Diversity, competition, extinction: the ecophysiology of language change

Ricard V. Solé1,2,3,*, Bernat Corominas-Murtra1 and Jordi Fortuny4

1ICREA-Complex Systems Laboratory, Parc de Recerca Biomèdica de Barcelona, Universitat Pompeu Fabra, Dr Aiguader 88, 08003 Barcelona, Spain
2Santa Fe Institute, 1399 Hyde Park Road, NM 87501, USA
3Institut de Biologia Evolutiva, CSIC-UPF, Passeig Marítim de la Barceloneta, 37-49, 08003 Barcelona, Spain
4Centre de Lingüística Teòrica, Facultat de Lletres, Edifici B, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

As indicated early by Charles Darwin, languages behave and change very much like living species. They display high diversity, differentiate in space and time, emerge and disappear. A large body of literature has explored the role of information exchanges and communicative constraints in groups of agents under selective scenarios. These models have been very helpful in providing a rationale on how complex forms of communication emerge under evolutionary pressures. However, other patterns of large-scale organization can be described using mathematical methods ignoring communicative traits. These approaches consider shorter time scales and have been developed by exploiting both theoretical ecology and statistical physics methods. The models are reviewed here and include extinction, invasion, origination, spatial organization, coexistence and diversity as key concepts and are very simple in their defining rules. Such simplicity is used in order to catch the most fundamental laws of organization and those universal ingredients responsible for qualitative traits. The similarities between observed and predicted patterns indicate that an ecological theory of language is emerging, supporting (on a quantitative basis) its ecological nature, although key differences are also present. Here, we critically review some recent advances and outline their implications and limitations as well as highlight problems for future research.

Keywords: language dynamics; extinction; diversity; competition; phase transitions

1. INTRODUCTION

Languages and species share some remarkable commonalities. Such similarities did not escape the attention of Charles Darwin, who mentioned them a number of times in writings and letters (see Whitfield 2008). In The Descent of Man, Darwin explicitly says:

The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel. (Darwin 1871)

Languages indeed behave as some kind of living species (Mufwene 2001; Pagel 2009). They exhibit a large diversity: it is estimated that around 6000 different languages exist today in our modern world (Krauss 1992; McWhorter 2001; Nettle & Romaine 2002). Languages and genes are known to be correlated at both global (Cavalli-Sforza et al. 1988; Cavalli-Sforza 2002) and local (see Lansing et al. (2007) and references therein) population scales. As is the case with biodiversity estimates too, the actual language diversity is unknown, and estimates fluctuate up to around 10 000 different spoken languages. Needless to say, another element to consider is the internal diversity displayed by languages themselves, where—like subspecies—dialects abound.

Languages also display geographical variation: as occurs with species, they become more and more different under the presence of physical barriers. They come to life, as species appear by speciation. They also become extinct, and language extinction has become a major problem for our cultural heritage: as for endangered species, many languages are also on the verge of disappearance (Crystal 2000; Dalby 2003; Sutherland 2003; Mufwene 2004). Languages die with their last speaker: Crystal mentions the example of Ole Stig Andersen, a researcher looking in 1992 for the last
speaker of the West Caucasian language Ubuh. In the words of Andersen:

(The Ubuh) . . . died at day break, October 8th 1992, when the last speaker, Tevfik Esenç, passed away. I happened to arrive in his village that very same day, without appointment, to interview the Last Speaker, only to learn that he had died just a couple of hours earlier.

Crystal (2000)

This story dramatically illustrates the last breath of any extinct language. It dies as soon as its last speaker dies (or stops using it). It is also interesting to observe that the extinction risk and its correlation with geographical distribution is shared by both species and languages (Sutherland 2003).

Language change involves both evolutionary and ecological time scales. Most theoretical studies deal with large-scale evolution: how languages emerge and become shaped by natural selection (Bickerton 1990; Hawkins & Gell-Mann 1992; Deacon 1997; Parisi 1997; Cangelosi & Parisi 1998; Nowak & Krakauer 1999; Pinker 2000; Cangelosi 2001; Hauser et al. 2002; Kirby 2002; Wray 2002; Brightont et al. 2005; Kosmidis et al. 2005, 2006; Baxter et al. 2006; Szathmary 2006; Nolfi & Mirolli 2010; Oudeyer & Kaplan 2007; Lipson 2007; Christiansen & Chater 2008; Chater et al. 2009; Nolfi & Mirolli 2010). But languages also display changes within the short time scale of one or a few human generations. Actually, a great deal of what will happen to languages in the future is deeply related to their ecological nature. Demographic growth, the dominant role of cities in social and economic organization and globalization dynamics will largely shape the world’s languages (Graddol 2003).

Languages evolve under centuries of accumulated modifications (this is well illustrated by written texts; see Howe et al. (2001) and Bennett et al. (2003)) and undergo evolutionary bursts (Atkinson et al. 2008). On short time scales they can be described in terms of ecological systems. These rapid modifications affect language diversity, their internal differentiation and even their survival. Different studies using the perspective of statistical physics (Nettle 1999a–c; Benedetto et al. 2002; Ke et al. 2002, 2008; Stauffer & Schulze 2005; Wang & Minett 2005; Loreto & Steels 2007; de Oliveira et al. 2008; Zanette 2008) have been able to cope with these phenomena, showing that the basic trends of language dynamics share remarkable similarities with the spatiotemporal behaviour of complex ecosystems.

We will consider different levels of language organization, from words to languages as abstract entities. The models reviewed here explore the conditions under which words or languages can survive or disappear. The time scale is ecological; therefore, we assume that in short time scales the dynamics of change does not affect the structure of language itself and thus evolutionary models are not considered. Moreover, we do not intend to quantitatively reproduce observed patterns, although the predictions of the models can be tested in many cases from real data. Instead, the models we revise try to capture the logic of the underlying processes in a qualitative fashion. These models follow the spirit of statistical physics in trying to reduce the system’s complexity to its bare bones. They provide a powerful approximation that allows us to see global patterns that might not depend on the intrinsic nature of the components involved. They also help in highlighting the differences. As will be discussed below, languages also exhibit marked departures from ecological traits.

This review critically examines a set of models of increasing complexity. Specifically, we review recent advances within the fields of statistical physics and theoretical ecology relative to a better understanding of language dynamics. We begin with a very simple model describing word propagation within a population. Next, the effects and consequences of competition among linguistic variants, with special attention to those scenarios leading to language extinction. This is expanded by considering alternative scenarios allowing language coexistence to occur, either through bilingualism or spatial and social segregation. Although spatial coexistence under local competition is shared with ecosystems, bilingualism belongs to a different class of phenomenon. All these models involve a small number of interacting languages. The final part of the review deals with language diversity in space and time. Both a simple model of multi-lingual communities and available data on scaling laws in language diversity are presented. Once again, striking similarities and strong differences are found. A synthesis of these ideas and open problems is presented at the end, together with a table comparing language and the ecosystem’s properties.

