9-1-2012

Front Speed of Language Replacement

Joaquim Fort
Complex Systems Laboratory and Departament de Física, Universitat de Girona, 17071 Girona, Catalonia, Spain

Joaquim Pérez-Losada
Complex Systems Laboratory and Departament de Física, Universitat de Girona, 17071 Girona, Catalonia, Spain

Front speed of language replacement

Joaquim Fort and Joaquim Pérez-Losada

Complex Systems Laboratory and Departament de Física,

Universitat de Girona, 17071 Girona, Catalonia, Spain.

Abstract

We use two coupled equations to analyze the space-time dynamics of two interacting languages. Firstly, we introduce a cohabitation model, which is more appropriate for human populations than classical (non-cohabitation) models. Secondly, using numerical simulations we find the front speed of a new language spreading into a region where another language was previously used. Thirdly, for a special case we derive an analytical formula that makes it possible to check the validity of our numerical simulations. Finally, as an example, we find that the observed front speed for the spread of the English language into Wales in the period 1961-1981 is consistent with the model predictions. We also find that the effects of linguistic parameters are much more important than those of parameters related to population dispersal and reproduction. If the initial population densities of both languages are similar, they have no effect on the front speed. We outline the potential of the new model to analyze relationships between language replacement and genetic replacement.

I. INTRODUCTION

The relationship between culture and genes is an important topic in several fields [1–3]. Language evolution can transform a single language into several new ones in a time scale of about 1000 years [2]. However, sometimes the language spoken in a given area changes in a much shorter time scale because it is replaced by another language. Renfrew suggested two main mechanisms of language replacement [1, 2]. The first one is due to demic expansions (range expansions of populations under demographic pressure). The second one is called elite dominance (the conquest by a small minority that takes control of institutions and imposes its language). Some examples of mainly demic expansions are the Neolithic transition in Europe [4], and the modern European invasions of North America and Australia. Some examples of language replacement via elite dominance are Hungary and Turkey during the Middle Ages [2]. Usually demic expansions replace both the language and the genes [2]. In contrast, elite dominance processes may replace the language or not but they always have a negligible effect on the genes [2]. For this reason, genetic data can be used to infer language replacement via elite dominance, as in Armenia and Azerbaijan [3]. Sometimes elite dominance is preceded by system collapse (the loss of control by the central authority) [1, 2]. Some examples of system collapse are the fall of the Roman Empire in Britain (after which an Anglo-Saxon minority acquired power) and possibly the fall of the Mayan civilization around the 10th century AD [1, 2]. A third mechanism, language acquisition from neighboring populations, has been noted by Cavalli-Sforza, who has observed it in African pygmies (sometimes with gene replacement and sometimes without it) [2]. This third mechanism is likely to be of importance also when language use maps display a gradual shrinkage of the area where a language is spoken, due to its replacement by another language [5, 6]. Of course, in such cases the first and second mechanisms can be also partially responsible for language replacement. In spite of the importance of such processes, we are not aware of quantitative models of the geographical spread rate of spreading languages. The present paper aims to present such a model. In fact, very interesting quantitative models on language interaction do exist [7–11], but they do not seem to deal with the problem of explaining the geographical spread rate of the expanding language. In contrast, the model in the present paper makes it possible to quantify the effect of different processes and parameters on the speed of language replacement.

Because of social factors, a given language can become more efficient than another one in recruiting new users and as a consequence it can even exclude the other one, which becomes extinct. This replacement can be seen as a consequence of competition, a widespread concept in ecological dynamics, which can be also applied to language dynamics. A simple, non-spatial model to explain language competition was proposed by Abrams and Strogatz (AS), and it was shown to be consistent with several observations of language decay [7]. But languages also display changes in space, which were not considered by the original AS model [7]. Patriarca and co-workers took into account the effect of space (by adding diffusion terms to the Abrams-Strogatz model) and analyzed the effect of initial conditions, population growth and geographical barriers on the dynamics of two initially segregated languages [8, 9]. More general models have been also considered (see Refs. [10] and [11] for some recent reviews).

The plan of this paper is as follows. As in previous work on non-linguistic fronts [4, 12], we consider a two-dimensional (2D) space so that we can apply our results to populations on the Earth surface. We first review the AS model. Then we introduce cohabitation models, and explain why they are more realistic than classical (non-cohabitation) models to describe the dynamics of human populations in space and time. Accordingly, we build a cohabitation model for two linguistic populations coupled by the AS model of language competition. We derive the corresponding speed of the language front using numerical simulations. We find that for a simple case (marginal volatility) the model can be solved analytically and agrees with the numerical simulations. The language status enters as a parameter in our model, and it is shown to play an important role in the dynamics of the geographical distribution of languages. The roles of other linguistic and demographic parameters are also analyzed. We find that if the initial population densities of both languages are similar, they have no effect on the front speed. Finally, for the historical decay of Welsh in the period 1960-1980, the predicted front speed (by our cohabitation AS model) is seen to be consistent with the observed speed (as measured from language use maps).

II. THEORETICAL MODEL

In this section we first review previous models. Then we introduce our model, its linguistic and demographic parameters, and the ranges of their values.

