Assortative Mixing, Neutrality, and the Shape of Language Change

Henri Kauhanen, University of Manchester
henri.kauhanen@postgrad.manchester.ac.uk

LANGUAGE COMMUNITIES are characterized by three features: (1) they usually relax into states where a single linguistic variant **DOMINATES**; (2) they are nonetheless able to **SHIFT** from a state of dominance by one variant to a state of dominance by another; (3) when a shift occurs, the change is **WELL-BEHAVED**, usually approximable by a monotone S-curve. Replicator-type models where variants are endowed with differential fitness values can be used to generate such dynamics. However, the notion of fitness is problematic in the context of linguistics, where it is often unclear why some competing variants such as alternative pronunciations or grammatical rules should be fitter than others. Here we study the possibility that social network effects may drive **NEUTRAL EVOLUTION** in language change and account for the above three features. Casting the question in terms of a generalized Moran model where dynamics *of* the network and dynamics *on* the network are updated concurrently, we find that dominance, ability to shift, and well-behavedness of change are satisfied if the strength of assortativity in the network, as well as the mutation rate inherent in language adoption, fall in certain conducive regions.

§1 Model

Assume a population of N speakers, each of whom espouses exactly one of C possible linguistic **VARIANTS** $v_1, \ldots, v_C$ (e.g. alternative pronunciations, grammatical rules) at any time. Suppose that the speakers are the vertices V of an undirected graph (V, E) and that an edge $(i, j) \in E$ represents that speakers i and j can influence each other linguistically.

Let the population evolve in discrete time as follows: at each time step, one speaker is removed from the network uniformly at random, and a new one is inserted using a **REWIRING ALGORITHM**, whereby she acquires exactly $K < N$ connections (see §2). We assume that the new speaker i adopts one of the C variants from her neighbourhood $E(i)$ in a frequency-based (neutral) manner but subject to mutation with a small probability $\mu \ll 1$; specifically, we set

$$P(v_r) = (1 - \mu) f_{E(i)}(v_r) + \mu / C$$

for the probability to adopt v_r, where $f_{E(i)}(v_r)$ is the frequency of this variant in the neighbourhood $E(i)$. The speaker is assumed not to change her variant thereafter.

§2 Preferential attachment

To study the effect of assortative mixing by degree on the shape of linguistic trajectories, we define the following simple **PARAMETERIZED PREFERENTIAL ATTACHMENT** algorithm. Let $Q(d)$ be an ordered set (possibly empty) that contains those speakers with degree exactly d (i.e. having d connections), for $d = 0, 1, \ldots, N-1$. At each rewiring, we shuffle each $Q(d)$ to produce permuted sets $\widehat{Q}(d)$, and define $Q = \widehat{Q}(N-1) \circ \widehat{Q}(N-2) \circ \cdots \circ \widehat{Q}(1) \circ \widehat{Q}(0)$ where $\circ$ denotes concatenation; call this ordered, shuffled set Q the **QUEUE**. When a new speaker i is inserted into the network, we attach her to a speaker j picked from the queue as follows:

– with probability σ, j is the first speaker in the queue (the one with the smallest index and hence the highest degree);

– with probability $1 - \sigma$, j is drawn from the queue uniformly at random.

In either case, speaker j is removed from the queue and the process repeated until the new speaker i has received her K connections. The parameter σ, then, controls the strength of assortative mixing by degree in the social network (Figure 1).

§3 Dominance, shifts, and well-behavedness

Let $x_r(t)$ represent the proportion of speakers using variant v_r at time t, and call $\vec{x}(t) = (x_1(t), \ldots, x_C(t))$ the **FREQUENCY STATE** of the system. We propose to operationalize the dominance, shifting, and well-behavedness properties of a community using three quantities:

– **DOMINANCE TIME** D_δ: the proportion of time that the language community spends in states where a single linguistic variant dominates, i.e. that $x_r > 1 - \delta$ for some v_r and some suitably small δ;

– **NUMBER OF SHIFTS** S_δ: the number of shifts from one δ-dominant variant to another;

– **LINEAR FIT TO MEAN DISPLACEMENT** F: let

$$M(\tau) = \frac{1}{n - \tau} \sum_{t=1}^{n-\tau} d(\vec{x}(t), \vec{x}(t+\tau))$$

be the mean displacement function of a trajectory $\vec{x}(1), \ldots, \vec{x}(n)$, where d is the Euclidean distance in $\mathbb{R}^C$ and $\tau \ll n$. It can then be shown that well-behaved trajectories such as the one depicted in Figure 2 (top) give rise to functions $M(\tau)$ which are nearly linear in τ, whereas for ill-behaved trajectories M is nonlinear. We shall estimate this (non)linearity by fitting a linear model to empirical mean displacements from simulations and calculating the goodness of fit.

