A Bayesian Approach to the Naming Game Model

Gionni Marchetti, Marco Patriarca,* and Els Heinsalu
NICPB – National Institute of Chemical Physics and Biophysics, Tallinn, Estonia

Abstract. We present a novel Bayesian approach to semiotic dynamics, which is a cognitive analogue of the naming game model restricted to two conventions. The one-shot learning that characterizes the agent dynamics in the basic naming game is replaced by a word-learning process, in which agents learn a new word by generalizing from the evidence garnered through pairwise-interactions with other agents. The principle underlying the model is that agents — like humans — can learn from a few positive examples and that such a process is modeled in a Bayesian probabilistic framework. We show that the model presents some analogies but also crucial differences with respect to the dynamics of the basic two-convention naming game model. The model introduced aims at providing a starting point for the construction of a general framework for studying the combined effects of cognitive and social dynamics.

Keywords: Complex Systems, Language Dynamics, Bayesian Statistics, Cognitive Models, Consensus Dynamics, Semiotic Dynamics, naming game, Individual-Based Models

*marco.patriarca@kbfi.ee
I. INTRODUCTION

A basic question in complexity theory is how the interactions between the units of the system lead to the emergence of ordered states from initially disordered configurations [1, 2]. This general question concerns phenomena ranging from phase transitions in condensed matter systems and self-organization in living matter to the appearance of norm conventions and cultural paradigms in social systems. Various models were used in order to study social interactions and cooperation, e.g. models of condensed matter systems (such as spin systems), statistical mechanical models (e.g. based on the master equation), ecological competition models [1], many-agents game-theoretical models [3–5]. Opinion dynamics and cultural spreading models represent suitable theoretical frameworks for a quantitative description of the emergence of social consensus [2].

In this respect, the emergence of human language remains a challenging, multi-fold question, related in turn to biological, ecological, social, logical, and cognitive aspects [6–10]. Language dynamics [11, 12] has provided models describing phenomena of language competition and change that focus on the mutual interactions of linguistic traits (sounds, phonemes, grammatical rules, or languages understood as fixed entities) under the influence of ecological and social factors, modeling such interactions in analogy to biological competition and evolution.

However, the basic learning process of a word has a complex dynamics due to its cognitive dimension. In fact, learning a word means to learn a concept (understood as a pointer to a subset of objects, see Refs. [10, 13, 14]) and a linguistic label —for example the name of the object— used for communicating the concept. The double concept↔name nature of words has been studied through semiotic dynamics models, such as the models of Hurford [15] and Nowak [16] (see also [17, 18]) and the naming game (NG) model [19, 20].

In the basic version of the model of Nowak [16], the language spoken by each agent \( i (i = 1, \ldots, N) \) is defined by two personal matrices, representing the links of a bipartite network joining \( Q \) names and \( R \) concepts: (1) an active matrix \( U^{(i)} \) representing the concept → name links, where the element \( U^{(i)}_{q,r} \) \((q \in (1,Q), r \in (1,R))\) gives the probability that agent \( i \) will utter the \( q \)th name to communicate the \( r \)th concept; (2) the passive matrix \( H^{(i)} \), representing the name → concept links, in which the element \( H^{(i)}_{q,r} \) represents the probability that an agent interprets the \( q \)th name as referring to the \( r \)th concept. In the models of Hurford and in the model of Nowak, the languages of each individual evolve with time according to a game-theoretical dynamics, with agents gaining a reproductive advantage if their matrices have a higher communication efficiency. These studies have achieved interesting results, such as the emergence of non-ambiguous one-to-one links between objects and sounds, and explain why homonyms are more frequent than synonyms [15–18].

In the NG model [19, 20] there is only one concept \((R = 1)\) that can be linked to a set of \( Q > 1 \) different names. The model can be reformulated through the agents’ lists \( L_i \) of the name↔concept connections known to each agent \( i \). In the case of two-conventions models, where the conventions are the names \( A \) and \( B \), the list of the \( i \)th agent can be \( L_i = \emptyset \) (no connection), \( L_i = (A) \) or \( (B) \) (one name is known), or \( L_i = (A,B) \) (both name↔concept connections are known).

Extending semiotic dynamics models is not trivial and already two-opinion variants of the NG model, taking into account committed groups, show a remarkable phase diagram [21]; and trying to describe actual cognitive effects requires entirely new features [22]. This paper presents a minimal model to study the interplay of the cognitive and social dynamical dimensions, assuming for simplicity the two-conventions NG model as a semiotic framework [20, 23] and making a cognitive generalization within the experimentally validated Bayesian framework of [10] (see also Refs. [13, 14, 24–27]). In that framework, an individual can learn a concept from a small number of examples, a most remarkable feature of human learning [10, 28, 29], to be contrasted with machine learning algorithms, which require a large amount of examples for generalizing successfully [30–32].

The paper is organized as follows. The new model is introduced in Sec. II. In Sec. III, we present and discuss the features of the semiotic dynamics emerging from the numerical simulations and quantitatively compare them with those of the two-conventions NG model. Future directions in the study of the interplay of the cognitive and the social dynamics are outlined in Sec. V.

II. A BAYESIAN LEARNING APPROACH TO THE NAMING GAME

A. The two-conventions naming game model

Before introducing the new model, we recall the basic 2-conventions NG model [33], in which there is a single concept \( C \), corresponding to an external object, and two possible names (synonyms) \( A \) and \( B \) for referring to \( C \). Thus, the possibility of homonymy is excluded [23]. Each agent \( i \) is equipped with the list \( L_i \) of the names known to the agent. We assume that at \( t = 0 \) each agent \( i \) knows either \( A \) or \( B \) and has therefore a list \( L_i = (A) \) or \( L_i = (B) \),
respectively. During a pair-wise interaction, an agent can act as a speaker, when conveying a word to another agent, or as a hearer, when receiving a word from a speaker. One can think of an agent conveying a word as uttering a name, e.g. $A$, while pointing at an external object, corresponding to concept $C$: thus, the hearer records not only the name $A$ but also the name↔concept association between $A$ and $C$. At a later time $t > 0$, the list $L_i$ of the $i$th agent can contain one or both names, i.e., $L_i = (A), (B), \text{ or } (A, B)$.

The system evolves according to the following update rules [23]:

1. Two agents $i$ and $j$, the speaker and the hearer, respectively, are randomly selected.
2. The speaker $i$ randomly extracts a name (here either $A$ or $B$) from the list $L_i$ and conveys it to the hearer $j$. Depending on the state of agent $j$, the communication is usually described as:
   
   (a) Success: the conveyed name is present also in the hearer’s list $L_j$, i.e. also agent $j$ knows its meaning; then the two agents erase the other name from their lists, if present.
   
   (b) Failure: the conveyed name is not present in the hearer’s list $L_j$; then agent $j$ records and adds it to the list $L_j$.
3. Time is increased of one step, $t \to t + 1$, and the simulation is reiterated from the first point above.

An example of unsuccessful and one of successful communication are schematized in the left panel (A) of Fig. 1, see Ref. [19] for more examples. Despite its simple structure, the basic NG model describes the emergence of consensus about which name to use, which is reached for any (disordered) initial configuration [34].

![Basic Naming Game](A)

![Bayesian Naming Game](B)

**FIG. 1.** Comparison of the basic and Bayesian NG model. Panel (A): basic 2-conventions NG model. In a communication failure (upper figure), the name conveyed, $B$ in the example, is not present in the list of the hearer, who adds it to the list. In a communication success (lower figure), the word $B$ is already present in the hearer’s list and both agents erase $A$ from their lists.

