Influence of geography on language competition

Marco Patriarca
National Institute of Chemical Physics and Biophysics, Tallinn

Marco Patriarca and Els Heinsalu, **Influence of geography on language competition**,
Physica A 388 (2009) 174
doi:10.1016/j.physa.2008.09.034
http://dx.doi.org/10.1016/j.physa.2008.09.034
arXiv.org:0807.3100

marco.patriarca [at] kbfi.ee

Summary

- Overview of models of language dynamics

- Constructing a **geographical model of culture transmission from**
1) a model of language competition + 2) a model of human dispersal

Statistical mechanics of language

Language as a complex system

- a single language can be considered as a complex phenomenon emerging from the interaction of many individuals
- a maximum number of languages estimated to be 10 000 – 15 000 different languages (~ 6000-7000 different languages documented)
- interdisciplinary field interacting with almost all other sciences

Social interest: disappearance of languages

- in the period 1490-1990 50% of the languages has disappeared
- it has been predicted that in the end of the 21st century about 90% of the languages spoken today will have disappeared

Interests

- adding quantitative tools/precision to historical and linguistic sciences
- automatic recognition and comparison of texts

Quantitative facts about languages

Languages as fixed species

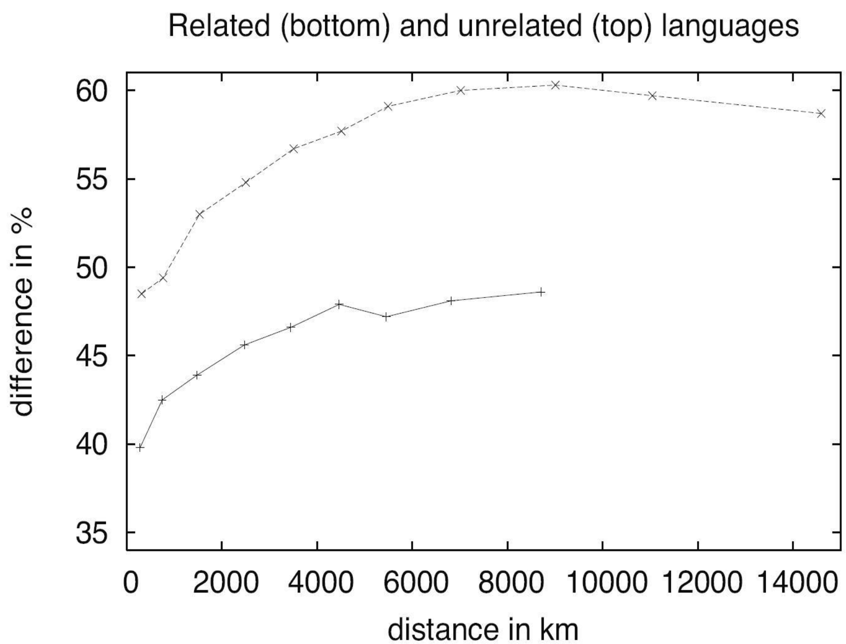
- D.M. Abrams, S.H. Strogatz, **Modelling the dynamics of language death**, Nature 424, 900 (2003)
- J. Pinasco, L. Romanelli, **Coexistence of languages is possible**, Physica A 361 (2006) 355

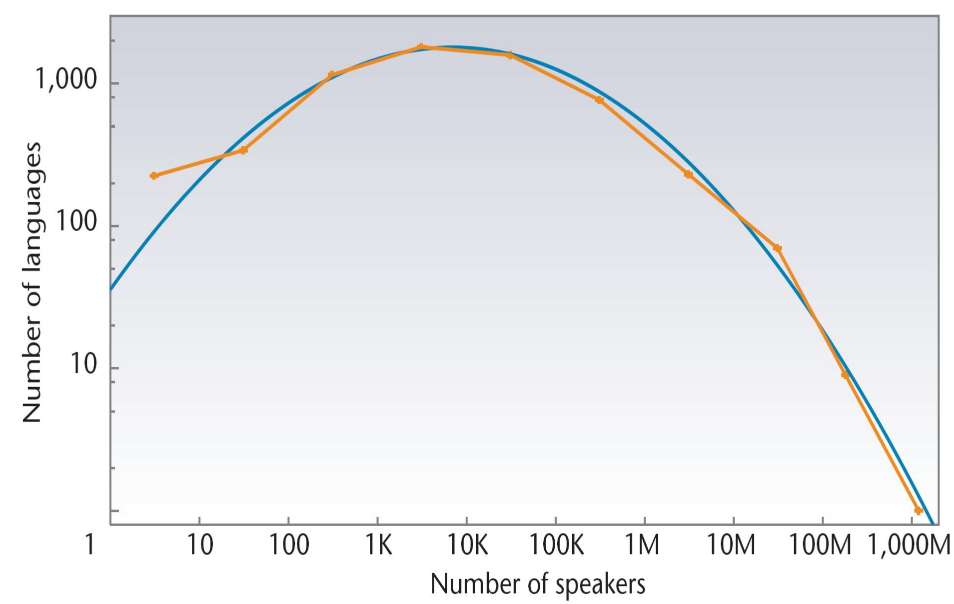
Size distribution of languages

- D. Stauffer, C. Schulze, **Microscopic and macroscopic simulation of competition between languages**, Phys. Life Rev. 2 (2005) 89.
- M.A. Nowak and D.C. Krakauer, **The evolution of language**, Proc. Natl. Acad. Sci. 96, 8028 (1999)
- Viviane M. de Oliveira, M.A.F. Gomes, I.R. Tsang, **Theoretical model for the evolution of the linguistic diversity**, Physica A 361 (2006) 361–370

Relation between diversity and geographical distance

- E.W. Holman, C. Schulze, D. Stauffer, S. Wichman, **On the relation between structural diversity and geographical distance among languages: Observations and computer simulations**, Linguistic Typology 11 (2007), 393–421





Bilingual communities

- J. Mira, A. Paredes, **Interlinguistic similarity and language death dynamics**, Europhys. Lett. 69 (2005) 1031.
- W. S.-Y. Wang, J. W. Minett, **The invasion of language: emergence, change and death,** Trends in Ecology and Evolution 20 (2005) 263.
- J. W. Minett, W. S.-Y. Wang, **Modelling endangered languages: The effects of bilingualism and social structure**, Lingua (2007)
- X. Castello, V. M. Eguiluz, M. S. Miguel, **Ordering dynamics with two non-excluding options: bilingualism in language competition**, New J. Phys. 8 (2006) 306.
- X. Castello, R. Toivonen, V. M. Eguiluz, J. Saramaki, K. Kaski, M. S. Miguel, **Anomalous lifetime distributions and topological traps in ordering dynamics**, EPL, 79 (2007) 66006

Geographical background

- C. Schulze, D. Stauffer, **Competition of languages in the presence of a barrier**, Physica A 379 (2007) 661. URL arXiv:physics/0702031
- D. Stauffer, X. Castello, V. M. Eguiluz, M. S. Miguel, **Microscopic Abrams-Strogatz model of language competition**, Physica A 374 (2007) 835.
- M. Patriarca, T. Leppänen, **Modeling language competition**, Physica A 338 (2004) 296
- C. Di Chio and P. Di Chio, **Simulation model for the evolution of language with spatial topology**
- Marco Patriarca and Els Heinsalu, **Influence of geography on language competition,** Physica A 388 (2009) 174

Historical conditions

- C. Schulze, D. Stauffer, **Language simulation after a conquest**, arXiv:0707.0072

Game-theoretical approaches, the Naming Game, "semiotic dynamics"

- K. Kosmidis, A. Kalampokis, P. Argyrakis, **Language evolution and population dynamics in a system of two interacting species**, Physica A 353 (2005) 595.
- A. Baronchelli, L. Dall'Asta, A. Barrat, V. Loreto, **Topology induced coarsening in language games**, arXiv:physics/0512045.
- A. Lipowski, D. Lipowska, **Computational approach to the emergence and evolution of language evolutionary naming game model**, arXiv:0801.1658.

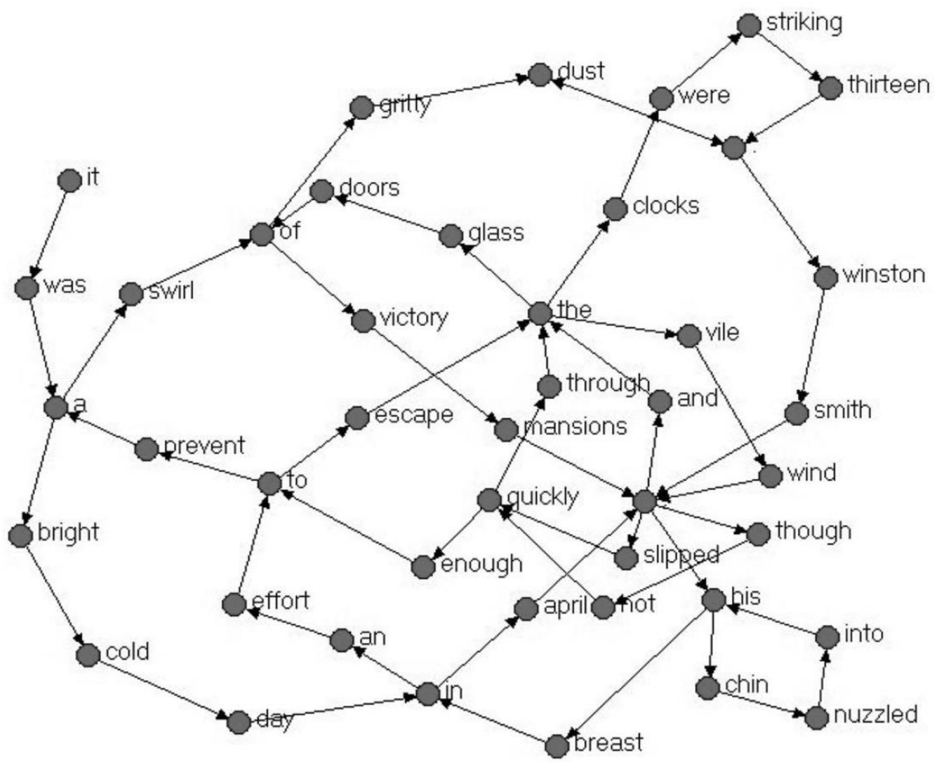


FIG. 1. Illustration of the language network for the first 60 words of Orwell's *1984*.

