Bilingual societies: from data to agent-based modeling

Thomas Louf, José J. Ramasco and David Sánchez

IFISC (CSIC-UIB) Palma de Mallorca – Spain.

thomaslouf@ifisc.uib-csic.es

UNIT OF EXCELLENCE MARÍA DE MAEZTU

AGENCIA ESTATAL DE INVESTIGACIÓN

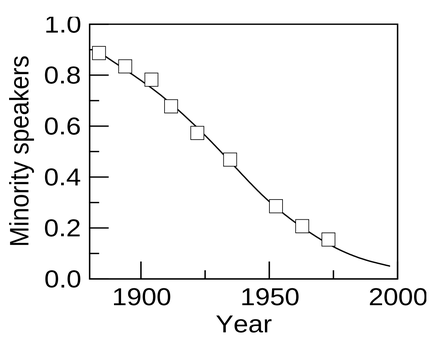
Abstract

A great part of the cultural diversity contained within the languages of the world is at risk, as many languages have become endangered in the last decades[1]. To preserve this diversity, one must first understand what drives the death of languages, and what could enable their coexistence. This is why we seek here to uncover the mechanisms at work in language shift through agent-based modeling[2]. A large scale empirical study of the patterns of spatial mixing of languages in multilingual societies using Twitter data unveils an interesting diversity. It ranges from an almost complete mixing of languages to their complete separation. To understand how these different states can arise, we propose a model in which coexistence of languages may be reached when learning the other language is facilitated and when bilinguals favor the use of the endangered language. Simulations carried out in a meta-population framework highlight the importance of spatial interactions arising from the mobility of people to explain the stability of a border between two linguistic regions

Motivation

Example of language dynamics model: prediction of language death according to Abrams and Strogatz[3]

Aim: Explore variety of multilingual societies to evaluate models under a new light.
How: Data-driven approach with geo-tagged tweets → spatial distributions of language use in multilingual societies.



A diversity of multilingual societies?

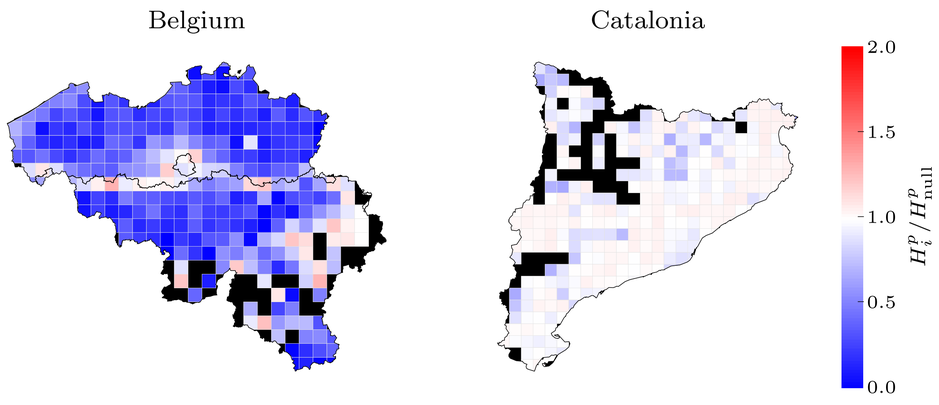
Data analysis → $N_{L,i}$: number of individuals of each language group L in each square cell i.

Measure segregation as departure from null model, in which individuals of each language group are distributed like the whole population → proportion entropy:

$$H_i^p = -\sum_L \frac{N_{L,i}}{N_i} \log\left(\frac{N_{L,i}}{N_i}\right)$$

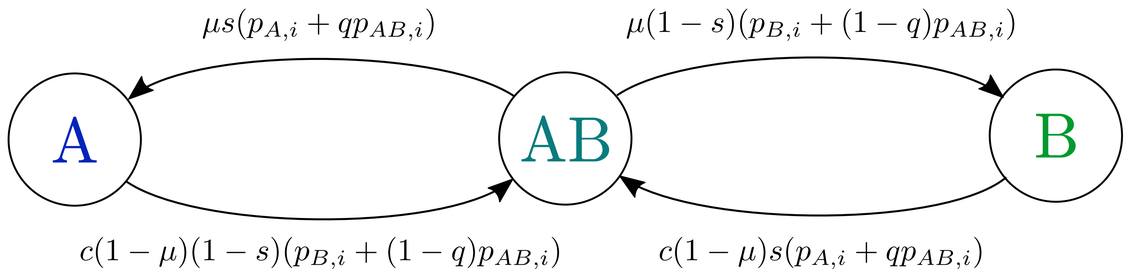
to be compared to its value in the null model → relative entropy to measure mixing in i .