2. LEXICAL DIFFUSION

The potential set of words used by a speaker’s community is listed in dictionaries (Miller 1991). They capture a given time snapshot of the available vocabulary, but in reality speakers use only part of the possible words: many are technical and thus used only by a given group and many are seldom used. Many words are actually extinct, since no one is using them. On the other hand, it is also true that dictionaries do not include all words used by the community and also that new words are likely to be created constantly within populations and their origins have been sometimes recorded (Chantrell 2002). Many of them are new uses of previous words or recombinations and sometimes they come from technology. One of the challenges of current theories of language dynamics is understanding how words originate, change and spread within and between populations, eventually being fixed or extinct. In this context, the appearance of a new word has been compared with a mutation (Cavalli-Sforza & Feldman 1981).

As occurs with mutational events in standard population genetics, new words or sounds can disappear, randomly fluctuate or become fixed. In this context, the idea that words, grammatical constructions or sounds can spread through a given population was originally formulated by William Wang. This was proposed
in order to explain how \textit{lexical diffusion} (i.e. the spread across the lexicon) occurs (Wang 1969). Such a process requires the diffusion of the innovation from speaker to speaker (Wang & Minett 2005).

\textbf{2.1. Logistic spreading}

A very first modelling approximation to lexical diffusion in populations should account for the spread of words as a consequence of learning processes (Shen 1997; Wang \textit{et al.} 2004; Wang & Minett 2005). Such a model should be able to establish the conditions favouring word fixation. As a first approximation, let us assume that each item is incorporated independently (Shen 1997; Nowak \textit{et al.} 2000). If \( x_i \) indicates the fraction of the population knowing the word \( W_i \), the population dynamics of such a word reads

\[ \frac{dx_i}{dt} = R_i x_i(1 - x_i) - x_i, \]

with \( i = 1, \ldots, n \). The first term on the right-hand side of equation (2.1) introduces the way words are learned. The second deals with deaths of individuals at a fixed rate (here normalized to 1). The way words are learned involves a nonlinear term where the interactions between those individuals knowing \( W_i \) (a fraction \( x_i \)) and those ignoring it (a fraction \( 1 - x_i \)) are present. The parameter \( R_i \) introduces the rate at which learning takes place.

Two possible equilibrium points are allowed, obtained from \( \frac{dx_i}{dt} = 0 \). The first is \( x_i^* = 0 \) and the second is

\[ x_i^* = 1 - \frac{1}{R_i}. \]

The first corresponds to the extinction of \( W_i \) (or its inability to propagate) whereas the second involves a stable population knowing \( W_i \). The stability of these fixed points is determined by the sign of

\[ \lambda(x_i^*) = \left( \frac{\partial x_i}{\partial x_i} \right)_{x_i^*}. \]

If \( \lambda(x_*) < 0 \) the point is stable and will be unstable otherwise (Kaplan & Glass 1995; Strogatz 2001).

The larger the value of \( R_i \), the higher the number of individuals using the word. We can see that, for a word to be maintained in the population lexicon, we require the following inequality to be fulfilled:

\[ R_i > 1. \]

This means that there is a threshold in the rate of word propagation to sustain a stable population. By displaying the stable population \( x^* \) against \( R_i \) (figure 1a) we observe a well-defined phase transition phenomenon: a sharp change occurs at \( R_i^* = 1 \), the critical point separating the two possible phases. The subcritical phase \( R_i < 1 \) will inevitably lead to the loss of the word.

The dynamical pattern displayed by a successful propagating word follows a so-called S-shaped curve (see Niyogi (2006) and references therein concerning the gradualness and abruptness of linguistic change). This can be easily seen by integrating the previous model. Let us first note that the original equation (2.1) can be re-written as a logistic one, namely

\[ \frac{dx}{dt} = (R_i - 1)x_i \left( 1 - \frac{x_i}{x_i^*} \right). \]

which, for an initial condition \( x_i(0) \) at \( t = 0 \), gives a solution

\[ x_i(t) = \frac{x_i(0)e^{(R_i-1)t}}{1 + x_i(0)(e^{(R_i-1)t} - 1)}. \]
This curve is known to increase exponentially at low population values, describing a scenario where words rapidly propagate, followed by a slow down as the number of potential learners decays. The accelerated, exponential growth has been dubbed the snowball effect (Wang & Minett 2005) and such curves have been fitted to available data (Wang 1969). Therefore, a central property of linguistic change, namely its gradualness, can be derived as an epiphenomenon from the dynamical patterns of successful propagation in the case of lexical diffusion. A further issue would be to explore whether the gradualness of grammatical (phonological, morphological and syntactical) change can be derived from equations similar to those that model the diffusion of words. It must be noted, from a different perspective, that the logistic trajectory of linguistic change may be favoured by ‘the underlying dynamics of individual learners’, as argued by Niyogi (2006, p. 167).

The previous toy model of word dynamics within populations is an oversimplification, but it illustrates fairly well a key aspect of language dynamics, which is also observed in ecology (Solé & Bascompte 2006): thresholds exist and play a role (Nowak & Krakauer 1999). They remind us that, beyond the gradual nature of change that we perceive through our lives (mainly affecting the lexicon), sudden changes are also likely to occur. An important aspect not taken explicitly into account by the previous model is the process of word generation and modification. Words are originated within populations through different types of processes. They also become incorporated by invasion from foreign languages. Once again, the processes of word invasion and origination recapitulate somehow the mechanisms of change in biological populations.

