General three-state model with biased population replacement: Analytical solution and application to language dynamics

Francesca Colaiori,[1,2] Claudio Castellano,[1,2] Christine F. Cuskley,[1,2] Vittorio Loreto,[2,1,3] Martina Pugliese,[2,1] and Francesca Tria[3]

[1]*Istituto dei Sistemi Complessi (ISC-CNR), via dei Taurini 19, I-00185 Roma, Italy*

[2]*Dipartimento di Fisica, Sapienza Università di Roma, P.le A. Moro 2, I-00185 Roma, Italy*

[3]*ISI Foundation, Via Alassio 11/C, Torino, Italy*

(Received 11 August 2014; published 9 January 2015)

Empirical evidence shows that the rate of irregular usage of English verbs exhibits discontinuity as a function of their frequency: the most frequent verbs tend to be totally irregular. We aim to qualitatively understand the origin of this feature by studying simple agent-based models of language dynamics, where each agent adopts an inflectional state for a verb and may change it upon interaction with other agents. At the same time, agents are replaced at some rate by new agents adopting the regular form. In models with only two inflectional states (regular and irregular), we observe that either all verbs regularize irrespective of their frequency, or a continuous transition occurs between a low-frequency state, where the lemma becomes fully regular, and a high-frequency one, where both forms coexist. Introducing a third (mixed) state, wherein agents may use either form, we find that a third, qualitatively different behavior may emerge, namely, a discontinuous transition in frequency. We introduce and solve analytically a very general class of three-state models that allows us to fully understand these behaviors in a unified framework. Realistic sets of interaction rules, including the well-known naming game (NG) model, result in a discontinuous transition, in agreement with recent empirical findings. We also point out that the distinction between speaker and hearer in the interaction has no effect on the collective behavior. The results for the general three-state model, although discussed in terms of language dynamics, are widely applicable.

DOI: 10.1103/PhysRevE.91.012808 PACS number(s): 89.65.−s, 87.23.Ge, 05.70.Ln

I. INTRODUCTION

Language is structured by rules [1,2]—but linguistic rules often have exceptions. This fact kindled a longstanding debate in cognitive science centering on how individual learners accommodate rule sets rife with exceptions (e.g., see Refs. [3,4]). However, how exceptions arise and evolve over time within a language system remains a largely open question.

The study of the English past tense is widely used as an exemplar of the interplay between rules and exceptions [5–8]. A recent study of historical corpus data [9] looks at rules in the language system rather than individual learners, shedding light on the relationship between the verb frequency and regularity (Fig. 1). Each verb in the language can be characterized by I, the fraction of irregular past tense tokens over the total number of tokens in the past tense, and ν, its frequency of usage. An interesting transition is found in the behavior of I as a function of ν: regular verbs dominate the low-frequency range, while most irregular verbs are located at higher frequencies (see also Refs. [10,11]). For intermediate values of ν fully regular ($I = 0$) and fully irregular ($I = 1$) verbs coexist. Only a small subset of verbs exhibit both regular and irregular forms ($0 < I < 1$) and occur primarily in a rather narrow range of frequencies between the dominant regular and irregular states.

The work presented in this paper takes a theoretical approach to the relationship between rules and exceptions in a population of interacting speakers. We investigate the dynamics of a set of very simple agent-based models aimed at describing the fundamental mechanisms by which rules and exceptions may be shared or disappear in a population, in the same spirit of the naming game (NG) [12,13] investigations of the emergence of shared naming conventions. We consider a single lemma and examine two-state and three-state models. In a two-state model, an individual has an internal inflectional inventory, which can contain either the regular (R) or irregular (I) inflection for the lemma. In a three-state model, the inflectional state can be either R, I, or mixed (M). The mixed state represents intraspeaker variation [14], where an individual may accommodate both regular and irregular forms for a single word [15]. For example, there is evidence that regular and irregular past-tense verb forms are simultaneously known and potentially used in seemingly free variation within a single speaker (e.g., both *sneaked* and *snuck* are acceptable [15]). Agents are endowed at the start of the dynamics with some inflectional state for the word and engage at rate ν in pairwise interactions, i.e., one of them utters the verb under consideration and the other listens to it. During interaction, the inflectional states of the speaker and hearer change according to a predefined set of interaction rules. In addition to interaction events, we implement a *replacement* mechanism: at rate r an agent is replaced with a "child" who engages in *overregularization*: these "child" agents assume the regular inflection applies to all words in the vocabulary, representing a known bias of child learners [16–18]. Replacement represents turnover in the population: as a child learner enters, an adult leaves (i.e., is replaced), so that the population size remains constant.

We first approach some specific models analytically, within the framework of mean-field theory. This analytical approach

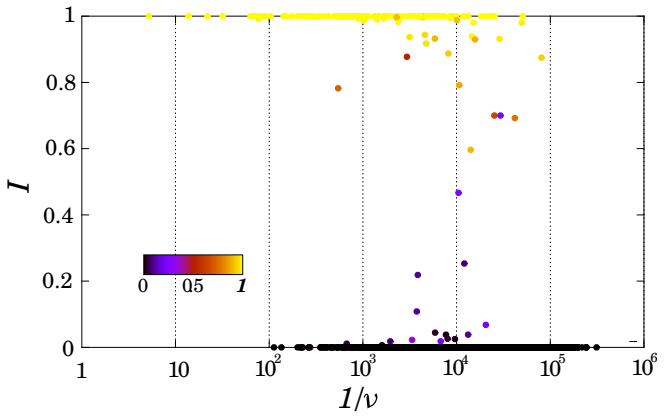

FIG. 1. (Color online) Plot of I, the fraction of times a verb is used in the irregular past tense form as a function of the inverse of its frequency of usage ν, for 2840 English verbs in the CoHA corpus for the decade 1980–1989. The color (shade) of each symbol represents the average value of I across 16 decades between 1830 and 1989. (Adapted from Ref. [9].)

allows for quantitative predictions regarding which state a population of speakers will reach given a particular set of interaction rules and the type of transition which may occur depending on the ratio between the frequency and replacement rates. It turns out that different interaction rules lead to qualitatively different types of behavior. Two-state models lead to total regularization or to a continuous transition as a function of r/ν. Three-state models, on the other hand, have the potential to exhibit a discontinuous transition with some highly frequent, mostly irregular words, reminiscent of findings in empirical data. To understand these variations and understand their origin, we introduce a very general three-state model, encompassing all possible sets of interaction rules that do not favor either the regular or irregular state in their outcome. We provide a formal solution for any value of the model parameters and study in detail the conditions under which no transition (i.e., total regularization) occurs, and the conditions under which we observe a continuous or discontinuous transition. Our analysis also shows that assuming asymmetric influence of the speaker over the hearer in the interaction has no effect on the collective behavior. The results for this general model are discussed in terms of language dynamics, but their applicability is fully general: they give the solution for any three-state model of population dynamics with unbiased interaction rules and biased replacement.

II. TWO-STATE MODELS

Let us start by considering two-state models, where each individual can be found in one of two possible rule states, regular (R) or irregular (I). While in this case the two states represent regular and irregular inflections, this framework can represent any set of binary options (e.g., the choice between two possible words to name an object, two alternative opinions on a given topic, etc.). As a specific example, the Abrams-Strogatz model [19] used a two-state approach to examine the dynamics of endangered languages, providing a prime example of the dynamics of languages in competition more generally.

Using this general two-state model approach, we focus here on the binary options of regular (R) or irregular (I) inflection for a single word, characterized by its frequency of usage, ν. At each interaction step, two agents are selected at random (i.e., mixing is homogeneous) and assigned the role of either speaker or hearer. With probability ν they engage in a pairwise interaction (i.e., the speaker utters the verb under consideration to the hearer), which affects their inflectional states according to specific interaction rules. Then, with probability r, one individual in the population is replaced with a "child" having its inflectional state set to R. This part of the dynamics mimics the turnover of some segment of the adult population into child learners at rate r, keeping the population size N fixed, and assuming over-regularization behavior in new learners. In this way the population replacement is biased toward one of the two options, in this case, regularity. The quantity $1/r$ can be interpreted as the life expectancy of an individual in the population.

Among the possible two-state interaction rules, we consider the sets presented in Table I. In the *irregular-biased* (A) and *regular-biased* (B) models, the speaker and the hearer roles are symmetric; in other words, which agent identifies as speaker or hearer is irrelevant, but the presence of an $I(R)$ state in the interaction leads the rules. In A(B) an agent switches to the $I(R)$ state whenever interacting with a partner in the $I(R)$ state, regardless of which agent is the speaker and which the hearer. In these cases the speaker can affect the hearer's state, and the hearer can also affect the speaker's. In the *speaker-leads* model (C), the roles are not symmetric: the speaker never changes its state and the hearer always adopts the state of the speaker.

TABLE I. Interaction rules for the two-state models. The two columns on the left refer to the status of speaker and hearer prior to interaction. The next three pairs of columns refer to the status of speaker and hearer after the interaction in the three models.

