0.10) HISI (0.32) HENU (0.15) WOMO (1.00) FETA (1.00)
o-4: HEWI (0.15) LEKA (1.00) BIRA (1.00) BABE (1.00)
o-5: SUGU (0.50) VATO (0.15) ROLO (0.30) GUDU (0.27)
     LUKU (0.25) PEKI (1.00)
o-6: HEWA (0.20) BOMI (0.30) NOBA (1.00) MAHO (0.99)
     NELI (1.00) NAVU (0.20)
o-7: NITA (0.15) RIBE (0.10) POBI (0.70) NORU (1.00) FUDA (1.00)
o-8: KUKA (0.22) MAVO (0.25) TAPE (0.27) FAPE (1.00) BAWU (1.00)
o-9: GAHA (0.15) RETA (0.39) PASO (0.49) WAHU (0.37) PIGI (1.00)
o-10: WOVO (0.15) NEVU (1.00) NUPI (0.57) GALA (0.36)
     PEDE (0.58) VUKU (0.10)
```

5 Handling Homonymy

So far it was assumed that the probability that agents use (indepently) the same name for the same object is zero and so homonyms cannot arise. We now relax that constraint and do a simulation where the number of names is fixed and equal to the number of objects (10). In this case there will be quite a few homonyms initially, because agents randomly associate names from this limited set to objects, and they have to disentangle the homonyms later to reach optimal coding. We start with the most successful parameter setting from earlier experiments: $\Delta_{inc} = 0.1, \Delta_{inh} = 0.2, \Delta_{dec} = 0$.

Results are shown in figure 4. We see that communicative success hovers around 50 %. The same name can occur more than once in the inventory explaining that the (non-zero strength) inventory size is higher than 10, even though the total number of names is 10.

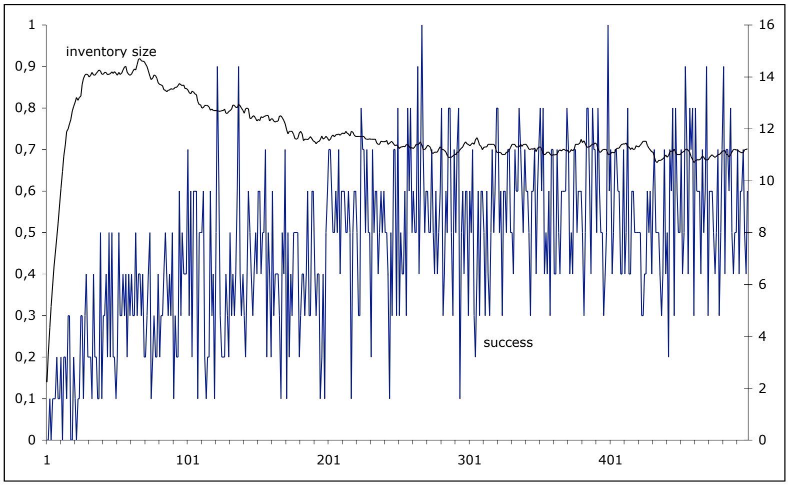

Figure 4: The evolution in average inventory size in the case of homonymy and with $\Delta dec = 0$. Total number of possible names and number of possible objects is 10. There are 10 agents in the population.

The following shows the mapping from objects to names and from names to objects obtained at the end of this experiment. The majority of objects has more than one associated name and all names except one refer to more than one object.

```
o-1: DE (1.00)
o-2: DOO (0.05) BI (0.10) BA (0.23) BOO (0.60)
o-3: BO (0.20) BE (0.70) BOO (0.10)
o-4: BOO (0.46) BE (0.30) DI (0.31)
o-5: DA (0.06) BE (0.05) DO (0.80)
o-6: BA (0.24) BO (0.80)
o-7: DA (0.05) BI (0.20) BOO (0.05) BA (0.25) DO (0.40)
o-8: DI (0.06) DA (0.90)
o-9: DOO (0.90) BA (0.07)
o-10: DI (0.21) BI (0.80)
```

```
DOO: o-2 (0.05) o-9 (0.90)
```

```
DO: o-7 (0.40) o-5 (0.80)
DI: o-10 (0.21) o-8 (0.06) o-4 (0.31)
DE: o-1 (1.00)
DA: o-7 (0.05) o-5 (0.06) o-8 (0.90)
BOO: o-4 (0.46) o-7 (0.05) o-2 (0.60) o-3 (0.10)
BO: o-3 (0.20) o-6 (0.80)
BI: o-7 (0.20) o-2 (0.10) o-10 (0.80)
BE: o-5 (0.05) o-4 (0.30) o-3 (0.70)
BA: o-6 (0.24) o-7 (0.25) o-2 (0.23) o-9 (0.07)
```

In the next experiment, the Δ_{dec} parameter is set to 0.1, which means that damping takes effect. Results are shown in figure 5. The agents now easily settle on an optimal inventory. So damping is absolutely necessary in the case of homonymy.