Maps of language mixing in Belgium and Catalonia, two bilingual societies displaying completely different mixing patterns:



Our model

Agents in one of three states: monolinguals A or B, or bilinguals AB.
Transition probabilities:

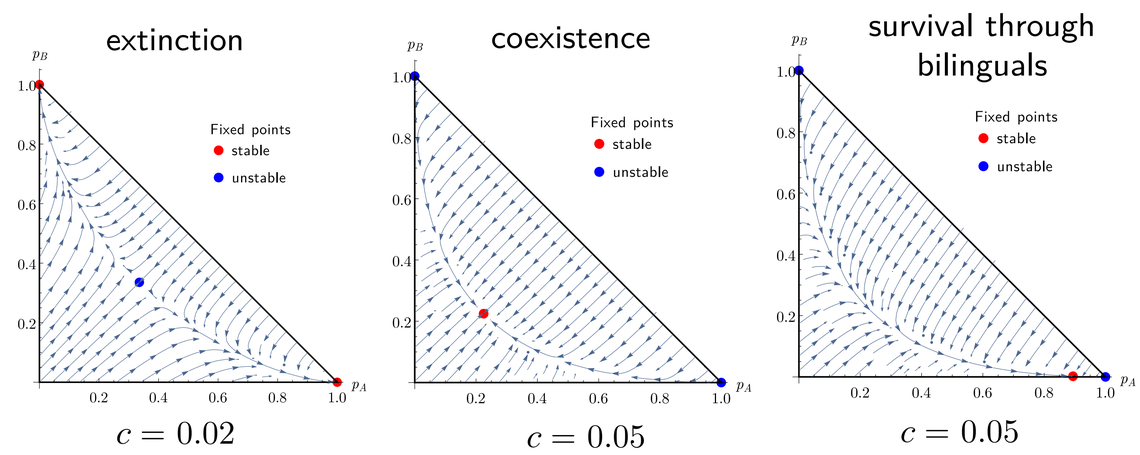


- s the prestige of A
- $p_{L,i}$ the proportion of A around agent i
- μ the mortality rate (fixed at 0.02 in the following)
- c adjusting for the learning rate
- q the bilinguals' preference to speak A

Results: single population

First we solve the system of master equations corresponding to our agent-based model set in mean field.

Stability analysis → **stable coexistence is possible**!



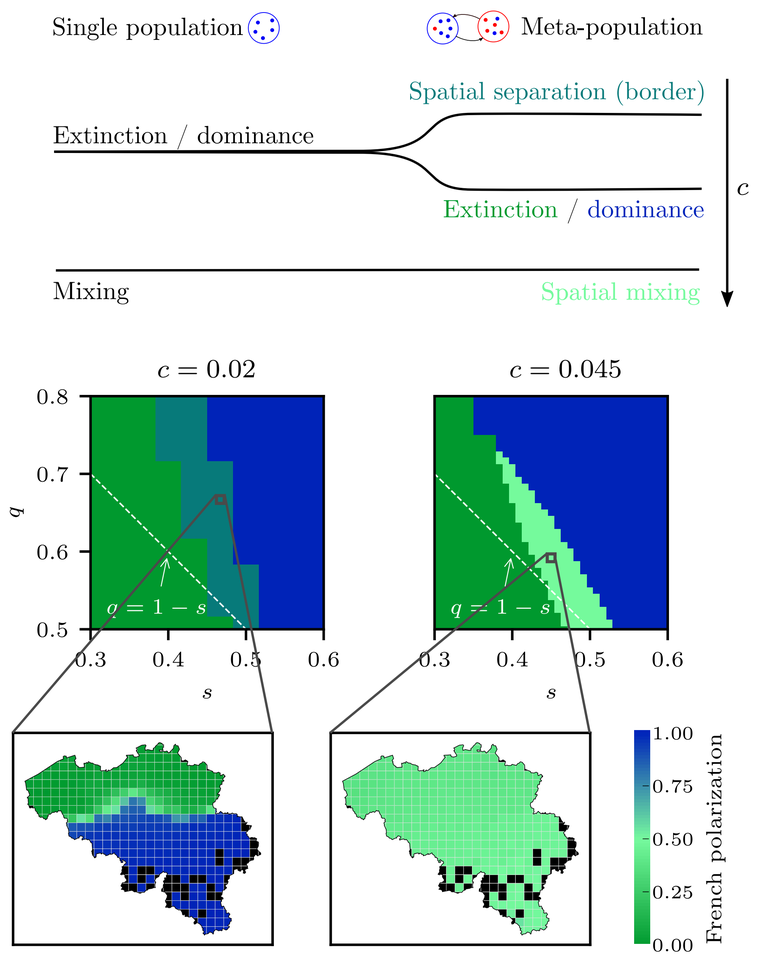
Results: meta-population

To investigate spatial patterns, simulation of agent-based model in meta-population.

Agents within language group, with residence and work cell, all interact at home then at work.

→ Iterated until convergence to stable state, enabling phase space exploration.

Belgium example: convergence to **stable situation with border** possible **as well as total mixing**, depends mostly on ease to learn other language, controlled by c



References

[1] M Krauss, The world's languages in crisis. *Language* **68**, 4–10 (1992).
[2] X Castelló, L Loureiro-Porto, M San Miguel, Agent-based models of language competition. *Int. J. Sociol. Lang.* **2013**, 21–51 (2013)
[3] DM Abrams, SH Strogatz, Modelling the dynamics of language death. *Nature* **424**, 900–900 (2003).