Before		After					
		Model A: irregular-biased		Model B: regular-biased		Model C: speaker leads	
Speaker	Hearer	Speaker	Hearer	Speaker	Hearer	Speaker	Hearer
R	R	R	R	R	R	R	R
R	I	I	I	R	R	R	R
I	R	I	I	R	R	I	I
I	I	I	I	I	I	I	I

The *irregular-biased* model is perfectly equivalent to one of the most fundamental models of nonequilibrium statistical physics: the contact process [20]. The temporal behavior of this model is easily understood by writing down the mean-field evolution equation for the density ρ_I of individuals in the I state (the density of R individuals, ρ_R, being trivially $1-\rho_I$)

$$\dot{\rho}_I = -r\rho_I + 2\nu\rho_I(1-\rho_I). \tag{1}$$

Equation (1) is solved straightforwardly and yields, for any initial configuration $\rho_I(0)$, $\rho_R(0) = 1-\rho_I(0)$,

$$\rho_I(t) = \frac{2/n-1}{2/n + \left(\frac{2/n-1}{\rho_I(0)} - 2/n\right)e^{-r(2/n-1)t}}, \tag{2}$$

where $n = r/\nu$.

For long time scales, the system reaches (for any initial configuration) a stationary state that exhibits a continuous transition for a critical value $n_c = 2$, between a fully regular state ($\rho_I = 0$) for $n > n_c$, and a state with individuals in both the R and the I state ($\rho_I > 0$):

$$\rho_I = \begin{cases} 0 & n \geqslant n_c \\ 1 - \frac{n}{2} & 0 \leqslant n < n_c. \end{cases} \tag{3}$$

The solutions for the *regular-biased* and *speaker-leads* models are obtained from Eq. (2) by simply replacing $n \to -n$ and $n \to \infty$, respectively, and both result in an exponential relaxation to the stationary fully regular state ($\rho_I = 0$) for any physical value of n. Unlike in case A, in these two cases the interaction rules are biased in favor of R (case B) or unbiased (C) and they cannot compensate for the increase in R states due to replacement, leading to a fully regular absorbing state.

As it will be demonstrated in Sec. IV A, no two-state model can give rise to a discontinuous transition between the fully regular state and a state with $\rho_I > 0$. Empirical data, however, exhibit such a discontinuous transition, and research shows that speakers can accommodate regular and irregular forms simultaneously [15]. For these reasons, we now turn our attention to a more complex modeling scheme that integrates a third, mixed state (M), wherein agents accommodate either the R or I. We will show that introducing this psychologically plausible, mixed state, a qualitatively different behavior appears, namely, a discontinuous transition in regularity, reminiscent of empirical data.

III. THREE-STATE MODELS

In three-state models there are still only two alternative inflections that can be applied to a word (R and I) during an interaction event, but internally, each individual can be in one of the three possible states: R (regular), I (irregular), and M (mixed). In the mixed state the individual can accommodate both R and I forms; this accounts for agents undecided on which is the correct form to use, or that consider both the regular and the irregular form acceptable.

The study of three-state models has a long history in the investigation of language dynamics (for a review see Ref. [21]). In particular, Wang and Minett proposed [22,23] deterministic models for the competition of two languages, which included a third potential state of bilingual individuals. Castelló *et al.* [24] proposed a modified version of the voter model [25,26] to examine language that included bilingual individuals, the so-called AB model. Synonymy, the possibility for having multiple potential names for a single meaning (much like multiple inflections for a single verb as in the mixed state) has also been examined in the classic naming game (NG) model [12,27]. The naming game and its variants have examined structures of increasing complexity, often including agents who can have multiple internal states, from the categorization of colors [28,29] to basic syntactic structures [30].

TABLE II. Interaction rules for the three examples of three-state models. The two columns on the left refer to the status of speaker and hearer prior to the interaction. The following three pairs of columns refer to the status of speaker and hearer after the interaction for the three different rule sets: the naming game (NG), the continuous transition (CT) model, and the no transition (NT) model. When two alternative inflections are possible because of a mixed state, the probability of each of them is 1/2. In the only case with three alternative outcomes (I,I) and (R,R) occur each with probability 1/4, while (M,M) occurs probability 1/2.

Before		After					
		Model NG		Model CT		Model NT	
Speaker	Hearer	Speaker	Hearer	Speaker	Hearer	Speaker	Hearer
R	*R*	*R*	*R*	*R*	*R*	*R*	*R*
R	*I*	*R*	*M*	*R*	*M*	*M*	*M*
R	*M*	*R*	*R*	*R*	*M*	*R*	*R*
						M	*M*
I	*R*	*I*	*M*	*I*	*M*	*M*	*M*
I	*I*	*I*	*I*	*I*	*I*	*I*	*I*
I	M	*I*	*I*	*I*	*M*	*I*	*I*
						M	*M*
M	*R*	*M*	*M*	*R*	*R*	*R*	*R*
		R	*R*	*I*	*M*	*M*	*M*
M	*I*	*I*	*I*	*I*	*I*	*I*	*I*
		M	*M*	*R*	*M*	*M*	*M*
M	*M*	*I*	*I*	*R*	*M*	*I*	*I*
		R	*R*	*I*	*M*	*R*	*R*
						M	*M*

We now study the dynamics of three specific examples in the class of three-state models, with different microscopic rules leading to qualitatively different behaviors (see Table II). The first set of rules is known as the naming game.

A. The naming game with biased replacement: A three-state model with a discontinuous transition

The interaction dynamics of the naming game with three states are as follows: first, at each time step a speaker and a hearer are selected at random. With the probability ν they interact; the speaker conveys to the hearer either the R or I form depending on his inventory (if in the mixed state he utters R or I with equal probability). If the hearer's inventory contains the inflection used in the utterance, both agents update their inventories keeping only the form involved in

the interaction. Otherwise, the hearer adds the form to his inventory (thus switching to the mixed state). Table II (first four columns) summarizes these interaction rules. In addition to these rules, the population turnover is implemented as in the previous two-state models: at each time step an individual is selected at random and, with probability r, is replaced by a new individual in state R.

In a generic three-state model, two densities are needed to specify the global state of the system. We choose ρ_I and ρ_R, the density of the mixed state being $\rho_M = 1 - \rho_R - \rho_I$. The mean-field equations for the naming game with biased replacement are

$$\begin{aligned} \dot{\rho}_I &= -r\rho_I + \nu\big(1 + \rho_R^2 - 2\rho_R - \rho_I\big) \\ \dot{\rho}_R &= r(1 - \rho_R) + \nu\big(1 + \rho_I^2 - 2\rho_I - \rho_R\big). \end{aligned} \tag{4}$$

Notice that the equations for the usual naming game with three states [12,31] are recovered by setting $r = 0$ and $\nu = 1$. Imposing the stationarity condition, after some algebra one finds that the density of individuals in the irregular state is given by the fourth order equation

$$\rho_I(1+n)^3 = \rho_I^2(\rho_I - 2)^2, \tag{5}$$

where $n = r/\nu$. One solution is, for any n, the trivial value $(\rho_I^{(1)}, \rho_R^{(1)}) = (0,1)$, corresponding to the fully regular state. Regarding the three other solutions, since $(1+n)^3$ is always larger than 1, it follows that one solution $(\rho_I^{(4)}, \rho_R^{(4)})$ is always real but unphysical (being larger than 1), while the two others are complex for $n > (2\sqrt[3]{4}/3 - 1) \approx 0.0583$. Below this critical value (corresponding to a saddle-node bifurcation), these two solutions are real and physical:

$$\rho_I^{(2)} = \frac{4}{3}\left[1 + \cos\left(\frac{\cos^{-1}\left[\frac{27}{16}(1+n)^3 - 1\right]}{3} + \frac{4\pi}{3}\right)\right], \tag{6}$$

$$\rho_I^{(3)} = \frac{4}{3}\left[1 + \cos\left(\frac{\cos^{-1}\left[\frac{27}{16}(1+n)^3 - 1\right]}{3} + \frac{2\pi}{3}\right)\right], \tag{7}$$

with the stationary value of ρ_R given by

$$\rho_R^* = 1 - \sqrt{(1+n)\rho_I^*}\,. \tag{8}$$

For $n = 0$ the solutions converge to the values found for the usual naming game [12,31]: $(\rho_I^{(2)}, \rho_R^{(2)}) = (1,0)$, and $(\rho_I^{(3)}, \rho_R^{(3)}) = ((3-\sqrt{5})/2, (3-\sqrt{5})/2)$.