A. Abrams-Strogatz (AS) model

The model due to Abrams and Strogatz (AS) studies the competition between two languages. In the AS model, the rate of change of the number density of speakers of the newcomer language (N) is given by [7]

$$\frac{dp_N}{dt} = \kappa \left[s \, p_I \left(\frac{p_N}{p_N + p_I} \right)^\alpha - (1 - s)p_N \left(\frac{p_I}{p_N + p_I} \right)^\alpha \right], \tag{1}$$

and an analogous equation is written for the indigenous language (I),

$$\frac{dp_I}{dt} = \kappa \left[-s \, p_I \left(\frac{p_N}{p_N + p_I} \right)^\alpha + (1 - s)p_N \left(\frac{p_I}{p_N + p_I} \right)^\alpha \right], \tag{2}$$

where p_N and p_I stand for the population number density (number of speakers/km^2) of the newcomer and indigenous languages, respectively, t is the time (measured in yr) and κ is a constant (with units yr^{-1}) [13]. The first terms in Eqs. (1)-(2) correspond to the conversion of I-speakers to language N, and the last terms to the conversion of N-speakers to the language I. In this model, parameter s is called the status of language N, with $0 \leq s \leq 1$, whereas the status of language I is $1 - s$. We see from Eqs. (1)-(2) that the higher the status of a language, the more new speakers it will recruit per unit time. Thus, the status of a language reflects its attractiveness (by aggregating multiple factors affecting its perceived utility, social prestige, etc.). Note from Eqs. (1)-(2) that: (i) if $s > 0.5$, the status of language N is higher than that of I; (ii) if $s = 0.5$, both languages are socially equivalent (i.e., they have the same status); (iii) if $s < 0.5$, the status of I is higher than that of N. The parameter α is a measure of the degree of resistance of speakers to change their language. This can be seen most easily as follows. Obviously $\frac{p_N}{p_N + p_I} < 1$, thus if $\alpha > 1$ then the power $\left(\frac{p_N}{p_N + p_I} \right)^\alpha$ in Eqs. (1)-(2) will be smaller than if $\alpha < 1$ (for any given values of p_N and p_I). Therefore $\alpha > 1$ corresponds to a high resistance of speakers to language change (i.e., to a low language volatility), $\alpha < 1$ to a low resistance (high volatility) and $\alpha = 1$ is the so-called marginal volatility case [14]. Models arising from equations such as (1)-(2) have been favorably compared to the observed dynamics for several languages [7].

For the sake of clarity, it is worth to mention that: (i) Abrams and Strogatz [7] used population fractions (not population densities) in their original formulation. (ii) For the special case of a constant total population density $p_N + p_I$ (as in Ref. [7] but not in our case), it is equivalent to use population densities (as in Eqs. (1)-(2) above) or population

fractions (as in Ref. [7]), because dividing both sides of Eqs. (1)-(2) by $p_N + p_I$ then yields the AS formulation [7]. (iii) In the latter case (constant total population density $p_N + p_I$), the AS model for $\alpha = 1$ and $s = 1$ corresponds to the Lotka-Volterra interaction, widely used in Ecology [15]. (iv) In order to take into account population dispersal in space (as done in Refs. [8–10] and below), it is much simpler to use population densities rather than population fractions (this will become obvious from the equations we introduce below), and this is why we use population densities (rather than fractions) throughout this paper.

B. Single-population dispersal model

Most well-known models are based on the following equation for the change in population density [12]

$$p(x,y,t+T) - p(x,y,t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x+\triangle_x, y+\triangle_y, t)\phi(\triangle_x, \triangle_y) d\triangle_x d\triangle_y - p(x,y,t) + R[p(x,y,t)],$$

(3)

where $p(x,y,t)$ is the population density at the location (x,y) and time t, and T is the time interval between two successive dispersal events. The first term on the right-hand side is the population reaching the location (x,y) at time $t+T$ from location $(x+\triangle_x, y+\triangle_y)$ at time t, added up over all possible values of the displacement $(\triangle_x, \triangle_y)$. The second term in the right corresponds to people leaving an area centered at (x,y). The last term $R[p(x,y,t)]$ corresponds to net reproduction (births minus deaths).

More precisely, the dispersal kernel $\phi(\triangle_x, \triangle_y)$ is the probability per unit area that an individual who was at $(x+\triangle_x, y+\triangle_y, t)$ moves to $(x,y,t+T)$. In Appendix A we explain how the classical model due to Fisher (sometimes called the reaction-diffusion approach) can be obtained from Eq. (3) as a special case.

A major drawback of Eq. (3) is the following. According to Eq. (3), newborn individuals can appear at (x,y) (last term) while their parents migrate away from (x,y) (second term on the right-hand side). In other words, in such models some parents leave their newborn children alone. However, newborn humans cannot survive away from their parents. This inconsistency can be avoided using so-called cohabitation models, which are based on the following equation instead of (3) [16],

$$p(x, y, t + T) = R_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x + \triangle_x, y + \triangle_y, t) \, \phi(\triangle_x, \triangle_y) d\triangle_x \, d\triangle_y, \qquad (4)$$

where R_0 is the net fecundity or reproductive rate (this is a net rate, i.e. it includes the effect of mortality). In this model, the time interval T is equal to one generation (defined as the mean age difference between a parent and her/his children). Then, according to Eq. (4) children cannot appear away from their parents' location. For this reason, for human populations Eq. (4) is more reasonable than Eq. (3). This point has been discussed in detail elsewhere (see especially Fig. 1 in Ref. [17] and Sec. 8 in Ref. [18]).

Strictly, Eq. (4) is valid only at sufficiently low values of the population density $p(x, y, t)$, because biological populations cannot reproduce without bound. Thus, if for some value of $(x, y, t + T)$ the result for $p(x, y, t + T)$ computed from Eq. (4) is larger than $p_{\max}$, then it is simply replaced by $p_{\max}$ [18] (more complex models are sometimes used [17], but they are not necessary here).