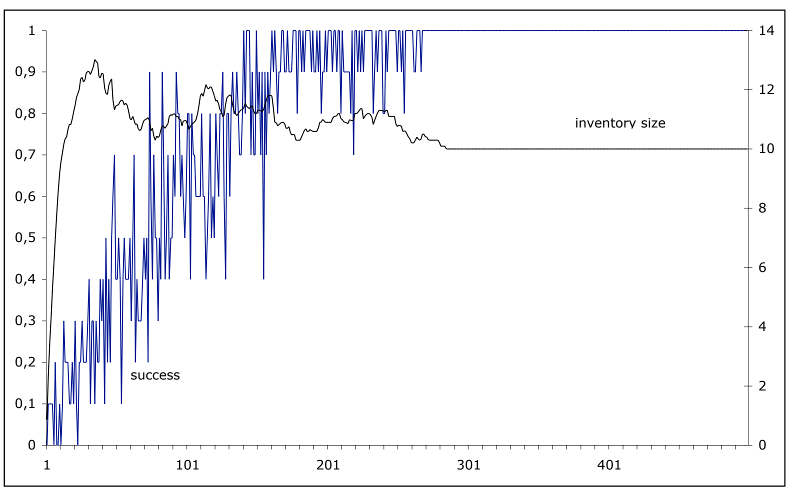

Figure 5: The evolution in non-zero strength inventory size in the case of homonymy and with $\Delta dec = 0.1$. Total number of possible names and number of possible objects is 10.

The resulting mappings are as follows:

```
DOO: o-9 (1.00)
DO: o-2 (1.00)
DI: o-5 (1.00)
DE: o-1 (1.00)
DA: o-3 (1.00)
BOO: o-4 (1.00)
BO: o-6 (1.00)
BI: o-7 (1.00)
BE: o-8 (1.00)
BA: o-10 (1.00)

o-1: DE (1.00)
o-2: DO (1.00)
o-3: DA (1.00)
o-4: BOO (1.00)
o-5: DI (1.00)
o-6: BO (1.00)
o-7: BI (1.00)
o-8: BE (1.00)
o-9: DOO (1.00)
o-10: BA (1.00)
```

We see that there is now not only communicative success but also an optimal coding.

6 The Effect of Context

Natural communication always takes place in a specific context that enormously constrains the possible set of referents, in fact it would not be doable if all the possible unique objects we know would always be possible referents. Consequently, synonyms and homonyms can be tolerated, as long as the meanings do not occur in the same contexts. For example, the word "table" has two quite different meanings in the sentence "the restaurant owner asked us to sit at another table" versus "she was using a spreadsheet but did not have all figures to fill in the table". But the multiple meanings do not get in the way of interpretation as they take place in very different contexts.

The context also helps in reaching consensus. Suppose a hearer has two possible objects o_1 and o_2 associated with the same name n. If o_1 does not occur in the present context, then only o_2 is a possible referent, even though it may have a lower strength than o_1. So the hearer may have a successful game, and learn more about the association used by the speaker without communicative failure.

In order to make use of context, the decoding behavior of the hearer needs to be adopted:

Given a speaker s and a hearer h and the payoff matrix $M = |\mathcal{I}_{s,t}| \times |\mathcal{I}_{h,t}|$ is constructed with a row for each $r_i \in \mathcal{I}_{s,t}$ and a column for each $r_j \in \mathcal{I}_{h,t}$. $M(i,j) = 1$ iff $r_i = \langle o_i, n_i, \gamma_i \rangle$, $r_j = \langle o_j, n_j, \gamma_j \rangle$ with $n_i = n_j$ and $o_i = o_j$, otherwise it is 0.