The physical stationary solutions are represented in Fig. 2 as a function of n. The stability of the generic solution (ρ_I^*, ρ_R^*) as a function of n is investigated by looking at the eigenvalues and eigenvectors of the stability matrix, defined through the equations:

$$\frac{d}{dt}\begin{bmatrix}\delta\rho_I \\ \delta\rho_R\end{bmatrix} = \begin{bmatrix}-r-\nu & 2\nu(\rho_R^* - 1) \\ 2\nu(\rho_I^* - 1) & -r-\nu\end{bmatrix}\begin{bmatrix}\delta\rho_I \\ \delta\rho_R\end{bmatrix}. \tag{9}$$

The eigenvalues are given by

$$\lambda_{1,2} = \pm 2\nu\sqrt{(\rho_I^* - 1)(\rho_R^* - 1)} - (r+\nu). \tag{10}$$

Figure 3 reports the complete phase flow in the space (ρ_I, ρ_R) for $n = 0$ and $n = 0.025$. For $(\rho_I^{(1)}, \rho_R^{(1)}) = (0,1)$ both eigenvalues are negative: the fully regular state is always attractive and stable and for $n > (2\sqrt[3]{4}/3 - 1) \approx 0.0583$ it is the only physical solution. For $n < (2\sqrt[3]{4}/3 - 1) \approx 0.0583$ the two other physical solutions appear. $(\rho_I^{(2)}, \rho_R^{(2)})$ is always stable and attractive and, for $n > 0$, it always corresponds to states for which $\rho_I + \rho_R < 1$. Let us now focus on the $(\rho_I^{(3)}, \rho_R^{(3)})$ solution. This solution corresponds to a saddle point (the red circles in Fig. 3) since it has one positive and one

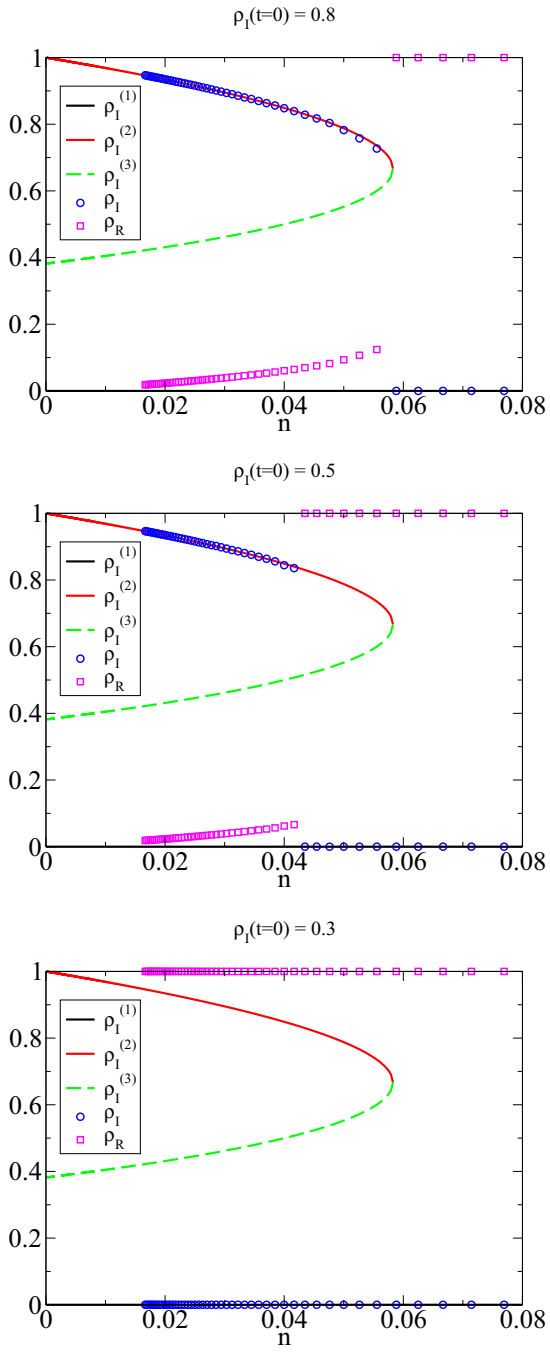

FIG. 2. (Color online) Phase diagram of the NG model as a function of $n = r/\nu$. The black line indicates the stationary solution $\rho_I^{(1)}$, while green (light gray) and red (dark gray) curves indicate solutions $\rho_I^{(2)}$ (stable) and $\rho_I^{(3)}$ (unstable), respectively. Symbols are the stationary values of the fraction of irregulars ρ_I and regulars ρ_R from numerical simulations with different initial conditions: from top to bottom $\rho_I(t=0) = 0.8, 0.5$, and 0.3. The three panels show the dependence of the stationary state on the initial condition.

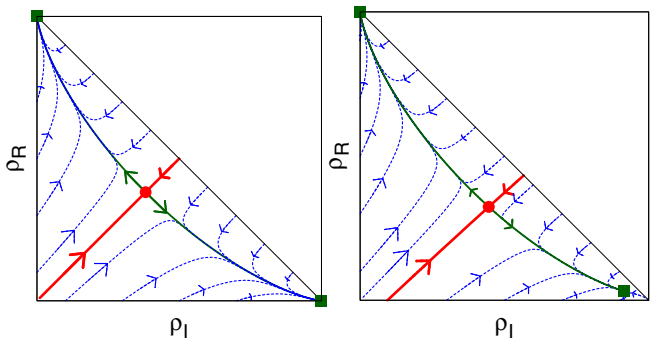

FIG. 3. (Color online) Phase-space (ρ_I, ρ_R) for the NG model with biased population replacement illustrating fixed points, separatrix, and phase flows (represented by blue dashed lines) for (left) $n = 0$ and (right) $n = 0.025$. The green squares are the attractive stable solution (1) and (2), while the red circles are the solution (3) that correspond to a saddle point with an attractive and a repulsive direction. Only the fraction of the physical phase space with $\rho_I + \rho_R < 1$ is represented.

negative eigenvalue. The separatrix in the attractive direction corresponds to the eigenvector associated to the negative eigenvalue:

$$\rho_R = \rho_R^{(3)} + m\big(\rho_I - \rho_I^{(3)}\big), \tag{11}$$

with $m = \sqrt{(1-\rho_I^{(3)})/\sqrt{(1+n)\rho_I^{(3)}}}$ (the thick red solid line in the figure). The other separatrix is locally approximated in the neighborhood of $(\rho_I^{(3)}, \rho_R^{(3)})$, by

$$\rho_R = \rho_R^{(3)} - m\big(\rho_I - \rho_I^{(3)}\big) \tag{12}$$

(the thin green solid line in the Fig. 3). In Appendix A we report the explicit expressions for the limit case $n = 0$, i.e., the original naming game. The model exhibits a discontinuous transition between a phase (high values of $n = r/\nu$) where, whatever the initial condition, the stationary state is fully regular and a phase (low values of n) where both the fully regular state [solution (1)] and a state with a large fraction of irregulars [solution (2)] are stable. As depicted in Fig. 3, the initial condition determines which one of the two states is asymptotically reached. In particular, if the initial condition is above the separatrix corresponding to the attractive direction (thick red line) all individuals converge to the fully regular state [solution (1)]; on the other hand, if the initial condition is below that separatrix the system converges to the solution (2) where a fraction $\rho_I^{(2)}$ ($\rho_I^{(2)} = 1$ for the case with no replacement, $n = 0$) of irregulars coexists dynamically with regular and mixed individuals. Only initial conditions exactly on the separatrix (thick red line) lead to convergence to solution (3). The predictions of the MF theory are confirmed by numerical simulations of the actual agent-based model (Fig. 2).

In summary, the naming game with biased population replacement exhibits a discontinuous transition as a function of n. The discontinuity also implies a dependence of the final steady state on the initial condition, which provides a theoretical justification for the observation of a range of frequencies where both fully regular and mostly irregular verbs exist (see Fig. 1). For every frequency in this interval, verbs will converge to the fully regular state (1) or to the mostly irregular state (2), depending on the initial values of ρ_R and ρ_I. This phenomenology is in agreement the empirical findings reported in [9] of the existence of a discontinuous transition between regular and irregular forms as a function of the frequency of usage.

B. Model CT: A three-state model with a continuous transition

We now consider a model with the interaction rules presented in Table II, columns CT, which differs from the NG case essentially because the hearer never discards the mixed state. The mean-field equations for the evolution of the densities in this case are written as

$$\begin{aligned}\dot\rho_I &= -r\rho_I + \nu\big\{-\rho_I\rho_R + \tfrac{1}{2}\rho_R[1-(\rho_I+\rho_R)] \\ &\quad + \tfrac{1}{2}[1-(\rho_I+\rho_R)]^2\big\} \\ \dot\rho_R &= r(1-\rho_R) + \nu\big\{-\rho_I\rho_R + \tfrac{1}{2}\rho_I[1-(\rho_I+\rho_R)] \\ &\quad + \tfrac{1}{2}[1-(\rho_I+\rho_R)]^2\big\}.\end{aligned} \tag{13}$$

By imposing the stationarity condition $\dot\rho_I = \dot\rho_R = 0$ and summing and subtracting the two equations it is possible to reduce them to

$$\begin{aligned} r(1-y) - \frac{\nu}{2}y(1-x) &= 0 \\ r(1-x) + \frac{\nu}{2}\big(y^2 - 3x + 2\big) &= 0,\end{aligned} \tag{14}$$

where we have introduced the auxiliary variables $x = \rho_R + \rho_I$ and $y = \rho_R - \rho_I$. From the second equation one obtains $x = (n + 1 + y^2/2)/(n + 3/2)$. Inserting this expression into the first equation, one is left with a third-order algebraic equation, which always admits three real solutions. One of them is, for any n, the fully regular solution $(x, y) = (1, 1)$. The two others are always unphysical (being larger than 1 or smaller than -1) except for $n \leqslant n_c = (\sqrt{17} - 3)/4 \approx 0.2807$. In such a case another physical solution appears,

$$y = 2\sqrt{-Q}\cos\left(\frac{\theta + 4\pi}{3}\right), \tag{15}$$

where $Q = [2(n+1) - (2n+1)(2n+3)]/3$, $R = -n(2n+3)$, and $\theta = \cos^{-1}[R/\sqrt{-Q^3}]$. The expressions for ρ_R and ρ_I are obtained from the relations $\rho_R = (x + y)/2$ and $\rho_I = (x - y)/2$. The physical stationary solutions are reported in Fig. 4, along with the results of the numerical simulations of the agent-based model that well fit the theoretical predictions. The plot illustrates that the transition occurring at $n_c = (\sqrt{17} - 3)/4$ is continuous.