Panel (B): Bayesian NG model; in order to convey an example “+” to the hearer in association with name $A$, the speaker must have already generalized concept $C$ in association with $A$, represented here by the label $[A]$. In a communication failure (upper figure), the hearer computes the Bayes probability $p$ and the result is a $p < 1/2$; then the only outcome is that the hearer records the example (reinforcement). In the Bayesian NG, there are two ways, in which the communication can be successful. The first way (lower figure) is when $p \geq 1/2$: the hearer generalizes $C$ in association with $A$ and attaches the label $[A]$ to the inventory. The second way (not shown) is the agreement process, analogous to that of the basic NG, when both agents had already generalized concept $C$ in association with name $A$ and remove label $[B]$ from the inventory $[+ + + \ldots]_B$ if present. See text for further details.

**B. Toward a Bayesian naming game model**

From a cognitive perspective, a “communication failure” of the NG model can be understood as a learning process, in which the hearer learns a new word. It is a “one-shot learning process”, because it takes place instantaneously (in a single time step) and independently of the the agent’s history (i.e. of the previous knowledge of the agent).
However, modeling an actual learning process should take into account the agents’ experience, based on the previous observations (the data already acquired) as well as the uncertain/incomplete character naturally accompanying any learning process.

Here, the one-shot learning is replaced by a process that can describe basic but realistic situations, such as the prototypical “linguistic games” [35]. For example, consider a “lecture game”, in which a lecturer (speaker) utters the name $A$ of an object and shows a real example “$+$” of the object to a student (hearer), repeating this process a few times. Then, the teacher can e.g. (a) show another example and ask the student to name the object; (b) utter the same name and ask the student to show an example of that object; or (c) do both things (uttering the name and showing the object) and ask the student whether the name$\leftrightarrow$object correspondence is correct. The student will not be able to answer correctly if not after having received some examples, enabling the student to generalize the concept $C$ corresponding to the object in association to name $A$. To model these and similar learning processes, we need a criterion enabling the hearer to assess the degree of equivalence between the new example and the examples recorded previously.

The starting point for the replacement of the one-shot learning is Bayes’ theorem. According to Bayes’ theorem, the posterior probability $p(h|X)$ that the generic hypothesis $h$ is the true hypothesis, after observing a new evidence $X$, reads [36, 37],

$$p(h|X) = \frac{p(X|h) p(h)}{p(X)} . \tag{1}$$

Here, the prior probability $p(h)$ gives the probability of occurrence of the hypothesis $h$ before observing the data and $p(X|h)$ gives the probability of observing $X$ if $h$ is given. Finally, $p(X)$ gives the normalization constraint; in the applications it can be evaluated as $p(X) = \sum_{h'} p(X|h') p(h')$, where \{h$'$\} $\in$ $H$ represents the set of hypotheses, within the hypothesis space $H$.

The next step is to find a way to compute explicitly the posterior probability $p(h|X)$, through a representation of the concepts and their relative examples in a suitable hypothesis space $H$ of the possible extensions of a given concept $C$, constituted by the mutually exclusive and exhaustive hypotheses $h$. Following the experimentally verified Bayesian statistical framework of Refs. [10, 28], we adopt the paradigmatic representation of a concept as a geometrical shape. For example, the concept of “healthy level” of an individual in terms of the levels of cholesterol $x$ and insulin $y$, defined by the ranges $x_a \leq x \leq x_b$ and $y_a \leq y \leq y_b$, where $x_i$ and $y_i$ ($i = a, b$) are suitable values, represents a rectangle in the Euclidean $x$-$y$ plane $\mathbb{R}^2$. Examples of healthy levels of specific individuals 1, 2, ... correspond to points $(x_1, y_1), (x_2, y_2), \ldots \in \mathbb{R}^2$. In the following, we assume that a hypothesis $h$ is represented by a rectangular region in $\mathbb{R}^2$. Figure 2 shows four positive examples, denoted by the symbol “$+$”, associated to four different points of the plane, consistent with (i.e. contained in) three different hypotheses, shown as rectangles.

![FIG. 2. Three different hypotheses represented as axis-parallel rectangles in $\mathbb{R}^2$ and four positive examples “$+$” that are all consistent with the three hypotheses. The set of all the rectangles that can be drawn in the plane constitutes the hypothesis space $H$.](image-url)

The problem of learning a word is now recast into an equivalent problem, consisting in acquiring the ability to
infer whether a new example \( z \) recorded, corresponding to a new point “+” in \( \mathbb{R}^2 \), corresponds to the concept \( C \), after having seen a small set of positive examples “+” of \( C \). More precisely, let \( X = \{(x_1, y_1), \ldots, (x_n, y_n)\} \) be a sequence of \( n \) examples of the true concept \( C \), already observed by the hearer, and \( z = (z_1, z_2) \) the new example. The learner does not know the true concept \( C \), i.e. the exact shape of the rectangle associated to \( C \), but can compute the generalization function \( p(z \in C|X) \) by integrating the predictions of all hypotheses \( h \), weighted by their posterior probabilities \( p(h|X) \):

\[
p(z \in C|X) = \int_{h \in H} p(z \in C|h) p(h|X) \, dh. \quad (2)
\]

Clearly, \( p(z \in C|X) = 1 \) if \( z \in h \) and 0 otherwise. By means of the Bayes’ theorem (1), one can obtain the right Bayesian probability for the problem at hand. A successful generalization is then defined quantitatively by introducing a threshold \( p^* \), representing an acceptance probability: an agent will generalize if the Bayesian probability \( p(z \in C|X) \geq p^* \). The value \( p^* = 1/2 \) is assumed, as in Ref. [28].

We assume that an Erlang prior characterizes the agents’ background knowledge. For a rectangle in \( \mathbb{R}^2 \) defined by the tuple \((l_1, l_2, s_1, s_2)\), where \( l_1, l_2 \) are the Cartesian coordinates of its lower-left corner and \( s_i \) its sides along dimension \( i = 1, 2 \), the Erlang prior density is [10, 28]

\[
p_E = s_1s_2 \exp \left\{ - \left( \frac{s_1}{\sigma_1} + \frac{s_2}{\sigma_2} \right) \right\}, \quad (3)
\]

where the parameters \( \sigma_i \) represent the actual sizes of the concept, i.e. they are the sides of the concept rectangle \( C \) along dimension \( i \). The choice of a specific informative prior, such as the Erlang prior, is well motivated by the fact that in the real world individuals have always some prior knowledge or expectation. In fact, a Bayesian learning framework with an Erlang prior of the form (3) well describes experimental observations of learning processes of human beings [28]. The final expression used below for computing the Bayesian probability \( p \) that, given the set of previous examples \( X \), the new example \( z \) falls in the same category of concept \( C \), reads [28]

\[
p(z \in C|X) \approx \frac{\exp \left\{ - \left( \frac{d_1}{\sigma_1} + \frac{d_2}{\sigma_2} \right) \right\}}{\left( 1 + \frac{d_1}{\sigma_1} \right)^2 \left( 1 + \frac{d_2}{\sigma_2} \right)^2}. \quad (4)
\]

Here \( r_i \) (\( i = 1, 2 \)) is an estimate of the extension of the set of examples along direction \( i \), given by the maximum mutual distance along dimension \( i \) between the examples of \( X \); \( d_i \) measures an effective distance between the new example \( z \) and the previously recorded examples, i.e., \( d_i = 0 \) if \( z_i \) falls inside the value range of the examples of \( X \) along dimension \( i \), otherwise \( d_i \) is the distance between \( z \) and the nearest example in \( X \) along the dimension \( i \). Equation (4) is actually a “quick-and-dirty” approximation that is reasonably good, except for \( n \leq 3 \) and \( r_i \leq \sigma/10 \), estimating the actual generalization function within a 10% error, see Refs. [10, 28] for details. Despite these approximations, Eq. (4) will ensure that our computational model, described in the next section, retains the main features of the Bayesian learning framework. It is to be noticed that for the validity of the Bayesian framework, it is crucial that the examples are drawn randomly from the concept (strong sampling assumption), i.e. they are extracted from a probability density that is uniform in the rectangle corresponding to the true concept [28]. This definition of generalization is now applied below to word-learning.