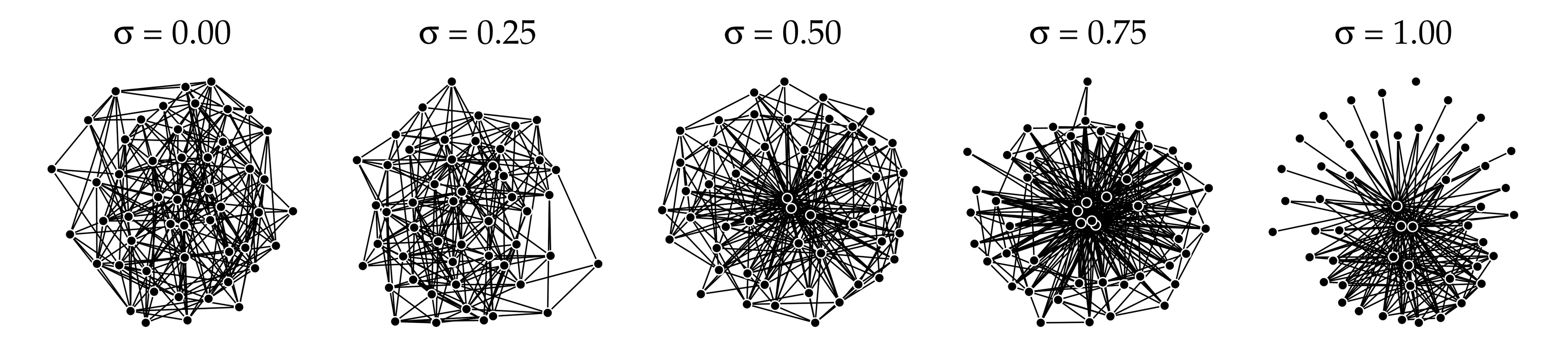

Figure 1: Networks ($N = 50$, $K = 10$) with different strengths of assortative mixing by degree σ.

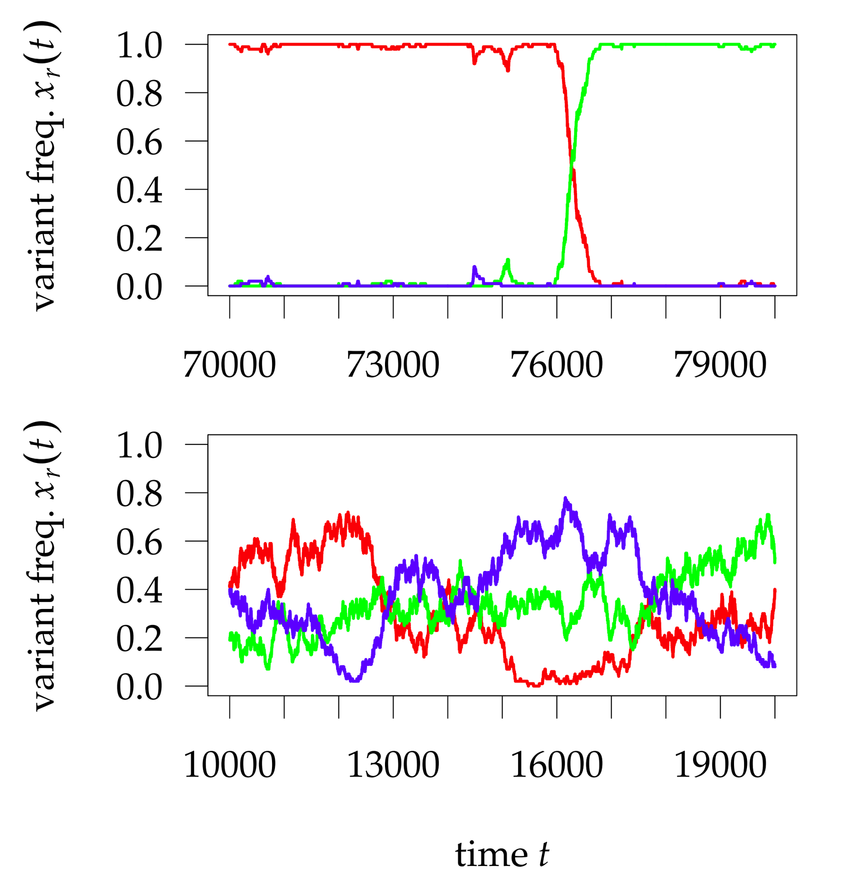

Figure 2: Portions of two trajectories from two simulations with three competing variants (red, green, and blue), depicted as time series of the variant frequencies $x_r(t)$. Top: a well-behaved trajectory with $\sigma = 1$ and $\mu = 0.005$. Bottom: an ill-behaved trajectory with $\sigma = 0$ and $\mu = 0.05$.