C. Model NT: A three-state model without a transition

Let us now focus on another set of rules, reported in Table II, column NT. Consider the case in which an individual in the mixed state M is undecided about which one of R and I is acceptable, and let the interactions between speaker and hearer be symmetric. When an individual in a state I interacts with one in a state R both become confused on which form is right and therefore both switch to M. When an individual in a state M interacts with one in a state I or R, the outcome of

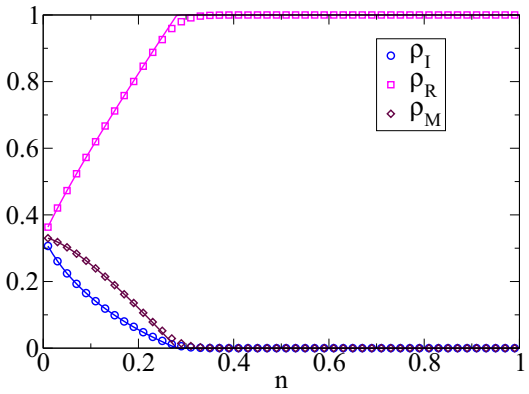

FIG. 4. (Color online) Stationary value of the fraction of irregulars ρ_I, regulars ρ_R, and mixed ρ_M, as a function of n, for the rules of model CT. Results of numerical simulations (symbols) are compared with the analytical predictions (solid lines). The rounding of the transition in simulation results is due to finite-size effects.

the interaction depends on which form the individual in the mixed state uses: if it is the same one used by his partner (with probability $1/2$) then nothing changes, if it is the alternative one (with probability $1/2$) then the partner becomes confused and also switches to the mixed state M. Interactions among individuals both in the same state I or R do not produce any change. The outcome of an interaction among two individuals in a state M depends on which form they use: if they both use the regular inflection (with probability $1/4$) they both switch to status R, if they both use the irregular inflection (probability $1/4$) they both switch to status I, if they use different inflections (probability $1/2$) they both remain in the mixed state M, coherently with the rules regulating the outcomes of the previous interactions.

It is easily seen that under this rule set the system converges to the fully regular state for any value of the ratio r/ν. The MF equations are

$$\begin{aligned}\dot{\rho}_I &= -r\rho_I + \nu\{-2\rho_I\rho_R + \tfrac{1}{2}[1-(\rho_I+\rho_R)]^2\}\\ \dot{\rho}_R &= r(1-\rho_R) + \nu\{-2\rho_I\rho_R + \tfrac{1}{2}[1-(\rho_I+\rho_R)]^2\}.\end{aligned} \tag{16}$$

Summing and subtracting the first equation from the second, one obtains

$$\begin{aligned}\dot{y}/\nu &= n(1-y)\\ \dot{x}/\nu &= n(1-x) - 2x + y^2 + 1,\end{aligned} \tag{17}$$

which shows that, for $n > 0$, the fully regular state $\rho_R = 1$, $\rho_I = 0$ is the only possible stationary state. The conclusion is that, for any value of ν (and $n > 0$), this set of rules always leads to a completely regular state for all individuals.

Unlike the previous models this one is discontinuous in the limit $n \to 0$: the model with replacement does not converge in the limit of vanishing replacement to the model with $n = 0$. It is easy to see that in the case $n = 0$ from Eq. (17) $\dot{y} \equiv 0$; therefore, the unbalance y_0 between ρ_I and ρ_R present in the initial condition is preserved during the dynamics, while x converges to $(1 + y_0)/2$. Therefore, the stationary state is continuously dependent on the initial condition, and given by $\rho_R = (1 + y_0)^2/4$, $\rho_I = (1 - y_0)^2/4$, which gives the fully regular solution only as long as the system is initiated in the fully regular state.

To conclude this section, we observe that in three-state models different microscopic interaction rules give rise to qualitatively different behaviors. In the next section we present a general approach to the modeling schemes presented so far, clarifying why they give rise to different phenomenologies.

IV. GENERAL THEORY

In this section we present a very general three-state model that provides a unified framework for generic sets of interaction rules, and we solve it analytically within the mean-field approximation. This framework allows us to comprehend the origin of the different behaviors found in the specific models investigated in the previous sections, providing a complete understanding of the global phenomenology of three-state models. We start by considering a general two-state model first, as this elucidates why the more complex three-state model is needed and how it behaves. We then consider a very general three-state model, encompassing all models considered before as particular cases. This approach will clarify a number of general points. In particular, it will show how the nature of the transition for both two and three-state models depends on the microscopic rules and clarify the role of asymmetries in the behavior of the speaker and hearer in the communication process.

A. General two-state model

Each individual is either in state R or I. At rate r each individual is replaced by one in state R. At rate ν an interaction occurs among two randomly selected individuals, the speaker and the hearer. We indicate the state of the pair of individuals in interaction as (X,Y), where X is the state of the speaker and Y of the hearer. As reasonable, we assume that nothing happens if the two individuals are in the same state $[(R,R) \to (R,R), (I,I) \to (I,I)]$. We first consider the case of deterministic rules, i.e., the state at the end of the interaction is fully determined by the initial state. Starting with the state (R,I) we parametrize the interaction rule by means of the coefficient γ_{RI}, which gives the variation in the number of individuals in state I. For example, for the interaction $[(R,I) \to (R,R)]$ $\gamma_{RI} = -1$, while for $[(R,I) \to (I,I)]$ $\gamma_{RI} = 1$. Analogously, when the initial state is (I,R) the rule is parametrized by γ_{IR}.

The mean-field equation for this process is simply

$$\dot{\rho}_I = -r\rho_I + \nu(\gamma_{RI} + \gamma_{IR})\rho_I(1-\rho_I). \tag{18}$$

Obviously, $\rho_R = 1 - \rho_I$. It follows immediately from Eq. (18) that assuming distinct asymmetric roles between speaker and hearer has no effect whatsoever on the collective behavior, since only the cumulative coefficient $\gamma = \gamma_{RI} + \gamma_{IR}$ enters the equation. The distinction between hearer and speaker is therefore *irrelevant* and any model defined by an asymmetric set of rules behaves exactly as its symmetrized version. This observation allows us to specify any two-state model by means of just one parameter, γ, with values between -2 and 2.

The general solution of Eq. (18) is obtained by replacing n with $2n/\gamma$ in Eqs. (2) and (3). The sign of γ determines the nature of the transition: for $\gamma > 0$ (rules biased in favor of I)

there exists a continuous transition with $n_c = \gamma$, while for $\gamma \leqslant 0$ (rules unbiased or biased against I) there is no transition, and the fully regular state is the only stationary solution. Therefore, we conclude in general that the system is driven toward a fully regular state unless a bias in the interactions compensates for the increase in the R population due to replacement.

In the most general case, the outcome of each interaction is decided probabilistically. In this case, γ_{IR} and γ_{RI} are defined as the *average* increase in I states in the interaction, each of them assuming any real value in $[-1,1]$. Correspondingly, γ can assume any real value in $[-2,2]$. It is immediate to realize that the dynamics is again described by Eq. (18) and all the above conclusions hold.

B. General three-state model

As demonstrated explicitly in the case of two-state models, asymmetric interaction rules (such as those in Table II) produce exactly the same behavior as their symmetrized version also in three-state models. It is therefore possible to express all possible interaction rules among individuals in a way analogous of what we have done for the two-state case. Let us define γ_i, ϕ_i, and δ_i as the average variation in the number of individuals in state I, R, or M, respectively, occurring when an interaction of type i takes place. We will denote with $i = 1$ the interactions with initial state (I,R) and (R,I), $i = 2$ for (R,M) and (M,R), $i = 3$ for (I,M) and (M,I), and $i = 4$ for (M,M). As above, we are assuming that no change occurs when two individuals in state R (or two individual in state I) interact. The interaction is instead nontrivial in the case (M,M). Conservation of the number of individuals implies $\gamma_i + \phi_i + \delta_i = 0$, for any $i = 1,2,3,4$, which reduces the number of independent parameters from 12 to 8. Notice that these quantities may be noninteger when we allow for different possible final states from a given initial state (each with a given probability).

