Physica A 388 (2009) 174–186

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

# Influence of geography on language competition

Marco Patriarca *, Els Heinsalu

*National Institute of Chemical Physics and Biophysics, Rävala 10, 15042 Tallinn, Estonia*
*Institute of Physics, University of Tartu, Tähe 4, Tartu 51010, Estonia*

ARTICLE INFO

*Article history:*
Received 18 July 2008
Received in revised form 8 September 2008
Available online 9 October 2008

*PACS:*
89.65.Ef
87.23.Ge
89.65.-s
89.75.-k

*Keywords:*
Language competition
Culture diffusion
Reaction-diffusion

ABSTRACT

Competition between languages or cultural traits diffusing in the same geographical area is studied combining the model of Abrams and Strogatz with a model of human dispersal on an inhomogeneous substrate. Also, the effect of population growth is discussed. It is shown through numerical simulations that the final configuration of the languages can be strongly affected by geographical and historical factors. These factors are not related to the dynamics of culture transmission, but rather to initial population distributions as well as geographical boundaries and inhomogeneities, which modulate the diffusion process.



## 1. Introduction

Currently, statistical mechanics and stochastic models are employed to investigate a wide range of topics, not only in physical sciences, but also beyond, e.g. in biology as well as in social and historical sciences. Examples include studies of financial time series [1], population dynamics [2], and archeology [3]. Recently, also various problems in linguistics have been approached using methods imported from the theory of complex systems and statistical mechanics; see Refs. [4–8] for an overview.

In the evolution and dispersal of biological species the importance of geography is well known [9]. The goal of the present paper is to study how purely geographical and historical constraints can affect the dynamics of different languages or cultural traits competing in a region. Previously the influence of geography on language dynamics has been investigated e.g. in Refs. [10–13].

We start from the Abrams–Strogatz (AS) model [14] of two competing languages. The languages are competing in the sense that at any time speakers can switch to the other language, as a consequence of the interaction between speakers of language 1 and 2. The evolution of languages is neglected on the time scale considered, so that the model is formally similar to a model of population dynamics of two biological species. In order to take into account population growth, dispersal, and the effect of geographical inhomogeneities, we introduce in Section 2 a more general model. It is then applied to some (idealized) examples concerning the influence of initial conditions (Section 3), boundary conditions (Section 4), and geographical barriers (Sections 5 and 6). These examples show how extending a 0-dimensional (i.e. homogeneous) model of culture transmission to physical space gives rise to new, unexpected effects. The model used in the present article describes the speakers communities at a coarse-grained level. Other more detailed models, such as those by Nowak [15] and by Schulze and Stauffer [8], take into account the dynamics of single speakers as well as language evolution. Some different

* Corresponding author at: National Institute of Chemical Physics and Biophysics, Rävala 10, 15042 Tallinn, Estonia. Fax: +372 73 755 79.
*E-mail addresses:* marco.patriarca@gmail.com, marco.patriarca@yahoo.com (M. Patriarca), els@kbfi.ee (E. Heinsalu).

0378-4371/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2008.09.034

models also address the language dynamics; for example the naming game model of Steels [16] investigates the emergence of a common vocabulary from the interaction between individuals. Recently it has been studied also with various types of social topologies [17,18] or taking into account population growth [19,20].

It is also to be noticed that, in general, bilingual communities have an important role in the dynamics of language competition, as discussed in various recent reports [21–25]. However, in the present paper we adhere to the approximate scheme underlying the AS model which neglects the bilinguals, yet providing a reasonably good description of some empirical data sets of endangered languages [14].

## 2. The model

In this section we formulate the model of culture dispersal and competition. We are interested in a description regarding a historical context,[1] in which two populations using different languages (or with different cultural traits) 1 and 2 disperse across a region. Simultaneously, the competition between the languages takes place and possibly the populations also grow. As mentioned, we start from the AS language competition model. A diffusion term, describing population dispersal, and an advection term, taking into account geographical inhomogeneities, are then added; population growth is modeled through a logistic term.

### *2.1. The Abrams–Strogatz model*

The AS model was first introduced in Ref. [14] in order to describe the time evolution of the population size of various endangered languages. It can be formulated through the following equations,

$$\begin{aligned}\frac{\mathrm{d}N_1}{\mathrm{d}t} &= R(N_1, N_2) = \frac{s_1}{\tau} N_1^a N_2 - \frac{s_2}{\tau} N_2^a N_1, \\ \frac{\mathrm{d}N_2}{\mathrm{d}t} &= -R(N_1, N_2) = -\frac{s_1}{\tau} N_1^a N_2 + \frac{s_2}{\tau} N_2^a N_1.\end{aligned} \tag{1}$$

Here $N_i(t)$ $(i = 1, 2)$ represents the fraction of speakers of language $i$ and $N_1(t) + N_2(t) = 1$. The quantity $k_1 = s_1/\tau$ is the rate constant for the switch of a speaker of language 2 to language 1, and vice versa for $k_2 = s_2/\tau$. The dimensionless parameter $s_i$, referred to as *language status*, represents the *attractiveness* of language $i$; it can result from a combination of factors such as, e.g. the language prestige and usefulness. Here we follow the normalization convention $s_1 + s_2 = 1$. As a consequence $s_1/\tau + s_2/\tau \equiv 1/\tau$, i.e., $\tau$ represents the time scale. For the coefficient $a$ the value $a = 1.3$ is assumed.

In the analogy with a population dynamics model, it should be noticed that the reaction term $R(N_1, N_2)$ in Eqs. (1) contains a positive and a negative contribution, depending on both populations $N_1$ and $N_2$, and representing an advantage and disadvantage due to the encounter with an individual of the other "species". For any $a \neq 1$, speakers of language 1 and 2 behave symmetrically to each other as prey and predator at the same time [2]. While this situation is not usual in biology, it can be justified for the interaction between two cultural traits [14].

The analysis shows that the AS model has one unstable and two stable equilibrium points. The latter ones correspond to one of the languages surviving and the other one disappearing. Which final state will be reached depends on the initial population sizes $N_i(t_0)$, the status parameters $s_i$, as well as on the value of $a$. The critical values of parameters defining the unstable equilibrium point can be obtained from Eqs. (1) setting the rate term $R$ equal to zero; it follows that

$$\frac{N_1^*}{N_2^*} = \left(\frac{s_2}{s_1}\right)^{1/(a-1)}. \tag{2}$$

When the ratio $N_1^*/N_2^* > (s_2/s_1)^{1/(a-1)}$ at some time $t = t'$, then $R(N_1(t), N_2(t)) > 0$ at any later time $t > t'$ and $N_2 \to 0$ for $t \to \infty$, while $N_1 \to N' \equiv N_1(t') + N_2(t')$. The opposite takes place if $N_1^*/N_2^*$ is smaller than the right-hand side of condition (2) at any time $t'$.

### *2.2. Generalized model*

Population and culture spreading may be affected by a wide range of geographical factors, e.g. due to physical barriers such as water boundaries and mountains or geophysical features such as type of ground and spatial distribution of resources [9]. While the underlying mechanisms determining the influence of geographical factors are in general complex, in a first approximation their overall effect can be described at a coarse-grained level in the framework of statistical mechanics. In fact, dispersal of human populations in a geographical environment recalls the diffusion of Brownian particles modulated by an external field or an inhomogeneous substrate. For example, human dispersal in neolithic Europe [26] and during the

[1] Our goal in this paper is not to reconstruct any real historical situation.

early colonization of South-America [27] has been studied previously using advection-diffusion equations with a logistic term taking into account population growth (i.e., employing the so-called Fisher equation).