The decode function is now defined as:
$decode_{h,t}(n_k, \mathcal{C}) = o_k$ iff $r_k = \langle o_k, n_k, \gamma_k \rangle \in \mathcal{I}_{h,t}$ and $o_k \in \mathcal{O}$, and $\forall r_j \in \mathcal{I}_{h,t}, r_j = \langle o_k, n_k, \gamma_k \rangle$ and $o_j \neq o_k \rightarrow \gamma_j < \gamma_k$.

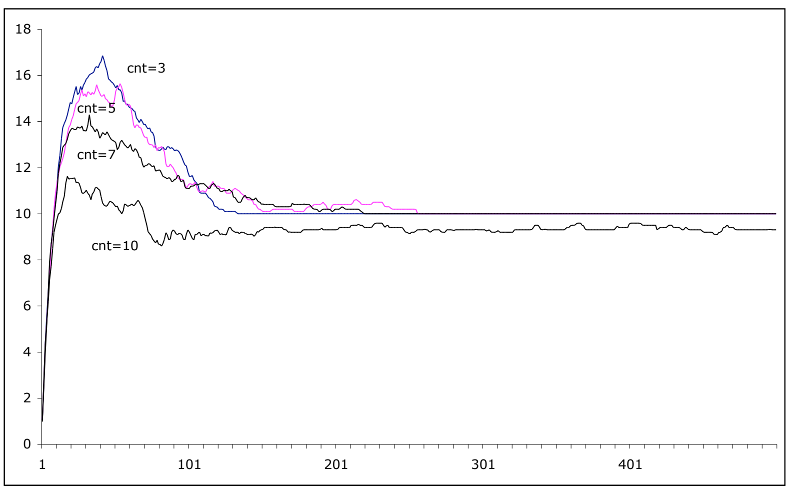

Figure 6: The effect of different context sizes in an experiment with 10 agents, 10 meanings, 5000 games, and potential for homonymy (same conditions as in 5).

The effect of the context is illustrated with a series of experiments where agents self-organise their inventories starting with the same conditions as in the previous section, namely there is a fixed set of possible names equal to the set of objects in the domain. Figure 6 shows the results of different simulations with a growing size of the context: from 3, 5, 7, to 10 (maximum number). We see that a context of size 7 causes the quickest convergence to an optimal coding. In the case of a context of size 10, the agents use much more homonymy and have not reached the optimal coding after 5000 games. With a context of 3, there are the most synonyms initially but they get resolved quickly.

7 Other Issues

There are many aspects of the Naming Game that could be investigated further (and many of them have already been investigated in various computer simulations). We just look briefly at the behavior of the Naming Game as an open system using enforcement, lateral inhibition, and damping: $\Delta_{inc} = 0.1, \Delta_{inh} = 0.2, \Delta_{dim} = 0.1$ and with no limitation on vocabulary size.

An Agent Learns an Existing Inventory

A first experiment (figure 7) shows that a new agent entering the population easily acquires the existing inventory. A population of 10 agents first evolved an inventory (the first 5000 games) and then a new agent is introduced without any knowledge of the existing set of conventions.

This experiment shows that the model is adequate for handling the influx of a new agent. The acquisition of a lexicon by the agent does not require any kind of change to the mechanisms that are already in the Naming Game model.

Flow in the population

Next the set of agents $\mathcal{A}$ is made open in the sense that either new members enter at a certain rate θ_{in} or leave at the rate θ_{out}, where both can be positive. The rate is defined in terms of number of games. $\theta_{in} = 1/10$ means a new agent every 10 games.

The effect of a steady influx of agents starting from a population of 2 and with $\theta_{in} = 1/1000$ is shown in figure 8. The communication system remains

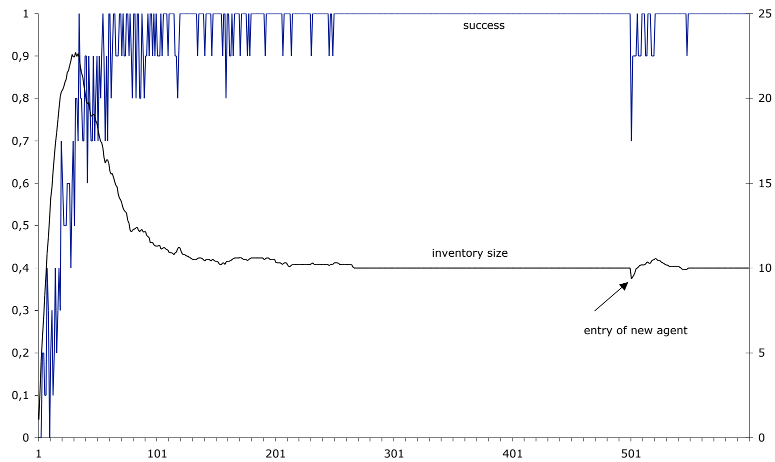

Figure 7: The first 5000 games, a population of 10 agents self-organises an inventory for 10 objects. Then a new agent is introduced in the population, giving a glitch both in communicative success and coding optimality.

