

MODELLING LANGUAGE COMPETITION: BILINGUALISM AND COMPLEX SOCIAL NETWORKS

XAVIER CASTELLÓ, VÍCTOR M. EGUÍLUZ, and MAXI SAN MIGUEL
*IFISC, Institut de Física Interdisciplinària i Sistemes Complexos (CSIC-Universitat de les
Illes Balears),
Palma de Mallorca, E-07122, Spain
xavi@ifisc.uib.es*

LUCÍA LOUREIRO-PORTO
*Dep. Filologia Espanyola, Moderna i Llatina, Universitat de les Illes Balears,
Palma de Mallorca, E-07122, Spain
lucia.loureiro@uib.es*

RIITTA TOIVONEN, J. SARAMÄKI and K. KASKI
*Laboratory of Computational Engineering, Helsinki University of Technology,
Helsinki, P.O. Box 9203, 02015 HUT, Finland
rtoivone@lce.hut.fi*

In the general context of dynamics of social consensus, we study an agent based model for the competition between two socially equivalent languages, addressing the role of bilingualism and social structure. In a regular network, we study the formation of linguistic domains and their interaction across the boundaries. We also analyse the dynamics on a small world network and on a network with community structure. In all cases, a final scenario of dominance of one language and extinction of the other is obtained (*dominance-extinction state*). In comparison with the regular network, smaller times for extinction are found in the small world network. In the network with communities instead, the average time for extinction does not give a characteristic time for the dynamics, and metastable states are observed at all time scales.

1. Introduction

Language competition occurs today worldwide. Different languages coexist within many societies and the fate of a high number of them in the future is worrying: most of the 6000 languages spoken today are in danger, with around 50% of them facing extinction in the current century. Even more striking is the distribution of speakers, since 4% of the languages are spoken by 96% of the world population, while 25% have fewer than 1000 speakers. New pidgins and creoles are also emerging, but their number is relatively small compared with the language loss rate (Crystal, 2000). In this scenario, and beyond Weinreich's *Languages in*

Contact (Weinreich, 1953), numerous sociolinguistic studies have been published in order to: (1) reveal the level of endangerment of specific languages (Tsunoda, 2005); (2) find a common pattern that might relate language choice to ethnicity, community identity or the like (O' Driscoll, 2001); and (3) claim the role played by social networks in the dynamics of language competition, which has given rise to a monographic issue of the *International Journal of Sociology* (De Bot & Stoessel, 2002a).

Abrams and Strogatz model for the dynamics of endangered languages (Abrams & Strogatz, 2003) has triggered a coherent effort to understand the mechanisms of language dynamics outside the traditional linguistic research. Their study considers a two-state society, that is, one in which there are speakers of either a language A or a language B. This seminal work belongs to the general class of studies of population dynamics based on nonlinear ordinary differential equations for the populations of speakers of different languages. In addition, other studies implement discrete agent based models with speakers of many languages (Stauffer & Schulze, 2005) or few languages (Stauffer, Castello, Eguiluz, & San Miguel, 2007), as reviewed in (Schulze & Stauffer, 2006).

Language competition, then, belongs to the general class of processes that can be modelled by the interaction of heterogeneous agents as an example of collective phenomena in problems of social consensus (San Miguel, Eguiluz, Toral, & Klemm, 2005). In this respect, a specific feature of language dynamics is that agents can share two of the social options that are chosen by the agents in the consensus dynamics. In the present work, these are the bilingual agents, that is, agents that use both language A and B, who have been claimed to play a relevant role in the evolution of multilingual societies (Wang & Minett, 2005).

In this work we are interested in the emergent phenomena appearing as a result of a self-organised dynamics in the case of two equally prestigious competing languages. With the aim of elucidating possible mechanisms that could stabilise the coexistence of these languages, we wish to discuss the role of bilingual individuals and, following Milroy, the effects of social structure (Milroy, 2001) in the process of language competition. To this end, and along the lines of the original proposal by Minett and Wang (Wang & Minett, 2005), we study an agent based model that incorporates bilingual agents on different networks: a regular network, a small world network, and a social type network with community structure (Toivonen et al., 2006). We compare the results obtained with our work on the agent-based version of Abrams-Strogatz two-state model (Stauffer et al., 2007). This way we will provide a quantitative analysis that is wanting in the field of sociolinguistics, as noted by de Bot and Stoessel (De Bot & Stoessel, 2002b).