### C. The Bayesian word-learning model

Based on the Bayesian learning framework discussed above, in this section we introduce a minimal Bayesian individual-based model of word-learning. For the sake of clarity, in analogy with the basic NG model, we study the emergence of consensus in the simple situation, in which two names \( A \) and \( B \) can be used for referring to the same concept \( C \) in pair-wise interactions among \( N \) agents.

At variance with the NG model, here in each basic pair-wise interaction an agent \( i \), acting as a speaker, conveys an example “+” of concept \( C \), in association with either name \( A \) or \( B \), to another agent \( j \), who acts as hearer \((i, j = 1, \ldots, N)\). In order to be able to communicate concept \( C \) uttering a name, e.g. name \( A \), the speaker \( i \) must have already generalized concept \( C \) in association with name \( A \). This is signalled by the presence of name \( A \) in the list \( L_i \). On the other hand, the hearer \( j \) always records the example received in the respective inventory, in the example the inventory \([+ + + \ldots]_A \).

The state of a generic agent \( i \) at time \( t \) is defined by
• the list \( L_i \), to which a name is added whenever agent \( i \) generalizes concept \( C \) in association with that name; agent \( i \) can use any name in \( L_i \) to communicate \( C \);

• two inventories \([+++]_A\) and \([+++]_B\), containing the examples “+” of concept \( C \) received from the other agents in association with name \( A \) and \( B \), respectively.

It is assumed that initially each agent knows one word: a fraction \( n_A(0) \) of the agents know concept \( C \) in association with name \( A \) and the remaining fraction \( n_B(0) = 1 - n_A(0) \) in association with name \( B \) — no agent knows both words, \( n_{AB}(0) = 0 \). We will examine three different initial conditions:

- symmetric initial conditions (SIC):
  \[ n_A(0) = n_B(0) = 0.5 \]

- asymmetric initial conditions (AIC):
  \[ n_A(0) = 0.3, \quad n_B(0) = 0.7 \]

- reversed case of AIC (AICr):
  \[ n_A(0) = 0.7, \quad n_B(0) = 0.3 \]

Initially, each agent \( i \), within the fraction \( n_A(0) \) of agents that know name \( A \), is assigned \( n_{ex,A} = 4 \) examples “+” of concept \( C \) in association with name \( A \), but no examples in association with the other name \( B \), so that agent \( i \) has an \( A \)-inventory \([+++]_A\) and an empty \( B \)-inventory \( [\cdot]_B \). The complementary situation holds for the other agents that know only name \( B \), who initially receive \( n_{ex,B} = 4 \) examples of concept \( C \) in association with name \( B \) but none in association with \( A \). This choice, somehow arbitrary, is dictated by the condition that Eq. (4) becomes a good approximation for \( n > 3 \) \cite{10}.

Examples are points uniformly generated inside the fixed rectangle corresponding to the true concept \( C \), here assumed to be a rectangle with lower left corner coordinates \((0,0)\) and sizes \( \sigma_1 = 3 \) and \( \sigma_2 = 1 \) along the \( x \) and \( y \) axis, respectively. Results are independent of the assumed numerical values; in particular, no appreciable variation in the convergence times \( t_{conv} \) is observed as the rectangle area is varied, which is consistent with the strong sampling assumption, on which the Bayesian learning framework rests; see Ref. \cite{10} and Sec. III.

Furthermore, we introduce an element of asymmetry between the names \( A \) and \( B \), related to the word-learning process: different minimum numbers of examples \( n_{ex,A}^* = 5 \) and \( n_{ex,B}^* = 6 \) will be used, which are needed by agents to generalize concept \( C \) in association with \( A \) and \( B \), respectively. This is equivalent to assume that concept \( C \) is slightly easier to learn in association with name \( A \) than \( B \). Such an asymmetry plays a relevant role in the model dynamics in differentiating the Bayesian generalization functions \( p_A \) and \( p_B \) from each other, see Sec. IV.

The dynamics of the model can be summarized by the following update rules:

1. A pair of agents \( i \) and \( j \), acting as speaker and hearer, respectively, are randomly chosen among the agents.

2. The speaker selects randomly: (a) a name from the list \( L_i \) (or selects the name present if \( L_i \) contains a single name), for example \( A \) (analogous steps follow if the word \( B \) is selected); (b) an example \( z \) among those contained in the corresponding inventory \([+++]_A \) — :
   then the speaker \( i \) conveys the example extracted \( z \) in association with (e.g. uttering) the name selected \( A \) to the hearer \( j \).

3. The hearer adds the new example \( z \) (in association with \( A \)) to the inventory \([+++]_A \). This reinforcement process of the hearer’s knowledge always takes place.

4. Instead, the next step depends on the state of the hearer:

   (a) Generalization. If the selected name, \( A \) in the example, is not present in the hearer’s list \( L_j \), then the hearer \( j \) computes the relative Bayesian probability \( p_A = p(z \in C|X_A) \) that the new example \( z \) falls in the same category of concept \( C \), using the examples previously recorded in association with \( A \), i.e. from the set of examples \( X_A \in [+++]_A \). If \( p_A \geq 0.5 \), the hearer has managed to generalize concept \( C \) and connects the inventory \([+++]_A \) to \( A \); this is done by adding name \( A \) to the list \( L_j \). Starting from this moment, agent \( j \) can communicate concept \( C \) to other agents by conveying an example taken from the inventory \([+++]_A \) while uttering the name \( A \). If \( p_A < 0.5 \), the hearer has not managed to generalize the concept and nothing more happens (the reinforcement of the previous point is the only event taking place).

   (b) Agreement. The name uttered by the speaker, \( A \) in the example, is present in the hearer’s list \( L_j \), meaning that that agent \( j \) has already generalized concept \( C \) in association with name \( A \) and has connected the corresponding inventory \([+++]_A \) to \( A \). In this case, the hearer and the speaker proceed to make an agreement — analogous to that of the NG model, leaving \( A \) in their lists \( L_i \) and \( L_j \) and removing \( B \) is present. No examples contained in any inventory are removed.
TABLE I. Pair-wise interactions in the Bayesian NG model. The speaker (S) conveys a name $A$ or $B$ to the hearer (H) together with an example taken from the speaker’s inventory, $[+ + + . . .]_A$ or $[+ + + . . .]_B$, respectively — this happens with a branching probability $q = 0.5$ if the speaker has the list $(A, B)$ and knows the meaning of both names. The outcome can be: (1) a reinforcement (only); (2) generalization of concept $C$, if the Bayes probability is $p > 1/2$; (3) an agreement between hearer and speaker, if both agents know the meaning of the conveyed name. Even if not indicated, reinforcement takes place also in cases (2) and (3).