C. Coupled-populations dispersal model

In this subsection we combine the cohabitation approach in the previous subsection with the AS linguistic model. Let $p_N(\overrightarrow{r}, t)$ stand for the population number density of the newcomers, and let $p_I(\overrightarrow{r}, t)$ stand for the population number density of the indigenous population per unit area centered at position $\overrightarrow{r} \equiv (x, y)$ and time t. We introduce into the cohabitation Eq. (4) the interaction (1)-(2) with $\frac{dp_i}{dt} \simeq \kappa \left[p_i(t + T) - p(t) \right]$ (κ is just a constant that sets the time scale [13], so we can choose $\kappa = 1/T$). This yields

$$p_N(\overrightarrow{r}, t + T) = R_{0N}$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[p_N(\overrightarrow{r} + \overrightarrow{\triangle}, t) + s \, p_I(\overrightarrow{r} + \overrightarrow{\triangle}, t) \left(\frac{p_N(\overrightarrow{r} + \overrightarrow{\triangle}, t)}{p_N(\overrightarrow{r} + \overrightarrow{\triangle}, t) + p_I(\overrightarrow{r} + \overrightarrow{\triangle}, t)} \right)^\alpha \right.$$

$$\left. - (1 - s) p_N(\overrightarrow{r} + \overrightarrow{\triangle}, t) \left(\frac{p_I(\overrightarrow{r} + \overrightarrow{\triangle}, t)}{p_N(\overrightarrow{r} + \overrightarrow{\triangle}, t) + p_I(\overrightarrow{r} + \overrightarrow{\triangle}, t)} \right)^\alpha \right] \phi_N(\triangle) d\triangle_x \, d\triangle_y, \quad (5)$$

$$p_I(\overrightarrow{r}, t+T) = R_{0I}$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[p_I(\overrightarrow{r} + \overrightarrow{\Delta}, t) - s\, p_I(\overrightarrow{r} + \overrightarrow{\Delta}, t) \left(\frac{p_N(\overrightarrow{r} + \overrightarrow{\Delta}, t)}{p_N(\overrightarrow{r} + \overrightarrow{\Delta}, t) + p_I(\overrightarrow{r} + \overrightarrow{\Delta}, t)} \right)^\alpha \right.$$

$$\left. + (1-s) p_N(\overrightarrow{r} + \overrightarrow{\Delta}, t) \left(\frac{p_I(\overrightarrow{r} + \overrightarrow{\Delta}, t)}{p_N(\overrightarrow{r} + \overrightarrow{\Delta}, t) + p_I(\overrightarrow{r} + \overrightarrow{\Delta}, t)} \right)^\alpha \right] \phi_I(\triangle) d\triangle_x \, d\triangle_y, \quad (6)$$

where $\phi_N(\triangle)$ and $\phi_I(\triangle)$ are the dispersal kernels of the newcomer and indigenous populations, respectively. For mathematical simplicity, in the present paper we will assume that during the time interval $T = 1$ generation each individual either moves a distance d or remains at rest (with d the same for all individuals). Let the probability to remain at rest be p_{eN} and p_{eI} for the N and I populations, respectively (p_e is called the persistence in demography). This simple mobility rule will make it possible to focus our attention on the effect of linguistic competition (for a thorough discussion of more complicated dispersal models, see Ref. [17]). The model parameters are T, d, p_{eN}, p_{eI}, the net reproductive rates R_{0i} ($i = N, I$), the social status s of language N ($0 \le s \le 1$), the resistance to language change α, and the carrying capacities $p_{\max i}$.

D. Parameter values

For the sake of definiteness, let us consider a specific example, namely the competition between English and Welsh. In agreement with the definition used in language use statistics and maps [5], we define N as the population able to speak only English and I as that able to speak both English and Welsh. Note that the model above has a time step $T = 1$ generation, so for example the linguistic conversion $I \rightarrow N$ quantified by Eqs. (1)-(2) corresponds to some individuals being I and their children being N (not to a given $I-$individual becoming N). The observed values of the parameters, as well as the ranges used in the simulations, are reported in Table 1. They have been obtained as follows. The status of English in Wales was estimated from empirical data by Abrams and Strogatz as $s = 0.6$ [7]. By fitting their model to data for several other endangered languages, they also estimated the ranges $0.26 < s < 0.74$ and $1.0 < \alpha < 1.6$ [7]. In order to calculate the observed speed of English into Wales (see below), the only source of information we are aware of is Ref. [5], which displays maps but only for the period from year 1961 (Fig. 4.1 in [5]) to 1981 (Fig. 4.3 in [5]).

The cohabitation time T (sometimes called generation time) is at least about 20 yr [19, 21], so we note that the 20-yr time span during which the front speed can be measured from the maps (1961-1981) is similar in magnitude to one step in the model. Of course, ideally we would like to have additional maps, making it possible to estimate the front speed during a longer time interval. Nevertheless, we think that our model can be applied for several reasons: (i) It is reasonable to expect that the language front of the English language began before 1961 and continued after 1981 (we have checked the latter case using a 2001 map [22], but it cannot be used to compute speeds because it plots limits of different percentages of Welsh speakers than the 1981 and 1961 maps in Ref. [5]); (ii) In the simulations we have observed that a constant front speed is attained after the first simulation step (details on this issue are given in the next section); (iii) Data for longer time intervals can become available in the future (and allow to determine whether the speed of this front is constant or not) but even in such an instance, we think that our model can be useful as a first attempt to quantify and model the front speed.