### 2.2. Multi-dimensional diffusion

Several modifications and extensions of the previous model have been suggested (Wang et al. 2004). They include considering multiple words involved in the diffusion process. This scenario would take into account the idea that words interact among them in multiple ways, and their diffusion can be constrained or enhanced by these interactions (Wang & Minett 2005). The resulting model describes the dynamics of a given novelty $x_i$ and its previous form $y_j$ (these can correspond to two words or sounds). Assuming conservation of their relative abundances, i.e. $x_i + y_j = 1$, it is possible to show that a set of equations

$$
\frac{dx_i}{dt} = (1 - x_i) \left[ \alpha_{ii}x_i + \sum_{j \neq i} \alpha_{ij}y_j \right]
$$

with $i, j = 1, \ldots, N$, describes the lexical diffusion process. The matrix elements $\alpha_{ij}$ introduce the coupling rate between a pair $(i, j)$ of words. It is interpreted as the rate at which adoption of the new word $i$ is induced by the frequency of other novel forms of word $j$. As it is formulated, the stable states are all given by $x_i = 1$ and thus (not surprisingly) there is no place for extinction of the novelty, although there exists some evidence for such a scenario, where new items spread initially but eventually decay (Ogura 1993). An interesting extension of this problem could take into account both positive and negative interactions. In this way, not only facilitation (as given by the positive interactions) but also competition would be considered. In other words, it seems reasonable to think that some words should be incompatible with others. This actually matches the problem of species invasion and assembly in multi-species communities (Levins 1968; Case 1990, 1991; Solé et al. 2002). For an exotic species invading a given community to succeed, some community-level constraints need to be satisfied. It would be interesting to see whether similar rules apply to the ups and downs of word spreading.

As in the previous subsection, it seems fair to us to pose the question of whether or not grammatical change can be modelled using equations similar to those explored in the study of lexical diffusion. As to multi-dimensional diffusion, it may be worth considering in future research whether the diffusion of a grammatical object such as a morphological paradigm or a syntactic structure can be described with an equation analogous to equation (2.7). It is also worth noting the existence of implicational universals (Greenberg 1963), which have the shape given a grammatical property $z$ in a language $L$, we always find a property $y$ in $L$, as well as the cross-linguistic observation that certain properties tend to entail other properties with overwhelmingly greater than chance frequency, to put it in Greenberg’s famous words. That is, cross-linguistic grammatical change cannot be perfectly mapped into a pure diffusion process: certain properties entail or tend to entail the presence or absence of certain properties, as some words may favour or ban the existence of others.

### 2.3. Naming games

A related problem which also involves the generation and spread of words is the so-called naming game. The original formulation and implementation of this problem was proposed by Luc Steels as a model for the emergence of a shared vocabulary within a population of agents (Steels 2001, 2003, 2005; see also Nolfi & Mirolli 2010). Originally, this approach involved communication between two embodied communicating agents. These agents (figure 1b) are able to visually identify objects from their environment and assign them to randomly generated names, which are then sent to the other agent in a speaker–hearer kind of interaction. Exchanges receive a pay-off every time the same word is used by both agents to name a given object. This is done by means of a trial and error process where failures are common at the beginning, as a common lexicon slowly emerges. Specifically, the set of rules are:

- The speaker selects an object.
- The speaker chooses a word describing the object from its inventory of word–object pairs. If it does not have a word then it invents one for the object. The speaker transmits the word–object pair to the listener.
— If the listener has the word–object pair then the
transmission is a success. Both agents remove all
other words describing the object from their inven-
tory and keep only the single common word.
— If the listener does not have the word–object pair,
then the listener will add this new word to its inven-
tory. And this is recorded as a failure.

Eventually, a shared, stable repertoire gets fixed. The
basic rules can be easily mapped into a toy model
(the naming game model) involving many agents, by
using a statistical physics approach (Baronchelli et al.
2006, 2008). Both hardware and simulated implementa-
tions display an S-shaped growth of the vocabulary,
although interesting differences arise when we take
into account spatial effects and the pattern of relations
between agents, describable as a complex network
(Steels & Mcintyre 1999; Dall’Asta et al. 2006; Lu

3. COMPETITION AND EXTINCTION

Languages are spoken by individuals, and the number
of speakers provides a measure of language breadth.
Because of both economic and social factors, a given
language can become more efficient than others in
recruiting new users and as a consequence it can reach
a larger fraction or even exclude the second language,
which becomes extinct.1 This replacement would be a
consequence of competition, one of the most essential
components of ecological dynamics, which can be
applied to language dynamics too. Early models of
two-species competition define the basic formal scenario
where species interactions under limited resources occur
(Case 1999). The standard model is provided by the
classical Lotka–Volterra equations, namely
\[
\frac{dx}{dt} = \mu_1 x(1 - x - \beta_{12} y)
\]
and
\[
\frac{dy}{dt} = \mu_2 y(1 - y - \beta_{21} x),
\]
where \(x\) and \(y\) indicate the (normalized) populations
of competing species, \(\mu_i\) indicates their (per capita)
growth rates and the coefficients \(\beta_{ij}\) are the rates of
interspecific competition. We can see that for \(\beta_{ij} = 0\)
two independent logistic equations would be obtained,
whereas for non-zero competition two possible scenarios
are at work.

Understanding language competition dynamics is
clearly important: if the exclusion scenario is also at
work, then competition can imply extinction. Moreover,
theoretical models can help in defining useful strategies
for language preservation and revitalization (Fishman
1991, 2001). Steady language decline has been observed
in some cases, when population records of speakers are
available. This is illustrated in figure 2, where the
decay over time of four different languages is depicted.

1Species and languages also become extinct under external events
(such as asteroid impacts or climate change). Sudden death of a
language can occur owing to a volcanic eruption killing the small
population of speakers or (more often) as a consequence of genocide
(Nettle & Romaine 2002).

All these languages were used by a minority of speakers,
competing with a dominant tongue that was gradually
adopted by speakers as the less used ones were aban-
donned. This type of increasing return is common in
ecomics, where positive feedbacks and amplification
phenomena are common (Arthur 1994).

A simple model was proposed by Abrams and
Strogatz (AS), which has been shown to provide a
rationale for the shape of language decay (Abrams &
Strogatz 2003; Stauffer et al. 2007). The model is
based on the assumption that two languages are com-
peting for a given population of potential speakers
(the limiting resource) where we will indicate as \(x\) and
\(y\) the relative frequency of each population (assuming
that individuals are monolinguals, see below). The
dynamics is governed by the following differential equation:
\[
\frac{dx}{dt} = yP_{as}[y \rightarrow x] - xP_{as}[x \rightarrow y],
\]
where it is assumed that \(P_{as}[x \rightarrow y] = 0\) if \(x = 0\) and
also constant population \((x + y = 1)\). The transition
probabilities depend on two parameters. The specific
model reads
\[
\frac{dx}{dt} = s(1 - x)x^a - (1 - s)x(1 - x)^a,
\]
where the \(s\) parameter indicates the so-called social
status of the language. Two extreme equilibrium
states are easily found after imposing \(\frac{dx}{dt} = 0\).
These are \(x^* = 0\) (zero population) and \(x^* = 1\) (all
speakers use the language). In our case, the stability
criterion gives \(\lambda(0) = s - 1 < 0\) and \(\lambda(1) = -s < 0\)
and thus both are stable attractors.
Together with the exclusion points \(x = 0\) and 1, there
is a third equilibrium point, which can be obtained from
\[
sx^{a-1} = (1 - x^*)^{a-1}(1 - s),
\]
and, after some algebra, one finds that
\[
x^* = \left[1 + \left(\frac{s}{1 - s}\right)^{1/(a-1)}\right]^{-1}.
\]
Given the stable character of the other two fixed
points, \(x^*\) can only be unstable and thus no coexistence
is allowed.