With this parametrization of the dynamical rules we can write the rate equations for the evolution of the system in the most general case:

$$\begin{aligned}\dot\rho_I/\nu &= -n\rho_I + \{2\gamma_1\rho_I\rho_R + 2\gamma_2\rho_R[1-(\rho_I+\rho_R)] \\ &\quad + 2\gamma_3\rho_I[1-(\rho_I+\rho_R)] + \gamma_4[1-(\rho_I+\rho_R)]^2\} \\ \dot\rho_R/\nu &= n(1-\rho_R) + \{2\phi_1\rho_I\rho_R + 2\phi_2\rho_R[1-(\rho_I+\rho_R)] \\ &\quad + 2\phi_3\rho_I[1-(\rho_I+\rho_R)] + \phi_4[1-(\rho_I+\rho_R)]^2\},\end{aligned} \tag{19}$$

where $n = r/\nu$ is the rate of the replacement process relative to the frequency of interaction.

We now focus on the case of unbiased interactions that do not favor either the regular or the irregular form of the verb. In other words, the interaction rules are perfectly symmetric under the exchange between R and I; the only mechanism that breaks the symmetry between the regular and the irregular form is replacement, which favors the diffusion of the former. This choice is based on the observation that, while in the two-state case unbiased interaction rules always drive the system toward the fully regular state, the existence of a mixed state allows the survival of the irregular form even for some R-I symmetric interactions. This can be deduced from the three models presented in the previous section, which all have unbiased rules yet exhibit three qualitatively different behaviors. Hence symmetric interaction rules are general enough to lead to continuous or discontinuous transitions or to the absence of any transition.

The assumption of unbiased interaction implies the following additional relations among the parameters:

$$\phi_1 = \gamma_1, \quad \phi_2 = \gamma_3, \quad \phi_3 = \gamma_2, \quad \phi_4 = \gamma_4\,, \tag{20}$$

thus reducing the number of free parameters to 4 (we choose to use the four γ_i). The values of these parameters are not arbitrary. An (R,I) interaction cannot produce an increase in the number of individuals in the I state, since the same change must occur also for individuals in the R state, since $\gamma_1 = \phi_1$; hence, $\gamma_1 \leqslant 0$. In the same interaction the number of irregulars cannot decrease by more than 1; $-1 \leqslant \gamma_1$. With similar considerations it is not difficult to verify that the γ_i parameters are bounded as follows:

$$-1 \leqslant \gamma_1 \leqslant 0, \;\; 0 \leqslant \gamma_2 \leqslant 2, \;\; -1 \leqslant \gamma_3 \leqslant 1, \;\; 0 \leqslant \gamma_4 \leqslant 1. \tag{21}$$

Introducing the relation Eqs. (20) in Eqs. (19) it is possible to write Eqs. (19) in a particularly simple form by defining the auxiliary quantities $x = \rho_R + \rho_I$ and $y = \rho_R - \rho_I$, which represent the fraction of individuals in an unmixed state, and the excess fraction of R states with respect to I states, respectively:

$$\begin{aligned}\dot x/\nu &= ax^2 + cy^2 + 2dx + f \\ \dot y/\nu &= (2(\gamma_3-\gamma_2) - n)y - 2(\gamma_3-\gamma_2)xy + n,\end{aligned} \tag{22}$$

where

$$\begin{aligned} a &= \gamma_1 - 2(\gamma_2+\gamma_3) + 2\gamma_4 \qquad c = -\gamma_1 \\ d &= (\gamma_2+\gamma_3) - 2\gamma_4 - n/2 \qquad f = 2\gamma_4 + n. \end{aligned} \tag{23}$$

Physically sensible solutions must be in the range $0 \leqslant x \leqslant 1$ and $-x \leqslant y \leqslant x$. Notice that $c \geqslant 0$, $f \geqslant 0$, while the sign of a and d may vary and the coefficients are related by

$$a + c + 2d + f = 0\,. \tag{24}$$

Equation (24) implies that, for any choice of the parameters, $(x,y) = (1,1)$ [which corresponds to the fully regular state $(\rho_R = 1, \rho_I = 0)$] is always a stationary solution of Eq. (22). By imposing stationarity in Eqs. (22), other stationary solutions can be determined. For specific values of the parameters γ_i it is always straightforward to solve analytically for the stationary solutions of Eqs. (22) (which boils down to the solution of a third-order algebraic equation) and study their behavior as a function of n.

In the following we derive instead in full generality conditions on the parameters γ_i for the existence, as a function of n, of a continuous transition, a discontinuous one or no transition at all.

1. Case $\gamma_3 = \gamma_2$: No transition

Let us first consider the special class of models with $\gamma_3 = \gamma_2$. In this case the second of Eqs. (22) trivially yields at stationarity $y = 1$, giving the two solutions $(x_1,y_1) = (1,1)$ and $(x_2,y_2) = [-(2d+a)/a, 1]$. The second solution is always unphysical. Indeed, from Eq. (24) it follows that $-(2d+a) > 0$. Hence, x_2 is smaller than 0 if $a < 0$. On the other hand,

TABLE III. The table summarizes the values of the set of γ_i parameters for the three models studied in the previous section.

	Model NT	Model CT	Model NG
γ_1	-1	$-1/2$,	$-1/2$
γ_2	0	1/4	0
γ_3	0	0	1/2
γ_4	1/2	1/2	1

for $\gamma_3 = \gamma_2$ the value of $\gamma_3 + \gamma_2$ is always positive, since $\gamma_2 \geqslant 0$. This implies $-(2d + a) > a$ so that $x_2 > 1$ if $a > 0$. We conclude that when $\gamma_3 = \gamma_2$ the only possible stationary state is the fully regular one: no transition may occur. As shown in Table III the NT model considered in the previous section falls in this class of models, featuring $\gamma_2 = \gamma_3 = 0$. We will see below that also the degenerate case $\gamma_1 = 0$ implies no transition, irrespective of the value of the other parameters.

2. *Case $\gamma_3 \neq \gamma_2$: Existence of a transition*

Assuming now $\gamma_3 \neq \gamma_2$, and imposing stationarity in Eq. (22), we get

$$\begin{aligned} ax^2 + cy^2 + 2dx + f &= 0[\mathcal{C}_1] \\ y[x - (1 - \epsilon)] &= \epsilon[\mathcal{C}_2], \end{aligned} \tag{25}$$

where $\epsilon = n/[2(\gamma_3 - \gamma_2)]$. The solutions of Eq. (25) are given by the intersections between two conic sections $\mathcal{C}_1$ and $\mathcal{C}_2$. $\mathcal{C}_2$ is a hyperbola with asymptotes $y = 0$, and $x = 1 - \epsilon$. Depending on the sign of ϵ, the two branches of the hyperbola lie in different quadrants with respect to the asymptotes (see Fig. 5). In the limit $\epsilon \to 0$, which corresponds to the case with no replacement, the hyperbola degenerates into the pair of asymptotes, $y = 0$ and $x = 1$. The conic section $\mathcal{C}_1$ is an ellipse for $a > 0$ and an hyperbola for $a < 0$, turning into a parabola for $a = 0$, with axes in all cases parallel to the Cartesian coordinate system. The intersections of $\mathcal{C}_1$ with the $y = 0$ axis are, for $a \neq 0$,

$$x_1 = \frac{-d - \Delta}{a}, \quad x_2 = \frac{-d + \Delta}{a}, \tag{26}$$

where $\Delta = \sqrt{d^2 - af} = \sqrt{(d + f)^2 + cf} > 0$. It is not difficult to show that x_1 is always physical (between 0 and 1) while x_2 is always unphysical (see Appendix B). In particular, $x_2 > 1$ if $a > 0$ and $x_2 < 0$ if $a < 0$. In the limit $a \to 0$, $x_1 \to -f/(2d)$ (which is the intersection point for the parabolic case $a = 0$) while $x_2 \to \pm\infty$, depending on the sign of a. Hence, despite the change in the global behavior for different values of a, $\mathcal{C}_1$ has always a similar shape in the region of physical interest $0 \leqslant x \leqslant 1$: it crosses the $y = 0$ axis for $x = x_1$ and is concave toward the right, passing through the two points $(x, y) = (1, 1)$ and $(x, y) = (1, -1)$ for any value of the parameters. Notice that also $\mathcal{C}_2$ always goes through the point $(x, y) = (1, 1)$, so that the fully regular state is always a stationary solution.

We now investigate in full generality the possible existence of other stationary solutions, i.e., other intersections in the physical region. Only the case $c = 0$ ($\gamma_1 = 0$) needs to be treated separately, because the conic section $\mathcal{C}_1$ degenerates into a pair of lines: $(x - 1)[x + (2d + a)/a] = 0$, one ($x = 1$) on the boundary of the physical region, the other outside it. The fully regular solution $(x, y) = (1, 1)$ is then the only stationary state, and there is no transition. We will assume $c \neq 0$ in what follows.