Language as a web

- A. P. Masucci and G. J. Rodgers, **Network properties of written human language**, Phys. Rev. E 74, 026102 (2006)
- Maria Markosova, **Network model of human language**, Physica A 387 (2008) 661–666

Statistical Properties of written text

- Mantegna, R. N. and Buldyrev, S. V. and Goldberg, A. L. and Havlin, S. and Peng, C.K. and Simons, M. and Stanley, **Linguistic features of non-coding DNA sequences**, H. E., Phys. Rev. Lett. 73 (1994) 3169

Possible topics for linguistic geography/historical linguistics

There are many features of language which are transmitted and evolve, which can be studied with statistical mechanical models from the geographical or historical point of view

- *phonetics* – sounds of human language
- *phonology* – patterns of a language's basic sounds
- *morphology* – internal structure of words
- *syntax* – word combination into grammatical sentences
- *semantics* – meaning of words (lexical semantics), fixed word combinations (phraseology), combination into sentences
- *pragmatics* – how utterances are used in communicative acts
- *discourse analysis* – how sentences are organised into texts
- *alphabets*

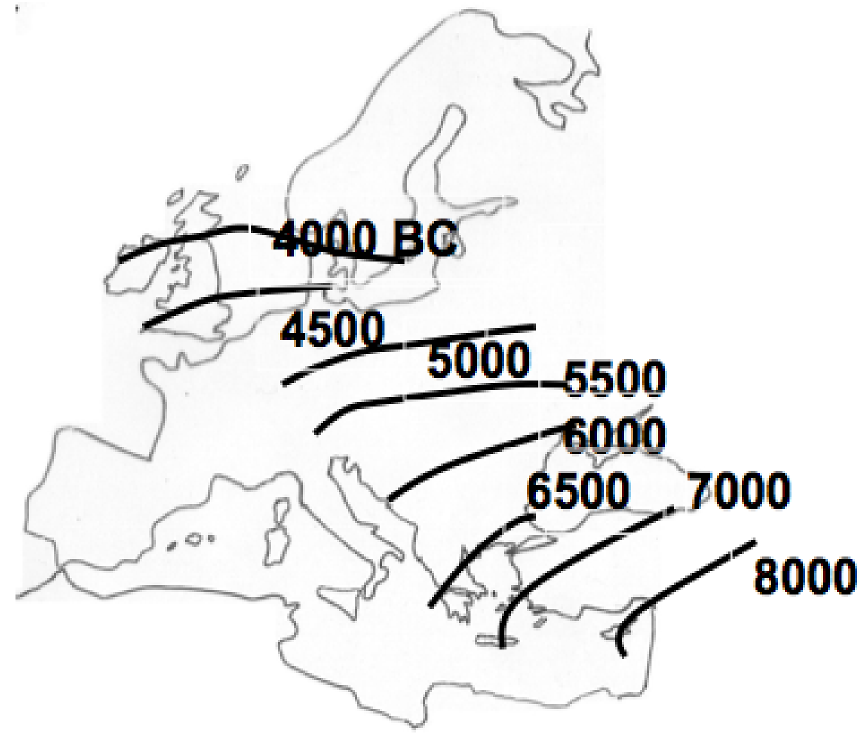
Elements of a geographical model of culture transmission

1) Population dispersal (diffusion/advection)

2) Population growth

3) Cultural transmission

Human population dynamics



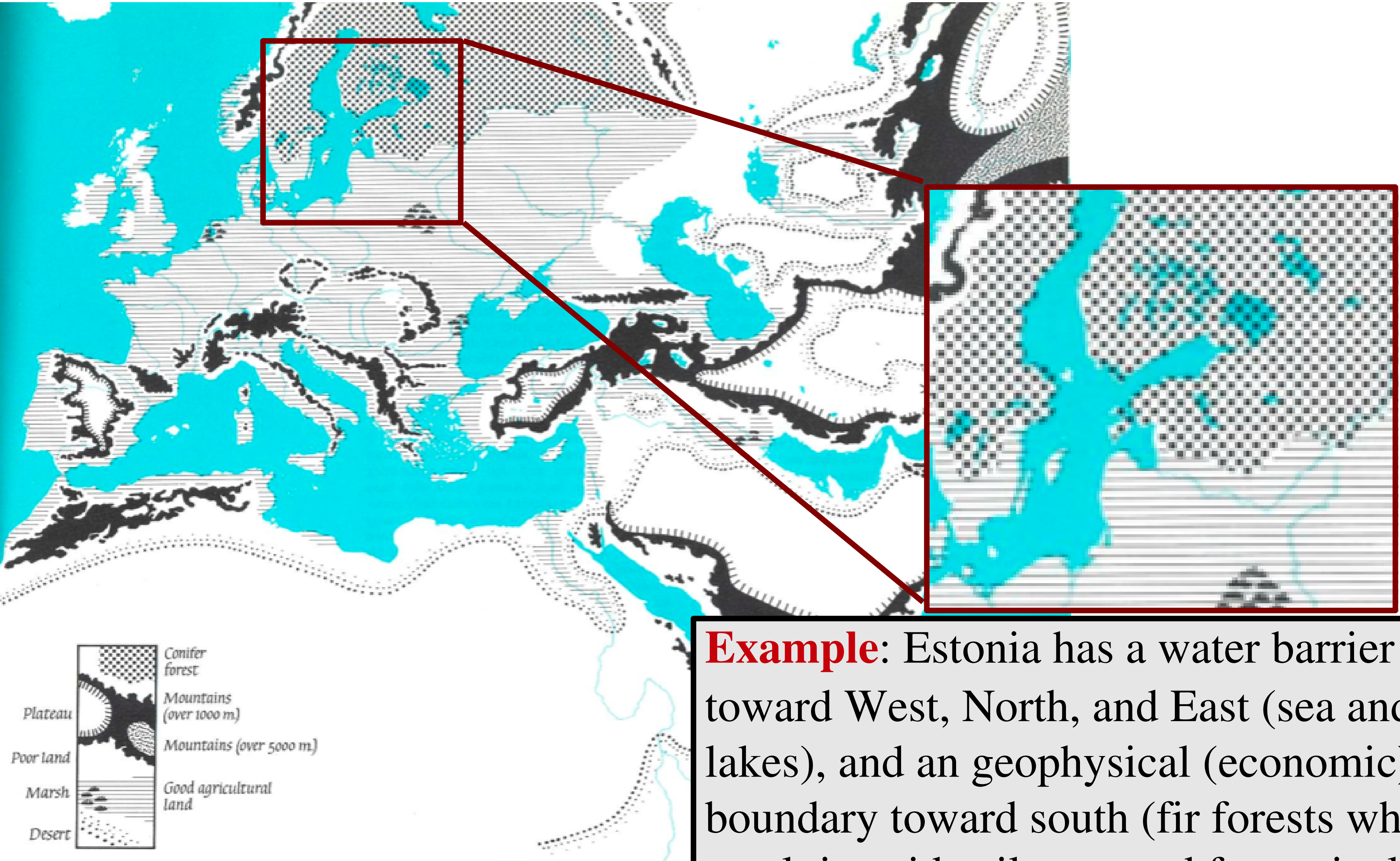
Diffusion equation + Verhulst growth term

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + \alpha n\left(1 - \frac{n}{k}\right)$$

The Fisher equation has been applied to human neolithic expansion by L.L. Cavalli-Sforza and A.J. Ammerman to describe the first use of agriculture instead of gathering and hunting. Archeology shows this expansion lasted 4000 years starting from the near east about 8.000 BC with propagation rate ~ 1 km/year and ending in north-west Europe.

A.J. Ammerman, L.L Cavalli-Sforza, **The Neolithic Transition and the Genetics of Populations in Europe**, Princeton University Press, Princeton, 1984.

Example of geographical inhomogeneities: type of ground/flora

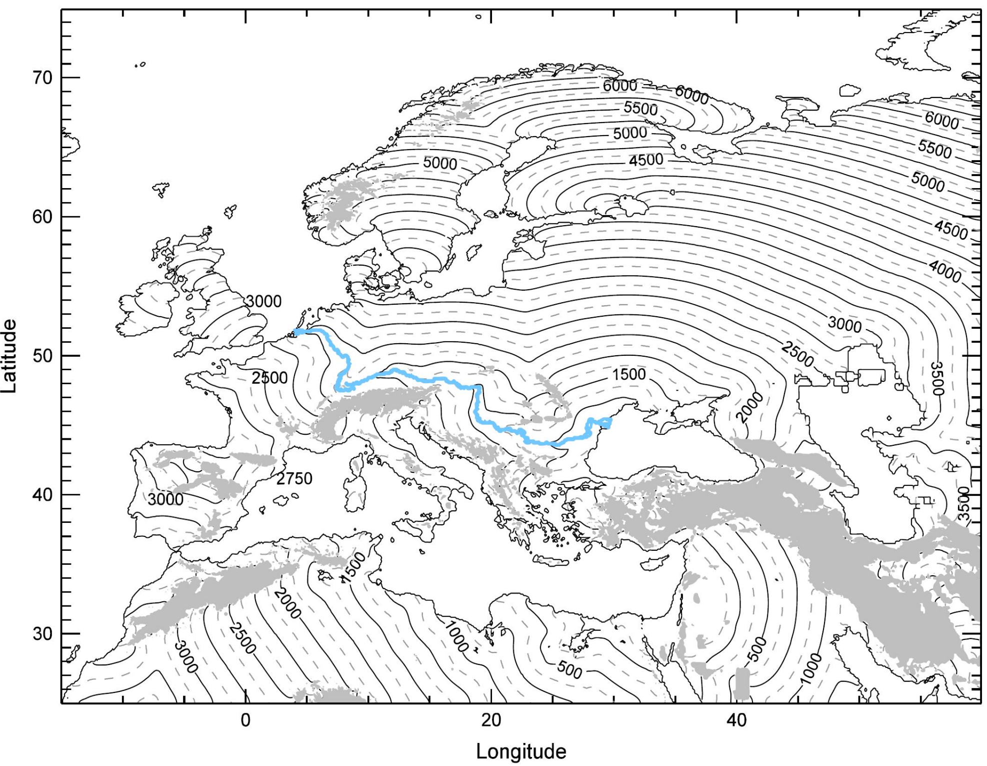


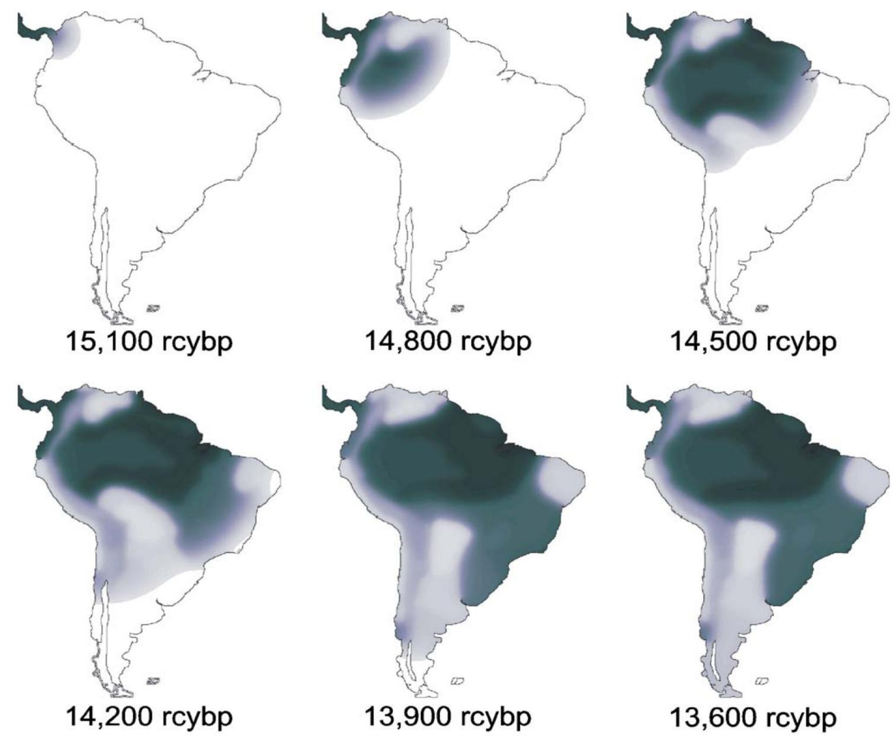
Example: Estonia has a water barrier toward West, North, and East (sea and lakes), and an geophysical (economic) boundary toward south (fir forests which result in acid soil not good for agriculture).