Accordingly to the fact that the front speed can be measured in the period from year 1961 to 1981, we have used the population numbers of Wales [20] at a time interval including this period to estimate the ratio between the numbers of final and initial populations during a interval of $T = 1$ generation. This yields the range $R_{0i} = 1.057 - 1.081$ in Table 1 (in the calculation of this range, we have also taken into account the range of T in Table 1). Note that, e.g., the reproductive parameter R_{0N} should not be estimated from the final and initial numbers of speakers of language N, because such a calculation would also include the effect of language shift but R_{0N} in the model corresponds to the effect of population growth only (deaths and births). We have also used census data [20] and the area of Wales to estimate the carrying capacities of the population densities, $p_{\max i}$ (Table 1). Values for the persistence p_{ei} and mobility d are rather difficult to estimate for a one-generation time interval, but there are some useful data for modern populations [17, 23]. The value of d has been calculated from the kernel (i.e., the distance distribution of displacements) in Ref. [17] excluding those of zero distance (because the latter are taken into account in the value of the persistency p_{ei}). The minimum and maximum values for p_{ei} in Table 1 were obtained from the two subsamples (females and males, respectively) in Ref. [23], and the minimum and maximum values for d were estimated by computing the standard error of the mean for

d from the kernel in Ref. [17].

Parameter (units)	Characteristic value	Minimum	Maximum	Refs.
s (dimensionless)	0.6	0.26	0.74	[7]
α (dimensionless)	1.31	1.0	1.6	[7]
T (yr)	25	20	37	[19, 21]
R_{0i} (dimensionless) $(i = N, I)$	1.064	1.057	1.081	[5, 20]
$p_{\max i}$ (people/km^2) $(i = N, I)$	50	20	100	[20]
p_{ei} (dimensionless) $(i = N, I)$	0.77	0.72	0.82	[17, 23]
d (km)	35	32	38	[17, 23]

Table 1. Model parameters and their ranges.

III.   FRONT SPEED CALCULATION

It is very difficult to find the speed of front solutions to Eqs. (5) - (6) analytically in general. Therefore we have integrated them numerically by considering a 2D grid with 1000·1000 nodes. Initially the indigenous population (I) is restricted to a central area of the grid (representing the surface of Wales), where $p_I(x, y, 0) = p_{\max I}$, and $p_I(x, y, 0) = 0$ elsewhere (the front speed does not depend on the size of the central area considered). For the newcomer population (N), initially $p_N(x, y, 0) = p_{\max N}$ everywhere except at the central area (which is occupied by the $I-$population, thus $p_N(x, y, 0) = 0$ at the central area). At each time step (corresponding to $T = 1$ year), we compute the new population number densities $p_N(x, y, t+T)$ and $p_I(x, y, t+T)$ at all nodes of the 2D lattice in a 3-step process: linguistic interaction, population dispersal and population growth (the latter includes reproduction and deaths). First, the interaction term is calculated using the AS model (1)-(2) at every node (x, y). In the dispersal step, as already mentioned above, for simplicity we used the following procedure. A fraction p_{ei} of each the population ($i = N,I$) stays at the original node, and the remaining fraction is distributed equally among the nearest neighbors, i.e., a fraction $(1 - p_{ei})/4$ jumps a distance $\pm d$ along each horizontal or vertical direction. Finally, we compute the new $i-$population density($i = N, I$) due to reproduction at every node by multiplying $p_i(x, y, t)$ (obtained from the previous step) by the factor R_{0i}, unless a value $p_i > p_{\max i}$ is obtained; in such a case we set $p_i = p_{\max i}$ (to avoid biologically unrealistic

population densities over the saturation value implied by the environment). This 3-step cycle can then be repeated, each cycle corresponding to 1 generation. At the end of each cycle (or time step), the profile of the number density of speakers of language I (or N) can be plotted, and by comparing such plots we find the speed moved by the front. In this way, we have obtained the simulation front speeds in Fig. 1. In the simulations we observed that the speed between generation 1 and generation 2 was already constant (i.e., the same as between generations 2 and 3, 3 and 4, etc.). We also observed that for realistic parameter values (Table 1), if the newcomer language N has a higher status (i.e., if $s > 0.5$), behind the $N-$front the indigenous language I eventually disappears and is replaced by language N (so in the final state $p_{total} = p_N + p_I = p_N$ everywhere), except if $\alpha = 1$ and $s = 0.5$. This exception is not surprising because in this very special case ($\alpha = 1$ and $s = 0.5$) the two terms on the right-hand sides of Eqs. (1)-(2) cancel out, so there is no conversion of speakers from one language into another (i.e., the very special model $\alpha = 1$ and $s = 0.5$ is not useful to describe interacting languages).

In order to check our simulations, we will take advantage of the fact that, for the so-called marginal volatility case ($\alpha = 1$) the front speed for the set of cohabitation Eqs. (5)-(6) can be calculated analytically. The final result is (see Appendix B)

$$c = \frac{1}{T} \min_{\lambda > 0} \frac{\ln \left[R_{0N} 2s \left(p_{eN} + \frac{1 - p_{eN}}{2} \left[\cosh(\lambda d) + 1 \right] \right) \right]}{\lambda}. \tag{7}$$

We note that the language front speed decreases with increasing values of the cohabitation phase T (as expected intuitively). This also happens in the simulations for all values of α, simply because any simulation run yields a speed in km/gen (with 1 generation corresponding to 1 simulation step), and we transform these units into km/yr by dividing by the value of T measured in years (for completeness, Appendix C shows that in the reaction-diffusion approximation the speed also diminishes with increasing values of T).