| S-List | Name (before) conveyed | H-List | Branching probability $q$ | Process | Condition | S-List (after) | H-List (after) |
|--------|------------------------|--------|---------------------------|---------|-----------|---------------|---------------|
| (A)    | $A$                    | (A)    | $q = 1$                   | Reinforcement always | $A$       | $A$           |               |
| (A)    | $A$                    | (B)    | $q = 1$                   | Reinforcement $p_A < 1/2$ | (A)       | (B)           |               |
| (A)    | $A$                    | (A, B) | $q = 1$                   | Agreement always | (A)       | (A)           |               |
| (B)    | $B$                    | (A)    | $q = 1$                   | Reinforcement $p_B < 1/2$ | (B)       | (A)           |               |
| (B)    | $B$                    | (B)    | $q = 1$                   | Reinforcement always | (B)       | (B)           |               |
| (B)    | $B$                    | (A, B) | $q = 1$                   | Agreement always | (B)       | (B)           |               |
| (A, B) | $A$                    | (A)    | $q = 1/2$                 | Agreement always | (A)       | (A)           |               |
| (A, B) | $A$                    | (B)    | $q = 1/2$                 | Reinforcement $p_B < 1/2$ | (A, B)    | (A)           |               |
| (A, B) | $A$                    | (B)    | $q = 1/2$                 | Learning $p_B ≥ 1/2$ | (A, B)    | (A, B)        |               |
| (A, B) | $B$                    | (B)    | $q = 1/2$                 | Agreement always | (B)       | (B)           |               |
| (A, B) | $A$                    | (A, B) | $q = 1/2$                 | Agreement always | (A)       | (A)           |               |

5. Time is updated, $t → t + 1$, and the simulation is reiterated from the first point above.

Two examples of Bayesian word-learning process, a successful and an unsuccessful one, are illustrated in the cartoon in the right panel (B) of Fig. 1. Table I lists the possible encounter situations, together with the corresponding relevant probabilities.

Notice that an agent $i$ can enter a pair-wise interaction with a non-empty inventory of examples, e.g. $[+ + + . . .]_A$, associated to name $A$, without being able to use name $A$ to convey examples to other agents, i.e., without the name $A$ in the list $L_i$ due to not having generalized concept $C$ in association with $A$. Those examples can have different origins: (1) in the initial conditions, when $n_{ex,A}$ randomly extracted examples associated to $A$ and $n_{ex,B}$ to $B$ are assigned to each agent; (2) in previous interactions, in which the examples were conveyed by other agents; (3) in an agreement about convention $B$, which removed label $A$ from the list $L_i$ while leaving all the corresponding examples in the inventory associated to name $A$. In the latter case, the inventory $[+ + + . . .]_A$ may be “ready” for a generalization process, since it contains a sufficient number of examples, i.e., agent $i$ will probably be able to generalize as soon as another example is conveyed by an agent. This situation is not as peculiar as it may look at first sight. In fact, there is a linguistic analogue in the case where a speaker that loses the habit to use a certain word (or a language) $A$ can regain it promptly, if exposed to $A$ again.

Notice also that without the agreement dynamics scheme introduced in the model, borrowed from the basic NG model, the population fraction $n_{AB}$ of individuals who know both $A$ and $B$ ($n_A + n_B + n_{AB} = 1$) would be growing, until eventually $n_{AB} = 1$.

III. RESULTS

In this section we study numerically the Bayesian NG model introduced above and discuss its main features. We limit ourselves to study the model dynamics on a fully-connected network.
In the new learning scheme, which replaces the one-shot learning of the two-conventions NG model, an individual generalizes concept $C$ on a suitable time scale $\Delta t > 1$, rather than during a single interaction. However, a few examples are sufficient for an agent to generalize concept $C$, as in a realistic concept-learning process. This is visible from the Bayesian probabilities $p_A$ and $p_B$ computed by agents in the role of hearer, according to Eq. (4), once at least $n_{ex,A}^* = 5$ and $n_{ex,B}^* = 6$ examples “+”, respectively, have been stored in the inventory associated to the name $A$ and $B$: Figure 3 shows the histograms of the $p_A$’s and $p_B$’s computed from the initial time until consensus for a single run with $N = 2000$ agents and starting with SIC. The low frequencies at small values of $p_A$ and $p_B$ and the highest frequencies at values close to unity are due to the fact that the Bayesian probabilities reach values $p_A \approx p_B \approx 1$ very fast, after a few learning attempts, consistently with the size principle, on which the Bayesian learning paradigm, and in turn Eq. (4), are based [10].

![Figure 3](image.png)

**FIG. 3.** Histograms of the Bayesian probabilities $p_A, p_B$ computed by agents during their learning attempts during a single run (for $N = 2000$ agents, starting with SIC; $n_{ex,A}^* = 5$, $n_{ex,B}^* = 6$).

In order to visualize how the system approaches consensus, it is useful to consider some global observables, such as the fractions $n_A(t)$, $n_B(t)$, and $n_{AB}(t)$ of agents that have generalized concept $C$ in association with name $A$ only, name $B$ only, or both names $A$ and $B$, respectively, or the success rate $S(t)$. The dynamics of a population of $N = 1000$ agents (panels (A) and (B)) using different initial conditions, SIC, AIC, and AICr, and that of a population of $N = 100$ agents starting with SIC (panels (C) and (D)) are shown in Fig. 4.

Panel (A) of Fig. 4 shows only the population fractions corresponding to the name found at consensus, for the sake of clarity (the remaining population fractions eventually go to zero). For asymmetrical initial condition (AIC or AICr), it is the initial majority that determines the convention found at consensus (that is $B$ for AIC and $A$ for AICr). If the system starts from SIC, the convention $A$, for which agents can generalize earlier ($n_{ex,A}^* = 5 < n_{ex,B}^* = 6$), is always found at consensus — in this case it is the asymmetry in the thresholds $n_{ex,A}^*$ and $n_{ex,B}^*$, characterizing the Bayesian learning process, to determine consensus.

Panel (B) of Fig. 4 shows the success rate $S(t = t_k)$, representing the average over different runs of the instantaneous success rate $S_k$ of the $k$th interaction at time $t_k$, defined as follows: $S_k = 1$ in case of agreement between the two agents or when a successful learning of the hearer takes places, following a Bayes probability $p > 1/2$; or $S_k = 0$ in case of unsuccessful generalization, when $p < 1/2$ and only reinforcement takes place. The success rate $S(t)$ varies between $S(0) \approx (n_A(0))^2 + (n_B(0))^2$, due to the respective fractions of agents that initially know the two conventions $A$ and $B$, to $S \approx 1$ at consensus, following a typical S-shaped curve of learning processes [38]. In the case of SIC, the initial value is $S(0) \approx 0.5^2 + 0.5^2 = 0.5$, while for AIC or AICr the initial value is $S(0) \approx (0.3)^2 + (0.7)^2 \approx 0.58$.