Importance of geographical inhomogeneities: rivers and mountains

K. Davison, P. Dolukhanov, G.R. Sarson, A. Shukurov, **The role of waterways in the spread of the Neolithic**, Journal of Archaeological Science 33 (2006) 641-652

L. A. Martino, A. Osella, C. Dorso, J.L. Lanata, **Fisher equation for anisotropic diffusion: Simulating South American human dispersals,** Phys. Rev. E 76, 031923 2007





$$\frac{\partial n}{\partial t} = -\frac{\partial}{\partial x}[F(x)n] + \frac{\partial}{\partial x}[D(x)\frac{\partial n}{\partial x}] - \alpha n(1-\frac{n}{K})$$

Diffusion + geographical inhomogeneities

Logistic term

Generalized diffusion of culture

Archeology and various authors remark that the population dynamics alone is too slow a process in order to account for the observed archaeological data.
This suggests, besides a demic diffusion, also a **transfer of ideas**, i.e., diffusion of ideas.

G. Vogl, *Diffusion and Brownian motion analogies in the migration of atoms, animals, men and ideas,* Diff. Fund., 2:2.1, 2005.

$\alpha \sim 0.03$ years^{-1}
$D \sim \lambda^2 / \tau$
$\lambda \sim$ distance between birth-place and marriage-place
$\tau \sim 1$ generation

$$v \sim 2\sqrt{D\alpha}$$

Hunter-gatherer: **30-50 Km**
Farmers (low population density): **10-20 Km**
Farmers (high population density): **5-10 Km**

Competition Models: Model of Abrams and Strogatz

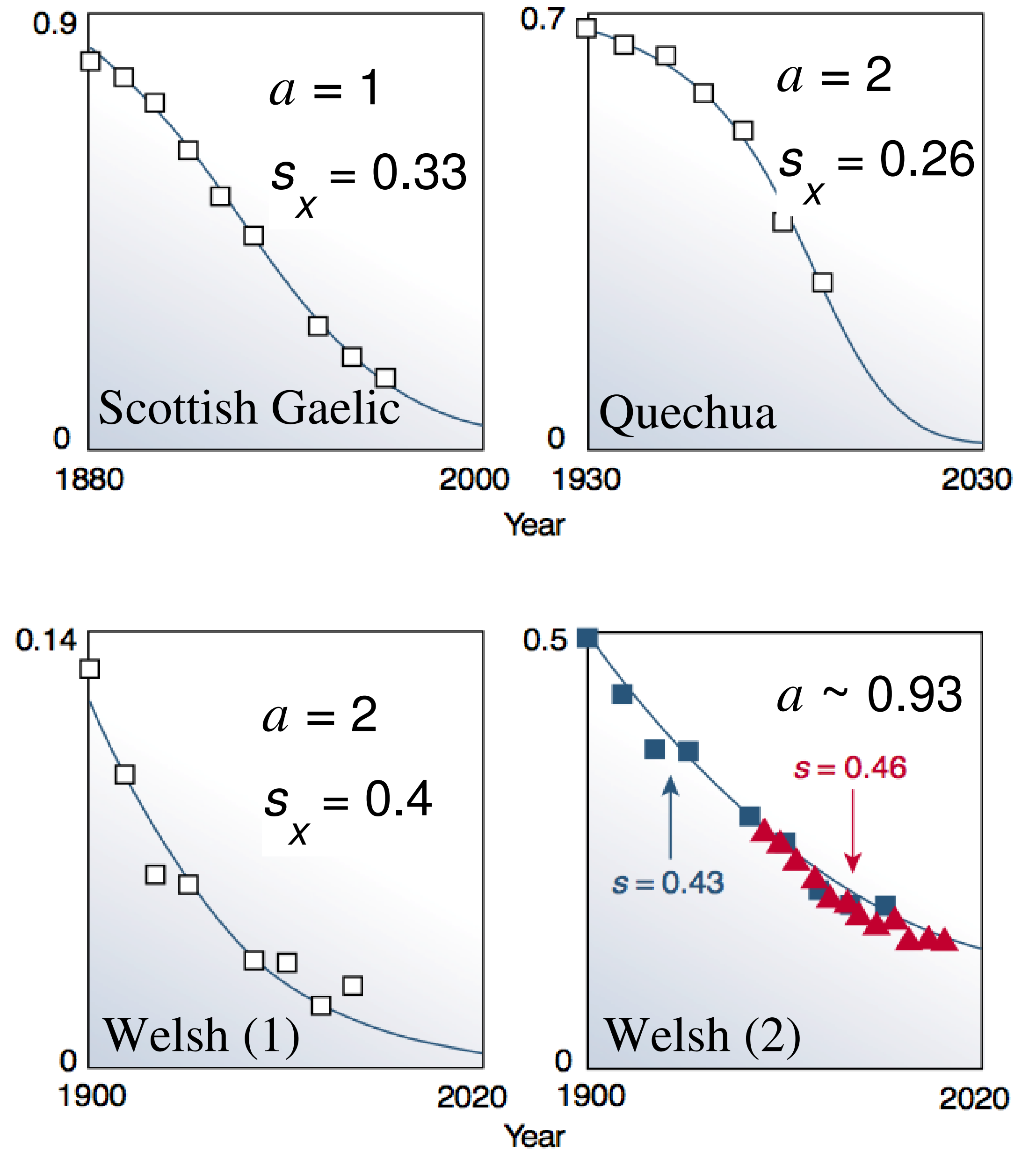
x = population of speakers 1
y = population of speakers 2

$$dx/dt = k\,[\,s_x\,x^a\,y - s_y\,y^a\,x\,]$$
$$dy/dt = k\,[\,-s_x\,x^a\,y + s_y\,y^a\,x\,]$$