intact and is steadily propagating to the new members (even if they occasionally invent new words which do not propagate much in the population). Each time a new agent comes in, communicative success drops and the number of non-zero associations temporarily goes up, but then the system gets back to an optimal state with steady communicative success. As the number of objects grows, it will be harder to establish an optimal coding for a new object, simply because there is less chance that this new object is chosen as the topic in a game.

Flow rate in the domain

The set of objects in the domain $\mathcal{O}$ can also be open in the sense that either new objects can become relevant at a certain rate ρ_{in} or become irrelevant at the rate ρ_{out} where both can be positive. The rate is again defined in terms of number of games. $\rho_{in} = 1/10$ means a new object every 10 games. An

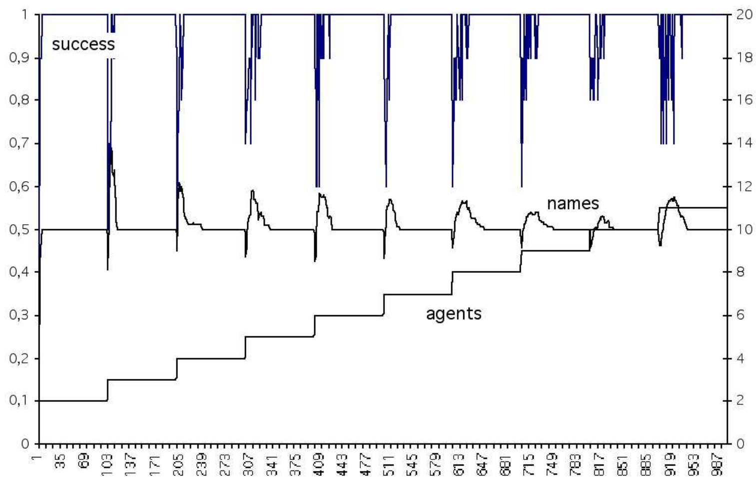

Figure 8: The evolution in communicative success and non-zero strength inventory size for a growing population starting from 2 with rate $\theta_i n = 1/1000$. 10,000 games are shown.

example of a run with the model for 10 agents, starting from an empty domain and with $\rho_{in} = 1/500$ is given in figure 9. We see that the population can initially cope well. After each disturbance the number of objects gets back to the optimal value. As there are more and more objects, the probability that they are chosen as a topic diminishes, there is no longer enough time to spread words for them and the population can no longer cope with the influx.

The inventory at the end of this run is shown below. The most recent objects have still several competing names whereas the earlier objects have unique names.

```
o-1: GEKO (1.00)
o-2: BEKI (1.00)
o-3: GALI (1.00)
o-4: NISA (1.00)
o-5: DOHO (1.00)
o-6: PABE (1.00)
o-7: KIFO (1.00)
o-8: REHI (1.00)
o-9: FERE (0.98)
```

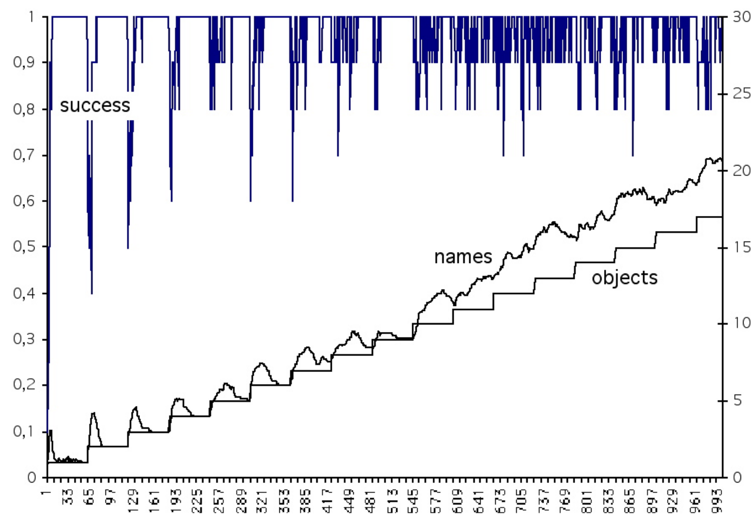

Figure 9: The evolution in communicative success and non-zero strength inventory size for a population of 10 agents and a domain $\mathcal{O}$ growing at the rate $\rho_{in} = 1/600$. 10,000 games are shown.