In order to extend the AS model to take into account the geographical inhomogeneities, we merge it with the two-dimensional geographical model of human dispersal and growth proposed in Refs. [26,27]. The corresponding evolution equations read,

$$\begin{aligned} \frac{\partial f_1}{\partial t} &= R(f_1, f_2) - \nabla \cdot (\mathbf{F} f_1) + \nabla \cdot (D \nabla f_1) + \alpha f_1 \left(1 - \frac{f_1 + f_2}{K}\right), \\ \frac{\partial f_2}{\partial t} &= -R(f_1, f_2) - \nabla \cdot (\mathbf{F} f_2) + \nabla \cdot (D \nabla f_2) + \alpha f_2 \left(1 - \frac{f_1 + f_2}{K}\right), \end{aligned} \tag{3}$$

with the reaction term given by

$$R(f_1, f_2) = k \left( s_1 f_1^a f_2 - s_2 f_2^a f_1 \right). \tag{4}$$

The quantity $f_i = f_i(x, y, t)$ represents the population density of speakers of language $i$, that is the number of speakers of language $i$ per unit area. The constant $k$ in Eq. (4) is an effective rate constant and $s_i$ still represents the status of language $i$. In Eqs. (3) population movement is described by the advection term containing the external force field $\mathbf{F}(x, y) = (F_x(x, y), F_y(x, y))$ and by the diffusion term with the diffusion coefficient $D = D(x, y)$. In general $\mathbf{F}$ and $D$ are related; their explicit form depends on the problem considered. One possibility is to set $\mathbf{F}(x, y) = 0$ and describe the inhomogeneous character of the substrate through a space-dependent diffusion coefficient $D(x, y)$ [26,27]. In the model systems studied below we ascribe the inhomogeneous character of dispersal to the external force $\mathbf{F} = \mathbf{F}(x, y)$, while $D$ is kept constant. For illustrative purposes and in analogy with Brownian motion, the force field is expressed as the gradient of a potential, $\mathbf{F}(x, y) = -\nabla U(x, y)$. The logistic terms with Malthus rate $\alpha$ and carrying capacity $K$ in Eqs. (3) take into account the population growth.

According to Eqs. (3), populations 1 and 2 disperse independently, whereas the densities $f_1$ and $f_2$ are coupled only through the reaction term $R$ and through the logistic terms, which introduce a negative competitive coupling proportional to $-f_1 f_2$. Also, it should be noticed that Eqs. (3) describe two populations with identical dispersal and growth properties, differing only in the way the respective language interact, according to the term $R$ given by Eq. (4). As a consequence, the total population density $f = f_1 + f_2$ follows a diffusion-advection-growth process without culture transmission, obtained by summing equations (3),

$$\frac{\partial f}{\partial t} = -\nabla \cdot [\mathbf{F} f] + \nabla \cdot (D \nabla f) + \alpha f \left(1 - \frac{f}{K}\right). \tag{5}$$

In order to solve Eqs. (3) numerically, one can approximate the derivatives through the corresponding finite differences, replacing the problem in the continuous time and space variables $(t, x, y)$ with the one on a lattice $(k, m, n)$ defined by the discrete variables $t_k = k\delta t$, $x_m = m\delta x$, and $y_n = n\delta y$, respectively, where $k, m, n$ are integers, while $\delta t$, $\delta x$, $\delta y$ are the lattice steps. By using the explicit Euler integration scheme [28], for constant $D$ one obtains from (3) the finite-difference equations

$$\frac{\delta_k f_i}{\delta t} = \pm R(f_1, f_2) - \frac{\delta_m (F_x f_i)}{\delta x} - \frac{\delta_n (F_y f_i)}{\delta y} + D \left[ \frac{\delta_m^2 f_i}{\delta x^2} + \frac{\delta_n^2 f_i}{\delta y^2} \right] + \alpha f_i \left(1 - \frac{f}{K}\right), \quad i = 1, 2. \tag{6}$$

Here the term $+R$ corresponds to $i = 1$ and $-R$ to $i = 2$. The time-difference operator $\delta_k$ when applied to a generic function $\phi(k, m, n)$ gives $\delta_k \phi(k, m, n) = \phi(k + 1, m, n) - \phi(k, m, n)$. Concerning the $x$-difference operators, they are defined as $\delta_m \phi(k, m, n) = [\phi(k, m + 1, n) - \phi(k, m - 1, n)]/2$ and $\delta_m^2 \phi(k, m, n) = \phi(k, m + 1, n) + \phi(k, m - 1, n) - 2\phi(k, m, n)$; analogous definitions apply to the $y$-difference operators $\delta_n$ and $\delta_n^2$. If the finite-difference equation (6) are used to solve numerically the continuous problem, suitable constraints have to be imposed on the values of the steps $\delta t$, $\delta x$, $\delta y$ in order to limit the numerical integration error [28].

One can also consider Eqs. (6) as a general lattice model of dispersal, growth, and cultural interaction of two populations, with no reference to the continuous equations (3). Such a discrete model is employed in the examples presented in the forthcoming sections, even if for convenience we will refer also to the corresponding continuous limit given by Eqs. (3).

## 3. Influence of initial conditions

The minimal spatial version of the AS model is obtained from Eqs. (3) taking a constant diffusion coefficient $D$ and neglecting the growth and advection terms. Similarly to the original AS model, it admits two stable equilibrium solutions, corresponding to one of the two languages surviving and the other one disappearing, as can be shown by linear stability analysis. Depending on the initial conditions, there exist also unstable equilibrium solutions. In the presence of a local spatial noise the unstable solutions are washed out by the random fluctuations [12]. However, the minimal spatial model (with no noise) presents also some new effects with respect to the corresponding homogeneous version. In this section we discuss the influence of initial conditions.

### 3.1. Cultural interaction and dispersal without growth

Differently from a homogeneous model, such as the AS model, in the spatial model the evolution in time and space depends on the particular form of the initial population densities $f_i(\mathbf{r}, t_0)$ (c.f. Section 2.1). While this is a standard mathematical property, it has a relevant meaning in terms of geographical and historical conditions. We consider the discretized equations (6) on a square lattice. Reflecting boundary conditions are assumed, i.e., no speaker can exit (or enter) the simulation area, corresponding to a zero population current through the boundaries. Also, we set $\mathbf{F} = 0$ (no advection) and $\alpha = 0$ (no population growth). Here and in the other simulations, units are chosen to have $D = 1$. Such a model represents a simplified version of a region that is isolated and geographically homogeneous. We use a lattice of size $50 \times 50$ with $\Delta x = \Delta y = 1$; the time step is $\Delta t = 0.01$ and the reaction constant $k = 2000$.

In order to check the consistency of the minimal spatial model, we have verified the AS model predictions by choosing uniform initial conditions, $f_i(x, y, t_0) = \text{const}$. In this case the diffusion terms in Eqs. (3) or (6) are zero and the distributions remain uniform at any time $t$; this is the only case in which integrating over the space coordinates exactly gives back the AS model (1). Uniform initial conditions for $f_1$ and $f_2$ represent a historical moment when populations 1 and 2 were broadly distributed across the territory. Population 1 is observed to disappear whenever the initial density ratio $f_1/f_2$ is smaller than the critical value $N_1^*/N_2^*$ given by condition (2).

In the non-uniform case, we have found for various values of parameters that a broader initial distribution represents a disadvantage when population growth is negligible ($\alpha \approx 0$). We illustrate this effect for two languages with status $s_1 = 0.55$ and $s_2 = 1 - s_1 = 0.45$ ($s_1 > s_2$) and equal initial population sizes. Since $\alpha = 0$, the total number of speakers is conserved and for simplicity we normalize it to one, $N_1(t) + N_2(t) = 1$. Thus $N_i(t)$ represents here the fraction of speakers of language $i$ at time $t$, and our choice corresponds to $N_1(t_0) = N_2(t_0) = 1/2$. With such parameters and initial conditions, language 1 would be clearly favored in the uniform case (or in the AS model) while language 2 would disappear. However, this does not necessarily happen when space dimensions are taken into account. Let us investigate the situation when populations 1 and 2 are initially distributed according to Gaussian densities,

$$f_i(x, y, t_0) = \frac{N_i(t_0)}{2\pi\sigma_i^2} \exp\left[-\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma_i^2}\right], \quad i = 1, 2. \tag{7}$$