$$x + y = 1$$

$$a = ?$$

fraction of speakers versus year



D. M. Abrams and S. H.Strogatz, **Modeling the dynamics of language death**, Nature 424, 900 (2003)

Competition Models: Model of Pinasco and Romanelli

$$dx/dt = \beta\, x\, y + \alpha_x\, x\,(1 - x / K_x)$$

$$dy/dt = -\beta\, x\, y + \alpha_y\, y\,(1 - y / K_y)$$

α_x, α_y = Malthus growth rates

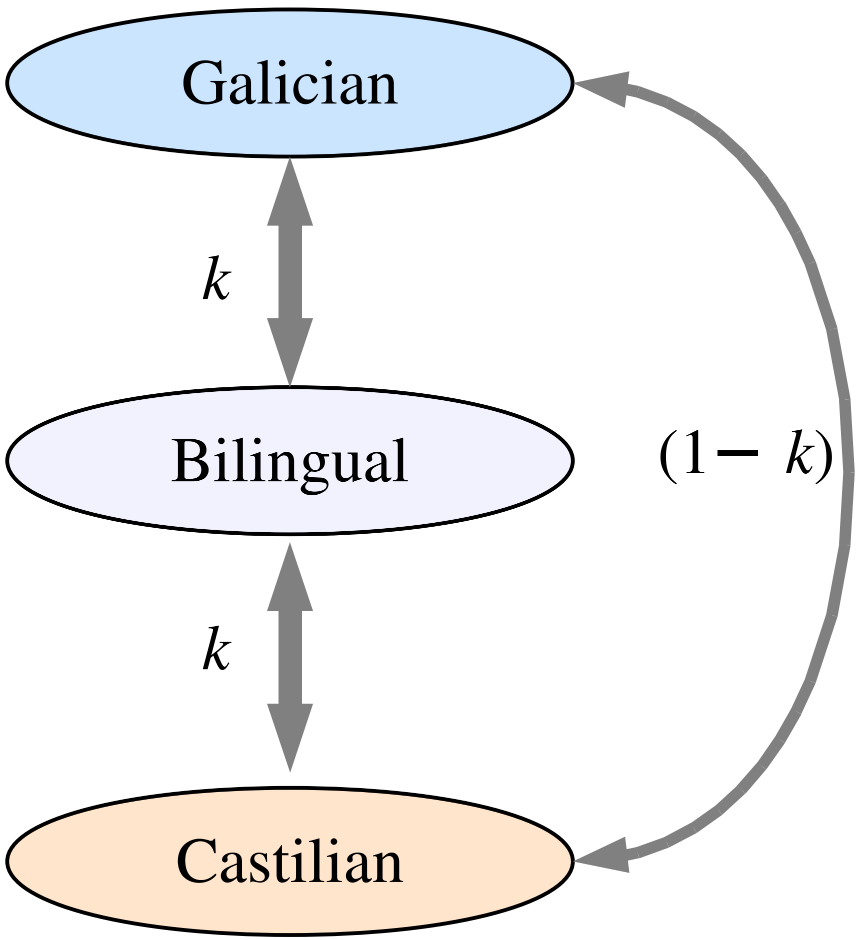
K_y, K_x = carrying capacities

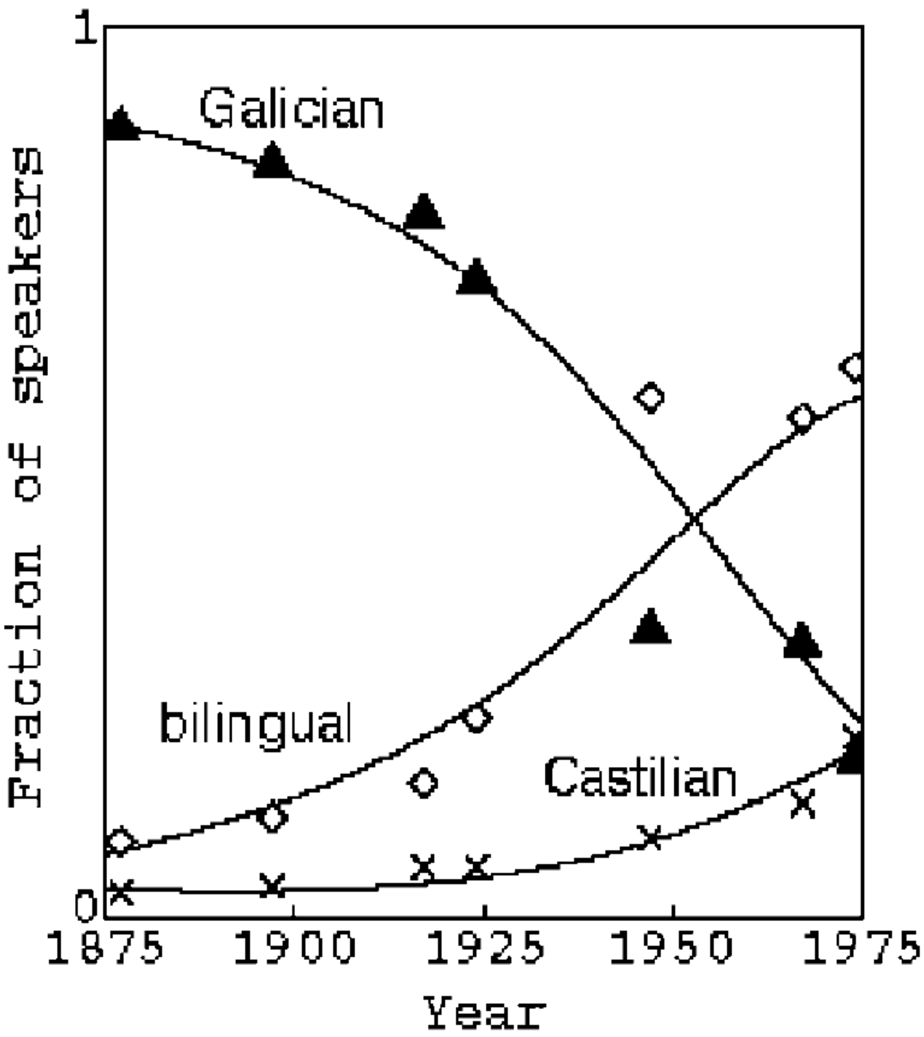
β = rate constant = $k\,(s_x - s_y)$ for $a = 1$

This model has a **stable equilibrium point (x', y') with $x' > 0$ and $y' > 0$**

J. Pinasco, L. Romanelli, **Coexistence of languages is possible**, Physica A 361 (2006) 355

Competition Models: Model of Mira and Paredes





- J. Mira, A. Paredes, **Interlinguistic similarity and language death dynamics**, Europhys.Lett. 69 (2005) 1031
- Seminario de sociolinguistica, Usos linguısticos en Galicia (Real Academia Galega, Santiago) 1995.

Constructing a geographical model of culture transmission

$$\frac{\partial f_1}{\partial t} = +\overbrace{R(f_1, f_2)}^{\text{Cultural interaction}} - \overbrace{\nabla \cdot [\mathbf{F}_1 f_1] + \nabla(D_1 \nabla f_1)}^{\text{Dispersal}} + \ldots$$

$$\frac{\partial f_2}{\partial t} = -R(f_1, f_2) - \underbrace{\nabla \cdot [\mathbf{F}_2 f_2] + \nabla(D_2 \nabla f_2)} + \ldots$$

Model 1

Political borders and geographical boundaries: modify the reaction term to model the *asymmetrical* influence of the other language

Model 2

Modify the diffusion terms in order to model geographical inhomogeneities *felt in the same way* by both populations

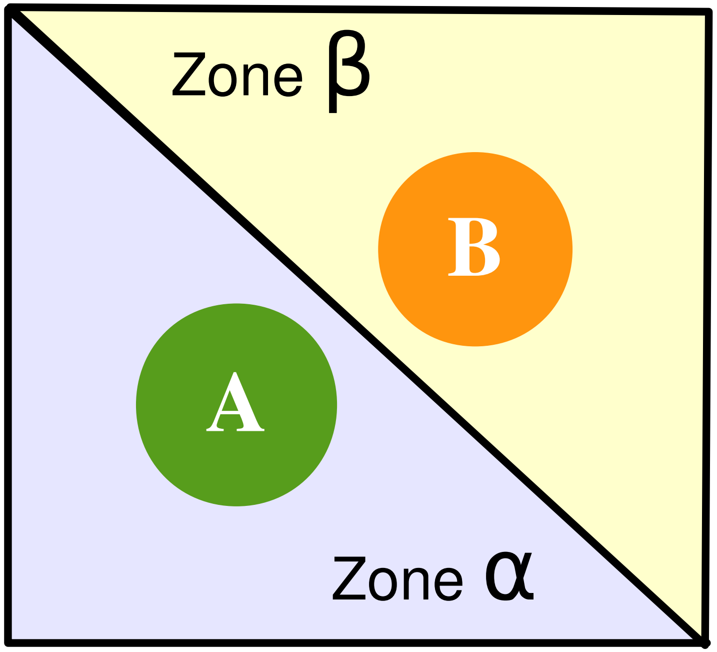
Model 1: Homogeneous dispersal with asymmetrical influence zone

$$\frac{\partial f_A(x,y,t)}{\partial t} = D_A \Delta f_A(x,y,t) + R(x,y,t) \,,$$
$$\frac{\partial f_B(x,y,t)}{\partial t} = D_B \Delta f_B(x,y,t) - R(x,y,t) \,,$$
$$R(x,y,t) = \lambda_0 \{F_A(x,y,t)\, f_B(x,y,t) - F_B(x,y,t)\, f_A(x,y,t)\}$$

Simulation area with populations A and B in the zones α and β

$F_A(x,y,t)$ is not the total population A but only that in the same zone of the position (x,y): it measures the influence of population A only in the zone α.

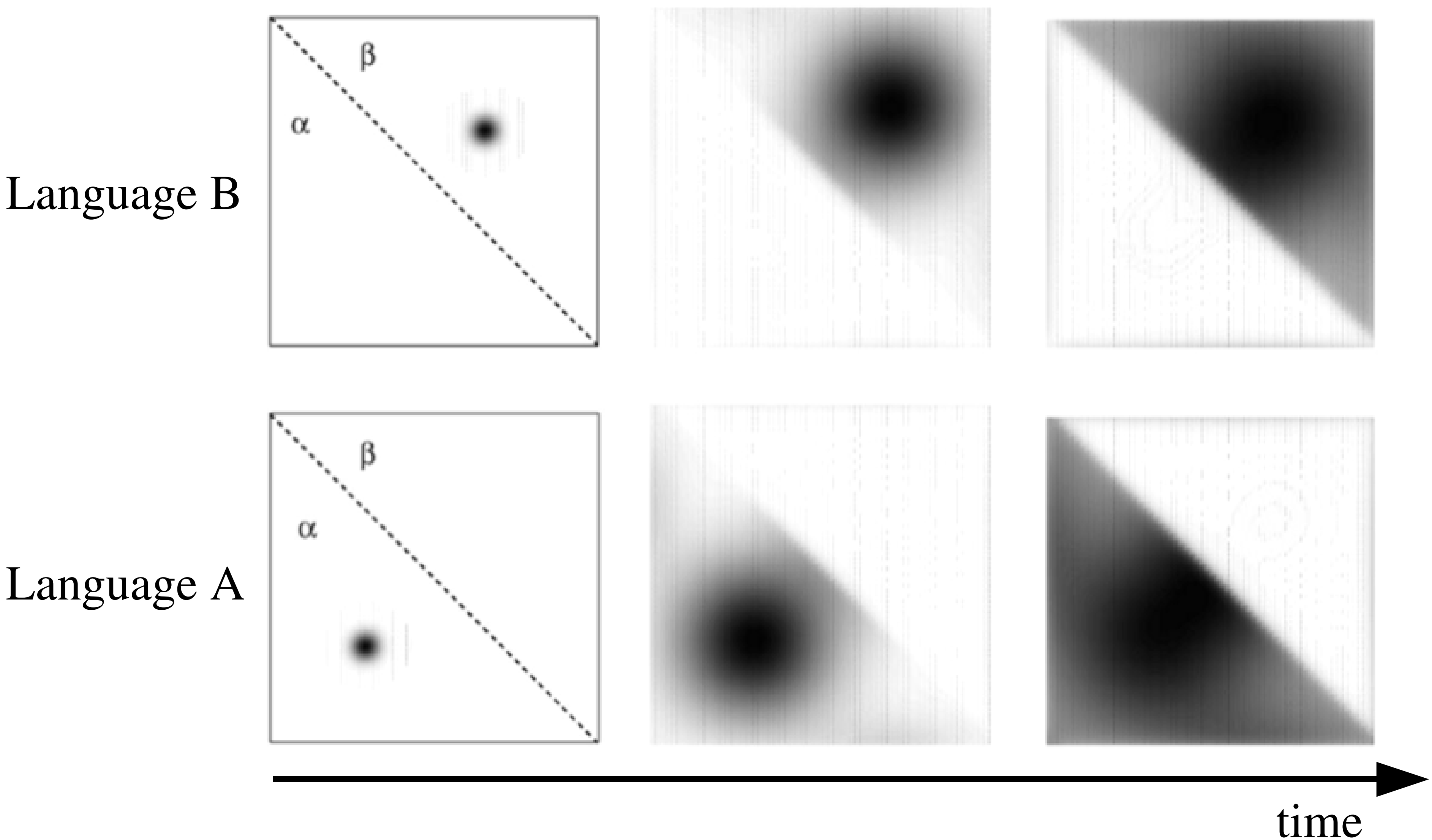
Analogously F_B measures the local influence of population B.



M. Patriarca, T. Leppänen, **Modeling language competition,** Physica A 338, 296 (2004)

The model introduces **an asymmetry in the influence zone**: who is e.g. in zone α is influenced only by speakers in the same zone.

This implies memory effects, that is a dependence on the initial conditions: who occupies first a region, which is difficult to access, has good chances to maintain it.