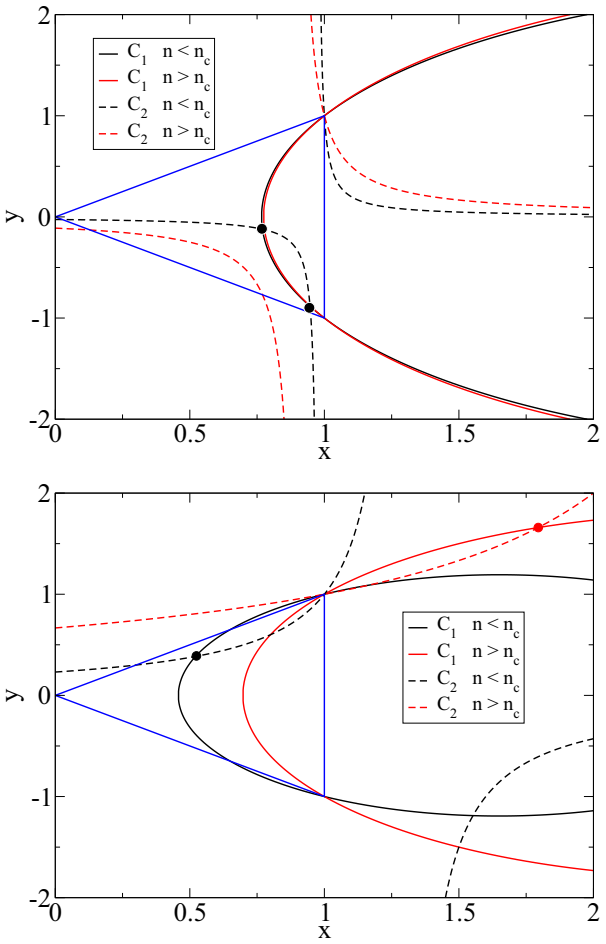

FIG. 5. (Color online) Plot of the two conic sections $\mathcal{C}_1$ and $\mathcal{C}_2$ for $\gamma_3 > \gamma_2$ (top) and for $\gamma_3 < \gamma_2$ (bottom), illustrating the intersections above [red (gray) curves] and below (black curves) the critical n_c. The triangle is the region of physical interest. The circles denote the intersections other than the fully regular solution.

$\gamma_3 > \gamma_2$: *Discontinuous transition.* For $\gamma_3 > \gamma_2$ ($\epsilon > 0$), the hyperbola $\mathcal{C}_2$ lies in the upper-right and lower-left quadrants with respect to the asymptotes (see Fig. 5). Notice that the vertical asymptote is always at $x < 1$. Apart from the fully regular solution, one intersection occurs always for $x < 0$ or $x > 1$ and is thus unphysical. Two other intersections may instead occur between $\mathcal{C}_1$ and the lower-left branch of $\mathcal{C}_2$. For large n these intersections do not exist as it can be recognized by observing that for $\epsilon > 1$ the lower branch of $\mathcal{C}_2$ has $x < 0$ while $\mathcal{C}_1$ has $x > x_1 > 0$. However, when $n \to 0$, $\mathcal{C}_2$ shrinks toward its asymptotes $x = 1 - \epsilon$ and $y = 0$, while $\mathcal{C}_1$ does not change much. At some critical value $n = n_c$, $\mathcal{C}_2$ starts intersecting $\mathcal{C}_1$, so that for $n < n_c$ there are two intersections (see Fig. 5), both in the physical region because the derivative of $\mathcal{C}_1$ for $x = 1$ is $-[1 + (d + f)/c] < -1$. In this case the system undergoes a discontinuous transition at n_c. Notice when these two solutions

exist they are in the region $y \leqslant 0$, implying that $\rho_I \geqslant \rho_R$, i.e., the fraction of individuals in the I state is larger than the fraction of those in the R state. As indicated in Table III, the model NG, which features $\gamma_3 - \gamma_2 = 1/2$, belong to this class and this explains why it undergoes a discontinuous transition.

$\gamma_3 < \gamma_2$*: Continuous transition.* For $\gamma_3 < \gamma_2$ ($\epsilon < 0$), the hyperbola $\mathcal{C}_2$ lies in the upper-left and lower-right quadrants (see Fig. 5). The lower branch has $x > 1 - \epsilon > 1$ and hence is unphysical. Therefore, at most two of the four solutions (the intersections of the upper branch with $\mathcal{C}_1$) are physical. One of them is the fully regular state. To investigate the location of the other intersection one can compare the slope of the two conic sections for $x = 1$. If the slope of $\mathcal{C}_1$ is larger than the slope of $\mathcal{C}_2$ (which happens for large n) the second intersection occurs for $x > 1$, it is unphysical and as a consequence the fully regular state is the only stationary solution. If n is reduced, the slope of $\mathcal{C}_1$ at $x = 1$ decreases while the slope of $\mathcal{C}_2$ grows. At a critical value n_c the two slopes are equal and for $n < n_c$ the second intersection becomes physical ($x < 1$). We conclude therefore that the system undergoes a continuous transition between a fully regular state and a state with coexisting regular and irregular individuals. Notice that in this case, since the physical intersections have $y \geqslant 0$, necessarily $\rho_I \leqslant \rho_R$. The value of n_c is easily determined by the condition that the two slopes are equal and turns out to be

$$n_c = (\gamma_1 - \gamma_2 - \gamma_3) \pm \sqrt{(\gamma_1 - \gamma_2 - \gamma_3)^2 + 4\gamma_1(\gamma_3 - \gamma_2)}, \quad (27)$$

which has always one positive value (the other being always negative), coherently with the fact that there is always a transition. Remarkably, the value of n_c does not depend at all on the coefficient γ_4, regulating the M-M interaction. The model CT in the previous section has $\gamma_3 = 0$, $\gamma_2 = 1/4$, and $\gamma_1 = -1/2$ (see Table III). These values explain why it undergoes a continuous transition at $n_c = (\sqrt{17} - 3)/4$.

3. *Stability analysis*

So far we have shown that below some critical value of n additional stationary solutions appear, beyond the fully regular solution. To complete the demonstration of the existence of phase transitions we must analyze their stability.

By linearizing Eqs. (22) around the solution (x^*, y^*), one gets

$$\frac{1}{2v}\frac{d}{dt}\begin{bmatrix}\delta x \\ \delta y\end{bmatrix} = \begin{bmatrix} ax^* + d & cy^* \\ -(\gamma_3 - \gamma_2)y^* & (\gamma_3 - \gamma_2)(1 - x^*) - \frac{n}{2}\end{bmatrix}\begin{bmatrix}\delta x \\ \delta y\end{bmatrix}. \quad (28)$$