Equation (7) has no analytical solution, but for given values of R_{0N}, p_{eN}, α, d, T and s it is easy to find its minimum numerically. In this way we have obtained the theoretical results for $\alpha = 1$ (line in Fig. 1).

IV.   RESULTS

In Fig. 1, there is good agreement between the analytical results from Eq. (7) (line $\alpha = 1$) and the corresponding simulations (circles). This makes us feel confident on the validity of our numerical simulations. Note that negative values of the speed in Fig. 1 simply indicate that the front propagates in the opposite direction, as expected because under the transformation $s \rightarrow 1 - s$ the dominant and dominated languages exchange their roles (see Eqs. (1)-(2)). This is also the reason why the curves in Fig. 1 are symmetric: all symbols in Fig. 1 have been obtained from simulations, which make it possible to determine the sign of the speed (in contrast, Eq. (7) is valid only for $c > 0$ (see Appendix C), so the line in Fig. 1 for $c < 0$ or $s < 0.5$ has been obtained simply changing the sign of the results for $s > 0.5$). In order to describe the specific case of English and Welsh it is enough to consider only values $s > 0.5$. Then, Fig. 1 shows that, for a given value of s, the speed becomes smaller if larger values of α (the resistance to language change) are considered, which is reasonable intuitively. On the other hand, for a given value of α, if the value of the relative status of the newcomer language s is higher, its invasion speed is faster, which is also as expected. In Fig. 1, the other model parameters are set to their characteristic values (Table 1). Figure 1 shows clearly that the value of the status s is important to predict the range expansion speed. In Fig. 1, we also see that for the status value obtained from the decay of the Welsh language in Ref. [7] ($s = 0.6$) and values of α in the range estimated by Abrams and Strogatz from endangered language data ($1.0 < \alpha < 1.6$), the predicted speed is in the range 0.1-0.4 km/yr. This result is consistent with the observed speed range (0.3-0.6 km/yr, shaded area in Fig. 1), which we estimated from maps of the Welsh language distribution in years 1961 and 1981 (maps in Figs. 4.1 and 4.3 in Ref. [5] display two clear fronts in Northern Wales, travelling about 7 km and 12 km during 20 yr in opposite directions, and shrinking the area occupied by the Welsh language). We think that this agreement between the predicted and the observed ranges (Fig. 1) is satisfactory, given the simplicity of our model and the fact the observed speed is not homogeneous due to cities, mountains, transport networks, etc. (for this reason, the observed speed range above (0.3-0.6 km/yr) is only a rough approximation to the average speed). Our model is rather simple in the sense that the roles of large cities, education policies, mass media, etc. [11, 24] are not explicitly included. For these reasons, we would like to stress our results should be

11

regarded only as initial step towards modelling a rather complex phenomenon, and we think that a detailed comparison to observations should await the development of more realistic (and complicated) models.

Figure 2 is a sensitivity analysis of the speed for the ranges of the model parameters. For each bar, a single parameter is varied, whereas the remaining parameters are kept at their characteristic values (Table 1). Besides the status s (Fig. 1), the resistance to language shift α is clearly the most important parameter (Fig. 2). This suggests the need to perform careful estimations of s and α when dealing with specific cases of language spread. According to Fig. 2, the length of the cohabitation phase T is of secondary importance, and the rest of parameters have a very small effect. Therefore, language choices (parameters s and α) seem to play the main role on the speed of language spread, and the length of the cohabitation phase (T) has a small relevance. Finally, population dispersal (parameters d, p_{eN} and p_{eI}) and net reproduction (R_{0N} and R_{0I}) seem to have little importance (Fig. 2).

Up to this point, all results (Figs. 1 and 2) have been obtained for the characteristic values of the maximum population densities (also called carrying capacities or saturation densities) in Table 1, namely $p_{\max N} = p_{\max I} = 50$ people/km^2 (which we estimated from census data). We have also performed simulations by varying the value of $p_{\max N} = p_{\max I}$ from 1 to 100 people/km^2. The results are shown in Fig. 3. For the marginal volatility case ($\alpha = 1$) the speed does not depend on the carrying capacity, both according to the simulations (rhombuses in Fig. 3) and to Eq. (7) (line in Fig. 3). Interestingly, the same happens for other values of α (recall that for $\alpha \neq 1$ we have no analytical formula but use the simulations, leading to the stars and circles in Fig. 3). Therefore, *as long as the population densities of the areas initially occupied by both languages are similar*, they have no effect on the language front speed (Fig. 3). We have performed additional simulations with a fixed value of $p_{\max I}$ and varying values of $p_{\max N}$ and vice versa, but we do not plot them here because we think that on average there were no substantial differences between both population densities (we reached this conclusion by computing population densities from census data [25] for some counties in Northern Wales where we measured the front speed (e.g., Denbighshire) and nearby English counties (e.g., Shropshire) in the period for which language maps are available (1961-1981)). This point could be interesting, however, in future work analyzing the local features of several examples of language substitution. Therefore, we summarize briefly those results in turn. For a fixed value of $p_{\max I}$, the speed increases

with $p_{\max N}$ (except if $\alpha = 1$, because in this case the speed is independent of $p_{\max N}$, see Eq. (7)). We think that this dependency is reasonable because larger values of $p_{\max N}$ correspond to more speakers that can spread the high-status language (either via cultural transmission, dispersal or reproduction). However, this effect is small (compared to those of s and α in Figs. 1 and 2): for example, the speed is 0.084 km/yr for $p_{\max N} = 25$ people/km^2 and 0.17 km/yr for $p_{\max N} = 100$ people/km^2 (with $p_{\max I} = 50$ people/km^2 in both cases, and all other parameters set at their characteristic values in Table 1). On the other hand, for a fixed value of $p_{\max N}$, the speed decreases with increasing values of $p_{\max I}$ (again except if $\alpha = 1$). This is also reasonable because for larger values of $p_{\max I}$, there are initially more low-status ($I-$language) speakers so their language I will be more difficult to replace. We stress that such dependencies are not observed in the special case $\alpha = 1$. This may indicate that the general AS model ($\alpha \neq 1$) is more realistic that the linear or marginal-volatility model $\alpha = 1$ (which includes the classical Lotka-Volterra model as a special case, see the text above Eq. (3)). However, let us stress that if the carrying capacities of the high-status and low-status languages are similar, the speed is the same regardless of the value of $p_{\max N} \simeq p_{\max I}$ (Fig. 3).