The model has been used to fit available data on
language decay (figure 2) and assumes a scenario of
minority languages competing with widely used
majority tongues. One clear implication of the stability
analysis is that the extinction of one of the competing
solutions is inevitable. The social parameter will influ-
ence which language will become extinct. Nonetheless,
linguistic diversification seems unavoidable: the
language that succeeds in the competition situation
will become more and more diverse as it extends
through time and space, and it may end up yielding
mutually unintelligible linguistic variants.

The AS model does not take into account that it is
probable that a fraction of individuals (under some cir-
cumstances) will become bilingual. This might not seem
so relevant, but bilingualism actually introduces a very
interesting ingredient to our view of language change, to be outlined in the next section.

4. COEXISTENCE AND BILINGUALISM

The previous model is simplified in many respects. By considering human populations as homogeneous systems, geographical effects and some idiosyncrasies of human language (not shared with ecosystems) are ignored. Spatial effects will be explored in the next section. Here we concentrate on a special property of human communities, namely the presence of individuals who are grammatically and communicatively competent in more than one language. Actually, a large fraction of humankind uses more than one tongue for communication. Historical reasons and the influence of modern invasions by languages such as English makes multi-lingualism an important ingredient to take into account.

The basic idea behind this approach is that the presence of bilingual speakers makes language coexistence likely to occur, provided that the two languages are close enough to each other. In this picture, three variables are used: as in the AS model, \( x \) and \( y \) will be the fraction of speakers using languages \( X \) and \( Y \). Moreover, a third group \( B \) using both languages has a size \( b \) in such a way that \( x + y + b = 1 \). Transitions are defined in similar ways (figure 3a). For example, changes in \( x \) would result from a kinetic equation,

\[
\frac{dx}{dt} = yP[y \rightarrow x] + bP[b \rightarrow x] - x(P[x \rightarrow y] + P[x \rightarrow b]),
\]

and the constant population constraint allows the model to be defined in terms of just two coupled equations, namely

\[
\frac{dx}{dt} = c((1 - x)(1 - \kappa)s_x(1 - y)^a - x(1 - s_x)(1 - x)^a)
\]

\[
\frac{dy}{dt} = c((1 - y)(1 - \kappa)s_y(1 - x)^a - y(1 - s_y)(1 - y)^a)
\]

Figure 2. The dynamics of language death. Here four different cases are represented: (a) Scottish Gaelic, (b) Quechua in Huanuco, Peru, (c) Welsh in Monmouthshire, Wales, and (d) Welsh in all of Wales, from historical data (filled squares) and a single modern census (open circles). Fitted curves show solutions of the Abrams–Strogatz model (schematically indicated in the upper plot). Redrawn from Abrams & Strogatz (2003).
respectively. The $k$ parameter as a measure of the probability that a monolingual speaker becomes bilingual, since $k$ is only an indicator of the degree of similarity among languages, and neglects the role of their social status. It is worth noting that many bilingual scenarios involve two highly differentiated languages, such as Basque and Castilian in northern Spain or Amazigh and Arabic in northern Africa.

How probable is it that the bilingual scenario will be relevant in the future? Recent model approaches suggest that maintaining a bilingual society necessarily requires the maintenance of status as a control parameter (Chapel et al. 2010). On the one hand, preserving language diversity in a globalized world will need active efforts when small populations of speakers are involved. But, on the other hand, we must also take into account current demographic trends (Graddol 2000), which will need to be incorporated into future models of language change. Against early predictions suggesting the dominant role of English as an exclusive language, the future looks multi-lingual. Different languages are gaining relevance as their social and economic status improves. Moreover, other interesting tendencies start to develop as some languages (such as English, Portuguese or Dutch) spread beyond their original geographic domains. They not only become mutualistic (as a bilingual speaker acquires a higher social status) but can also develop internal differentiation. We should expect in the future to see the emergence of (perhaps unintelligible) dialects of English, as happened with Latin.

5. SPATIAL DYNAMICS

The exclusion point resulting from the Lotka–Volterra equation and related models (such as AS’s model) implies that strong competition leads to diversity reduction. Within the context of population dynamics, such a result was challenged under the introduction of spatial degrees of freedom (Solé et al. 1993; see also Solé & Bascompte 2006 for a review of results). Spatial dynamics involves two basic components. One is the reaction term, describing how populations interact (for example the equations described above). The second describes how populations move through space. It is well known that space is responsible for the emergence of qualitative changes in dynamical patterns (Turing 1952; Murray 1989; Bascompte & Solé 2000; Dieckmann et al. 2000). Competition under spatial structure generates a completely novel result: since exclusion depends on initial conditions, the two potential attractors can be (locally) possible. Starting from random initial conditions, different species or languages can exclude each other at different locations.

The extension of the AS model to space was performed by Patriarca & Leppänen (2004), who used a reaction–diffusion framework. The model considers the local dynamics of the normalized densities of speakers using a given language at each point $r$ in space. If

$J. R. Soc. Interface (2010)$

and

$$\frac{dy}{dt} = c((1 - y)(1 - k)(1 - s_x)(1 - x) - y s_x (1 - y)),$$

(4.3)

where $\kappa \in [0, 1]$ is a new parameter measuring the degree of similarity among languages and the status of the languages is now indicated as $s_x$ and $s_y = 1 - s_x$, respectively. The $k$ parameter provides a measure of the likelihood that two single-language speakers can communicate with each other. It also affects the probability that a monolingual speaker becomes bilingual. We can easily check that the model reduces to the AS scenario for $\kappa = b = 0$.

Available data from language change in Northern Spain (Mira & Paredes 2005) provide a test of this model. Here the two languages are Castilian and Galician, both derived from Latin. These languages allow a relatively good mutual understanding and parameters are easily estimated. For this dataset, a best fit was obtained using $a = 1.5$, $s(Galician) = 0.26$, $c = 0.1$ and $\kappa = 0.8$. As we can see, the apparent decline of Galician is actually a consequence of a simultaneous increase of Castilian monolinguals and bilinguals.