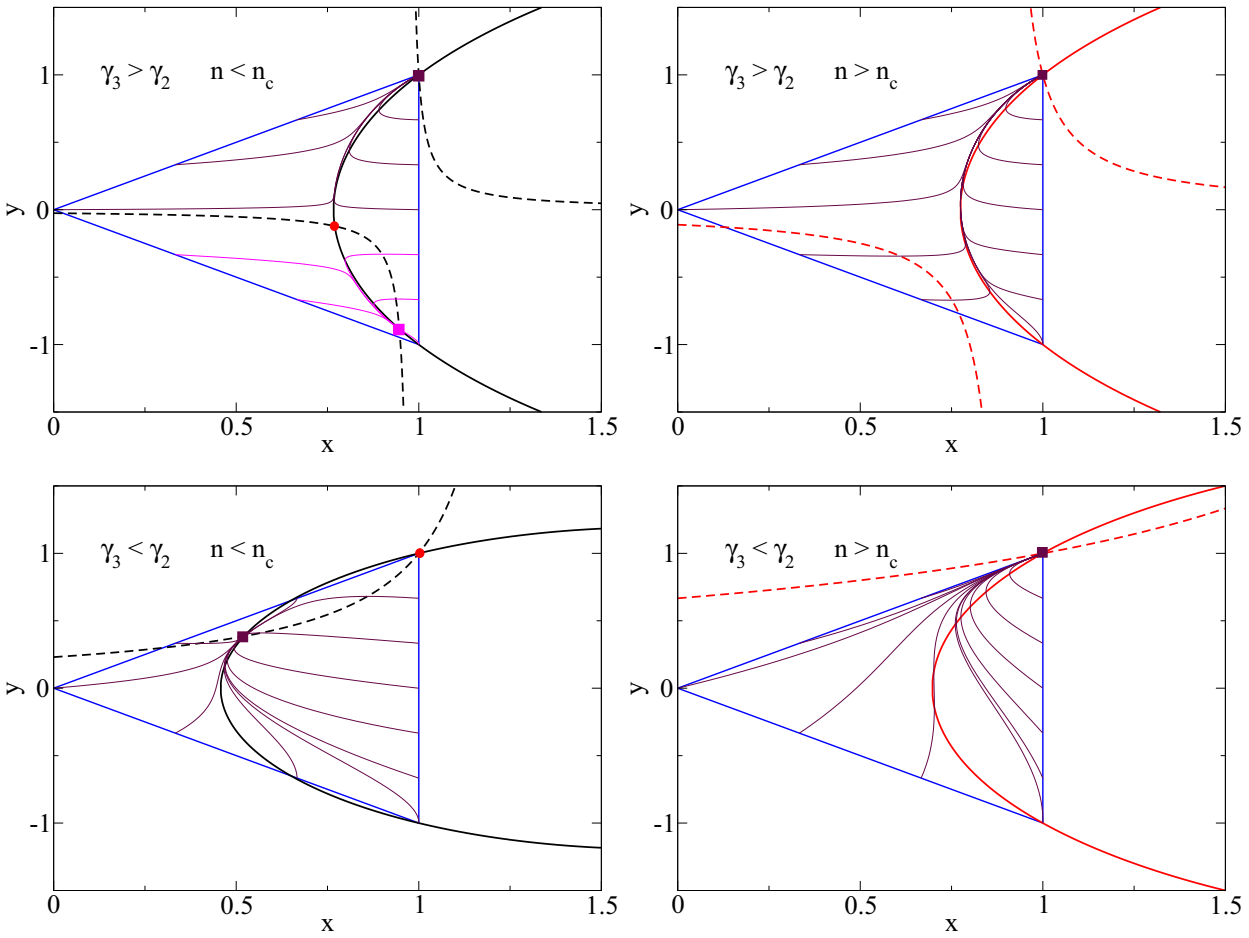

FIG. 6. (Color online) Phase flow in the (x, y) space for discontinuous (top) and continuous (bottom) transition both below (left) and above (right) the transition. The triangle is the region of physical interest. The thick curves are $\mathcal{C}_1$ (solid) and $\mathcal{C}_2$ (dashed). Symbols indicate stationary solutions. The red circles are saddle points with an attractive and a repulsive direction. The squares are attractive stable solutions. Thin solid lines are flow lines converging to the stable solution denoted by symbols with the same color (shade).

For the fully regular state $(x^*, y^*) = (1,1)$ the eigenvalues of the stability matrix M can be evaluated explicitly yielding

$$\lambda_{1,2} = [\mathrm{tr}(M) \pm \sqrt{\Delta_1}]/2, \tag{29}$$

where the trace of the matrix is $\mathrm{tr}(M) = \gamma_1 - \gamma_2 - \gamma_3 - n$ and $\Delta_1 = (\gamma_1 + \gamma_2 + \gamma_3)^2 - 8\gamma_1\gamma_2 \geqslant 0$. Notice that Δ_1 is independent of n and positive, implying that the two eigenvalues are always real.

In the case $\gamma_3 > \gamma_2$, both eigenvalues are negative, as can be deduced by considering that $\mathrm{tr}(M)$ (the sum of the eigenvalues) is always negative (being the sum of four negative terms) and $\det(M) = n^2/4 - n(\gamma_1 - \gamma_2 - \gamma_3)/2 - \gamma_1(\gamma_3 - \gamma_2)$ (the product of the eigenvalues) is always positive for $n > 0$. This last condition can be understood by considering that $\det(M)$ is a parabola with upward concavity and zeros for negative values of n [given by Eq. (27)]. Hence, the fully regular state is always stable. For small n, the two other solutions appearing in the physical region are a saddle and a stable fixed point [see Fig. 6]: a discontinuous transition occurs.

In the case $\gamma_3 < \gamma_2$, corresponding to a continuous transition, $\det(M)$ is positive for large n, but it changes sign for $n < n_c$ [where n_c is the only positive determination of Eq. (27)] thus implying that one of the eigenvalues becomes positive. Hence, below the transition the fully regular state corresponds to a saddle, while the other solution appearing in the physical region is stable [see Fig. 6].

Finally, we remark that the parameters corresponding to the naming game, considered in the previous section as an example of models with discontinuous transition, give $\Delta_1 = 0$, corresponding to two degenerate eigenvalues, and only one eigenvector. In this case the fully regular state is stable but *defective*: in principle arbitrarily small changes of the parameters may result in a standard stable fixed point if $\Delta_1 > 0$, or in a stable spiral when $\Delta_1 < 0$. This last case would change the physical picture, because spiralling trajectories would cross the boundary of the physical domain before reaching the fixed point, and therefore the physical system would be driven toward the boundary of the physical region without reaching the fixed point. However, this possibility is ruled out by the condition $\Delta_1 \geqslant 0$, which holds for physically sensible values of the parameters γ_i.

4. Wrap-up of the theory for the general model

The conclusions that follow from the theory presented in this section are very general and simple. The existence and the nature of a transition depends only on the sign of $\gamma_3 - \gamma_2$. If $\gamma_3 > \gamma_2$, as the frequency increases a discontinuous transition occurs between a fully regular state and a state with coexistence of fully regular and mostly irregular inflections. If $\gamma_3 < \gamma_2$, the transition is instead continuous. If $\gamma_3 = \gamma_2$, no transition occurs and the system always reaches a fully regular state.

It is important to notice that the condition $\gamma_3 > \gamma_2$ has a very natural interpretation in the context of language dynamics. It simply means that (recalling that $\gamma_2 = \phi_3$) an interaction between an individual in state M and one in state I will produce an increment in the use of the irregular inflection larger than the increment in the use of the regular inflection. The fact that this asymmetry alone is sufficient to give rise to a discontinuous transition is a strong indication of the relevance of these theoretical modeling efforts for the interpretation of empirical data.

V. CONCLUSIONS

This work has presented an investigation of agent-based models aimed at understanding the processes regulating the interplay between rules and exceptions in language dynamics. In particular, the models aim to investigate the observed behavior of verbs in natural language. Corpus data from natural language points to the existence of a discontinuous transition as a function of the frequency of usage: high-frequency items are highly irregular and low-frequency ones are regular, while in an intermediate frequency range coexistence between the two behaviors is observed.

In the minimal models considered each agent is endowed with an inventory, containing the possible inflections (regular or irregular) of a lemma. Two processes have the potential to change agents' inventories over time: interaction and replacement. In interaction, individuals influence each other, adding or deleting forms from their inventories according to a specific set of interaction rules. In replacement events, agents are substituted by new "child" individuals, who are automatically biased toward the regular form by being "born" with a regular inventory.

We analyze two classes of models. In the first one each individual may store in the inventory only one of the two competing inflections, either the regular or the irregular one. Three-state models instead integrate a mixed state, which represents an individual who finds both the regular and irregular forms acceptable. We solve these models analytically within the mean-field framework and confirm the results by means of numerical simulations. We first focus on a few specific models, including the naming game for language dynamics. The analysis reveals that the the global phenomenology changes qualitatively depending on the interaction rules: one can observe the absence of a transition with a move to total regularity, a continuous transition, or a discontinuous one. We then consider a very general three-state model, encompassing all previous examples as special cases, which allows the description of the previous models with a set of four minimal parameters describing the interaction rules. From this comprehensive approach several results follow:

(i) Asymmetries in the influence of the speaker and of the hearer in interaction do not play any role in the collective behavior of the system.

(ii) In two-state models the fully regular state is the only attractor unless the interaction rules are biased in favor of the irregular inflection; the three-state models have instead nontrivial behavior even when the rules are unbiased;

(iii) Allowing for a third state is crucial for the appearance of a discontinuous transition that cannot arise in two-state versions of the model.

(iv) In three-state models the quantity $\gamma_3 - \gamma_2$ rules the macroscopic behavior by changing the nature of the transition: when $\gamma_3 > \gamma_2$, a discontinuous transition is observed to a

state where irregular inflection is prevalent ($\rho_I > \rho_R$); in the opposite case, when $\gamma_3 > \gamma_2$ a continuous transition is observed, to a state with $0 < \rho_R < \rho_I$; in the case $\gamma_3 = \gamma_2$ there is no transition and the fully regular state is reached for any frequency.

(v) In the case $\gamma_3 > \gamma_2$, above the discontinuous transition the steady state depends on the initial condition: verbs with the same frequency can end up as fully regular or mostly irregular, similarly to what is observed in empirical data.

(vi) In the context of language dynamics, the condition $\gamma_3 > \gamma_2$ is satisfied by sets of more plausible rules, so that a discontinuous transition is to be expected.

This model provides a framework that could potentially be used to consider additional, more complex aspects of rule dynamics in language. In particular, empirical data shows that the growth of language contributes to the expansion of regularity [9], since a core aspect of a linguistic rule's utility is that it can be generatively applied to new forms (e.g., the past tense of the neologism *selfie* is uncontroversially *selfied*). Our model considers a word's frequency to be static over time; however, natural languages are living, and populations, vocabulary sizes, and turnover rates are not static. Furthermore, there are other cognitive mechanisms beyond child learner biases that may contribute to regularity dynamics. General memory constraints may contribute to the persistence of highly frequent irregular forms [32], and adult learners may possess qualitatively different regularization biases from children [33]. Moreover, the use of a disordered topology for the pattern of interaction, as opposed to the homogeneous mixing assumed by the MF approach, combined with the interaction among the different lemmas in the agents' inventories may lead to different patterns of regularization in frequency. Future models might also consider another key aspect in the persistence of irregularity: the notion that irregular forms are not always exceptions, but sometimes constitute subrules [34] (e.g., *foot-feet/goose-geese, sing-sang/ring-rang*). Our model provides a basic starting point from which to consider the complex dynamics underlying temporal trends of the rules that form the core of language.