This form of density represents a population symmetrically distributed around the average position $(x_i, y_i)$ with a standard deviation $\sigma_i$. The average positions are located in the center of the simulation area, $x_i = y_i = 25$ (see Fig. 1 top). As a mathematical remark, we notice that, given the symmetry of the initial configuration, using reflecting or periodic boundary conditions leads to perfectly equivalent results. For population 1 we assume $\sigma_1 = 10$ in both examples A and B, whereas for population 2 we assign $\sigma_2 = 1.75$ in example A and $\sigma_2 = 3$ in example B. The particular initial configurations assumed can be interpreted from a historical point of view as the sudden appearance of a high population density of speakers 2 in the center (of mass) of population 1.[2] One is then interested in predicting the final configuration, i.e., which language will eventually prevail. Surprisingly, even when the two initial population sizes are equal, it is not the language with a higher status which necessarily survives. This situation is illustrated by example A in Fig. 1. For clarity, the two population densities 1 and 2 are depicted separately in the respective columns. The snapshots at $t = t_0 = 0$ represent the initial conditions; the color intensity is proportional to the population density $f_i(x, y, t)$. Initially population 2 appears to be concentrated in the middle of the area, corresponding to $\sigma_2 = 1.75$, while population 1 is spread over a wider area, due to the larger value of $\sigma_1 = 10$. Therefore, comparing the function $R(f_1, f_2)$ given by Eq. (4), the $1 \to 2$ language switching is favored especially in the middle of the area, where the population density $f_2$ is largest and the $1 \to 2$ switching probability highest. This leads at $t = 10$ to a situation in which population 1 is concentrated in a ring-shaped region. Eventually, language 1 disappears despite its higher status $s_1 > s_2$ and the same initial population $N_1(t_0) = N_2(t_0)$, as shown by the snapshots at $t = 1000$, which represent with good approximation the asymptotic state.

Starting with a situation in which population 2 is initially more spread ($\sigma_2 = 3$), while all other parameters maintain the same values as in example A, the opposite final configuration is recovered, i.e., it is population 2 which now disappears, as shown by example B in Fig. 1. In Fig. 2 we compare the population size $N_1(t) \equiv 1 - N_2(t)$ for the same examples A and B of Fig. 1. As one can notice, in example B population 1 eventually prevails, even if $N_1$ becomes smaller than $N_2$ immediately at $t > t_0 = 0$ and remains such until $t \approx 400$. This is possible since language 1 has a higher status, $s_1 > s_2$. To estimate the conditions for the survival of language 1, we notice in Fig. 1 that population densities have become almost uniform at time $t \approx 100$, so that one can use the AS model. From Eq. (2) one obtains $N_1^*/N_2^* \approx 0.5122$ for $s_1 = 1 - s_2 = 0.55$, corresponding to a critical fraction $N_1^* \approx 0.3387$. In example B of Fig. 2 the surviving population $N_1(t) > N_1^*$ at any time $t$, while in the example A $N_1(t)$ crosses the line $N_1 = N_1^*$, beginning its irreversible decrease.

### 3.2. Cultural interaction and dispersal with growth

In this subsection we show that taking into account population growth, for sufficiently large values of the Malthus rate $\alpha$, a larger initial spreading can lead to the survival of the language, on the contrary to the situation with no growth, considered

[2] A historical conquest scenario has been recently studied in the framework of the bit-string model of language evolution by Schulze and Stauffer in Ref. [29].

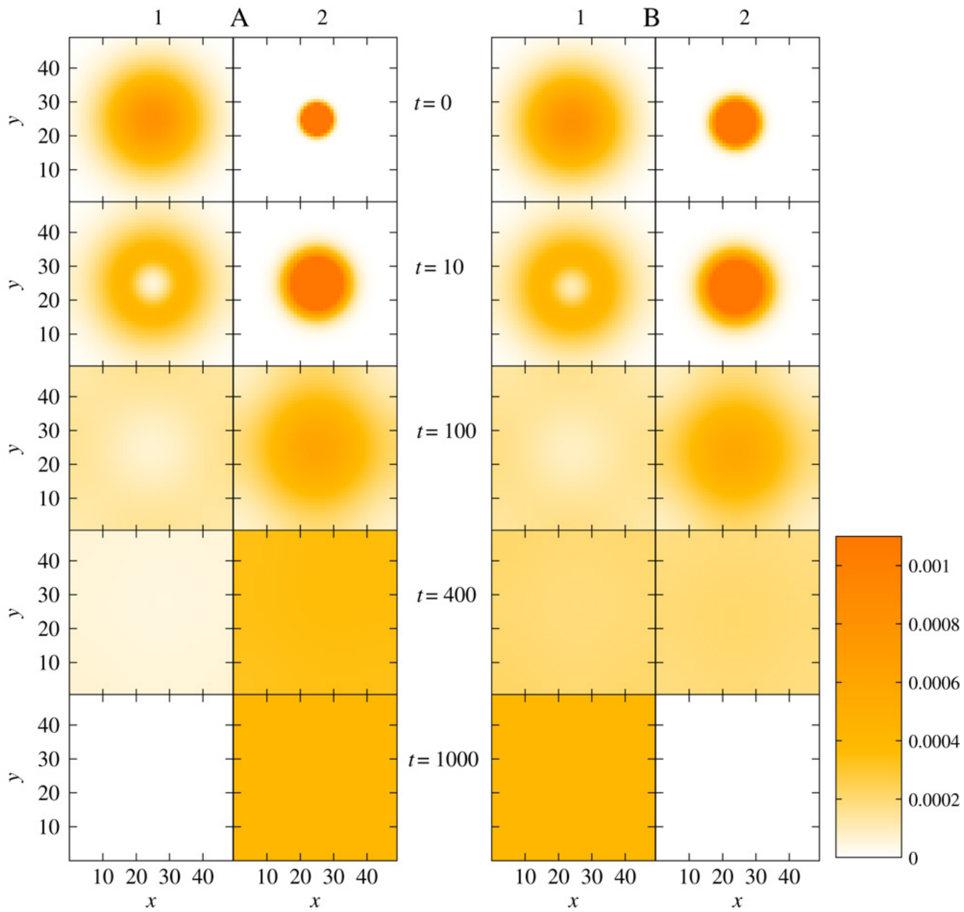


**Fig. 1.** Comparison of the evolution of population densities $f_1(x, y, t)$ and $f_2(x, y, t)$ (columns 1 and 2) for two languages with status $s_1 = 1 - s_2 = 0.55$ for different values of $\sigma_2$ (different widths) of the initial distribution $f_2(x, y, 0)$. The intensity of color is proportional to the population density. The initial population sizes are $N_1(0) = N_2(0) = 1/2$, population growth is neglected $[N_1(t) + N_2(t) = 1]$, and the initial distributions $f_i(x, y, 0)$ are given by Eqs. (7). In example A population 2 is initially more localized around the average position ($\sigma_2 = 1.75$) than that in example B ($\sigma_2 = 3$); in both examples A and B $\sigma_1 = 10$. See text for details.

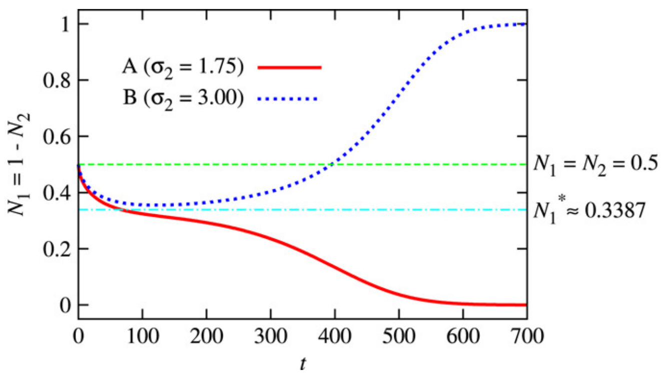


**Fig. 2.** Time evolution of the population size $N_1(t) = 1 - N_2(t)$ with a status $s_1 = 0.55 = 1 - s_2$ for the examples A (continuous line) and B (dotted line) of Fig. 1. The critical fraction $N_1^* = 0.3387$ given by the AS model for the survival of language 1 and the line corresponding to $N_1 = N_2$ are also drawn.

in Section 3.1. In Fig. 3 we present two illustrative examples, C and D, which differ from each other with respect to the initial spread of population 1.