V. CONCLUDING REMARKS

Abrams-Strogatz (AS) verified their model by fitting its predictions to several endangered languages [7]. Here we have generalized the AS model by taking into account: (i) population dispersal in space; (ii) population reproduction (births and deaths); and (iii) the cohabitation effect (i.e., the fact that newborn humans cannot survive without their parents).

We have applied our model to predict language substitution front speeds for two initially segregated languages, using realistic parameter values from observed data (no free or adjustable parameters have been used). According to our results, linguistic parameters seem to have more important effects on the spread rate than reproductive and dispersal parameters.

We have shown that a simple space-time model can reasonably account for the observed spatial speed of substitution of the Welsh language. For this case we have also noted that the average initial population densities of the two linguistic populations involved were rather similar. We have found that, in such an instance, the population density does not affect the speed of language replacement (for completeness, we have also discussed how cases with

substantially different initial population densities would affect the front speed).

Our model could be applied in the future to other examples of language replacement [6, 26], as well as to non-homogeneous geographies. It could be also generalized to take into account the effects of large cities, mobility networks, policies directed to avoid the extinction of endangered languages, etc. [11, 24]. For a given example of language replacement, our model can be also applied to specific regions. It will lead to faster language substitution speeds in those regions with higher values of the relative status of the dominating language (s in the model), lower values of the initial population density of speakers of the low-status language ($p_{\max I}$), etc.

It is interesting to note that the speed of substitution of the Welsh language (about 0.3-0.6 km/yr) is somewhat slower, but similar in magnitude, to the speed of the Neolithic transition in Europe (about 0.9-1.3 km/yr [4]). The three mechanisms at work (population dispersal, net reproduction and cultural transmission) are the same in both cases, but the cultural transmission terms are rather different: for language competition we have used the AS interaction (Eqs. (1)-(2)), as motivated by decay data for several endangered languages [7]. In contrast, for the transmission of agricultural techniques from farmers to hunter-gatherers we used an interaction derived from cultural transmission theory [27] (see Eq. (1) in Ref. [4]). The fact that the speed obtained in two such different systems is similar in magnitude seems interesting. Future work could analyze the front speed problem in several additional anthropological phenomena (if the necessary data are available), and compare the results to the two speeds quoted above. Presumably, the speed will be faster in cases for which the cultural trait that spreads is easy to imitate (simple techniques, fashion trends, etc.) but similar to the two cases mentioned for cultural traits that cannot be copied without a substantial learning effort (farming, language, etc.).

The relationship between culture and genes is a relevant issue in several disciplines [1–3]. Therefore, in future work it would be of interest to combine our spatial model of language replacement with spatial models of genetic replacement [28–30] in order to relate the degree of genetic replacement with that of language replacement as a function of different processes and parameter values.

14

VI. ACKNOWLEDGMENTS

VII. APPENDIX A. REACTION-DIFFUSION MODEL FOR A SINGLE POPULATION

If ϕ depends only on $\Delta \equiv \sqrt{\triangle_x^2 + \triangle_y^2}$ (isotropic kernel), Eq. (3) can be Taylor-expanded up to second order in space and first order in time to yield [12]

$$\frac{\partial p}{\partial t} = D\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right) + F(p), \tag{8}$$

where

$$D \equiv \frac{1}{4T} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi(\triangle_x, \triangle_y)\Delta^2 d\triangle_x \, d\triangle_y, \tag{9}$$

is the diffusion coefficient and $F(p) \equiv \partial R/\partial t$ is called the population growth function.

Equation (8) is called Fisher's equation [31, 32] if a logistic growth function is assumed,

$$F(p) = ap\left(1 - \frac{p}{p_{\max}}\right), \tag{10}$$

where a is called the initial growth rate and $p_{\max}$ the carrying capacity. Then front solutions to Eq. (8) travel with Fisher's speed, $c = 2\sqrt{aD}$ [31, 32]. Note that net reproduction vanishes $[F(p) = 0]$ if and only if $p = 0$ or $p = p_{\max}$.