Figure 3. Dynamics of language use under the presence of bilingual speakers. (a) The probabilities of jumps between two single-language speakers and between a bilingual speaker and one of the single-language speakers involved. (b) The fraction of speakers versus time in Galicia (north western Spain). The smooth curves (modified after Mira & Paredes 2005) are the results of fitting a modified AS model (see text).
\( \phi_x(r, t) \) and \( \phi_y(r, t) \) indicate the local densities of \( x \) and \( y \) at a given point and time, they read
\[
\frac{\partial \phi_x(r, t)}{\partial t} = F(\phi_x, \phi_y) + D_x \nabla^2 \phi_x(r, t)
\] (5.1)
and
\[
\frac{\partial \phi_y(r, t)}{\partial t} = -F(\phi_x, \phi_y) + D_y \nabla^2 \phi_y(r, t),
\] (5.2)
where \( F(\phi_x, \phi_y) \) is just the AS equation for the local densities
\[
F(\phi_x, \phi_y) = s_x \phi_y \phi_x^\mu - s_y \phi_x \phi_y^\nu,
\] (5.3)
where \( s_x, s_y \) indicate the status of each language. The \( D_i \)'s on the right-hand side of the equation are the so-called diffusion coefficients associated with the spreading process.

The previous equations can be numerically integrated (Dieckmann et al. 2000). We will illustrate this by using a one-dimensional spatial system (the generalization to two dimensions is straightforward). First, we discretize \( \partial \phi/\partial t \) as follows:
\[
\frac{\partial \phi_x(r, t)}{\partial t} \approx \frac{\phi_x(r, t + \Delta t) - \phi_x(r, t)}{\Delta t},
\] (5.4)
where \( r \) is the local position in the one-dimensional domain \( Z = [0, L] \) and \( \Delta t \) some characteristic time scale. Similarly, the discretization of the diffusion term is made as follows:
\[
\frac{\partial^2 \phi_x(r, t)}{\partial r^2} \approx \frac{\phi_x(r + \Delta r, t) + \phi_x(r - \Delta r, t) - 2\phi_x(r, t)}{\Delta r^2},
\] (5.5)
\( \Delta r \) being the corresponding characteristic spatial scale. Using these definitions, we obtain an equation for the time evolution of \( \phi_x(r, t) \),
\[
\phi_x(r, t + \Delta t) = \phi_x(r, t) + \left[ F(\phi_x, \phi_y) + \frac{D_x}{\Delta r^2} (\phi_x(r + \Delta r, t) + \phi_x(r - \Delta r, t) - 2\phi_x(r, t)) \right] \Delta t.
\]
Additionally, boundaries are to be included. These allow the impact of finite-size effects and geography on the dynamics and equilibrium states to be defined. The reasonable assumption is to use zero-flux (von Neumann) boundary conditions, namely
\[
\left( \frac{\partial \phi_x(r, t)}{\partial r} \right)_{r=0} = \left( \frac{\partial \phi_x(r, t)}{\partial r} \right)_{r=L} = 0.
\] (5.6)
In terms of our discretization, we would have \( \phi_x(0, t) - \phi_x(L, t) = 0 \) and \( \phi_y(0, t) - \phi_y(L, t) = 0 \).

The dynamics starts with two populations of speakers located in two different domains \( Z_a \) and \( Z_b \) (so that \( ZZ_a \cup ZZ_b \)). This is shown in figure 4a, where we display the initial condition. If we label as \( N_a^\mu \) and \( N_b^\mu \) the total populations of speakers in each domain \( \mu = 1, 2 \), at a given domain \( Z_\mu \) we would have
\[
N_a^\mu(t) = \int_{Z_\mu} \phi_x(r, t) \, dr,
\] (5.7)
starting from \( N_i = \frac{1}{2} \) following a Gaussian shape (see Patriarca & Leppänen 2004). As the dynamics proceeds, we can observe a tendency towards maintaining the spatial segregation. Each language ‘wins’ in its initial domain, and eventually both reach a homogeneous steady state within such a domain.

Generalizations to heterogeneous domains reveal that the previous patterns can be affected by both historical events and spatial inhomogeneities (Patriarca & Heinsalu 2008). However, the main message from this
treated as a binary string, i.e. length \( L \). A set of languages with \( M \) = \( 2^L \) for a given population of speakers can evolve towards coexistence. Specifically, some relevant problems such as diversity and its maintenance can be properly addressed by ignoring interactions. Following this picture, we consider in this section the lexical component of language viewed as a bag of words and how a set of languages competing for the population dynamics of change owing to abandonment can change if individuals modify some of their bits.

6. STRING MODELS OF LANGUAGE CHANGE

As already mentioned in §2, a collection of words provides the first definition of a language in terms of its lexicon. This of course ignores a crucial component of language: words interact in non-random ways and higher order levels of organization should be taken into account. However, as occurs with some theoretical models of diverse ecosystems (Sole & Bascompte 2006), some relevant problems such as diversity and its maintenance can be properly addressed by ignoring interactions. Following this picture, we consider in this section the lexical component of language viewed as a bag of words and how a set of languages competing for a given population of speakers can evolve towards a single, dominant tongue or instead a diverse set of coexisting languages.

A fruitful toy model of language change is provided by the string approximation (Stauffer et al. 2006; Zanette 2008). In this approach, each language \( L_i \) is treated as a binary string, i.e. \( L_i = (S_1, S_2, \ldots, S_L) \) of length \( L \). Here \( S_j \in \{0,1\} \) and, as defined, a finite but very large set of potential languages exists. Specifically, a set of languages \( \mathcal{L} \) is defined, namely

\[
\mathcal{L} = \{ L_1, L_2, \ldots, L_M \},
\]

with \( M = 2^L \). These languages can be located as the vertices of a hypercube, as shown in figure 5 for \( L = 3 \). Nodes (languages) are linked through arrows (in both directions) indicating that two connected languages differ in a single bit. This is a very small-sized system. As \( L \) increases, a combinatorial explosion of potential strings takes place.

6.1. Mean field model

A given language \( L_i \) is shared by a population of speakers, to be indicated as \( x_i \), and such that the total population of speakers using any language is normalized (i.e. \( \sum_i x_i = 1 \)). A mean field model for this class of description has been presented by Damian Zanette, using a number of simplifications that allow the qualitative behaviour of competing and mutating languages to be understood (Zanette 2008). A few basic assumptions are made in order to construct the model. First, a simple fitness function \( \phi(x) \) is defined. This function measures the likelihood of abandoning a language. This is a decreasing function of \( x \), and such that \( \phi(0) = 1 \) and \( \phi(1) = 0 \). Different choices are possible, including for example \( 1 - x \), \( 1 - x^2 \) or \( (1 - x)^2 \). On the other hand, mutations are also included: a given language can change if individuals modify some of their bits.