Model 2 : Homogeneous reaction term

- symmetrical dispersal
- asymmetrical dispersal

$$\frac{\partial f_1}{\partial t} = +R(f_1, f_2) - \nabla \cdot [\mathbf{F}_1 f_1] + \nabla(D_1 \nabla f_1) \ ,$$

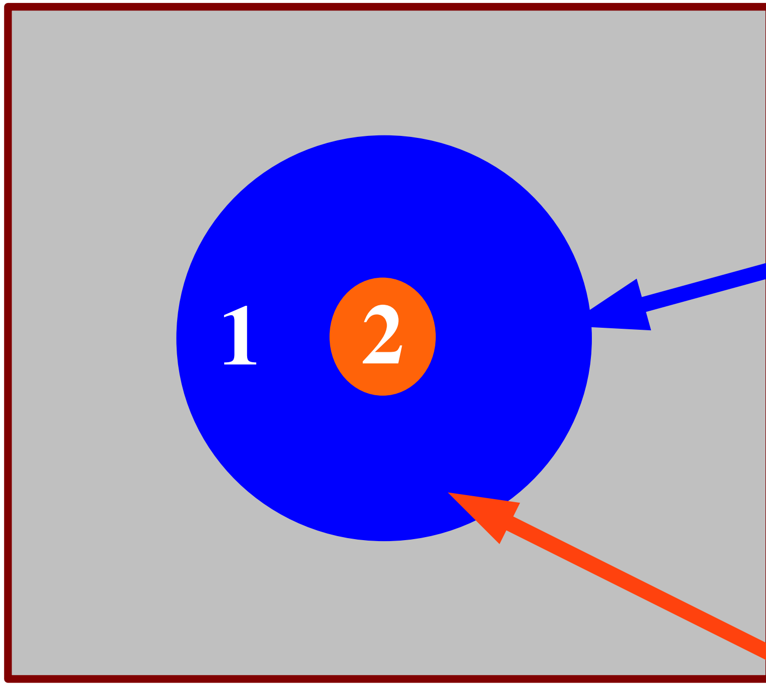
$$\frac{\partial f_2}{\partial t} = -R(f_1, f_2) - \nabla \cdot [\mathbf{F}_2 f_2] + \nabla(D_2 \nabla f_2) \ ,$$

$$R(f_1, f_2) = k \left[s_1 f_1^a \, f_2 - s_2 f_2^a \, f_1\right] \ ,$$

- Here cultural exchange, represented by the reaction term $R(f_1, f_2)$, is symmetrical in the two languages.

- Instead, there is a modulation of dispersal due to the external fields $F_i\,(x, y)$ and/or the diffusion coefficients $D_i(x, y)$.

Example 1: Influence of initial distributions



simulation area
with reflecting boundary conditions

population 1:

status $s_1 = 0.55$

initial fraction $x_1 = 0.5$

$\sigma_1 = 10$

population 2:

status $s_2 = 0.45$

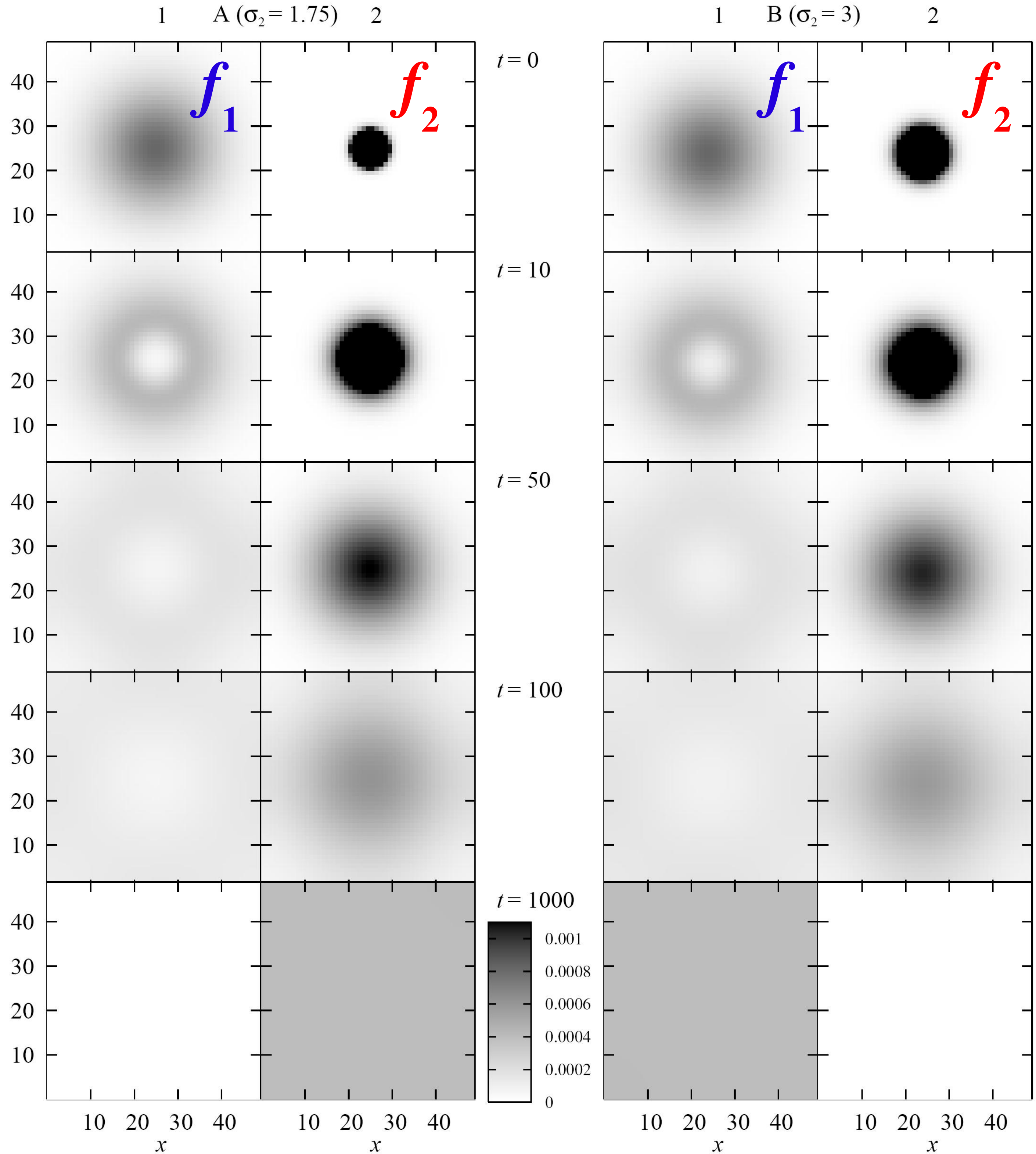
initial fraction $x_2 = 0.5$

σ_2 = 1.75 (case A)

3.0 (case B)

case A

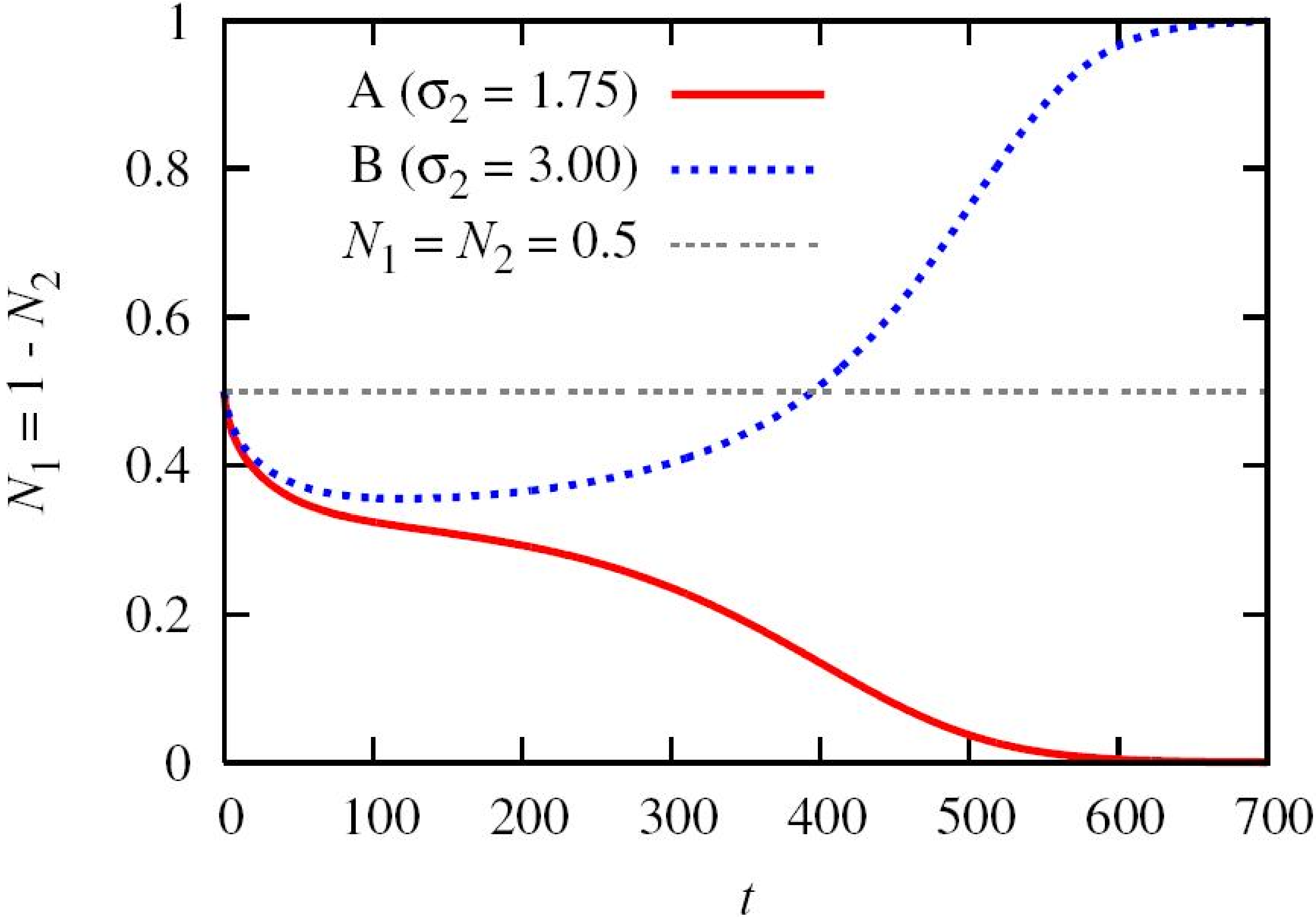
$\sigma_2 = 1.75$



case B

$\sigma_2 = 3$

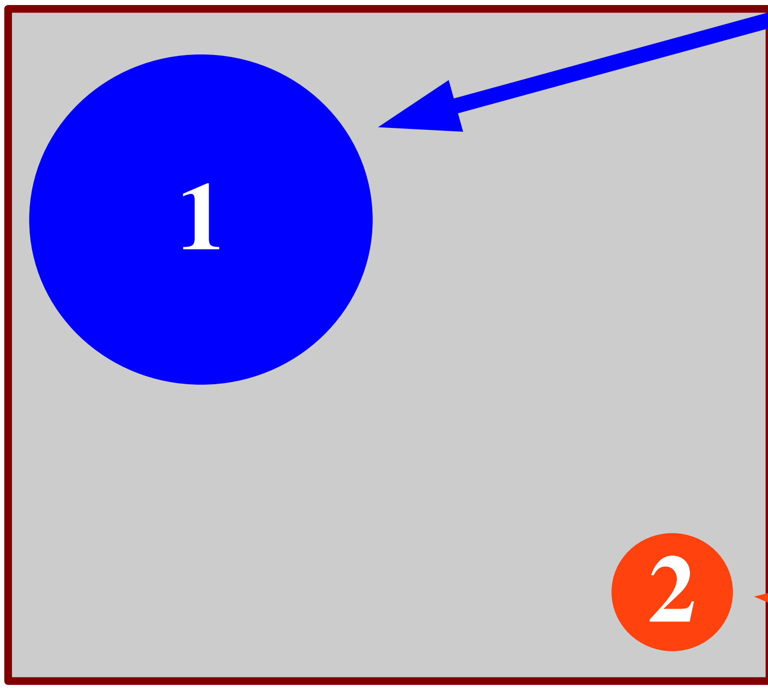
Population size versus time



Example 2: Influence of boundary conditions

The simulation area has either

- Reflecting Boundary Conditions (RBC) or
- Periodic Boundary Conditions (PBC)