VIII. APPENDIX B. LINEAR ANALYSIS ($\alpha = 1$)

Here we show that the front speed for the set of cohabitation Eqs. (5)-(6) can be calculated analytically for the so-called marginal volatility case ($\alpha = 1$). This case is very useful to check the validity of the numerical simulations (Figs. 1 and 3). As in Ref. [33], we assume that the invasion front of population N spreads in a region where the density of the indigenous species I is initially equal to its maximum possible value, $p_{\max I}$. Thus, in the leading edge of the invasion front we may write

$$p_N(\overrightarrow{r}, t) \simeq \varepsilon(\overrightarrow{r}, t) + O(2), \tag{11}$$

$$p_I(\overrightarrow{r}, t) \simeq p_{\max I} - \delta(\overrightarrow{r}, t) + O(2), \tag{12}$$

15

where $O(2)$ stands for second and higher-order terms,

$$\varepsilon(\overrightarrow{r},t) \ll p_{\max N}, \tag{13}$$

and

$$\delta(\overrightarrow{r},t) \ll p_{\max I}. \tag{14}$$

Therefore, for $\alpha = 1$ up to first order we have for the linguistic interaction term in Eqs. (5)-(6)

$$s\, p_I(\overrightarrow{r},t)\left(\frac{p_N(\overrightarrow{r},t)}{p_N(\overrightarrow{r},t)+p_I(\overrightarrow{r},t)}\right) - (1-s)p_N(\overrightarrow{r}+\overrightarrow{\Delta},t)\left(\frac{p_I(\overrightarrow{r},t)}{p_N(\overrightarrow{r},t)+p_I(\overrightarrow{r},t)}\right)$$

$$= (2s-1)\,p_I(\overrightarrow{r},t)\left(\frac{p_N(\overrightarrow{r},t)}{p_N(\overrightarrow{r},t)+p_I(\overrightarrow{r},t)}\right)$$

$$= (2s-1)\,(p_{\max I} - \delta(\overrightarrow{r},t) + O(2))\,(\varepsilon(\overrightarrow{r},t)+O(2))\left(\frac{1}{p_{\max I}}+O(1)\right)$$

$$\simeq (2s-1)\,p_N(\overrightarrow{r},t) + O(2)$$

This result is very useful here because it reduces Eq. (5) to an evolution equation in which the variable $p_I(\overrightarrow{r},t)$ no longer appears,

$$p_N(x,y,t+T) \simeq R_{0N}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} 2s\, p_N(x+\triangle_x, y+\triangle_y,t)\phi_N(\triangle)d\triangle_x\, d\triangle_y. \tag{15}$$

In agreement with the numerical simulations explained above, we simply assume that an individual will either remain at rest (with probability p_{eN}) or will move a distance d (with probability $1 - p_{eN}$) by jumping into point (x,y) from points $(x \pm d, y)$ or $(x, y \pm d)$. Then, in discrete space Eq. (15) is replaced by

$$p_N(x,y,t+T) = R_{0N}2s\,\{p_{eN}\,p_N(x,y,t)$$
$$+(1-p_{eN})\left[\tfrac{1}{4}p_N(x-d,y,t)+\tfrac{1}{4}p_N(x+d,y,t)\right.$$
$$\left.+\tfrac{1}{4}p_N(x,y-d,t)+\tfrac{1}{4}p_N(x,y+d,t)\right]\}. \tag{16}$$

In order to find the front speed, the simplest approach is the following. For each value of time t, in general the front profile $p_N(x,y,t)$ depends on both x and y (this is easily seen by considering, e.g., the case of a circular front). However, let us choose the $x-$axis parallel to the local velocity of the front [34]. Then, it is reasonable to assume that for $t \to \infty$ the curvature will become negligible and the front profile p_N will depend only on x (within this

local region). Let $c \equiv |c_x|$ stand for the front speed. In the local frame just introduced, $c_y = 0$ and the front profile $p_N(z)$ depends only on $z = x - ct$. We look for constant-shape solutions with the form $p_N = p_0 \exp[-\lambda z]$ as $z \to \infty$ (with $\lambda > 0$ for $c > 0$). Using this into Eq. (16) and assuming that the minimum speed is the one of the front [34] we obtain the speed given by Eq. (7).

IX.   APPENDIX C. REACTION-DIFFUSION MODEL FOR COUPLED POPULATIONS

In the case of two coupled populations, the model in Appendix A is easily generalized into (see the text above Eq. (5))

$$
\begin{cases}
\frac{\partial p_N}{\partial t} = D \left(\frac{\partial^2 p_N}{\partial x^2} + \frac{\partial^2 p_N}{\partial y^2} \right) + F(p_N) + \frac{1}{T} \left[s \, p_I \left(\frac{p_N}{p_N+p_I} \right)^\alpha - (1-s)p_N \left(\frac{p_I}{p_N+p_I} \right)^\alpha \right] \\
\frac{\partial p_I}{\partial t} = D \left(\frac{\partial^2 p_I}{\partial x^2} + \frac{\partial^2 p_I}{\partial y^2} \right) + F(p_I) + \frac{1}{T} \left[-s \, p_I \left(\frac{p_N}{p_N+p_I} \right)^\alpha + (1-s)p_N \left(\frac{p_I}{p_N+p_I} \right)^\alpha \right],
\end{cases}
\tag{17}
$$

and applying linear analysis for $\alpha = 1$ [18, 34] (as in Appendix B above), the front speed is found to be $c = 2\sqrt{\left(a + \frac{2s-1}{T}\right)D}$. We note that this front speed decreases with increasing values of the time T between two successive dispersal events (as expected intuitively), although in a different way that in the model in the main paper (see the text below Eq. (7)). This is not surprising, because the model in this appendix neglects the effects of the cohabitation phase and the dispersal kernel. In fact, even the interpretation of T is different: here T is not the length of a cohabitation phase (as in the model in the main paper) but just the time interval between two subsequent dispersal events (compare the text below Eq. (3) to that below Eq. (8)).

If there is no language interaction ($s = 0.5$, see the text above Eq. (7)) we recover Fisher's speed, namely $c = 2\sqrt{aD}$ [31, 32] (see Appendix A).

---

[1] C. Renfrew, *Archaeology and language: The puzzle of Indo-European origins* (Jonathan Cape, London, 1987).