2. The Bilinguals Model

We consider a model of two socially equivalent (i.e. equally prestigious) competing languages in which an agent i sits in a node within a network of N individuals

and has k_i neighbours. It can be in three possible states: A , agent using ^a language A ; B , agent using language B ; and AB , bilingual agent using both, A and B .

The state of an agent evolves according to the following rules: at each iteration we first choose one agent i at random, and, then, we compute the local densities of language users of each linguistic community in the neighbourhood of agent i : σ_i^l ($l=A, B, AB$; $i=1, N$; $\sigma_i^A + \sigma_i^B + \sigma_i^{AB} = 1$). The agent i changes its state of language use according to the following transition probabilities ^b:

$$\begin{aligned} p_{i,A \rightarrow AB} &= \frac{1}{2}\sigma_i^B \quad , & p_{i,B \rightarrow AB} &= \frac{1}{2}\sigma_i^A & (1) \\ p_{i,AB \rightarrow B} &= \frac{1}{2}(1 - \sigma_i^A) \quad , & p_{i,AB \rightarrow A} &= \frac{1}{2}(1 - \sigma_i^B). & (2) \end{aligned}$$

Equation (1) gives the probabilities for an agent to move away from a monolingual community to the bilingual community AB . They are proportional to the density of monolingual speakers of the other language in its neighbourhood. On the other hand, equation (2) gives the probabilities for an agent to move from the bilingual community towards one of the monolingual communities. Such probabilities are proportional to the density of speakers of the adopting language including bilinguals ($1 - \sigma_i^l = \sigma_i^j + \sigma_i^{AB}$, $l, j=A, B$; $l \neq j$). It is important to note that a change from being monolingual A to monolingual B or vice versa always implies an intermediate step through the bilingual community. The transition probabilities (1) and (2) are fully symmetric under the exchange of A and B , which is consistent with the fact that both languages are socially equivalent in terms of prestige.

We recover the agent-based version of Abrams-Strogatz two-state model when bilinguals are not present. In this model, an agent essentially imitates language use of a randomly chosen neighbour. The corresponding transition probabilities are the following:

$$p_{i,A \rightarrow B} = \frac{1}{2}\sigma_i^B \quad , \quad p_{i,B \rightarrow A} = \frac{1}{2}\sigma_i^A \quad (3)$$

For a quantitative description of the emergence and dynamics of linguistic spatial domains we use the ensemble average interface density $\langle \rho \rangle$ as an order parameter. This is defined as the density of links joining nodes in the network which are in different states (San Miguel et al., 2005). The ensemble average, indicated as $\langle \cdot \rangle$, denotes average over realizations of the stochastic dynamics starting from

^aNote that we always refer to language use rather than competence. Therefore, an agent using two languages might stop using one of them as this becomes less spoken in its social vicinity.

^bNon-equivalent languages were considered in the original version of the model (Wang & Minett, 2005). The prefactor $1/2$ corresponds to the special case of equivalence between A and B .

different random distributions of initial conditions. During the time evolution, the decrease of ρ from its initial value describes the ordering dynamics, where linguistic spatial domains, in which agents are in the same state, grow in time. The minimum value $\rho = 0$ corresponds to a stationary configuration in which all the agents belong to the same linguistic community.