Performing the numerical simulations, we have used a $100 \times 100$ simulation area with $\Delta x = \Delta y = 0.5$ and periodic boundary conditions, i.e., speakers exiting through a certain boundary reenter the area through the opposite boundary

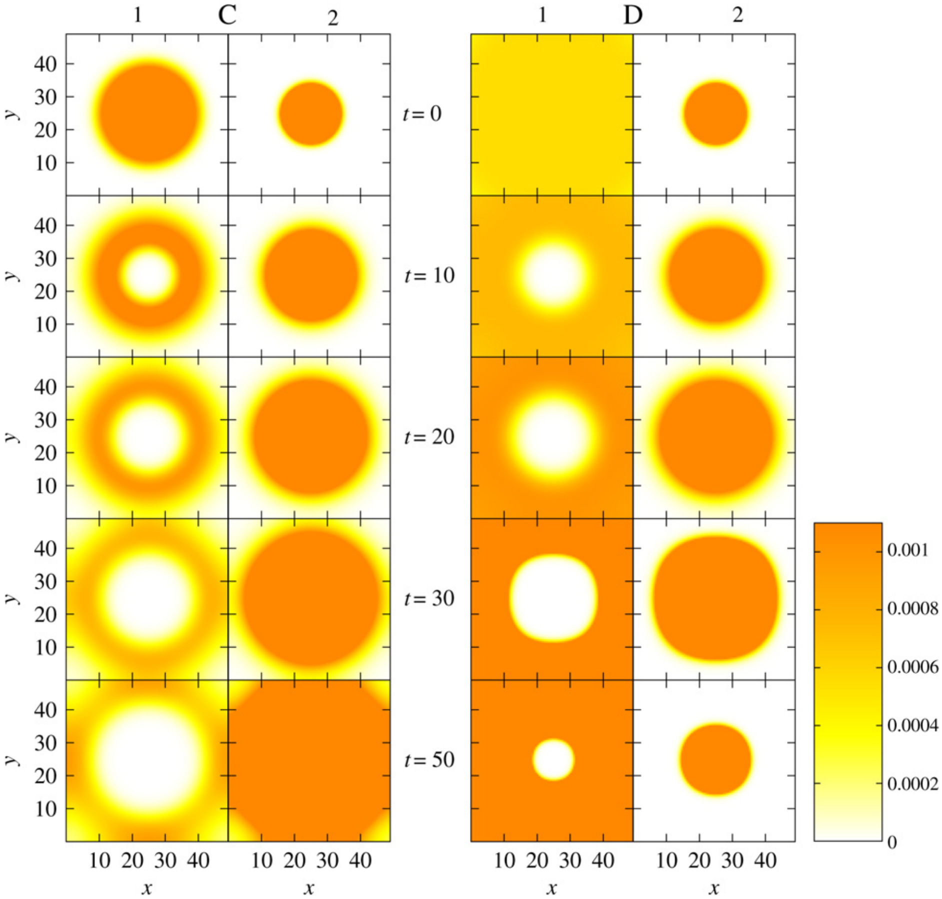


**Fig. 3.** Comparison of the evolution of population densities $f_1(x, y, t)$ and $f_2(x, y, t)$ (columns 1 and 2) for two languages with status $s_1 = 1 - s_2 = 0.55$ for different values of $R_1$ and $\sigma_1$ (different widths) of the initial distribution $f_1(x, y, t_0)$. In both examples C and D the initial population sizes are $N_1(0) = 1.36$ and $N_2(0) = 2.64$; populations grow with rate $\alpha = 0.03$ and carrying capacity $K = 0.1$. Example C: initial distribution $f_1(x, y, t_0)$ localized within a radius $R_1 = 15$ with $\sigma_1 = 3$. Example D: practically uniform initial distribution $f_1(x, y, t_0) \approx$ const ($R_1 = 25, \sigma_1 = 10$). See text for details.

(see also the discussion at the beginning of Section 4). The time step is $\Delta t = 0.001$ and the reaction constant $k = 1000$. Regarding the growth term, a rate $\alpha = 0.03$ and a carrying capacity $K = 0.1$ have been used.

The initial densities have been chosen of the form

$$f_i(x, y, t_0) = \begin{cases} \mathcal{M}_i, & r_i(x, y) < R_i, \\ \dfrac{2\mathcal{M}_i}{1 + \exp\left\{-[r_i(x, y) - R_i]^2/2\sigma_i^2\right\}}, & r_i(x, y) > R_i, \end{cases} \tag{8}$$

where $r_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ is the distance between position $(x, y)$ and the average position $(x_i, y_i)$ of the population density $f_i$. This function (8) defines a population initially localized mostly within a radius $R_i$ from the average position $(x_i, y_i)$. Differently from a Gaussian distribution, Eqs. (8) describe a population with an approximately constant density $f_i \approx \mathcal{M}_i$, decreasing to zero at large distances $r_i(x, y) > R_i$ on a scale $\sigma_i$. The average positions $(x_1, y_1)$ and $(x_2, y_2)$ of both populations have been chosen in the center of the simulation area, $x_i = y_i = 25$. In both examples C and D population 2 starts from the same initial density (8); the parameters are $\mathcal{M}_2 = 0.0163647$, $R_2 = 4$, and $\sigma_2 = 2$. Instead, population 1 starts from two different initial distributions: a more localized distribution in example C and a practically uniform distribution in example D. Such distributions have been obtained by choosing $\mathcal{M}_1 = 0.0020048$, $R_1 = 10$, and $\sigma_1 = 3$ in example C, while $\mathcal{M}_1 = 0.000548599$, $R_1 = 25$ and $\sigma_1 = 10$ have been used in example D (notice the large values of $R_1$ and $\sigma_1$ which make the initial density $f_1$ practically uniform). The parameters $\mathcal{M}_1$ and $\mathcal{M}_2$ have been determined in order to ensure that the density (8) is nowhere larger than the carrying capacity, $f_i(x, y, t_0) \leq K$; in both examples C and D, $N_1(t_0) = 1.36$ and $N_2(t_0) = 2.64$. The initial populations are shown in the snapshot at $t = t_0 = 0$ in Fig. 3.

As time goes by, from Fig. 3 one observes that, differently from what happens in examples A and B (Fig. 1), a larger initial spreading of a population favors its survival: the initially more localized population 1 is observed to disappear in example C; while if population 1 is initially (almost) uniformly spread (example D), it eventually survives, and it is population 2 that

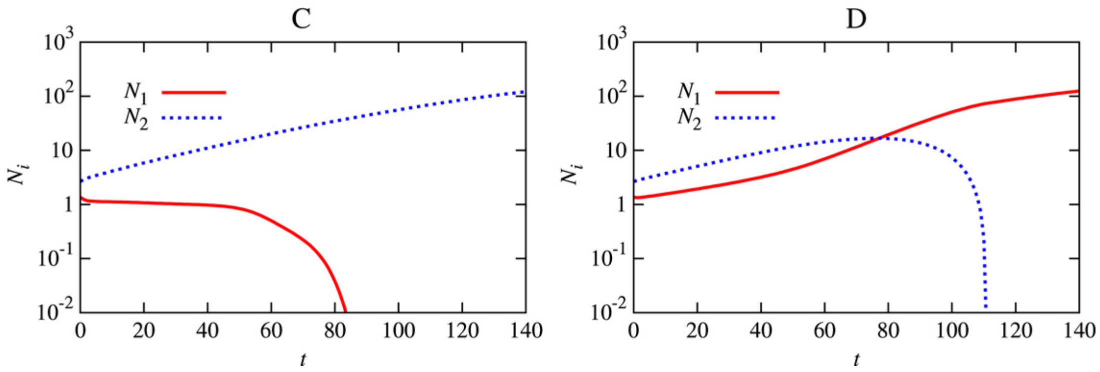


**Fig. 4.** Time evolution of the population sizes $N_1(t)$ and $N_2(t)$ corresponding to the examples C (left) and D (right) of Fig. 3.

disappears. This can be traced back to the growth of population 1 in the peripheral regions, where the population density $f_2$ is negligible. The time evolution of the population sizes $N_1(t)$ and $N_2(t)$ for examples C and D are depicted in Fig. 4.