[2] L. L. Cavalli-Sforza, P. Menozzi and A. Piazza, *The History and Geography of Human Genes* (Princeton University Press, Princeton, 1994), pp. 102- 103.

[3] I. Nasidze, T. Sarkisian, A. Kerimov and M. Stoneking, Hum. Genet. **112**, 255 (2003).

[4] J. Fort, Proc. Natl. Acad. Sci. USA **109**, 18669 (2012).

[5] J. Aitchison and H. Carter, *A geography of the Welsh language* (University of Wales Press, Cardiff, 1994).

[6] C. W. J. Withers, *Gaelic in Scotland 1698-1981. The geographical history of a language* (John Donald Publ., Edinburgh, 1984).

[7] D. M. Abrams and S. H. Strogatz, Nature **424**, 900 (2003).

[8] M. Patriarca and T. Leppänen, Physica A **338**, 296 (2004).

[9] M. Patriarca and E. Heinsalu, Physica A **388,** 174 (2009).

[10] R. V. Solé, B. Corominas-Murtra and J. Fortuny, J. R. Soc. Interface **7**, 1647 (2010).

[11] A. Kandler, Human Biology **81**, 181 (2009).

[12] J. Fort and V. Méndez, Phys. Rev. Lett. **82**, 867 (1999).

[13] D. M. Abrams, H. A. Yaple and R. J. Wiener, Phys. Rev. Lett. **107**, 088701 (2011).

[14] L. Chapel, X. Castelló, C. Bernard, G. Deffuant, V. M. Eguýluz , et al. PLoS ONE **5**, e8681 (2010).

[15] J. Murray, *Mathematical Biology* (Springer, Berlin, 2002).

[16] J. Fort, J. Pérez-Losada and N. Isern, Phys Rev E **76,** 031913 (2007).

[17] N. Isern, J. Fort and J. Pérez-Losada, J. Stat. Mechs. **2008**, P10012 (2008).

[18] J. Fort and T. Pujol, Rep Prog Phys **71,** 086001 (2008).

[19] B. Bogin, *The growth of humanity* (Wiley-Liss, New York), Fig. 2.5.

[20] *200 years of the Census in Wales* (Office for National Statistics, London, 2001), available at www.ons.gov.uk/

[21] J. Fort, D. Jana and J. Humet, Phys. Rev. E **70**, 031913 (2004), specially note [24] and the text above Eq. (26).

[22] H. M. Jones, *A statistical overview of the Welsh language* (Welsh Language Board, Cardiff, 2011).

[23] L. L. Cavalli-Sforza and W. F. Bodmer, *The Genetics of Human populations* (Dover, New York, 1999), table 8.7.

[24] C. Fernando, R.-L. Valijärvi and R. A. Goldstein, Human Biology **82**, 47 (2010).

[25] http://en.wikipedia.org/

[26] J. A. Fishman, *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages* (Mulitlingual matters, Cevedon, 1991).

18

[27] L. L. Cavalli-Sforza and M. W. Feldman, *Cultural transmission and evolution: a quantitative approach* (Princeton University Press, Princeton, 1981).

[28] S. Rendine, A. Piazza and L. L. Cavalli-Sforza, Am. Nat. **128**, 681 (1986).

[29] G. Barbujani, R. R. Sokal and N. L. Oden, Am. J. Phys. Anthropol. **96**, 109 (1995).

[30] M. Currat and L. Excoffier, Proc. Roy. Soc. B **272**, 679 (2005).

[31] R. Fisher, Ann. Eugenics **7**, 353 (1937).

[32] A. Kolmogorov, J. Petrovsky, and N. Piscounov, Moscow Univ. Bull. **1**, 1 (1937).

[33] V. Méndez, J. Fort and J. Farjas, Phys. Rev. E **60**, 5231 (1999), Sec. V.

[34] J. Fort and V. Méndez, Rep. Progr. Phys. **65**, 895 (2002).

FIGURE CAPTIONS

Fig. 1 Speed of linguistic front propagation as a function of the status parameter s, for three values of the resistance to language change α, according to our numerical simulations (symbols). For $\alpha = 1$ (linear or marginal volatility model) the front speed can be calculated using Eq. (7) and is also shown (line). It agrees with the corresponding simulations. The hatched area corresponds to the observed front speed of the English language into Wales (0.3-0.6 km/yr). The rest of parameters have been set to their characteristic values in Table 1.

Fig. 2 One-at-a-time sensitivity analysis for the model parameters. For each parameter indicated on the horizontal axis, we have used its range in the upper right corner and the characteristic values of the other parameters in Table 1. Besides the status (Fig. 1), the resistance to language change α has the most important effect on the speed of linguistic fronts.

Fig. 3 Speed of linguistic front propagation as a function of the population carrying capacity, for three values of the resistance to language change α, according to our numerical simulations (symbols). For $\alpha = 1$ (linear or marginal volatility model), the analytical front speed from Eq. (7) is also shown (line) and it agrees with the corresponding simulations (rhombuses). For all three values of α it is seen that the speed is independent of the carrying capacities of both populations provided that they are similar (as appropriate on average for the English and Welsh languages). The results for cases with $p_{\max N} \neq p_{\max I}$ are discussed in the main text. We have used the characteristic values of the other parameters in Table 1.

$1.057 < R_{0N} < 1.081$

$1.0 \quad < \alpha \quad < 1.6$

$0.72 < p_{eN} < 0.82$

$32 < d < 38$ km

$1.057 < R_{0I} < 1.081$

$0.72 < p_{eI} < 0.82$

$20 \quad < T \quad < 37$ yr

Fort & Pérez-Losada, Fig. 3