The mean field model considers the time evolution of populations assuming no spatial interactions. If we indicate \( x = (x_1, \ldots, x_M) \), the basic equations will be described in terms of two components,

\[
\frac{dx_i}{dt} = A_i(x) - M_i(x),
\]

where both language abandonment \( A_i(x) \) and mutation \( M_i(x) \) are introduced. Specifically, the following choices are made:

\[
A_i(x) = \rho \phi_i((\phi) - \phi(x_i)),
\]

for the population dynamics of change owing to abandonment. This is a replicator equation, where the speed of growth is defined by the difference between average fitness \( \langle \phi \rangle \), namely

\[
\langle \phi \rangle = \sum_{j=1}^{N} \phi(x_j)x_j,
\]

and the actual fitness \( \phi(x_i) \) of the \( i \)-language. Here \( \rho \) is the recruitment rate (assumed to be equal in all languages). What this fitness function introduces is a multiplicative effect: the more speakers who use a given language, the more probable that they keep using it and others join the same group. Conversely, if a given language is rare, its speakers might easily shift to some other, more common, language.
The second term includes all possible flows between ‘neighbouring’ languages. It is defined as

\[ M_i(x) = \frac{\mu}{L} \left( \sum_{j=1}^{N} W_{ij} x_j - x_i \sum_{j=1}^{N} W_{ji} \right). \]  

In this sum, we introduce the transition rates \( W_{ij} \) of mutating from language \( L_i \) to language \( L_j \) and vice versa. Only single mutations are allowed, and thus \( W_{ij} = 1 \) if the Hamming distance \( D(L_i, L_j) \) is exactly 1. More precisely, if

\[ D(L_i, L_j) = \sum_{k=1}^{L} |S^i_k - S^j_k| = 1. \]  

In other words, only nearest-neighbour movements through the hypercube are allowed. In summary, \( A(x) \) provides a description of competitive interactions whereas \( M(x) \) gives the contribution of small changes in the string composition. The background ‘mutation’ rate \( \mu \) is weighted by the matrix coefficients \( W_{ij} \) associated with the likelihood that each specific change will occur.

This model is a general description of the bit string approximation to language dynamics. However, the general solution cannot be found and we need to analyse simpler cases. An example is provided in the next section. Although the assumptions are rather strong, numerical models with more relaxed assumptions seem to confirm the basic results reported below.

### 6.2. Supersymmetric scenario

A solvable limit case with obvious interest to our discussion considers a population where a single language has a population \( x \) whereas all others have a small, identical size, i.e. \( x_i = (1-x)/(N-1) \). The main objective of defining such a supersymmetric model is making the previous system of equations collapse into a single differential equation, which we can then analyse. In particular, we want to determine when the \( x = 0 \) state will be observed, meaning that no single dominant language is stable.

We have the normalization condition, now defined by

\[ \sum_{j=1}^{N} x_j = x + \sum_{j=1}^{M-1} \left( \frac{1-x}{N-1} \right) = 1 \]  

(6.7)

(where we choose \( x \) to be the \( M \)th population, without loss of generality). In this case the average fitness reads

\[ \langle \phi \rangle = \phi(x)x + \sum_{j=1}^{M-1} \phi \left( \frac{1-x}{N-1} \right) \left( \frac{1-x}{N-1} \right). \]  

(6.8)

Using the special linear case \( \phi(x) = 1-x \), we obtain

\[ A(x) = \rho x (1-x) \left( x - \frac{1-x}{N-1} \right). \]  

(6.9)

The second term is easy to obtain: since \( x \) has (as any other language) exactly \( L \) nearest neighbours, and given the symmetry of our system, we have

\[ B(x) = \frac{\mu}{L} \left( L \frac{1-x}{N-1} - xL \right) = -\mu \left( \frac{N-1}{N-1} \right). \]  

(6.10)

And the final equation for \( x \) is, thus, for the large-\( N \) limit (i.e. when \( N \gg 1 \))

\[ \frac{dx}{dt} = \rho x^2 (1-x) - \mu x. \]  

(6.11)

This equation describes an interesting scenario where growth is not logistic, as happened with our previous model of word propagation. As we can see, the first term on the right-hand side involves a quadratic component, indicating a self-reinforcing phenomenon. This type of model is typical of systems exhibiting cooperative interactions and an important characteristic is its...
hyperbolic dynamics: instead of an exponential-like approximation to the equilibrium state, a very fast approach takes place.

The model has three equilibrium points: (i) the extinction state, \( x^* = 0 \), where the large language disappears and (ii) two fixed points, \( x^\pm \), defined as

\[
x^\pm = \frac{1}{2} \left( 1 \pm \sqrt{1 - 4\mu \rho} \right).
\]

As we can see, these two fixed points exist provided that \( \mu < \mu_c = \rho/4 \). Since three fixed points coexist in this domain of parameter space, and the trivial one (\( x^* = 0 \)) is stable, the other two points, namely \( x^- \) and \( x^+ \), must be unstable and stable, respectively. If \( \mu < \mu_c \), the upper branch \( x^+ \), corresponding to a monolingual solution, is stable.

In figure 3a we illustrate these results by means of the bifurcation diagram using \( \rho = 1 \) and different values of \( \mu \). In terms of the potential function we have

\[
\frac{dx}{dt} = -\frac{\partial \Phi_\mu(x)}{\partial x},
\]

where \( \Phi_\mu(x) = -\int(A(x) - B(x))dx \), which for our system reads

\[
\Phi_\mu(x) = -\rho \left( x^3 + \frac{x^4}{4} \right) + \mu \frac{x^2}{2}.
\]

In figure 5a–d three examples of this potential are shown, where we can see that the location of the equilibrium point is shifted from the monolanguage state to the diverse state as \( \mu \) is tuned. The corresponding phases in the \( (\mu, \rho) \) parameter space are shown in figure 5.

It is interesting to see that this model and its phase transition are somewhat connected to the error threshold problem associated with the dynamics of RNA viruses (Eigen et al. 1987; Domingo et al. 1995; Adami 1998; Solé & Goodwin 2001). For a single language to maintain its dominant position, it must be efficient in recruiting and keeping speakers. But it also needs to keep heterogeneity (resulting from ‘mutations’) at a reasonably low level. If changes go beyond a given threshold, there is a runaway effect that eventually pushes the system into a variety of coexisting sub-languages. An error threshold is thus at work, but in this case the transition is of first order. This result would indicate that, provided that a source of change is active and beyond threshold, the emergence of multiple unintelligible tongues should be expected.