## 4. Influence of boundary conditions

As shown for instance in Ref. [13] in the study of the three-state voter model, boundary conditions can have a crucial influence on the competition process between cultural traits. Here we investigate how dispersal properties of competing languages are affected by different boundary conditions. We compare reflecting versus periodic boundary conditions. Reflecting boundaries allow one to model a geographical area which is isolated, i.e., it is not possible for a speaker to enter or leave it. Periodic boundaries provide a numerically convenient way to simulate an open region within a finite simulation area. We found that if the growth rate is negligible ($\alpha \approx 0$) the vicinity of a reflecting boundary may favor the survival of a language. For high growth rates, the effect of different boundaries is less appreciable. Therefore, we discuss here the situation where population growth is neglected ($\alpha = 0$). In Fig. 5 we compare the time evolutions of population densities $f_i(x, y, t)$ for two languages in the presence of reflecting and periodic boundaries. For the language status, the values $s_1 = 1 - s_2 = 0.55$ have been assigned. In both examples the initial distributions $f_i(x, y, t_0)$ are assumed to have the Gaussian shape defined by (7), with $x_1 = 20$, $y_1 = 30$, $\sigma_1 = 3$, for population 1, and $x_2 = 45$, $y_2 = 5$, $\sigma_2 = 1$, for population 2; the normalized initial population sizes are $N_1(t_0) = 1 - N_2(t_0) = 0.37$. The size of the simulation area is $50 \times 50$ with $\Delta x = \Delta y = 1$, the time step is $\Delta t = 0.01$, and the reaction constant $k = 1000$. From Eq. (2) one obtains, for $s_1 = 1 - s_2 = 0.55$, that the critical fraction ensuring survival of language 1 is $N_1^* \approx 0.36$ ($N_2^* = 1 - N_1^* \approx 0.64$). Thus, the initial population fractions $N_1(t_0) = 0.37$ and $N_2(t_0) = 0.63$ used here would give a slight advantage to population 1 in the uniform case. Instead, as one can see from Fig. 5, language 1 disappears with reflecting boundary conditions (Fig. 5, left), while it prevails if periodic boundary conditions are used (Fig. 5, right). This effect is due to the fact that reflecting boundaries bounce back a part of population 2 located near the boundary increasing the corresponding density $f_2$. With open or periodic boundaries, population 2 would spread and its density $f_2$ decrease. This in turn would lower the term representing the $2 \to 1$ language switching rate in the function $R(f_1, f_2)$. Instead, a higher density $f_2$ favors the switching of speakers 1 to language 2. In Fig. 6 the population size $N_1(t) = 1 - N_2(t)$ for the two examples with reflecting and periodic boundary conditions are plotted.

## 5. Geographical barrier

In real situations the coexistence of more than one language in neighboring areas for long times is often observed [30]. Some models, such as the AS model, describe well situations in which after a relatively short time only one of the competing languages survives. In order to cover the situation of various languages existing together, different models are needed. For instance, more than one language can survive in the same area according to the model studied in Ref. [31], if population growth is taken into account. Another mechanism was suggested in Ref. [32], which includes bilinguals and assumes similar competing languages. In the framework of the bit-string language evolution model, the influence of a barrier on language diffusion and evolution was considered by Schulze and Stauffer, who showed that two different languages can develop and exist on the opposite sides of the barrier [11,5]. Another possibility is that the observed equilibrium between two languages is actually a nonequilibrium state with an underlying slow dynamics; numerical experiments of models, which can represent the interaction of two different language communities also taking bilinguals into account, show the existence of metastable states characterized by power law life time distributions [23].

In the present paper we concentrate on the role of purely geographical factors in the competition between languages. The scheme employed here differs from that used in Ref. [10], where the coexistence of two languages in neighboring regions was made possible by a barrier (a geographical boundary or a political border) affecting the form of the switching rate $R(f_1, f_2)$. Instead, the mechanism described in the examples presented in this and the following sections is based on the presence of

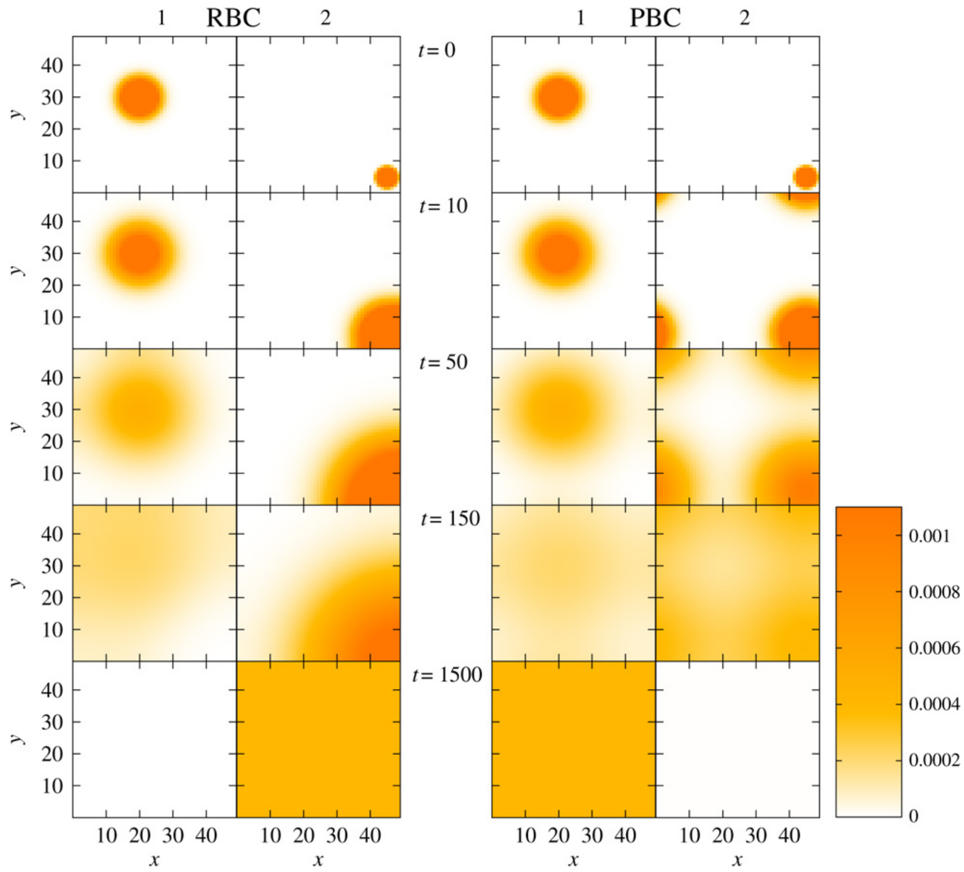


**Fig. 5.** Comparison of the evolution of population densities $f_1(x, y, t)$ and $f_2(x, y, t)$ (columns 1 and 2 respectively) with language status $s_1 = 0.45 = 1 - s_2$ for reflecting (RBC, left part) and periodic (PBC, right part) boundary conditions. The only difference between the two examples is in the boundary conditions. Population growth is neglected. See text for details.

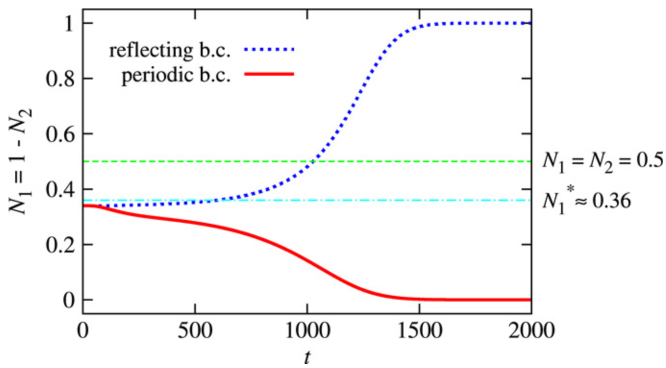


**Fig. 6.** Time evolution of the population size $N_1(t) = 1 - N_2(t)$ for the examples with reflecting and periodic boundary conditions of Fig. 5.

a geographical barrier which influences *solely* population dispersal, while the cultural interaction remains of the same form as in the AS model.

As a first example of geographical inhomogeneity, we consider a barrier, representing e.g. a mountain chain, which divides the accessible area into two regions. For the sake of simplicity we model the problem in one dimension and assume that the populations of speakers of language 1 and 2 are initially localized on the opposite sides of the barrier, as depicted in Fig. 7 top; we do not take into account population growth ($\alpha = 0$). We are interested in the influence of the barrier on the time evolution and the asymptotic state of the system. As discussed below, the barrier allows an equilibrium state (not possible in the homogeneous model) corresponding to both languages surviving on the opposite sides of the barrier.