We now investigate how the modified Bayesian dynamics affects the convergence times to consensus. The study of the size-dependence of the convergence to consensus shows that there is a critical value $N^* \approx 500$ in the case of SIC, such that for $N \leq N^*$ there is a non-negligible probability that the final absorbing state is $B$. Panels (C) and (D) of Fig. 4, representing the results for a system starting with SIC and a smaller size $N = 100$, show the existence of two possible final absorbing states and that there are different times scales associated to the convergence to consensus: name $A$ is found at consensus in about 90% of cases and name $B$ in the remaining cases. The branching probability into $A$ or $B$ consensus is further investigated in panel (A) of Fig. 5, where we plot the branching probabilities $p_{e,A}, p_{e,B}$ versus the system sizes $N$. The nonlinear behavior (symmetrical sigmoid) signals the presence of finite-size effects, particularly clear for relatively small $N$-values. In fact, when the fluctuations in the system are larger, the system

---

*Figures and equations are placeholders and are not present in the text.*
FIG. 4. Average population fraction associated to the name shared in the final consensus state (upper panels (A) and (C)) and success rate $S(t)$ (lower panels (B) and (D)) versus time. Left panels (A) and (B): system with $N = 1000$ agents starting from different initial conditions, SIC, AIC, and AICr; averages done over 600 runs. Right panels (C) and (D): system with $N = 100$ agents starting from SIC; averages done over 1000 runs — notice that due to the smaller size $N = 100$, the system can converge to consensus both with name $A$ (in a fraction of cases $p_{e,A} \approx 0.9$) and with name $B$ ($p_{e,B} \approx 0.1$).

FIG. 5. Panel (A): probabilities $p_{e,A}$ and $p_{e,B}$ that the system reaches the consensus at $A$ and $B$ respectively, versus the system sizes $N$, obtained by averaging over 1000 runs of a system starting with SIC. Panel (B): average number of examples $\bar{n}_{ex,A}$ and $\bar{n}_{ex,B}$ recorded by an agent at consensus, for a system of $N = 50, 100, 500, 1000, 1500, 2000$ agents, starting with SIC, AIC, AICr. Averages are done over 600 runs.

size can play an important role in the dynamics of social systems, as an actual thermodynamic limit is only allowed for simulations of macroscopic physical systems [39].

The convergence time $t_{conv}$ follows a simple scaling rule with the system size $N$, related to the average number of examples $\bar{n}_{ex,A}, \bar{n}_{ex,B}$ relative to $A, B$ respectively, stored in the agents’ inventories at consensus. These values depend on the number of learning and reinforcement processes, and hence are related to the system size $N$. The average number of interactions undergone by the agents until the system reaches the consensus is given by the sum $\bar{n}_{int} = \bar{n}_{ex,A} + \bar{n}_{ex,B}$ [40]. One expects that

$$t_{conv} \approx \bar{n}_{int}N,$$ 

which suggests a linear scaling law ($t_{conv} \sim N$) for convergence time with the system size $N$ for all the possible initial conditions. A linear behavior is indeed confirmed by the numerical simulations with population sizes $N =$
TABLE II. Scaling laws $t_{\text{conv}} \sim N^\alpha$ with the system size $N$. Here the parameters are $n_{\text{ex},A}^* = 5$, $n_{\text{ex},B}^* = 6$ with initial conditions SIC, AIC and AICr. The average number of examples, $\bar{n}_{\text{ex},A}, \bar{n}_{\text{ex},B}$, stored at $t_{\text{conv}}$, are obtained averaging over 600 runs of a system with $N = 1000$ agents.

| $\alpha$ | $\bar{n}_{\text{ex},A}$ | $\bar{n}_{\text{ex},B}$ | outcome |
|----------|-----------------|-----------------|---------|
| SIC      | 1.06            | 20              | 8       | A, B   |
| AIC      | 1.08            | 3               | 19      | B      |
| AICr     | 1.09            | 18              | 3       | A      |

From the above mentioned scaling law, it is clear that the average number of examples stored by the agents at consensus plays an important role in the semiotic dynamics. In particular, it is found that if the final absorbing state is $A$ (or $B$), then $\bar{n}_{\text{ex},A} > \bar{n}_{\text{ex},B}$ ($\bar{n}_{\text{ex},B} > \bar{n}_{\text{ex},A}$). Moreover, the average number of examples, relative to the absorbing state, always increases monotonically with the system size while a size-independent behavior is observed in the opposite case, see the right panel (B) of Fig. 5.

Finally, we compare the convergence time of the Bayesian word-learning model, $t_{\text{conv}}$, with that of two-conventions NG model, $\bar{t}_{\text{conv}}$ [33], by studying the corresponding ratio $R = \frac{t_{\text{conv}}}{\bar{t}_{\text{conv}}}$ for common initial conditions and population sizes. When starting with SIC, the values of the convergence times obtained from the two models become of the same order by increasing $N$: $R$ decreases with $N$, reaching unity for $N = 10000$, see Fig. 6. In other words, the time scales of the two models become equivalent for relatively large system sizes, i.e., the learning processes of the two models perform equivalently and the Bayesian approach roughly gives rise to the one-shot learning that characterizes the two-conventions NG model. In the next section we discuss how the Bayesian model becomes asymptotically equivalent to the minimal NG model. The inset of Fig. 6 represents $R$ versus $N$, for $N < 2000$, given different starting configurations, with SIC, AIC and AICr, and different population sizes. In the following, we focus on the case of SIC.

FIG. 6. The ratio of the convergence times of the Bayesian word-learning model and the 2-conventions NG model, $R = \frac{t_{\text{conv}}}{\bar{t}_{\text{conv}}}$, versus the system size $N$, for a system starting with SIC. The inset illustrates the dependence of $R$ on different initial conditions. The curves are obtained averaging over 900 runs.
FIG. 7. Model scheme with two non-excluding options. Arrows indicate allowed transitions between the “bilingual” state \((A,B)\) and the “monolingual” states \(A\) and \(B\). Direct \(A \leftrightarrow B\) transitions are not allowed.

IV. STABILITY ANALYSIS

In this section we investigate the stability properties of the mean-field dynamics of the Bayesian NG model, in which statistical fluctuations and correlations are neglected. In the Bayesian NG model, as in the basic NG, agents can use two non-excluding options \(A\) and \(B\) to refer to the same concept \(C\). The main difference between the Bayesian model and the basic NG model is in the learning process: a one-shot learning process in the basic NG and a Bayesian process in the Bayesian NG model. In the latter case the presence of a name in the word list indicates that the agent has generalized the corresponding concept from a set of positive recorded examples.

The NG model belongs to the wide class of models with two non-excluding options \(A\) and the “monolingual” states (bilingual) state \((A,B)\), of bilingualism [41], in which transitions between state \((A,B)\) are allowed only through an intermediate (“bilingual”) state \((A,B)\), as schematized in Fig. 7. The mean-field equations for the fractions \(n_A(t)\) and \(n_B(t)\) can be obtained considering the gain and loss contributions of the transitions depicted in Fig. 7,

\[
\dot{n}_A = p_{AB \rightarrow A} n_{AB} - p_{A \rightarrow AB} n_A ,
\]

\[
\dot{n}_B = p_{AB \rightarrow B} n_{AB} - p_{B \rightarrow AB} n_B .
\]

(6)

Here \(\dot{n}_a(t) = dn_a(t)/dt\) and the quantities \(p_{a \rightarrow b}\) represent the respective transition rates per individual, corresponding to the arrows in Fig. 7 \((a, b = A, B, AB)\). The equation for \(n_{AB}(t)\) was omitted, since it is determined by the condition that the total number of agents is constant, \(n_A(t) + n_B(t) + n_{AB}(t) = 1\).