3. Results

3.1. Regular and small world networks

The bilinguals model has been extensively studied in two-dimensional regular networks, and small world networks (Castelló, Eguíluz, & San Miguel, 2006). In two-dimensional regular networks, and starting from a randomly distributed state of the agents, spatial domains of each monolingual community are formed and grow in size (Fig 3.1). This is known in the physics literature as *coarsening*. Meanwhile, domains of bilingual agents are never formed. Instead, bilingual agents place themselves in a narrow band between monolingual domains (Fig 3.1). Finally a finite size fluctuation drives the system to a dominance-extinction state, where all the agents become monolingual, while the other monolingual community together with the bilingual agents face extinction. Average interface density $\langle \rho \rangle$ decays as a power law $\langle \rho \rangle \sim t^{-\gamma}$, $\gamma \simeq 0.45$ (Castelló et al., 2006). This indicates that the growth law found for the bilinguals model is compatible with the well known exponent 0.5 associated with domain growth driven by mean curvature and surface tension reduction observed in SFKI (spin flip kinetic Ising model) (Gunton, San Miguel, & Sahni, 1983). The characteristic time to reach an extinction of one of the languages τ scales with system size as $\tau \sim N^{1.8}$. A very different behaviour is found for the agent based Abrams-Strogatz model, where bilingual agents are not present: coarsening is slower ($\langle \rho \rangle \sim (\ln t)^{-1}$) and driven by interfacial noise.

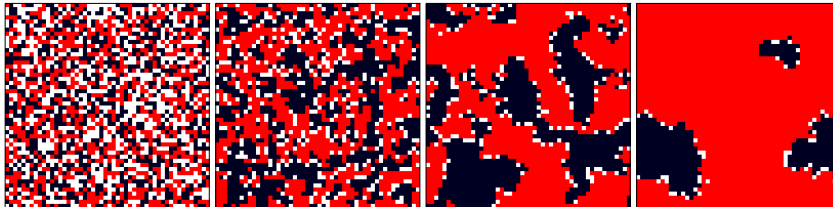


Figure 1. Formation and growth of monolingual domains. Starting from random initial conditions, snapshots of a typical simulation of the dynamics in a two-dimensional regular network of 2500 individuals. $t=0, 2, 20, 200$ from left to right. Grey: monolinguals A , black: monolinguals B , white: bilinguals.

To study the effect of long range social interactions, which are one of the ba-

sis characteristics of social networks, we next consider a small world network (Watts & Strogatz, 1998). There, $\tau \sim \ln N$ (Castelló et al., 2006). For the agent based Abrams-Strogatz model, the long range connections inhibit the formation and growth of monolingual domains by producing long-lived metastable states. Metastable states are those where we find dynamical coexistence of the two languages during a long but finite time, after which the system drops to a dominance-extinction state. However, in the bilinguals model, when we moved towards a small world network by adding long range connections to the two-dimensional regular network, bilingual agents destroy a metastable state of dynamical coexistence, and slow down coarsening, although domains keep growing in size. In addition, they speed up the decay to extinction of one of the languages due to finite size fluctuations (Castelló et al., 2006).

3.2. Social type network with community structure

Community structure is a prominent characteristic of real social networks which may crucially affect social dynamics, and in particular, language competition. A combination of random attachment with search for new contacts in the neighbourhood has proved fruitful in generating cohesive structures (Toivonen et al., 2006). We choose this model, because it produces well-known features of social networks, such as assortativity, broad degree distributions, and community structure.

The most important result regarding this topology (Castelló et al., 2007), is the behaviour of the characteristic time to reach a dominance-extinction state. To this end, we analyse the fraction $f(t)$ of runs still alive at any time t , i.e. the fraction of runs which have not reached the dominance-extinction state. We average over different realizations of the network, and several runs in each. For the agent based Abrams-Strogatz model, the fraction of alive runs decreases exponentially. Results are more interesting for the bilinguals model: $f(t)$ appears to have power law behaviour $f(t) \sim t^{-\alpha}$, $\alpha \approx 1.3$. Since the exponent $\alpha < 2$, the average decay time for the bilinguals model does not give a characteristic time scale, but alive realizations which have not reached the dominance-extinction state are found at any time scale.