population 1:

status $s_1 = 0.55$

initial fraction $x_1 = 0.37$

$\sigma_1 = 10$

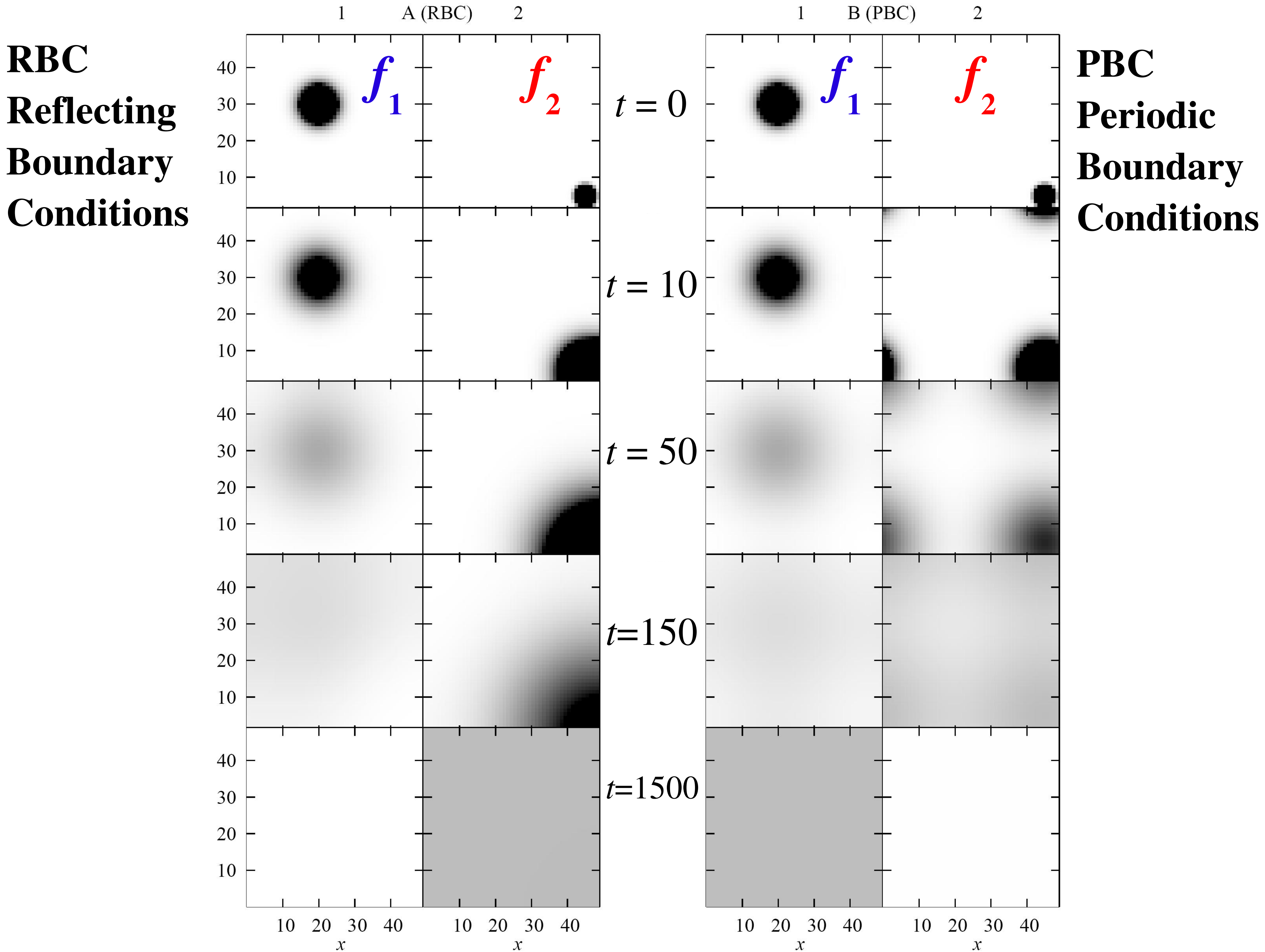
population 2:

status $s_2 = 0.45$

initial fraction $x_2 = 0.63$

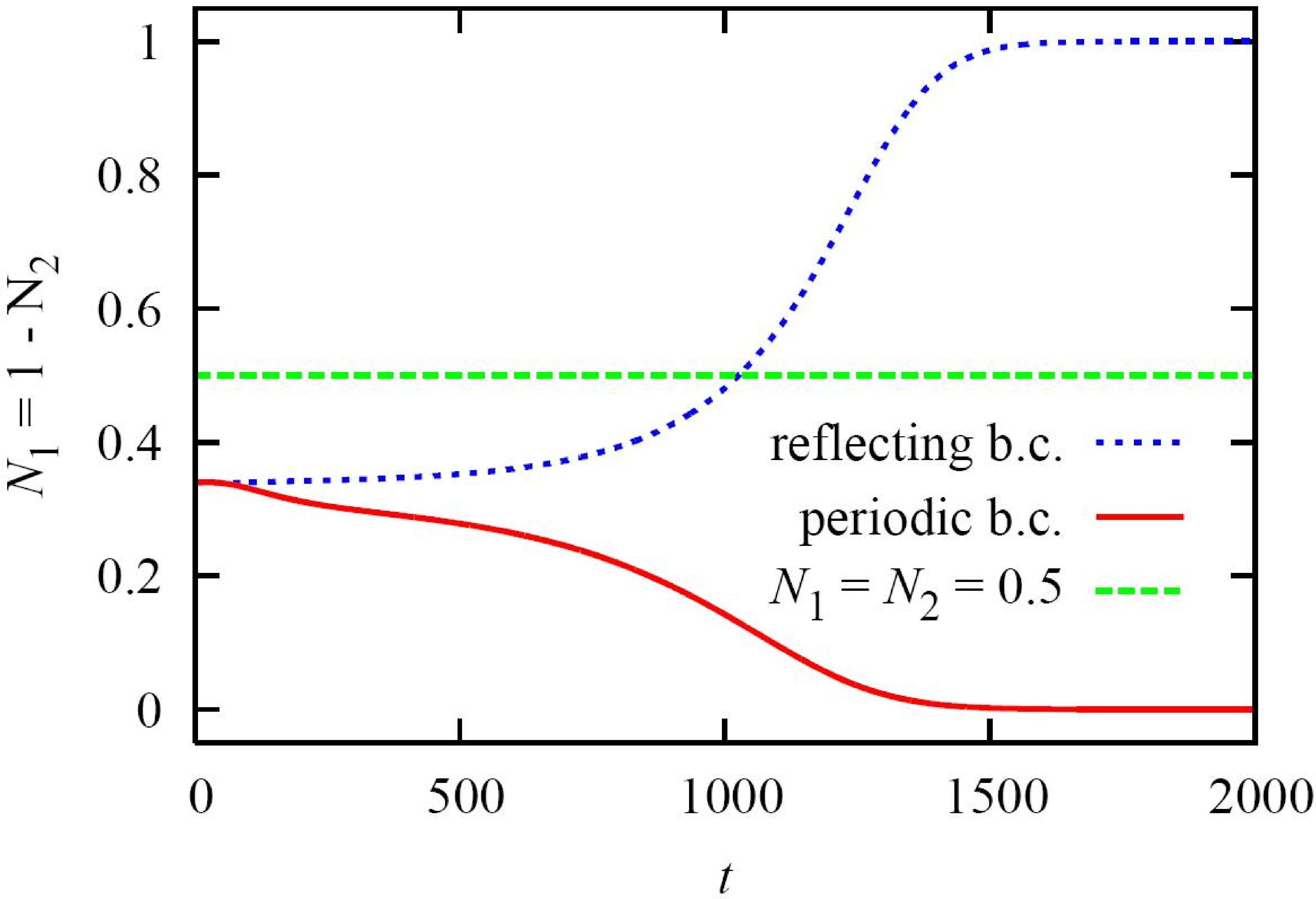
$\sigma_2 = 1$

critical fractions $x_2 = 0.66 = 1 - x_1$



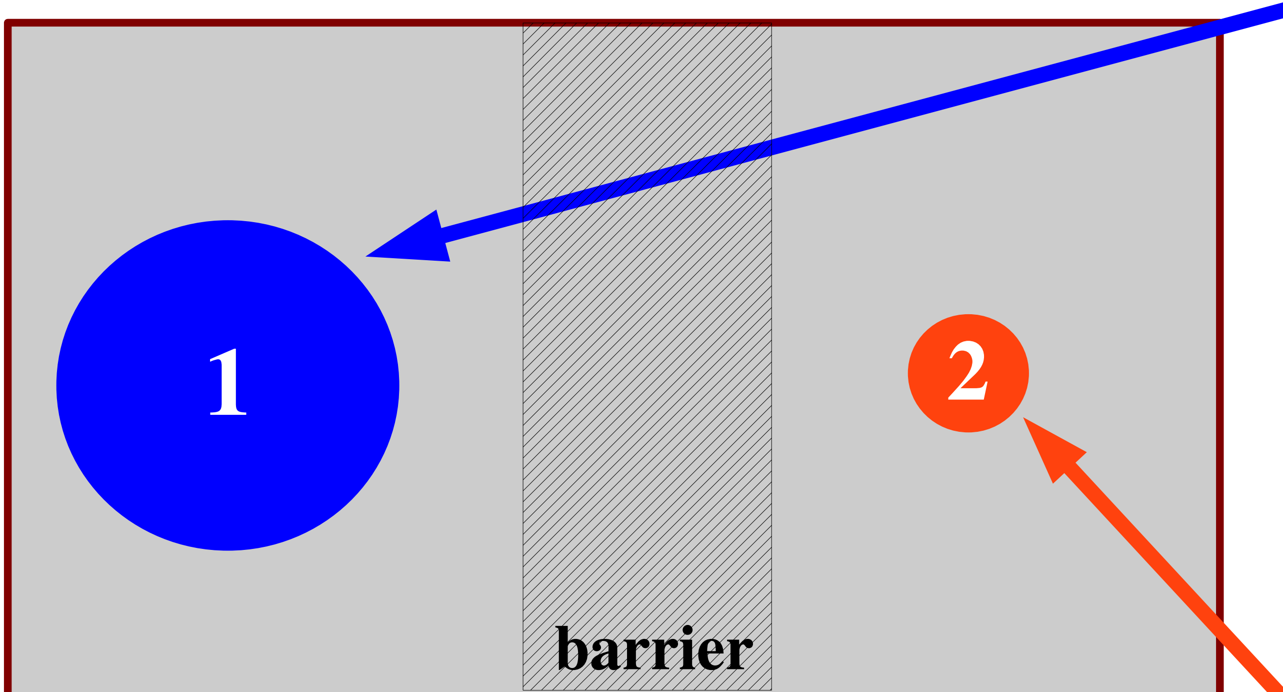
RBC
Reflecting
Boundary
Conditions
1
A (RBC)
2
1
B (PBC)
2
PBC
Periodic
Boundary
Conditions
f_1
f_2
f_1
f_2
$t = 0$
$t = 10$
$t = 50$
$t = 150$
$t = 1500$
10
20
30
40
x