String models of this type capture only one layer of word complexity. Perhaps future models will consider ways of introducing further internal layers of organization described in terms of superstrings. Such superstring models should be able to introduce semantics, phonology and other key features that are known to be relevant. An example in this direction is provided by models of the emergence of linguistic categories (Puglisi et al. 2008).

7. GLOBAL PATTERNS AND SCALING LAWS

Tracking the relative importance of languages and in particular their likelihood of becoming extinct requires having the appropriate censuses of number of speakers using each language. The statistical patterns displayed by languages in their spatial and demographic dimensions provide further clues for the presence of non-trivial links between language and ecology (Nettle 1998; Pagel & Mace 2004; Pagel 2009). These patterns also provide a large-scale picture of languages, not restricted to small geographical domains or countries. In this section we consider two such statistical patterns. It is important to note that, strictly speaking, this problem involves both ecological and evolutionary time scales. In a given ecosystem, the succession process leading to a mature, diverse community can be described in terms of ecological dynamics. At this level, invasion and network species interactions are both relevant. However, the composition of the local pool of species is the outcome of evolutionary dynamics.

Some spatial models of language change have been presented in order to explain the results shown below (see de Oliveira et al. 2006, 2008). The close correlation between species diversity and language richness, as reported by different studies (Mace & Pagel 1995; Moore et al. 2002; Gaston 2005), suggests that some rules of organization might be common. As an example, a large-scale study of correlations among biological species and cultural and linguistic diversity in Africa (Moore et al. 2002) revealed that one-third of language richness can be explained on the basis of environmental factors. These included rainfall and productivity, which were shown to affect the distributions of both species and languages. However, there are also important differences that need an explanation.

7.1. Species–area relations

One of the universal laws of ecological organization is the so-called species–area relation (Rosenzweig 1995). It establishes that the diversity \( D \) (measured as the number of different species) in a given area \( A \) follows a power law

\[
D \sim A^z,
\]

where the exponent \( z \) typically varies from \( z = 0.1 \) to 0.45. Interestingly, languages seem to follow similar trends. They exhibit an enormous diversity, strongly tied to geographical constraints. As is the case with species distributions, languages and their evolution are shaped by the presence of physical barriers, population sizes and contingencies of many kinds. In this context, differences are also clear: speciation in ecosystems can take place without the presence of physical barriers, whereas some type of population isolation seems necessary for one language to yield two different languages, i.e. two linguistic variants that are not fully inter intelligible. On the other hand, there is a continuous drift in both species and languages that makes them change. A second difference involves the way extinction occurs. Species become extinct once the last of its
languages spoken in 228 countries. The power-law fit to the data reveals the impact of other forces acting on language diversity. An example is the contrast between Europe and New Guinea (see Diamond (1997) and references therein). The former has 10^7 km^2 and contains just 0.1 per cent of the world’s population but 10^3 different languages. The singularity of New Guinea has been carefully analysed by many authors. Take, for example, Papua New Guinea, which contains just 0.1 per cent of the world’s population but more than 13 per cent of the world’s languages. It is geographically an extremely irregular landscape, which creates multiple opportunities for isolation. Moreover, 80 per cent of its land is covered by rainforests. Additionally, food production is continuous, with no food shortages and a good yield. Bilingualism is widespread, with most speakers of the dominant Tok Pisin also speaking some local language (being exposed to several). Given the high yields of food harvest together with biogeographical constraints, there has been little incentive to create large-scale trade. A consequence of such a scenario is a dynamic equilibrium far from language homogenization (see Nettle & Romaine 2002 for a review).

The species–area relation has been explained in a number of ways through models of population dynamics on two-dimensional domains. Beyond their differences, these models share the presence of stochastic dynamics involving multiplicative processes. In ecology, such processes are characterized by positive and negative demographical responses proportional to the current populations involved; a larger population will be more likely to increase, but also more likely to suffer the attack of a given parasite (and thus experience a rapid decline). Within language, the rich-get-richer effect is obvious, whereas there is no equivalent for the negative effects of ‘parasitic’ languages.

### 7.2. Language richness laws

A different measure of language diversity involves the language richness among different countries. If \( N(D) \) is the frequency of countries with \( D \) different languages each, we can plot the cumulative distribution \( N_r(D) \), defined as

\[
N_r(D) = \int_D^\infty N(D) \, dD.
\]  

(7.3)

The resulting plot is rather interesting (figure 8a): the distribution follows a two-regime scaling behaviour, i.e.

\[
N_r(D) \sim D^{-\beta},
\]  

(7.4)

with \( \beta = 0.6 \) for \( 6 < D < 60 \) and \( \beta = 1.1 \) for \( 60 < D < 700 \). What is revealed from this plot? The first domain has an associated power law with a small exponent (here \( N(D) \sim D^{-1.1} \)): many countries have a small language diversity. But once we cross a given threshold \( D \approx 60 \) the decay becomes faster. One possible interpretation is that countries with a very large diversity will find it more difficult to preserve their unity under the social differentiation associated with ethnic diversity (Gomes et al. 1999).

A related distribution is given by the number of languages \( n_l(N) \) with a population size of \( N \) speakers. In figure 8b we display a log–log plot of the dataset (after binning) which shows a log-normal behaviour, with an enhanced number of small-sized languages. This pattern (as well as the scaling with area) is reproduced by a simple model presented below.

### 7.3. Language diversity model

A simple spatial model has been proposed in de Oliveira et al. (2008) as an extension of previous work (de Oliveira et al. 2006; see also Silva & de Oliveira 2008). The model combines a stochastic cellular automaton approach with non-local rules and a bit-string implementation. Starting from an empty lattice \( \Omega \) of \( L \times L \) sites, each site \((i, j) \in \Omega\) is characterized by a random number \( 1 \leq K_{ij} \leq M \) (with uniform distribution) representing the maximum population of speakers achievable by the language occupying it (the carrying capacity). Only one language \( L \) can be present at a given site and (as in \( S_0 \)) is represented by a string \( L_t = (S_1, S_2, \ldots, S_L) \) of length \( L \). A seed \( L_1 \) is located at \( t = 0 \) at a given site \((a, b)\), thus having a population \( K_{ab} \).
Now dispersal to nearest neighbours in the lattice occurs, favouring the spread towards sites having higher $K_{ij}$. Moreover, at a given site the given language $L_k$ can change (mutate) to a new one with a probability $m_k = a/f(L_k)$. Here $f(L_k)$ is the fitness associated with $L_k$, here chosen as

$$f(L_k) = \sum_{i,j} K_{ij} \theta(L(i,j), L_k), \quad (7.5)$$

with $\theta(m, n) = 1$ if $m = n$ and zero otherwise. In other words, the fitness considers the total occupation of the lattice (in terms of speakers), and the likelihood of a language to mutate is thus size-dependent following an inverse law. In this way, we incorporate the well-known fact that the impact of mutations favours genetic drift. The previous rules allow a diverse set of languages to expand and eventually occupy the whole lattice. An example is shown in figure 8b for a small ($L = 50$) lattice. We can see how languages emerge and spread around, generating monolingual patches.