The difference between the agent based Abrams-Strogatz model and the bilinguals model is better understood by looking at snapshots of the dynamics (Fig. 3.2) which show the characteristic behaviour for each of the models, starting from random initial conditions ($t = 0$). In the former (left), the homogeneous domains of nodes with the same option appear to follow the community structure, but a particular community (topological region) may change the language adopted by the community rather quickly ($t = 50, 60, 70$). At variance with this behaviour, in the bilinguals model (right) spatial linguistic domains grow and homogenise steadily in a community without much fluctuation. For this dynamics, communities that have adopted a given language, and which are poorly linked to the rest of the network, take a long time to be invaded by the other language, act-

ing therefore as topological traps. As an example of this we show two long lived trapped metastable state at $t = 430$ and $t = 1000$, where the interface density stayed relatively stable for a prolonged period (~ 100 and ~ 1000 time steps, respectively). Again, these different behaviours reflect in the community structure two different interfacial dynamics: interfacial noise driven dynamics for the agent based Abrams-Strogatz model, and curvature driven dynamics for the bilinguals model with agents in the AB state at the interfaces.

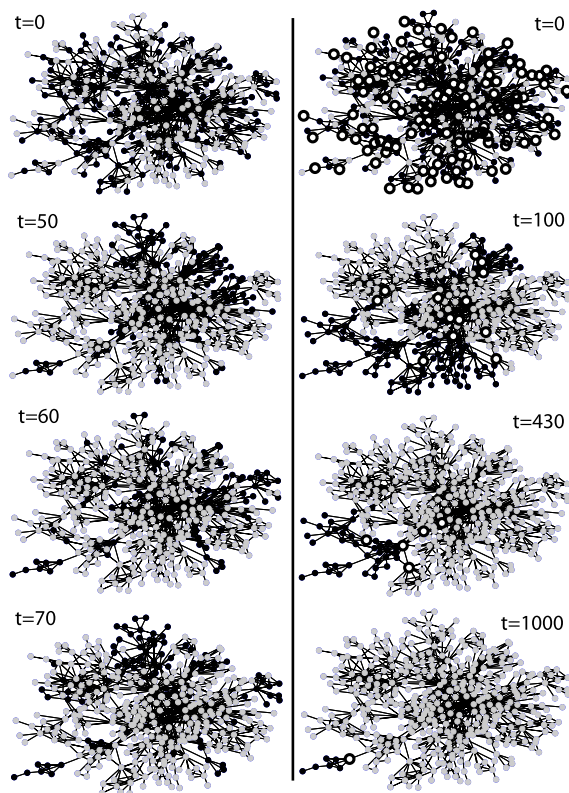


Figure 2. Snapshots of a single run of the dynamics, with nodes in state A in black, B in grey, and AB in white circled in black. Simulations start from random initial conditions. Left: Abrams-Strogatz agent-based model. Right: Bilinguals model (example of a simulation leading to metastable states).

4. Conclusion and further research

We have analysed the bilinguals model (in comparison to the agent-based version of Abrams-Strogatz model) in different topologies. Although the final state

of the system is always a homogeneous state where one of the languages faces extinction, the transient towards this final state depends crucially on the network structure. This comes to complement, from an agent based modelling point of view, the importance of social networks in the processes of language contact already claimed in (Milroy, 1987). Within the limitations and assumptions of the model, the analysis of the small world phenomenon, which is characteristic of the current interconnected societies, might be an ingredient which accelerates language extinction. This effect might be related to an overall globalisation process in which not only languages, but also whole cultures tend to homogeneity rather than diversity. However, the study of the dynamics in the social type network with communities, which is the network that mimics the features of real social networks in a closer way, shows that there exist metastable states at all time scales. This indicates that in presence of bilingual individuals, minority languages might survive within some communities for very long periods of time when the social network displays community structure.