Population size versus time



Example 3 : Influence of a geographical barrier

simulation area with a barrier in the middle



population 1:

status $s_1 = 0.6$

initial fraction $x_1 = 0.8$

$\sigma_1 = 1$

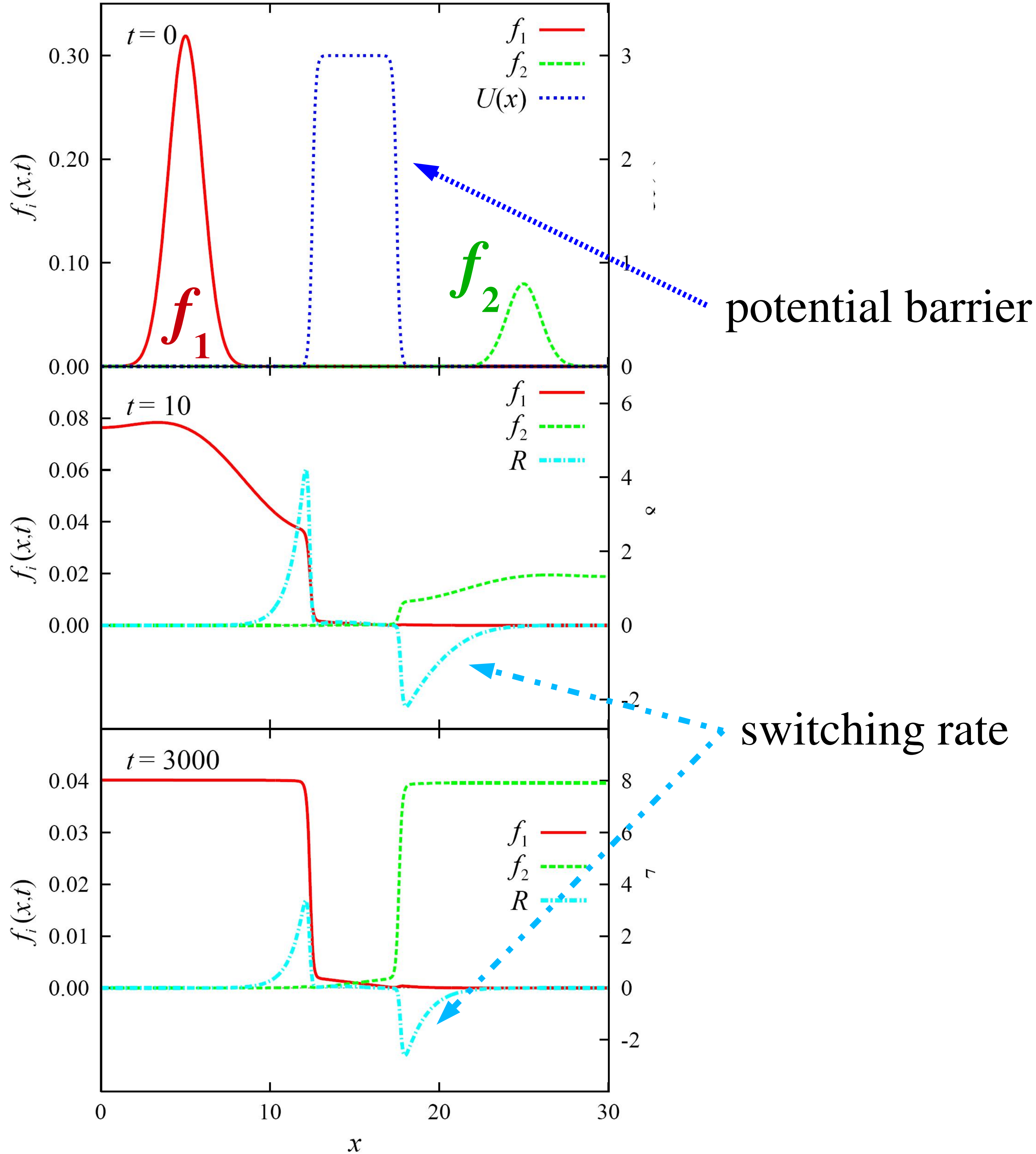
population 2:

status $s_2 = 0.4$

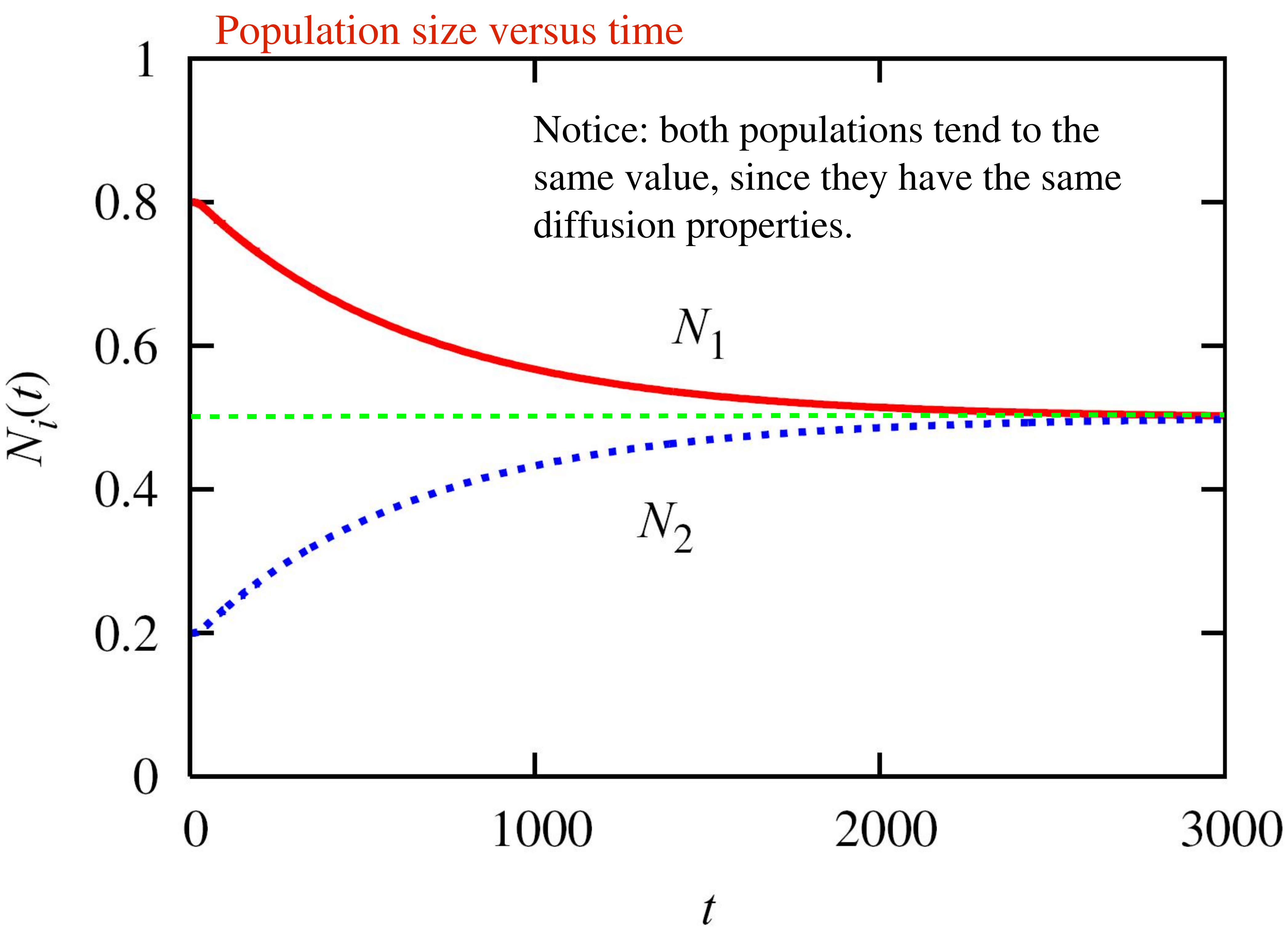
initial fraction $x_2 = 0.2$

$\sigma_2 = 1$

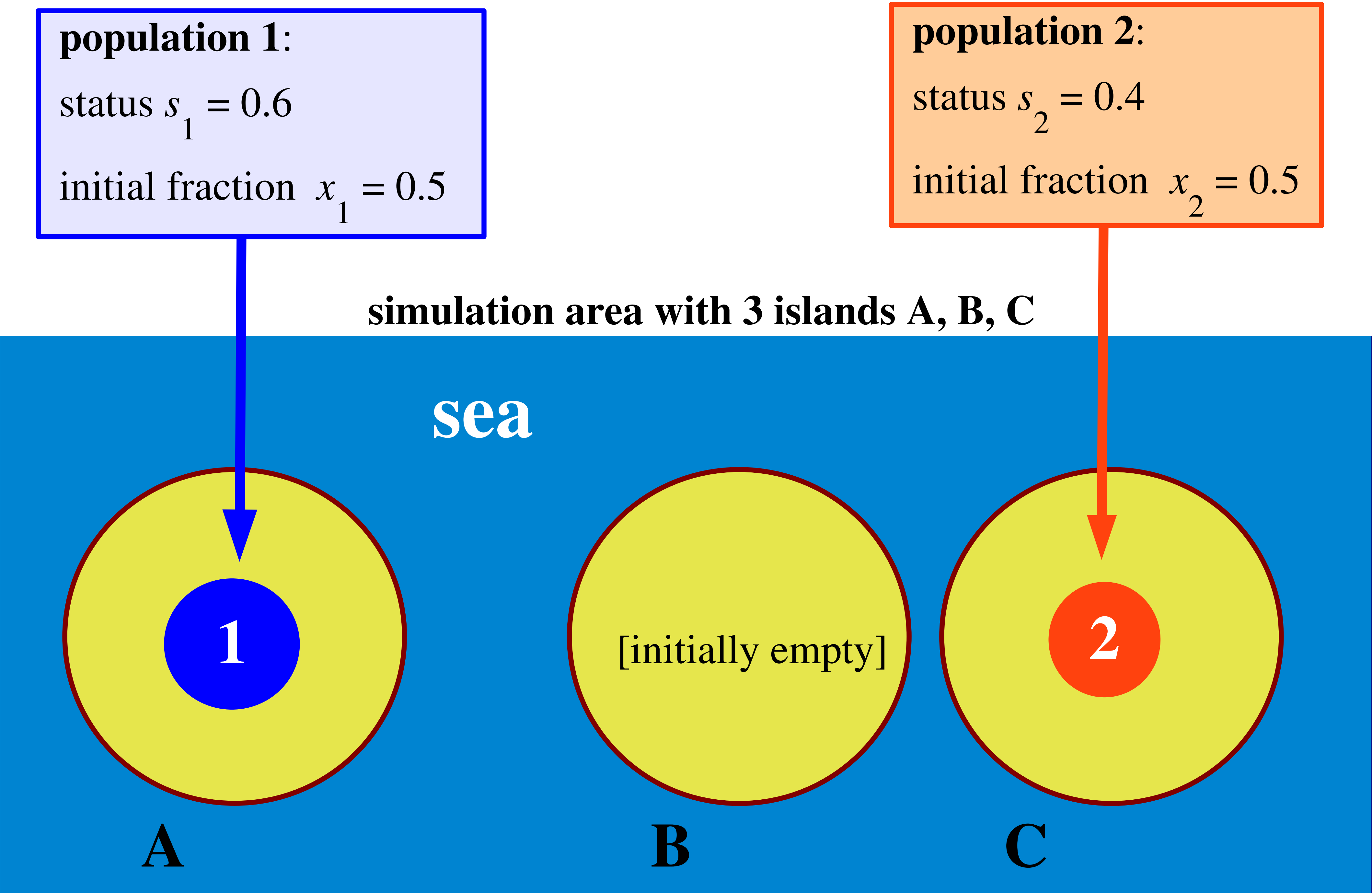
population 1 is highly advantaged both in initial population and status