As for the numerical simulation, population densities $f_i(x, t)$ are evolved according to Eqs. (3), where in one dimension $\nabla = \partial/\partial x$; numerical integration is performed through the Crank–Nicolson method [28], assuming reflecting boundary

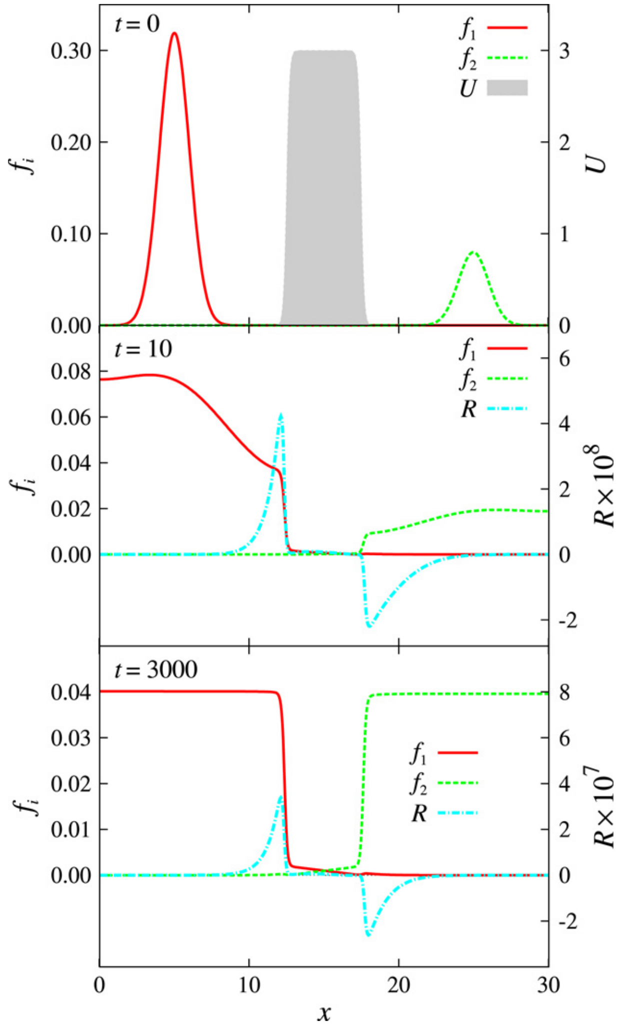


**Fig. 7.** Evolution of the speaker population density $f_1(x, t)$ (continuous line, left axis) and $f_2(x, t)$ (dashed line, left axis) in the presence of the barrier $U(x)$ drawn only at $t = 0$ (top, gray area, right axis); the initial population densities are given by Eqs. (10). Language 1 is favored both in status ($s_1 = 1-s_2 = 0.6$) and initial population [$N_1(t_0) = 1 - N_2(t_0) = 0.8$]. There is no population growth, $N_1(t) + N_2(t) = 1$. At times $t = 10$ and $t = 3000$ also the local reaction rate $R(f_1(x, t), f_2(x, t))$ is depicted (dashed-dotted line, right axis). In the asymptotic state ($t = 3000$) both languages survive, being localized on the opposite sides of the barrier. See text for further details.

conditions. A space step $\Delta x = 0.05$, a time step $\delta t = 0.001$, and a reaction constant $k = 200$, are used. The barrier is modeled through the force field $\mathbf{F} = F(x) = -\partial U(x)/\partial x$, where the potential $U(x)$, depicted in Fig. 7 top at $t = t_0 = 0$ (gray area), is the following:

$$U(x) = \frac{U_0}{1 + \exp[-(x - x_a)/\sigma_U] + \exp[(x - x_b)/\sigma_U]}. \tag{9}$$

This function represents a step of height $U_0$ located in the interval between $x_a$ and $x_b$ ($x_a < x_b$) and decreasing to zero outside of it on a scale length $\sigma_U$.

For the initial population densities $f_i(x, t_0)$ a Gaussian shape is assumed,

$$f_i(x, t_0) = \frac{N_i(t_0)}{\sqrt{2\pi}\,\sigma_i} \exp[-(x - x_i)^2/2\sigma_i^2], \quad i = 1, 2. \tag{10}$$

We choose $\sigma_1 = \sigma_2 = 1$ and for the average coordinates $x_1 = 5$ and $x_2 = 25$, i.e., $x_1$ and $x_2$ are located symmetrically with respect to the barrier and the reflecting boundaries at $x = 0$ and $x = 30$. The status of language 1 is $s_1 = 1 - s_2 = 0.6$ and the initial population fraction $N_1(t_0) = 1-N_2(t_0) = 0.8$. Thus, language 1 is highly favored, regarding both status and initial population size. In fact, in the absence of the potential barrier, language 2 is observed to disappear, as it is easy to guess. This also happens when the barrier is low or thin enough, corresponding to a high probability for speakers of language 1 to overcome the barrier and reach the other side. The problem considered is analogous to that of the escape problem of a

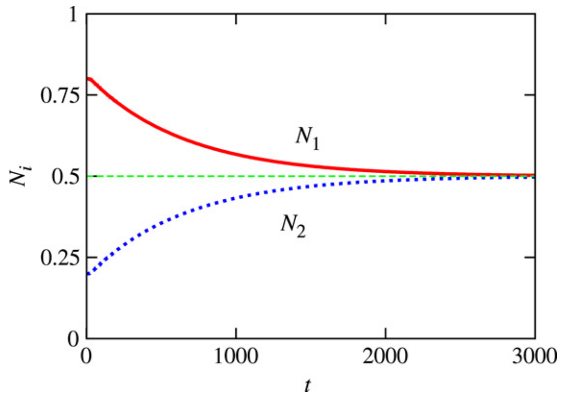


**Fig. 8.** Time evolution of population sizes $N_1(t)$ and $N_2(t)$ in the presence of a barrier for the example of Fig. 7.

Brownian particle overcoming a potential barrier. When the barrier parameters are such that the escape rate for speakers 1 to cross the barrier and reach the region on the right is high enough, language 1 eventually prevails both on the left and the right side. However, when the escape rate becomes small enough (e.g. the barrier is sufficiently high and/or wide) both languages can survive on the opposite sides of the barrier. Asymptotically, there are always only two possibilities: one language prevails on both sides of the barrier or the two languages survive localized on the opposite sides. An example of evolution with the second final scenario is illustrated by Fig. 7, in the presence of the potential barrier (9). The following parameter values have been used: $U_0 = 3$, $x_a = 12.5$, $x_b = 17.5$, and $\sigma_U = 0.1$. The time evolution of the population sizes $N_i(t)$ is depicted in Fig. 8.

The potential barrier modulates the diffusion toward both directions and in particular it decreases the flux of population 1 from the left toward the right region. This in turn causes the term $ks_1 f_1^a(x,t) f_2(x,t)$ in the reaction rate (4) to remain very small in the right region, so that the local $2 \to 1$ language switching rate is negligible respect to that of the complementary process $1 \to 2$. In Fig. 7 center ($t = 10$) and bottom ($t = 3000$) also the switching rate $R(f_1, f_2)$ is depicted with a dash-dotted line. The largest values of $|R|$ are located close to the barrier borders, where speakers coming from the other side meet the local speakers and switch to the local language. In Fig. 8 one can also notice that the asymptotic population sizes are equal $N_1(t \to \infty) = N_2(t \to \infty) = 1/2$, as a consequence of the identical dispersal properties assumed for the two populations and the symmetrical geometry of the system.

As mentioned already and as it is well known, the escape rate depends strongly on the barrier height $U_0$. Performing the simulations for different values of $U_0$ (keeping the other parameters fixed), we found that the critical value of $U_0$ for both languages to survive is $U_0^* \approx 2.1326$ (numerical uncertainty on the last digit). If $U_0 < U_0^*$ it is always language 1 which survives on both sides of the barrier.

The survival of the two languages, observed in the example discussed above, can be ascribed to the coincidence of suitable historical conditions, the initial localization of the two speakers communities on the opposite sides of the barrier, and geographical inhomogeneities, represented by the barrier.