```
o-10: PIBA (1.00)
o-11: DATI (1.00)
o-12: HOTO (1.00)
o-13: SAHI (1.00)
o-14: MUMI (1.00)
o-15: HERA (1.00)
o-16: BABE (0.05) VILA (0.84)
o-17: GASI (0.20) RERO (0.56)
o-18: BATO (0.99)
o-19: TITA (0.06) PUHE (0.06) LOBU (0.08) MEBU (0.81)
o-20: VURE (0.02) POGU (0.30) LEKA (0.45) NUTA (0.13) PEPI (0.11)
```

We refer to [8] for explorations of spatially distributed naming games and [7] for explorations how stochasticity affects the behavior of the emerging communication system.

8 Theoretical Challenges

We have presented in this paper the general problem that the Naming Game tries to solve (section 2), namely how can a population of agents arrive at a shared set of conventions for naming objects in their environment, and then a concrete model (section 3) that has been shown in computational simulations to (section 4 and 5) to be up to the task, even if all dimensions of the system are made open (section 6). It would be desirable to construct a theoretical approach to semiotic dynamics, in terms of theorems that show that the model presented is indeed capable to satisfy its requirements.

More concretely, the following theoretical challenges can then be posed:

1. *Communicative success.* The first step is to show that for the proposed agent behavior, communicative success increases and goes to hundred procent.

2. *Coherence.* A second theoretical challenge is to show that the system is optimal in the sense that the number of This will imply showing that there is a tendency to eliminate strict synonyms (more than one name for exactly the same object) and strict homonyms (more than one object for the same name). Both tendencies are observed in human natural languages.

3. *Population size.* What is the effect of population size on reaching coherence? It will obviously take longer to reach communicative success and coherence when the population is larger. Are there any limits?

4. *Object set size.* What is the effect of object size on reaching coherence and hence communicative success? Presumably, the more objects the longer it may take to develop a functioning communication system. Are there any limits?

5. *Population change.* What is the effect of population change? When new agents (without any inventory) enter the population, they have to learn the inventory in place. When agents leave the population some of their knowledge is lost. Will the population be able to culturally transmit the inventory? Are there limits to the rate of change?

6. *Object set change.* What happens if the set of objects is changing, so that new objects become added to $\mathcal{O}$ and others disappear? Are agents

able to expand their lexicon to cope with the change? Do names which have become irrelevant disappear, as they do in human languages?

7. *Transmission robustness.* What if there is some error rate in transmission of the word string, so that mismatches are possible? Will a coherent lexicon nevertheless form and maintain itself?

8. *Feedback robustness.* What if there is some error rate in feedback, i.e. in information on whether the game succeeded or failed? Will a coherent lexicon nevertheless form and maintain itself?

9. *Cognitive robustness.* What happens if there is an error rate or stochasticity in the behavior of the agents themselves?

10. *Role of the context.* The Naming Game incorporates the context as part of the definition. This means for example that an object can have more than one name depending on the context, or that the context can help to guess the intended topic. In other words, it is of interest to study the effect of context restriction on success and coherence, particularly when no context is considered, i.e. when for all games $\mathcal{C} = \mathcal{O}$, or conversely when the context contains only the topic.

11. *Population structure.* Obviously a structure in the population which determines the probability with which certain members interact with each other will have an impact on the evolving communication system.

12. *Drift.* A final important topic is to investigate under what circumstances there might be a change in the communication system purely driven by internal dynamics (as indeed seems to happen in natural languages), in other words without population change or object change.