The details of the possible pair-wise interactions in the Bayesian naming game are listed in Table I. From the various contributions, one obtains the master equation

\[
\dot{n}_A = -p_B n_A n_B + n_A^2 + \frac{3 - p_B}{2} n_A n_{AB} ,
\]

\[
\dot{n}_B = -p_A n_A n_B + n_B^2 + \frac{3 - p_A}{2} n_B n_{AB} ,
\]

(7)

which can be rewritten in the form (6) with transition rates per individual given by

\[
p_{A \rightarrow AB} = p_B n_B + \frac{1}{2} p_B n_{AB} ,
\]

\[
p_{B \rightarrow AB} = p_A n_A + \frac{1}{2} p_A n_{AB} ,
\]

(8)

\[
p_{AB \rightarrow A} = \frac{3}{2} n_A + n_{AB} ,
\]

\[
p_{AB \rightarrow B} = \frac{3}{2} n_B + n_{AB} .
\]

(9)

Equations (8) provide the transition rates of learning processes, while Eqs. (9) give the transition rates of agreement processes. Setting \(x \equiv n_A, y \equiv n_B,\) and \(z = n_{AB} \equiv 1 - x - y,\) the autonomous system (7) becomes

\[
\dot{x} = f_z(x, y) \equiv -p_B xy + (1 - x - y)^2 + \frac{1}{2} (3 - p_B) x (1 - x - y) ,
\]

(10)

\[
\dot{y} = f_y(x, y) \equiv -p_A xy + (1 - x - y)^2 + \frac{1}{2} (3 - p_A) y (1 - x - y) ,
\]

(11)

where \(v = (f_x(x, y), f_y(x, y))\) is the velocity field in the phase plane. For the following analysis, it is convenient to write the Bayesian probabilities \(p_A\) and \(p_B\) appearing in these equations as time-dependent parameters of the model.
we assume microscopic Bayesian probability in Eq. (4) over the possible dynamical realizations. For this reason, they have also but they are actually highly non-linear functions of the variables. In fact, they can be thought as averages of the microscopic Bayesian probability in Eq. (4) over the possible dynamical realizations. For this reason, they have also a complex non-local time-dependence on the previous history of the interactions between agents. For the moment, we assume \( p_A(t) = p_B(t) = p(t) \), returning later to the general case.

From the conditions defining the critical points, \( f_x(x, y) = f_y(x, y) = 0 \), one obtains \((x - y)z = 0\). Setting \( z = 0 \), one obtains two solutions that correspond to consensus in \( A \) or \( B \), given by \((x_1, y_1, z_1) = (1, 0, 0)\) and \((x_2, y_2, z_2) = (0, 1, 0)\). Instead, setting \((x - y) = 0\) leads to the equation

\[
2x^2 - (p + 5)x + 2 = 0,
\]

that has the solutions

\[
x_\pm = \frac{p + 5 \pm \sqrt{(p + 5)^2 - 16}}{4}.
\]  

For \( p \in (0, 1) \), the corresponding solutions \((x_\pm, x_\pm, 1 - 2x_\pm)\) are not suitable solutions, because \( z_\pm = 1 - 2x_\pm < 0 \).

This analysis is valid for \( p > 0 \). In fact, \( p = p(t) \) is a function of time and for a finite interval of time after the initial time one has that \( p = 0 \), which defines a different dynamical system. In the initial conditions used, \( z(0) = 0 \), which implies \( z(t) = 0 \), \( x(t) = x(0) \), and \( y(t) = y(0) \) at any later time \( t \) as long as \( p(t) = 0 \), since \( \dot{x}(t) = \dot{y}(t) = \dot{z}(t) = 0 \) (see Eq. (7)); in fact, the whole line \( x + y = 1 \) (for \( 0 < x, y < 1 \)) represents a continuous set of equilibrium points. The reason why in this model \( p(0) = 0 \) at \( t = 0 \) and also during a subsequent finite interval of time is twofold. First, agents do not have any examples associated to the name not known and they have to receive at least \( n^*_{ex,A} \) or \( n^*_{ex,B} \) examples, before being able to compute the corresponding Bayesian probability \( p_A(t) \) or \( p_B(t) \)—thus it is to be expected that \( p(t) = 0 \) meanwhile. Furthermore, even when agents can compute the Bayesian probabilities, the effective probability to generalize is actually zero, due to the threshold \( p^* = 0.5 \) for a generalization to take place. The existence of the (temporary) equilibrium points on the line \( x + y = 1 \) ends as soon as the parameter \( p(t) > p^* \) and, according to Eqs. (7), the two A- and B-consensus states become the only stable equilibrium points. The representative point in the \( x-y \)-plane is deemed to leave the initial conditions on the \( z = 1 - x - y = 0 \) line, due to the stochastic nature of the dynamics, which is not invariant under time reversal [42].

To determine the nature of the critical points \((x_1, y_1) = (1, 0)\) and \((x_2, y_2) = (0, 1)\), one needs to evaluate at the equilibrium points the \( 2 \times 2 \) Jacobian matrix \( A(x, y) = \{\partial_i f_j\} \), where \( i, j = x, y \). It is easy to show that both the critical points \((0, 1)\) and \((1, 0)\) are asymptotically stable [43].

As long as the general case \( p_A \neq p_B \), it can be shown that the trajectory of the system can point toward and eventually reach the consensus state with \( A \) or \( B \), depending on whether \( p_A(t^*) > p_B(t^*) \) or \( p_A(t^*) < p_B(t^*) \), where \( t^* > 0 \) is the critical time at which the representative point leaves the initial position.
The convention $A$ or $B$ is selected randomly, depending on various factors related to the specific realization of the system evolution, such as the numbers of examples $n_{ex,A}^*$ and $n_{ex,B}^*$ recorded by the agents until time $t$, their quality from the point of view of the generalization, and the initial asymmetry of the thresholds for generalizing. $n_{ex,A}^* \neq n_{ex,B}^*$. The asymmetrical thresholds $n_{ex,A}^* = 4 < n_{ex,B}^* = 5$ produce a bias toward consensus in $A$ and play a crucial role in the subsequent Bayesian semiotic dynamics; in fact, swapping the threshold values (setting $n_{ex,A}^* = 5 > n_{ex,B}^* = 4$), the approach to consensus occurs with the outcomes $A, B$ swapped.

We observed that for $N \geq N^* \approx 500$, the chances that the system converges to $(B)$ become negligible. This can be seen in panels (C) and (D) of Fig. 8, showing $\bar{n}_{ex,A}(t)$ and $\bar{n}_{ex,B}(t)$ versus time (averaged over the agents of the system) for single runs, a population of $N = 100$ agents, and SIC, for different runs that relax toward consensus $A$ and $B$, respectively. After an initial transient, in which $\bar{n}_{ex,A}(t) \approx \bar{n}_{ex,B}(t)$, they differ more and more from each other at times $t > t^*$. In turn, also $p_A$ and $p_B$ begin to differ significantly from each other, thus affecting the rate of depletion of the populations during the subsequent dynamics. For instance, if $p_A > p_B$, then $p_B \rightarrow AB > p_A \rightarrow AB$, see Eqs. (8), which means that the depletion of $n_B$ occurs faster than that of $n_A$. In turn, this favours the decay of the mixed states $(A, B)$ into the state $(A)$, see Eqs. (9), being $n_A > n_B$.

The asymmetry discussed above also affects the convergence times $t_{conv}^A$ and $t_{conv}^B$ and we find $t_{conv}^B > t_{conv}^A$ in all the numerical simulations. Despite the noise, such a trend is already appreciable in a single run, as shown in panels (A) and (B) of Fig. 8. The mean fractions $n_A(t), n_B(t)$, and $n_{AB}(t)$ versus time, obtained by averaging over many runs, result in less noisy outputs and provide a more clear picture of the difference, which is visible in Fig. 9, obtained using 600 runs starting with SIC and for $N = 100$ agents (panels (A) and (B)) and $N = 200$ agents (panels (C) and (D)). In addition, the convergence times depend on the system size, increasing with the number of agents $N$: compare the left panels (A) and (B), where $N = 100$ agents, with the right panels (C) and (D), where $N = 200$ agents.