In spite of its simplicity and strong assumptions, the model is able to capture several qualitative properties of both spatial and statistical power laws, similar to those presented above (de Oliveira et al. 2006, 2008). In some sense, we can conclude that the observed commonalities point towards shared system-level properties. This conclusion is partially true: the process of ecosystem building can be understood in terms of a spatial colonization of available patches. Each patch offers a given range of conditions that make it more or less suitable for the colonizer to persist. If colonization occurs locally, nearest patches will be occupied by best-fit competitors. In an ecological-like model, non-local colonization events will occur owing to the introduction of species from the regional pool (see Sole et al. 2002), but these events can also be interpreted as speciations. Perhaps the most obvious difference with ecological models is the assumption of a fitness trait that involves the whole population of the species. Such a non-local effect seems reasonable to assume when thinking of language as a vehicle of economic influence. Larger communities of speakers are likely to be much more efficient in further expanding.

8. DISCUSSION

Language dynamics has attracted the attention of physicists, computer scientists and theoretical biologists.

\footnote{In fact two opposite strategies can be observed in nature, particularly when looking at the colonization of habitat by plants, which can invest either in a few, well-protected seeds or in many, small ones. In the second case, most of the seeds will fail to survive.}
Table 1. A comparative list of features relating the organization and change of languages and species. The list of mechanisms is not exhaustive: it considers only mainstream phenomena. Some parallelisms between languages and species should be considered carefully. Although small invasions have a deep impact on an ecosystem’s organization, this factor rarely has a remarkable effect within large linguistic communities. This is arguably related to the tendency that an invading language displays both a low demographic weight and a low social status. It is also interesting to observe that mutualism, i.e. a cooperative strategy for survival that benefits two or more species, is completely absent in language dynamics. On the contrary, multi-lingualism as well as diglossia and related phenomena—see text—are features exclusive to language. Finally, we emphasize that analogies to food webs are difficult to define in the study of language contact. However, some kind of network abstraction to represent the socio-cultural relations among languages or communities of speakers is conceivable.

<table>
<thead>
<tr>
<th>Species</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes of living beings</td>
<td>Community-shared codes</td>
</tr>
<tr>
<td>Lack of interbreeding</td>
<td>Unintelligibility</td>
</tr>
<tr>
<td>Genetic/geographic isolation</td>
<td>Geographic barriers</td>
</tr>
<tr>
<td>Competition/external events</td>
<td>Competition/external events</td>
</tr>
<tr>
<td>Two-regime scaling</td>
<td>Scaling law</td>
</tr>
<tr>
<td>Gradual + punctuated</td>
<td>Dialects</td>
</tr>
<tr>
<td>Species–area law</td>
<td>Language–area scaling</td>
</tr>
<tr>
<td>Very important</td>
<td>Gradual + punctuated</td>
</tr>
<tr>
<td>Very important</td>
<td>Rare</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>Very important</td>
</tr>
</tbody>
</table>

...
community, it may seem that languages compete for the number of speakers only in a metaphorical sense. However, it is remarkable that the competition among languages and the competition among species can be mathematically modelled using similar methods. At this point, it is necessary to take into consideration the importance of the role of a given language as a social status parameter in language competition, provided that different languages may distribute differently in society, but not different species in an ecosystem. Moreover, competition among different languages in contact can be materialized in many different ways, depending on how a given culture conceives mono/multi-lingualism.

Although the ecological metaphor of language dynamics fits well with several important features, there are a number of important linguistic phenomena which have no equivalent in ecology. Some members of a community may be bilingual or multi-lingual, i.e. they may possess not only the traditional language of the community (namely, their mother tongue), but also other languages or dialects. Indeed, some members of a community may use different languages or dialects in different social spheres, a phenomenon called diglossia. It is also worth noting that, when speakers of multiple languages have to communicate and do not have the chance to learn each other’s language, they develop a simplified code, a pidgin, which may increase its degree of complexity over the years. However, when a group of children are exposed to a pidgin at the age when they acquire a language, they transform it into a full complex language, a creole (DeGraff (1999) and references therein). In this context, although some parallels have been traced between creolization and genetic hybridization in plants (Croft 2000) they do not seem well supported or even properly defined.

Another related and remarkable linguistic idiosyncrasy is the emergence of new languages ex nihilo. This is the case of the Nicaraguan sign language (Kegl et al. 1999) which spontaneously developed among deaf school children in western Nicaragua over a short period of time once deaf individuals (until then growing essentially isolated) could start communicating with each other. Starting from a very limited number of signs and unable to learn Spanish, it was found that the group rapidly developed a grammar, which became a complex language at the second ‘generation’, as soon as the next group of children learned it from the first one. A similar situation was analysed for the Al-Sayyid Bedouin sign language, which has arisen in the last 70 years within an isolated community (Sandler et al. 2005). This type of phenomenon highlights the role of the cognitive dimension of language, which makes it far more flexible than species behaviour. Indeed, nothing similar to multi-linguism, diglossia or the appearance of new languages (pidgins and creoles) is attested in non-linguistic ecological systems. Modelling such phenomena is still an open challenge.

In sum, as suggested by Darwin, both languages and ecosystems share some of their crucial features. These include spreading dynamics, the presence of dramatic thresholds and the role of space in favouring heterogeneity. In the language context, this space-driven enrichment can be interpreted in other ways than physical space, such as social distance. It is also true, however, that a close inspection of both systems reveals some no less interesting differences, particularly those related to the flexibility of individuals in acquiring several languages or the social, cultural or political factors that constantly interfere in linguistic phenomena. Future efforts towards a theory of language change might help us to understand our origins as a complex, social species and the future of language diversity.

We thank Guy Montag and the members of the Complex Systems Lab for useful discussions. This work has been supported by NWO research project Dependency in Universal Grammar, the Spanish MCIN Theoretical Linguistics 2009SGR1079 (JF), the James S. McDonnell Foundation (BCM) and by the Santa Fe Institute (RS).

REFERENCES


Lieberman, E., Michel, J.-B., Jackson, J., Tang, T. & Nowak, M. A. 2007 Quantifying the evolutionary dynamics of