9 Comparison to Other Work

Research on how a population of agents could develop a shared lexicon has been going on since the early nineties, pioneered by Hurford [1](see an early review in [6]). These investigations can be classified according to many different dimensions, but the following distinctions are useful:

1. One class of models takes immediately a **population approach** The language of the total population is characterised by two matrices which contains the probability that an agent uses a name for an object (in coding) or

interprets a name as referring to an object (in decoding) [1], [3], [2]. The potential for success of communication (sometimes called communicative accuracy) of an individual agent against all others or for the total population can then be calculated easily. Another class of models, including the one presented in this paper, takes an **agent-based approach**. Individual agents each with their own inventories play language games with other agents and communicative success and coherence are now observational measures. The two classes of models can be related to each other by computing the probability matrices for coding and decoding for the total population and use it to make theoretical estimates of success in communication, but there are many subtleties involved, particularly if we want to model learning processes which depend on what language samples are available to the learner. The population-based approach is typical for researchers interested in game-theoretic studies, whereas the agent-based approach is naturally adopted by those who want to develop real world applications, such as software agents or robots evolving shared communication systems.

2. Another dimension for classifying language models concerns the use of iterated transmission versus self-organisation as the main mode in which a population arrives at a shared inventory. In the **iterated transmission** approach, changes to the language are assumed to happen when one population of learners (or children) acquire the language of an earlier population of teachers (or parents). The teacher/learner mode is characteristic of iterated learning models (recently reviewed in [4]) and the parent/child mode is characteristic for genetic transmission models [2]. In the latter case, the inventory of each agent is encoded in terms of a 'genome', communicative accuracy is interpreted as the fitness function of an agent, and mutation and recombination determine how populations evolve over time. In the **self-organisation** approach, illustrated by the model presented in this paper, each agent has his own local inventory and adapts this inventory after every game. The shared inventory emerges as a side effect of the cumulation of these adaptations. There might be a flow in the population with new members entering or leaving (as in section 7) but this flow is not necessary to cause the creation of structure.

3. The next important dimension concerns what agents use as information to update their inventories. In the **observational learning** approach, agents use only positive evidence and attempt to induce the statistical properties of sampled verbal interactions. This approach is sometimes claimed to be more realistic with respect to human language learning. In the **active learning**

approach, agents use positive as well as negative evidence. For example, when a speaker names an object, with the goal of obtaining it, and the hearer hands over a different object, both of them can learn from this negative experience: the speaker learns that the hearer perhaps uses another name (or did not have one yet) and the hearer learns that one of his associations is incorrect. The model of the Naming game presented earlier uses an active mode of learning as soon as $\Delta_{dec} > 0$. We have seen moreover that this is crucial to handle the damping of homonymy.

10 Conclusions

The main objective of this paper was to introduce the Naming Game in a formal way and show the behavior of its main parameters based on computer simulations. Given these definitions it is now easier to study the game in an analytic fashion.

11 Acknowledgement

Research was funded by Sony CSL, ECAgents, and the ESF OMLL project.

References

[1] Hurford, J. (1989) Biological evolution of the Saussurean sign as a component of the language acquisition device. Lingua 77, 197-222.

[2] Nowak, M. A. and Krakauer, D. (1999). The evolution of language. Proceedings of the National Academy of Sciences of the United States of America, 96(14):8028–8033

[3] Oliphant, M. (1997). Formal Approaches to Innate and Learned Communication: Laying the Foundation for Language. PhD thesis, Department of Cognitive Science, University of California, San Diego.

[4] Smith, K. (2004) The evolution of vocabulary. Journal of Theoretical Biology. 228(1), p127-142

[5] Steels, L. (1996) Emergent Adaptive Lexicons. In: Maes, P. (ed.) (1996) *From Animals to Animats 4: Proceedings of the Fourth International Conference On Simulation of Adaptive Behavior*, The MIT Press, Cambridge Ma.

[6] Steels, L. (1997). The synthetic modeling of language origins. Evolution of Communication, 1(1):1–3

[7] Steels, L. and Kaplan, F. (1998). Stochasticity as a source of innovation in language games. In C. Adami and R. Belew and H. Kitano and C. Taylor, editors, Artificial Life VI. Los Angeles: MIT Press

[8] Steels, L. and McIntyre, A. (1999). Spatially Distributed Naming Games. Advances in Complex Systems, 1(4):301–323

[9] Steels, L. (2003) Evolving grounded communication for robots. Trends in Cognitive Science. Volume 7, Issue 7, July 2003 , pp. 308-312.