The possibility that the same system, starting with the same initial conditions and evolving with the same dynamical parameters, can reach either $A$ or $B$ is a consequence of the stochastic nature of dynamics. This does not happen for $N \geq N^*$, when both $\bar{n}_{ex,A}$ and $\bar{n}_{ex,B}$ reach some threshold values close to those observed at $t_{conv}$, which is clearly a value sufficient for the agents to generalizing concept $C$. In fact, the scaling law of $t_{conv}$ with $N$ shows that the sum of $\bar{n}_{ex,A}$ with $\bar{n}_{ex,B}$ becomes nearly constant for $N \geq N^*$, implying that the dynamics is uniquely determined, that is, the consensus always occurs at $A$ from SIC, once the agents have stored a threshold number of $\bar{n}_{ex,A}, \bar{n}_{ex,B}$. It is found that these threshold values correspond to $\bar{n}_{ex,A}^* = 21, \bar{n}_{ex,B}^* = 12$. Note that in $\bar{n}_{ex,A}^*, \bar{n}_{ex,B}^*$ we add values the four initial given examples stored in the agents’ inventories at the beginning. The reason is that the generalization function $p(t)$ outputs will effectively depend on them all. Therefore, at these threshold values, it would be very unlikely that $p_B > p_A$, and so it would be the same for the consensus at $B$.

Now we consider the Bayesian probabilities $p_A(t)$ and $p_B(t)$ computed by agents and the corresponding number
FIG. 10. Average values $\bar{p}_A(t)$ and $\bar{p}_B(t)$ computed using a (temporal) bin $\Delta t = 16 \times 10^3$, versus time from for a single run of a system reaching consensus at $A$. The convergence time is $t_{\text{conv}} \approx 160 \times 10^3$ and the population size is $N = 5000$. The inset shows the average number of learning attempts $n_{A}$, $n_{B}$ versus time for the same single run.

of learning attempts $n_{A}(t)$ and $n_{B}(t)$ made by agents at time $t$ to learn concept $C$ in association with word $A$ or $B$, respectively, i.e. the number of times that the agents compute $p_A$ or $p_B$ (only the case of a system starting with SIC is considered). We consider a single run of a system with $N = 5000$ agents and study the average values $\bar{p}_A(t)$, $\bar{p}_B(t)$, obtained by averaging $p_A(t)$ and $p_B(t)$ over the agents of the system. We also assume a coarse-grained view, consisting in an additional average of $\bar{p}_A(t)$, $\bar{p}_B(t)$, and $n_{A}$, $n_{B}$, over a a temporal bin $\Delta t = 16 \times 10^3$, in order to reduce random fluctuations. Figure 10 shows the time evolution of the average probabilities $\bar{p}_A(t)$ and $\bar{p}_B(t)$ in the time-range where data allow a good statistics. The probabilities grow monotonically and eventually reach the value one. While this points at an equivalence between the mean-field regime of the Bayesian naming game and that of the two-conventions NG model, in which agents learn at the first attempt (one-shoot learning), such an equivalence is suggested but not fully reproduced by the coarse-grained analysis. The time evolution of the number of learning attempts $n_{A}(t)$ and $n_{B}(t)$ shows that they are negligible both at the beginning and at the end of the dynamics — see inset in Fig. 10. This is due to the fact that at the beginning it is most likely that either interactions between agents with the same conventions take place (starting with SIC, each agent has a probability of 50% to interact with an agent having the same convention) or interactions between agents with different conventions but with still too small inventories to be able to generalize concept $C$, leading to reinforcement processes only. When approaching consensus, agents with one of the conventions constitute the large majority of the population and thus they are again most likely to interact through reinforcements only. Thus, the largest numbers of attempt to learn concept $C$ in association with $A$ and $B$ are expected to occur at the intermediate stage of the dynamics. In fact, $n_{A}(t)$ and $n_{B}(t)$ are observed to reach a maximum at $t \approx t_{\text{conv}}/2$ for any given system size $N$, as visible in the inset of Fig. 10. Notice that also the fraction of agents $n_{AB}$, who know both conventions and can communicate using both name $A$ and name $B$, possibly allowing other agents to generalize in association with name $A$ or $B$, reaches its maximum roughly at the same time.

V. CONCLUSION

We introduced a novel agent-based model that describes the appearance of linguistic consensus through a word-learning process. The work presented is exploratory in nature, concerning the minimal problem of a single concept that can be associated to two different possible names $A$ or $B$, but is aimed at providing a prototype of general framework for describing the interaction between the social and the cognitive dimension. To this aim, the model is constructed on the basis of the semiotic dynamics of the NG model and is then extended by adding a Bayesian cognitive process, mimicking human learning processes.

The model describes in a natural way (1) the uncertainty accompanying the first phase of a learning process, (2) the gradual reduction of the uncertainty as more and more examples are provided, and (3) the ability to learn from a few
examples. The semiotic dynamics of the synonyms is different from the basic NG, in that it depends on parameters that are of a strictly cognitive nature, such as the thresholds $n_{\star x}$ of the number of examples necessary before an agent can try to generalize and the acceptance threshold $p^{\star}$ for carrying out the generalization of a concept. The interplay between the asymmetry of the conventions $A$ and $B$, the system size, and the stochastic character of the time evolution, have dramatic consequences on the consensus dynamics: there is a critical time $t^{\star} > 0$, when the system begins to move in the phase-plane to eventually converge toward a consensus state; there is a critical system size $N^{\star}$, such that for $N < N^{\star}$ the system can end up in any of the two consensus states and the convergence times depends on $N$; there is an asymmetry in the branching probabilities that the system converges toward one of the two possible conventions and of the corresponding convergence times; the scaling laws of the convergence times versus $N$ differ from those observed in the basic NG model, because they depend on the learning experience of the agents.

The cognitive dimension offers additional possibilities for modelling in terms of specific cognitive parameters problems that are out of the reach of traditional social dynamics models. The model illustrated in this work represents a step toward a generalized Bayesian approach to social interactions, leading to cultural conventions.

Future work can address specific problems of current interest from the point of view of cognitive processes; or features relevant from the general standpoint of complexity theory. In the first case, it is possible to study in the cognitive dimension the semiotic dynamics of homonyms, synonyms, and innovation, e.g., the cognitive conditions leading to a name $A_1$, associated to a concept $C_1$, splitting into two names $A_1$ and $A_2$, associated to two related but distinct concepts $C_1$ and $C_2$, as more examples become available that make the two concepts eventually distinguishable from each other — a type of problems that cannot be tackled within models of cultural competition. In the second case, one can mention the classical problem of the interplay between a central information source (bias) and the local influences of individuals — this time in a cognitive framework.

Another question to be investigated within a cognitive framework would be the role of heterogeneity. In fact, heterogeneity is known to characterize most of the known complex systems at various levels — here the diversity could affect the dynamical parameters of e.g. the different competing names as well as those of the agents. Heterogeneity of individuals can lead to counter-intuitive effects, such as resonant behaviors [44, 45]. Furthermore, the complex, heterogeneous nature of a local underlying social network can change drastically the co-evolution and the time-scales of the conventions in competition with each other [46].

**ACKNOWLEDGEMENTS**

The authors acknowledge support from the Estonian Ministry of Education and Research through Institutional Research Funding IUT (IUT39-1), the Estonian Research Council through Grant PUT (PUT1356), and the ERDF (European Development Research Fund) CoE (Center of Excellence) program through Grant TK133.