Finally, it is very important to stress that while models and results are presented in this paper in terms of linguistic rule dynamics, they are fully general and apply to any system where individuals have three possible internal states and the population exhibits turnover. The results presented in this paper, and in particular the conditions determining the existence of a transition and its nature, may have strong implications not only for linguistic rules, but also for all those systems.

ACKNOWLEDGMENTS

This work was supported by the European Science Foundation as part of the DRUST project, a EUROCORES EuroUnderstanding programme.

APPENDIX A

In this appendix we report the stability analysis for the naming game without biased replacement, i.e., with $n = 0$ as discussed in Sec. III A. For the case $n = 0$ the stability matrix is given by

$$\frac{d}{dt}\begin{bmatrix}\delta\rho_I\\ \delta\rho_R\end{bmatrix} = \begin{bmatrix}-1 & 2(\rho_R-1)\\ 2(\rho_I-1) & -1\end{bmatrix}\begin{bmatrix}\delta\rho_I\\ \delta\rho_R\end{bmatrix}, \tag{A1}$$

whose eigenvalues are

$$\lambda_{1,2} = \pm\sqrt{4(\rho_I^*-1)(\rho_R^*-1)} - 1, \tag{A2}$$

where, as before, (ρ_I^*, ρ_R^*) indicates the generic stationary solution. For $(\rho_I^{(1)}, \rho_R^{(1)})$ and $(\rho_I^{(2)}, \rho_R^{(2)})$ both eigenvalues are negative and the solutions are both stable. For $(\rho_I^{(3)} \simeq 0.382, \rho_R^{(3)} \simeq 0.382)$ the eigenvalues are one positive ($\lambda_1 = 2.236$) and one negative ($\lambda_2 = -4.236$) and this corresponds to a saddle point with an attractive and a repulsive direction. The separatrix in Fig. 3 (left) in the attractive direction corresponds to the eigenvector associated to the negative eigenvalue: $\rho_R = \rho_I$ (the red line in figure). The other separatrix is locally approximated in the neighborhood of $(\rho_I^{(3)}, \rho_R^{(3)})$, by the eigenvector associated to the positive eigenvalue $\rho_R = -\rho_I + 2\rho_I^{(3)}$ (the green line in figure). The phase flow is such that if the initial condition is such that $\rho_I < \rho_R$ ($\rho_I > \rho_R$) the system will deterministically converge to the regular (irregular) state. On the other hand, if the initial condition is such that $\rho_I = \rho_R$ the system will converge to the stationary solution $(\rho_I^{(3)}, \rho_R^{(3)})$ (see also Refs. [12,31]).

APPENDIX B

In this appendix, we provide explicit proofs that the intersections x_1 and x_2 of the conic section $\mathcal{C}_1$ with the axis $y = 0$ [see Eq. (26)] are always physical and always unphysical, respectively.

Let us first consider the case $a > 0$. A crucial point to recognize is that, since $a + 2d = -(c + f) < 0$, if a is positive d must be negative. Hence, $\Delta = \sqrt{d^2 - af} \leqslant |d| = -d$. As a consequence, $x_1 = -(\Delta + d)/a \geqslant 0$. The alternative expression of $\Delta = \sqrt{(a+d)^2 + ac}$ implies, since both a and c are positive, $\Delta \geqslant |a + d|$. This means that $\Delta \geqslant -(a + d)$ so that $x_1 = -(\Delta + d)/a \leqslant 1$. It also implies $\Delta \geqslant (a + d)$, which, inserted into the expression $x_2 = (\Delta - d)/a \geqslant 1$, yields $x_2 \geqslant 1$.

The arguments are similar for $a < 0$. In this case, $\Delta = \sqrt{d^2 - af} \geqslant |d| \geqslant -d$ so that $x_1 = (-d - \Delta)/a \geqslant 0$. By the same token, $\Delta = \sqrt{d^2 - af} \geqslant d$, implying $x_2 = (\Delta - d)/a \leqslant 0$. Finally, to show that $x_1 \leqslant 1$ we start from $\Delta = \sqrt{(a+d)^2 + ac} \leqslant |a + d|$. The quantity $a + d$ is negative for $a < 0$, because $a + 2d < 0$ implies $2a + 2d < a < 0$. Hence, $\Delta \leqslant |a + d| = -(a + d)$, so that $x_1 = -(d + \Delta)/a < 1$.

[1] C. F. Hockett, Language **37**, 29 (1961).
[2] N. Chomsky, *Knowledge of Language: Its Nature, Origin, and Use* (Greenwood Publishing Group, Santa Barbara, CA, 1986).
[3] S. Pinker, *Words and Rules: The Ingredients of Language* (Basic Books, New York, 1999).
[4] J. L. McClelland and K. Patterson, Trends Cogn. Sci. **6**, 464 (2002).

[5] S. Pinker and M. T. Ullman, Trends Cogn. Sci. **6**, 456 (2002).
[6] J. L. Bybee and D. I. Slobin, Language **58**, 265 (1982).
[7] W. Marslen-Wilson and L. K. Tyler, Trends Cogn. Sci. **2**, 428 (1998).
[8] K. Plunkett and V. A. Marchman, Cognition **61**, 299 (1996).
[9] C. F. Cuskley, M. Pugliese, C. Castellano, F. Colaiori, V. Loreto, and F. Tria, PLoS ONE **9**, e102882 (2014).
[10] E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, and M. A. Nowak, Nature **449**, 713 (2007).
[11] R. Carroll, R. Svare, and J. C. Salmons, J. Hist. Linguist. **2** 153 (2012).
[12] A. Baronchelli, V. Loreto, and L. Steels, Int. J. Modern Phys. C **19**, 785 (2008).
[13] V. Loreto, A. Baronchelli, A. Mukherjee, A. Puglisi, and F. Tria, J. Stat. Mech. (2011) P04006.
[14] W. Labov, *Sociolinguistic Patterns*, Vol. 4 (University of Pennsylvania Press, Philadelphia, 1972).
[15] R. Dale and G. Lupyan, Adv. Complex Syst. **15**, 1150017 (2012).
[16] R. A. Berman, Papers Rep. Child Lang. Dev. **20**, 34 (1981).
[17] J. L. Bybee and D. I. Slobin, in *Papers from the 5th International Conference on Historical Linguistics*, Vol. 21 (John Benjamins Publishing Company, 1982).
[18] R. J. Maslen, A. L. Theakston, E. V. Lieven, and M. Tomasello, J. Speech Lang. Hear. Res. **47**, 1319 (2004).
[19] D. M. Abrams and S. H. Strogatz, Nature **424**, 900 (2003).
[20] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).
[21] C. Castellano, S. Fortunato, and V. Loreto, Rev. Modern Phys. **81**, 591 (2009).
[22] W. S.-Y. Wang and J. W. Minett, Trends Ecol. Evol. **20**, 263 (2005).
[23] J. W. Minett and W. S.-Y. Wang, Lingua **118**, 19 (2008).
[24] X. Castelló, V. M. Eguíluz, and M. San Miguel, New J. Phys. **8**, 308 (2006).
[25] P. Clifford and A. Sudbury, Biometrika **60**, 581 (1973).
[26] R. Holley and T. Liggett, Ann. Probabil. **3**, 643 (1975).
[27] A. Baronchelli, M. Felici, E. Caglioti, V. Loreto, and L. Steels, J. Stat. Mech. (2006) P06014.
[28] A. Puglisi, A. Baronchelli, and V. Loreto, Proc. Natl. Acad. Sci. U.S.A. **105**, 7936 (2008).
[29] A. Mukherjee, F. Tria, A. Baronchelli, A. Puglisi, and V. Loreto, PloS ONE **6**, e16677 (2011).
[30] F. Tria, B. Galantucci, and V. Loreto, PLoS ONE **7**, e37744 (2012).
[31] A. Baronchelli, L. Dall'Asta, A. Barrat, and V. Loreto, Phys. Rev. E **76**, 051102 (2007).
[32] C. Cuskley, C. Castellano, F. Colaiori, V. Loreto, M. Pugliese, and F. Tria, in *The Evolution of Language*, Proceedings of the Tenth International Conference EVOLANG 10, Vienna, Austria, 2014, edited by E. Cartmill, S. Roberts, H. Lyn, and H. Cornish (World Scientific, Singapore, 2014), pp. 417–418.
[33] C. L. Hudson Kam and E. L. Newport, Cogn. Psychol. **59**, 30 (2009).
[34] C. Yang, Ling. Var. Yearbook **5**, 265 (2005).