## 6. Immigration to an island

Let us now discuss, what happens if we consider the spreading of two languages toward the same initially empty region. To make an example, we study two islands A and C, initially colonized by populations speaking language 1 and 2, respectively. Language 1 has a higher status $s_1 = 1 - s_2 = 0.6$. Between the two islands A and C is located a third island B, which is empty. For simplicity, we choose for the islands a circular shape with identical radius $R$; the centers of the islands are located on a line. The situation is illustrated by Fig. 9 top.

The system has been studied on a rectangular simulation area of sides $L_x = 35$ and $L_y = 11$. The population densities were evolved through the explicit Euler algorithm (6) on a $350 \times 110$ lattice with steps $\delta x = \delta y = 0.1$, using a time step $\delta t = 0.001$ and a rate constant $k = 2000$.

Speakers can freely diffuse inside the islands, but have to overcome a potential barrier in order to cross the sea and reach other islands. The effective potential $U(x, y)$ modeling the barrier due to the sea also defines the island shapes; it is depicted in the lower part of Fig. 9 and is given by

$$U(x, y) = \begin{cases} U_0 \exp\left\{-[r_j(x, y) - R]^2/2\sigma_U^2\right\}, & r_j(x, y) < R, \\ U_0, & \text{otherwise}. \end{cases} \qquad (11)$$

Here $j = \text{A, B, C}$ labels the islands and $r_j(x, y) = \sqrt{(x - x_j)^2 + (y - y_j)^2}$ is the distance between the generic position $(x, y)$ and the center $(x_j, y_j)$ of island $j$; i.e., a value $r_j(x, y) < R$ for a given $j$ corresponds to a point $(x, y)$ inside island $j$, while if $r_j(x, y) > R$ for all $j = \text{A, B, C}$ the point $(x, y)$ is on the sea. The parameter values used are: $\sigma_U = 2$, defining the smoothness

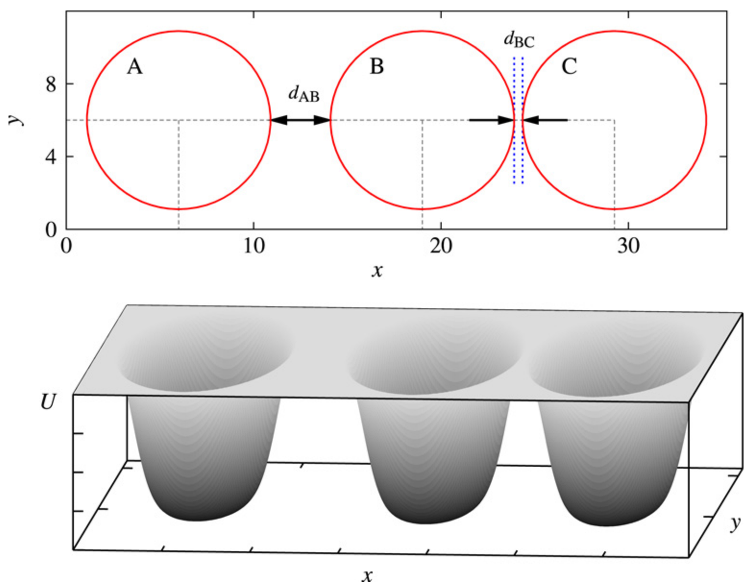


**Fig. 9.** Schematic map (top) and potential energy landscape representing the sea barrier (bottom) for the three-island configuration. Islands A, B, and C have a circular shape with the same radius, defined by the effective potential (11). Islands B and C are closer to each other than islands A and B. See text for further details.

of the potential step, $U_0 = 4$, for the barrier height, and $R = 5$, for the island radius; the coordinates of the centers of the island are $x_A = 6$, $x_B = 19$, $x_C = 29.25$, and $y_A = y_B = y_C = 6$. The location of the three islands can be recognized in Fig. 9 bottom as the three potential wells, while the sea is represented by the plateau $U(x, y) = U_0$. The initial population sizes on islands A and C have been assigned the same value $N_1(t_0) = N_2(t_0) = 1/2$. The population densities $f_i(x, y, t_0)$ have the same Gaussian shape (7), with $\sigma_1 = \sigma_2 = 2$ and average coordinates $x_i$ and $y_i$ coinciding with the center coordinates of the corresponding islands.

In order to find out if it is language 1 or 2 which will eventually be spoken on island B, we have studied the problem for various values of the distance $d_{BC}$ between island B and C, while keeping constant the other distance $d_{AB}$. It turns out that if the central island B is located symmetrically between islands A and C then it is language 1 that in the end will be used on island B due to its higher status; this also happens if $d_{AB}$ is larger (but not too much larger) than $d_{BC}$. Whether population 1 colonizes also island C depends on the barrier between B and C. If it is large enough, population 2 may still survive on island C, where it was initially, due to the effect described in the previous section which would transform island C into a refugium. If not, language 1 may prevail finally also on island C.

On the other hand, if it is much easier to cross the B–C rather than the A–B channel, e.g. if island B is much closer to island C than to A, then language 2 may spread on island B before language 1 and then maintain its superiority due to the barrier between A and B. This is the situation represented by the example in Figs. 10 and 11, corresponding to a distance $d_{AB} = 3$ between islands A and B much larger that the distance $d_{BC} = 0.25$ between islands B and C. Fig. 10 shows how populations 1 and 2 disperse over the neighboring island starting from their initial locations. Due to the geographical asymmetry favoring dispersal of population 2, there is a much larger flow of population 2 from island B to C than population 1 from island A to B. This in turn leads to a rapid spreading of language 2 on island B. Once language 2, with a lower status, dominates on island B, the wide sea barrier between island A and B will maintain the advantage gained by language 2 (see Section 5), through the same mechanism illustrated in the example in the previous section. Figs. 10 and 11 also show a small temporary presence of language 1 on the central island B.

## 7. Conclusion

Dispersal in space and time of two languages or cultural traits competing in the same geographical area has been studied through an extended language competition model based on the one proposed by Abrams and Strogatz. We have discussed how initial and boundary conditions, as well as geographical inhomogeneities, have a relevant (even drastic) meaning for language spreading and competition.

We have observed various examples where a language, which in the corresponding homogeneous model would disappear, actually survives. Namely, in a homogeneous model (without space dimensions), for given values of the parameters, the evolution of a population is determined by the initial population size. In the diffusion model studied here, a population of speakers, initially distributed in space in different ways, may evolve toward opposite asymptotic scenarios. Another result of our investigation shows that, when growth is negligible, a language whose dispersal is more affected by the restricting boundaries, is favored respect to the case where no boundaries are present. In the presence of geographical

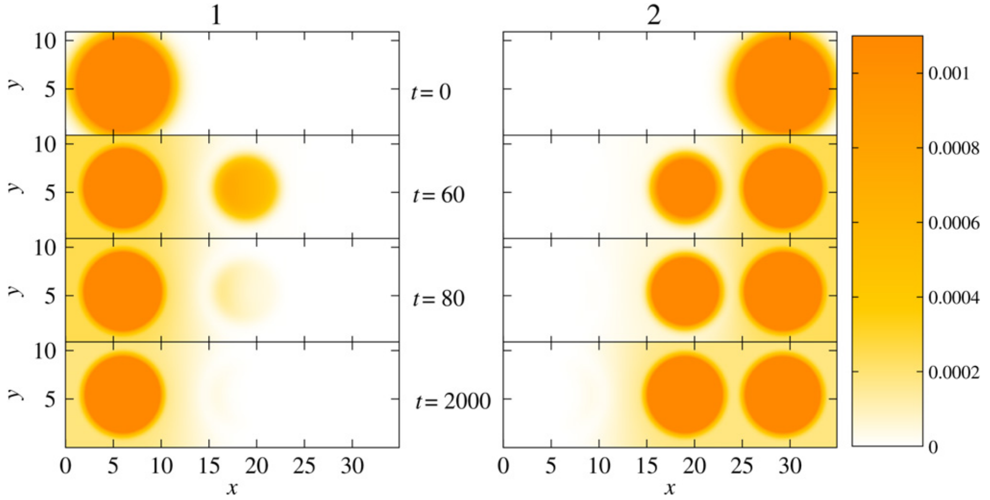


**Fig. 10.** Evolution of population density $f_1(x, y, t)$ (left column) and $f_2(x, y, t)$ (right column) in the geometry depicted in Fig. 9 with distances $d_{AB} = 3$ and $d_{BC} = 0.25$. The language status is $s_1 = 1 - s_2 = 0.6$ and the initial population sizes are equal, $N_1(t_0) = N_2(t_0) = 1/2$ [population growth is neglected, $N_1(t) + N_2(t) = 1$]. Notice the temporary presence of population 1 on (the central) island B at $t \approx 60$. Despite the lower status of language 2, geographical inhomogeneities favor its immigration to the central island B. See text for the details.