We are currently studying, in a two-dimensional regular network, the case of competition between non-equivalent languages ($s \neq 0.5$), and the effect of perturbing the linear transition probabilities of the model. In general, when one of the languages has a higher prestige, and starting from random initial conditions, the system always evolves towards the extinction of the less prestigious language. This happens in any of the two models studied in this paper. We are also studying the perturbation of the linear transition probabilities of the model. On the one hand, when perturbing them in such a way that agents change his language with a probability larger than in the linear case we get coexistence for some range of the exponent we use as parameter (we call it *volatility*), even in the case of non-equivalent languages. We call this regime *high volatility* regime (i.e., agents easily change language use). On the other hand, when perturbing the model in such a way that agents change his language with a probability smaller than linear, we got flatter linguistic borders and slower growth of linguistic domains, and the times of extinction in both models increase significantly (*low volatility* regime; i.e., agents have larger *inertia* to change language use).

We are also currently analysing a model we have proposed, where language is taken as a property of the social interaction (*link*) instead of a feature of the agent (*node*), getting a new perspective from what is usually assumed in agent based models regarding language competition. A new interpretation of *where* we find language in the network, and the emergence of different degrees of bilingualism in the interfaces between monolingual domains, are some of the novelties that we learn from this new approach to language competition.

Acknowledgements

We acknowledge financial support of the European Commission through the NESTComplexity project PATRES (043268).

References

- Abrams, D. M., & Strogatz, S. H. (2003). Modelling the dynamics of language death. *Nature*, *424*, 900.
- Castelló, X., Eguíluz, V. M., & San Miguel, M. (2006). Ordering dynamics with two non-excluding options: Bilingualism in language competition. *New Journal of Physics*, *8*, 308–322.
- Castelló, X., Toivonen, R., Eguíluz, V. M., Saramäki, J., Kaski, K., & San Miguel, M. (2007). Anomalous lifetime distributions and topological traps in ordering dynamics. *Europhysics Letters*, *79*, 66006.
- Crystal, D. (2000). *Language death*. Cambridge: Cambridge University Press.
- De Bot, K., & Stoessel, S. (Eds.). (2002a). *International journal of the sociology of language* (Vol. 153). Berlin and New York: Mouton de Gruyter.
- De Bot, K., & Stoessel, S. (2002b). Introduction: Language change and social networks. *International Journal of the Sociology of Language*, *153*, 1–7.
- Gunton, J. D., San Miguel, M., & Sahni, P. (1983). Phase transitions and critical phenomena. In C. Domb & J. L. Lebowith (Eds.), (Vol. 8, pp. 269–446). London: Academic Press.
- Milroy, L. (1987). *Language and social networks*. Oxford and New York: (2nd ed) Basil Blackwell.
- Milroy, L. (2001). Theories on maintenance and loss of minority languages. In J. Klatter-Folmer & P. van Avermaet (Eds.), (pp. 39–64). Münster and New York: Waxmann.
- O’ Driscoll, J. (2001). A face model of language change. *Multilingua*, *20*, 245.
- San Miguel, M., Eguíluz, V., Toral, R., & Klemm, K. (2005). Binary and multivariate stochastic models of consensus formation. *Computer in Science and Engineering*, *7*, 67–73.
- Schulze, C., & Stauffer, D. (2006). Recent developments in computer simulations of language competition. *Computing in Science and Engineering*, *8*, 60–67.
- Stauffer, D., Castello, X., Eguiluz, V. M., & San Miguel, M. (2007). Microscopic abrams-strogatz model of language competition. *Physica A*, *374*, 835–842.
- Stauffer, D., & Schulze, C. (2005). Microscopic and macroscopic simulation of competition between languages. *Physics of Life Reviews*, *2*, 89–116.
- Toivonen, R., Onnela, J., Saramäki, J., Hyvönen, J., Kertész, J., & Kaski, K. (2006). *Physica A*, *371*(2).
- Tsunoda, T. (2005). *Language endangerment and language revitalisation*. Berlin and New York: Mouton de Gruyter.
- Wang, W. S.-Y., & Minett, J. W. (2005). The invasion of language: emergence, change and death. *Trends in Ecology and Evolution*, *20*, 263–269.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440–442.
- Weinreich, U. (1953). *Languages in contact*. New York: Linguistic Circle of NY.