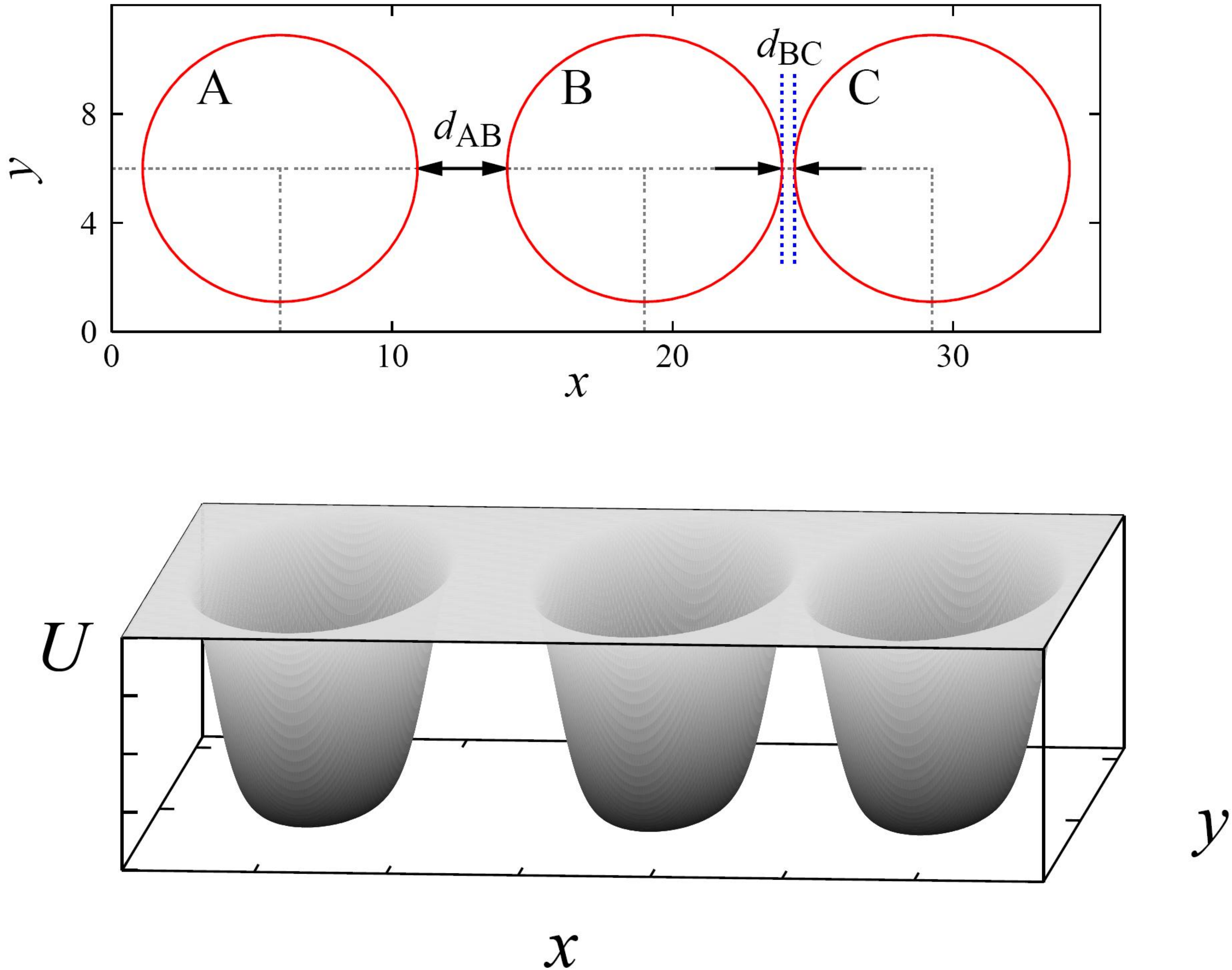


t = 0
f1
f2
U(x)
0.30
0.20
0.10
0.00
3
2
1
0
f1
f2
potential barrier
t = 10
f1
f2
R
0.08
0.06
0.04
0.02
0.00
6
4
2
0
-2
switching rate
t = 3000
f1
f2
R
0.04
0.03
0.02
0.01
0.00
8
6
4
2
0
-2
fi (x,t)
0
10
20
30
x

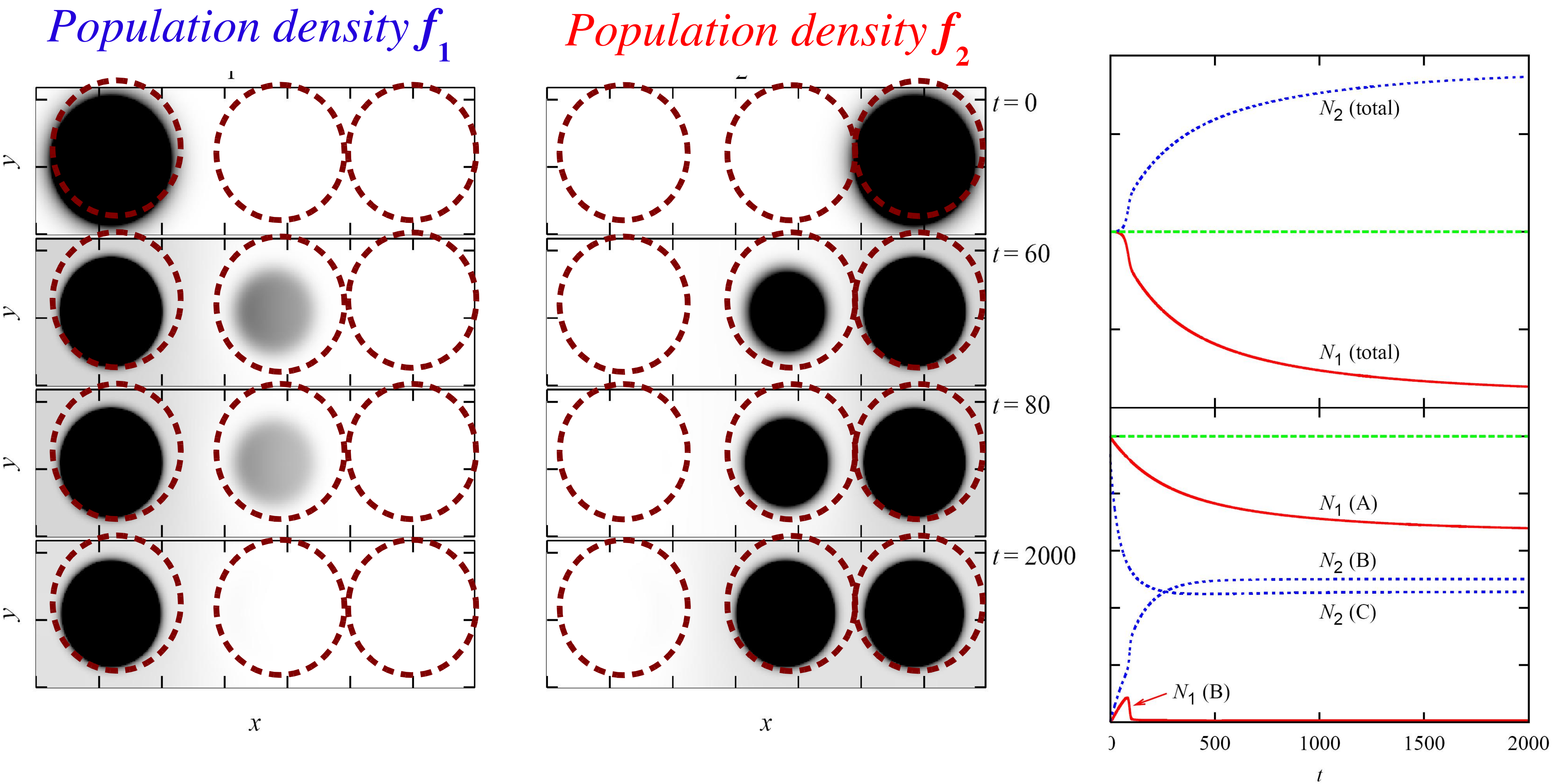

Population size versus time
1
0.8
0.6
0.4
0.2
0
0
1000
2000
3000
Notice: both populations tend to the same value, since they have the same diffusion properties.
N_1
N_2
$N_i(t)$
t

Example 4 : immigration to an island





A
B
C
d_{AB}
d_{BC}
y
8
4
0
0
10
20
30
x
U
y
x



Population density f_1
Population density f_2
$t = 0$
$t = 60$
$t = 80$
$t = 2000$
y
x
N_2 (total)
N_1 (total)
N_1 (A)
N_2 (B)
N_2 (C)
N_1 (B)
500
1000
1500
2000
t

Final Remarks

- Survival of languages/cultural traits with a lower status can be favored by
1. boundaries and borders (geographical constraints)
2. zones of high concentration (historical conditions)

- Different cultural traits can survive in different regions
1. if they feel different influence zones
2. if there is a barrier causing geographical isolation

- It is possible to study the diffusion of information (culture, languages) according to a diffusional scheme

- Some facts concerning the geographical evolution and distribution of languages could be explained quantitatively

- For the future work: random path approaches and discrete models (also convenient for computational reasons)