We also thank Andrea Baronchelli for providing useful remarks about the naming game model and the manuscript.

[1] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81:591, 2009.
[2] A. Baronchelli. The emergence of consensus: a primer. *R. Soc. open sci.*, 5:172189, 2018.
[3] C. Xia, S. Ding, C. Wang, J. Wang, and Z. Chen. Risk analysis and enhancement of cooperation yielded by the individual reputation in the spatial public goods game. *IEEE Systems Journal*, 11(3):1516–1525, Sep. 2017.
[4] Chengyi Xia, Xiaopeng Li, Zhen Wang, and Matjaž Perc. Doubly effects of information sharing on interdependent network reciprocity. *New Journal of Physics*, 20(7):075005, jul 2018.
[5] Yingchao Zhang, Juan Wang, Chenxi Ding, and Chengyi Xia. Impact of individual difference and investment heterogeneity on the collective cooperation in the spatial public goods game. *Knowledge-Based Systems*, 136:150 – 158, 2017.
[6] S. Mufwene. *The ecology of language evolution*. Cambridge University Press, Cambridge, 2001.
[7] R. Lass. *Historical Linguistics and Language Change*. Cambridge University Press, Cambridge, UK, 1997.
[8] C. Berruto. *Prima lezione di sociolinguistica*. Laterza, Roma & Bari, 2004.
[9] S. Edelman and H. Waterfall. Behavioral and computational aspects of language and its acquisition. *Phys. Life Rev.*, 4:253277, 2007.
[10] J. B Tenenbaum. *A Bayesian Framework For Concept Learning*. PhD thesis, MIT, 1999.
[11] S. Wichmann. The emerging field of language dynamics. *Language and Linguistics Compass*, 2/3:442, 2008.
[12] S. Wichmann. Teaching & learning guide for: The emerging field of language dynamics. *Language and Linguistics Compass*, 2, 2008.
[13] Joshua B. Tenenbaum and Fei Xu. Word learning as Bayesian inference. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, 2000.
[14] F. Xu and J. B. Tenenbaum. Word learning as Bayesian inference. *Psychological Review*, 114:245–272, 2007.
[15] J.R. Hurford. Biological evolution of the saussurean sign as a component of the language-acquisition device. Lingua, 77:187–222, 1989.
[16] Martin A. Nowak, Joshua B. Plotkin, and David C. Krakauer. The evolutionary language game. Journal of Theoretical Biology, 200(2):147 – 162, 1999.
[17] M.A. Nowak. Evolutionary biology of language. Philos. Trans. R. Soc. London B Biol. Sci., 355(1403):1615–22, 2000.
[18] Peter E. Trapa and Martin A. Nowak. Nash equilibria for an evolutionary language game. Journal of Mathematical Biology, 41(2):172–188, Aug 2000.
[19] Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. Sharp transition towards shared vocabularies in multi-agent systems. Journal of Statistical Mechanics: Theory and Experiment, 2006(06):P06014, 2006.
[20] Guanrong Chen and Yang Lou. Naming Game. Springer International Publishing, Switzerland, 2019.
[21] Jierui Xie, Jeffrey Emenheiser, Matthew Kirby, Sameet Sreenivasan, Boleslaw K. Szymanski, and Gyorgy Korniss. Evolution of opinions on social networks in the presence of competing committed groups. PLOS ONE, 7(3):1–9, 03 2012.
[22] Zhong-Yan Fan, Ying-Cheng Lai, and Wallace Kit-Sang Tang. Knowledge consensus in complex networks: the role of learning. arXiv:1809.00297, 2018.
[23] Andrea Baronchelli. A gentle introduction to the minimal naming game. Belgian Journal of Linguistics, 30(1):171–192, 2016.
[24] Joshua B. Tenenbaum and Thomas L. Griffiths. Generalization, similarity, and Bayesian inference. The Behavioral and brain sciences, 24:4:629–40; discussion 652–791, 2001.
[25] Thomas L. Griffiths and Joshua B. Tenenbaum. Optimal predictions in everyday cognition. Psychological Science, 17(9):767–773, 2006. PMID: 16984293.
[26] Amy Perfors, Joshua B. Tenenbaum, Thomas L. Griffiths, and Fei Xu. A tutorial introduction to Bayesian models of cognitive development. Cognition, 120:3:302–21, 2011.
[27] Brenden M. Lake, Ruslan Salakhutdinov, and Joshua B. Tenenbaum. Human-level concept learning through probabilistic program induction. Science, 350(6266):1332–1338, 2015.
[28] Joshua B. Tenenbaum. Bayesian modeling of human concept learning. In Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems II, pages 59–65, Cambridge, MA, USA, 1999. MIT Press.
[29] Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. How to grow a mind: Statistics, structure, and abstraction. Science, 331(6022):1279–1285, 2011.
[30] D. Barber. Bayesian Reasoning and Machine Learning. Cambridge University Press, Cambridge, UK, 2012.
[31] K.P. Murphy. Machine Learning: A Probabilistic Perspective. Adaptive Computation and Machine Learning series. MIT Press, Cambridge, MA, 2012.
[32] Theodoros Evgeniou, Massimiliano Pontil, and Tomaso Poggio. Statistical learning theory: A primer. International Journal of Computer Vision, 38(1):9–13, Jun 2000.
[33] Castelló, X., Baronchelli, A., and Loreto, V. Eur. Phys. J. B, 71(4):557–564, 2009.
[34] Andrea Baronchelli, Luca Dall’Asta, Alain Barrat, and Vittorio Loreto. Phys. Rev. E, 76:051102, Nov 2007.
[35] Ludwig Wittgenstein. Philosophical investigations. Macmillan, New York, 1953.
[36] Hanns Ludwig Harney. Bayesian Inference. Data Evaluation and Decision. Springer, 2003.
[37] H. Jeffreys. Theory of Probability. Clarendon Press, Oxford, 1939.
[38] Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. Journal of Statistical Mechanics: Theory and Experiment, 2006(06):P06014, 2006.
[39] R. Toral and C.J. Tessone. Finite size effects in the dynamics of opinion formation. Comm. Comp. Phys., 2:177, 2007.
[40] The $N_{ex,A} = N_{ex,B} = 4$ examples given initially to each agent are not accounted for by $\bar{N}_{ex,A}$ and $\bar{N}_{ex,B}$.
[41] M. Patriarca, X. Castelló, J.R. Uriarte, V.M. Eguíluz, and M. San Miguel. Modeling two-language competition dynamics. Adv. Comp. Syst., 15(3&4):1250048, 2012.
[42] Haye Hinrichsen. Non-equilibrium phase transitions. Physica A: Statistical Mechanics and its Applications, 369(1):1 – 28, 2006. Fundamental Problems in Statistical Physics.
[43] Steven H. Strogatz. Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering. Westview Press, 1995.
[44] C.J. Tessone and R. Toral. Diversity-induced resonance in a model for opinion formation. Eur. Phys. J. B, 71:549, 2009.
[45] T. Vaz Martins, R. Toral, and M.A. Santos. Divide and conquer: resonance induced by competitive interactions. Eur. Phys. J. B, 67:329–336, 2009.
[46] R. Toivonen, X. Castelló, V.M. Eguíluz, J. Saramäki, K. Kaski, and M. San Miguel. Broad lifetime distributions for ordering dynamics in complex networks. Phys. Rev. E, 79:016109, 2009.