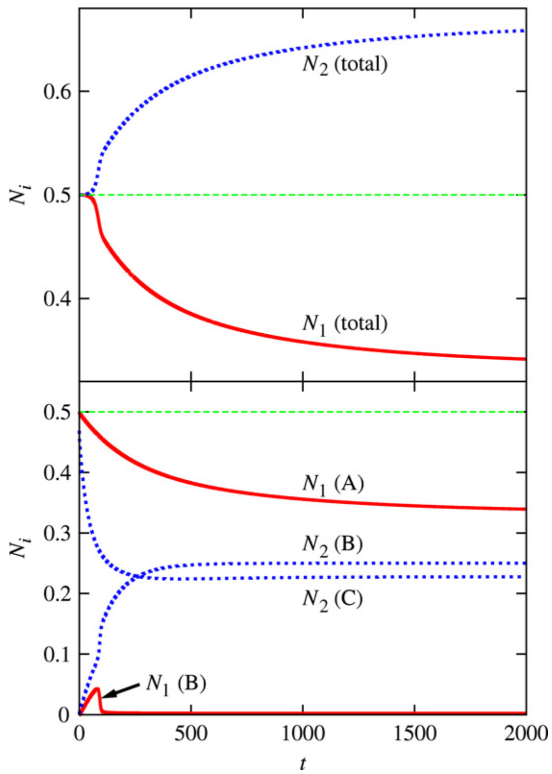


**Fig. 11.** Time evolution of the speaker community 1 (continuous line) and 2 (dotted line) for the three island example. Top: population sizes $N_1$ and $N_2$. Bottom: population 1 on island A and B [$N_1(A)$ and $N_1(B)$] and population 2 on island B and C [$N_2(B)$ and $N_2(C)$]; population 1 on island C as well as population 2 on island A are negligible and are not shown. Notice that until $t \approx 100$ both populations 1 and 2 are present on the central island B.

inhomogeneities, modeled as potential energy barriers, languages can survive in different regions, despite the possibly lower status and smaller initial population size. The effects discussed in the present paper are purely geographical, in the sense that they are related to the diffusion processes and to the modulation of diffusion due to inhomogeneities of the background. They influence culture spreading only indirectly and are not due to a change in the cultural interaction law, differently from the model introduced in Ref. [10].

The diffusion model and the highly idealized examples presented in this paper are intended as a first step toward a quantitative description of the space–time diffusion of language and cultural traits. Our study will hopefully be useful in solving some of the many challenging problems regarding language diversity, see e.g. Ref. [30]. At the same time, a more detailed understanding of the mechanisms underlying culture transmission could be valuable concerning the alarming rate of disappearance of cultural and linguistic diversity.

## Acknowledgements

This work was supported by the Estonian Science Foundation Grant No. 7466. M.P. gratefully acknowledges also financial support from the GIACS Workshop on *Language Simulations* (2006 September 11-14, Warsaw, Poland) and the participants of the workshop for useful discussions and constructive remarks.

## References

[1] J.-P. Bouchaud, The subtle nature of financial random walks, CHAOS 15 (2005) 026104.
[2] J.D. Murray, Mathematical biology I, in: An Introduction, Springer, New York, 2002.
[3] A.J. Ammerman, L.L. Cavalli-Sforza, The Neolithic Transition and the Genetics of Populations in Europe, Princeton University Press, Princeton, 1984.
[4] D. Stauffer, C. Schulze, Microscopic and macroscopic simulation of competition between languages, Phys. Life Rev. 2 (2005) 89.
[5] C. Schulze, D. Stauffer, Recent developments in computer simulations of language competition, Comp. Sci. Eng. 8 (2006) 60.
[6] V. Loreto, L. Steels, Emergence of language, Nature Phys. 3 (2007) 758.
[7] S. Wichmann, The emerging field of language dynamics, Language and Linguistics Compass 2–3 (2008) 442–455.
[8] C. Schulze, D. Stauffer, S. Wichmann, Birth, survival and death of languages by Monte Carlo simulation, Comm. Comp. Phys. 3 (2008) 271.
[9] M.V. Lomolino, B.R. Riddle, J.H. Brown, Biogeography, Sinauer Associates, Sunderland, Massachusetts, 2006.
[10] M. Patriarca, T. Leppänen, Modeling language competition, Physica A 338 (2004) 296.
[11] C. Schulze, D. Stauffer, Competition of languages in the presence of a barrier, Physica A 379 (2007) 661.
[12] D. Stauffer, X. Castelló, V.M. Eguíluz, M.S. Miguel, Microscopic Abrams–Strogatz model of language competition, Physica A 374 (2007) 835.
[13] T. Hadzibeganovic, D. Stauffer, C. Schulze, Boundary effects in a three-state modified voter model for languages, Physica A 387 (2008) 3242.
[14] D.M. Abrams, S.H. Strogatz, Modelling the dynamics of language death, Nature 424 (2003) 900.
[15] M.A. Nowak, N.L. Komarova, P. Niyogi, Computational and evolutionary aspects of language, Nature 417 (2002) 611.
[16] L. Steels, A self-organizing spatial vocabulary, Artif. Life 2 (1995) 319.
[17] A. Baronchelli, L. Dall'Asta, A. Barrat, V. Loreto, The role of topology on the dynamics of the naming game, Eur. Phys. J. Special Topics 143 (2007) 233.
[18] A. Baronchelli, L. Dall'Asta, A. Barrat, V. Loreto, In-depth analysis of the naming game dynamics: The homogeneous mixing case, Internat. J. Modern Phys. C 19 (2008) 785.
[19] A. Lipowski, D. Lipowska, Bio-linguistic transition and Baldwin effect in an evolutionary naming-game model, Internat. J. Modern Phys. C 19 (2008) 399.
[20] A. Lipowski, D. Lipowska, Computational approach to the emergence and evolution of language — evolutionary naming game model. arXiv:0801.1658.
[21] W.S.-Y. Wang, J.W. Minett, The invasion of language: Emergence, change and death, TRENDS in Ecology and Evolution 20 (2005) 263.
[22] X. Castelló, V.M. Eguíluz, M.S. Miguel, Ordering dynamics with two non-excluding options: Bilingualism in language competition, New J. Phys. 8 (2006) 306.
[23] X. Castelló, R. Toivonen, V.M. Eguíluz, J. Saramäki, K. Kaski, M.S. Miguel, Anomalous lifetime distributions and topological traps in ordering dynamics, Europhys. Lett. 79 (2007) 66006.
[24] J.W. Minett, W.S.-Y. Wang, Modelling endangered languages: The effects of bilingualism and social structure, Lingua 118 (2007) 19.
[25] J. Ke, T. Gong, W.S.-Y. Wang, Language change and social networks, Commun. Comput. Phys. 3 (2008) 935.
[26] K. Davison, P. Dolukhanov, G.R. Sarson, A. Shukurov, The role of waterways in the spread of the Neolithic, J. Arch. Sci. 33 (2006) 641.
[27] L.A. Martino, A. Osella, C. Dorso, J.L. Lanata, Fisher equation for anisotropic diffusion: Simulating South American human dispersals, Phys. Rev. E 76 (2007) 031923.
[28] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C : The Art of Scientific Computing, 2nd Edition, Cambridge University Press, Cambridge, 1992.
[29] C. Schulze, D. Stauffer, Language simulation after a conquest. arXiv:0707.0072.
[30] J. Nichols, Modeling ancient population structures and movement in linguistics, Annu. Rev. Anthropol. 26 (1997) 359–384.
[31] J. Pinasco, L. Romanelli, Coexistence of languages is possible, Physica A 361 (2006) 355.
[32] J. Mira, A. Paredes, Interlinguistic similarity and language death dynamics, Europhys. Lett. 69 (2005) 1031.