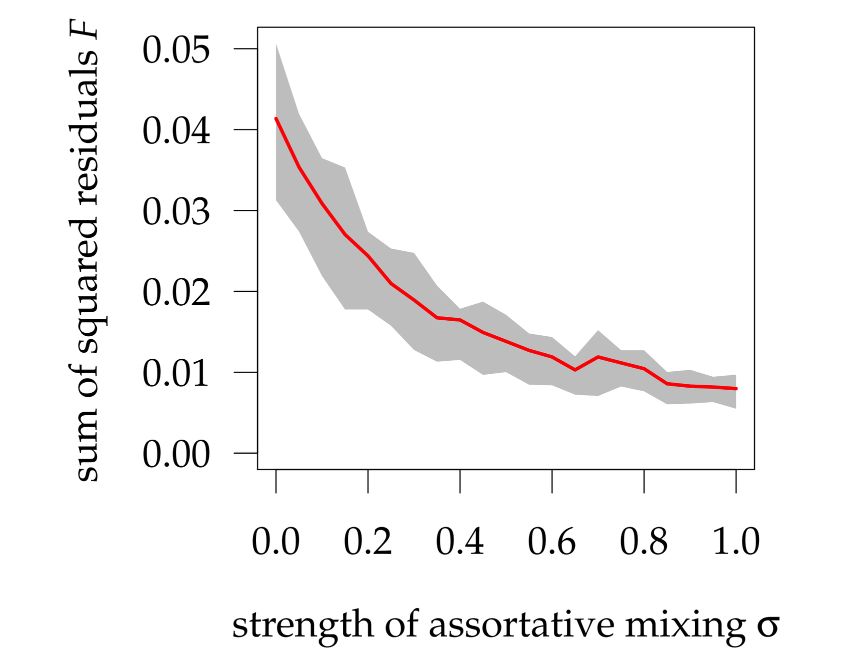

Figure 4: Goodness of linear fit to the empirical mean displacement function $M(\tau)$ of several simulation runs with varying strengths of assortative mixing σ; mutation rate is kept constant at $\mu = 0.005$. Means over 50 runs; grey area gives the interquartile range.

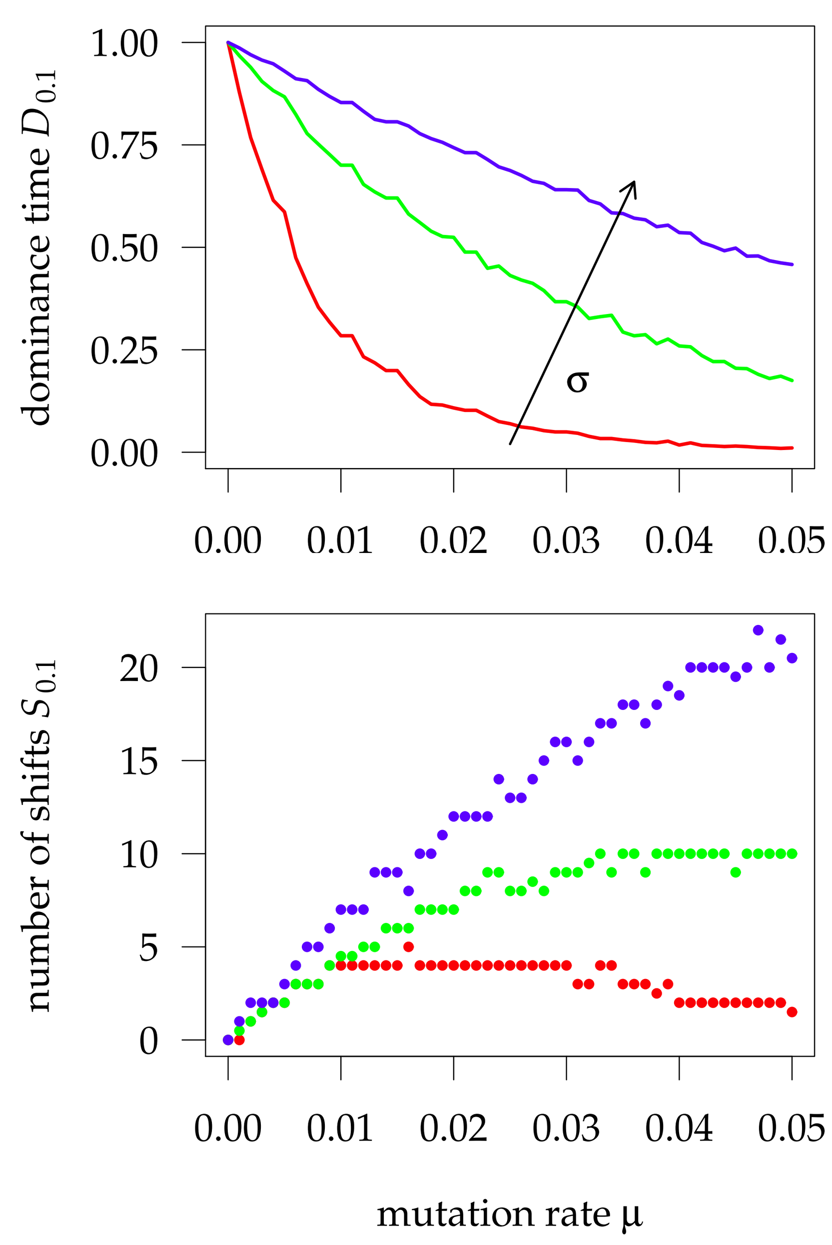

Figure 3: Dominance times (top) and numbers of shifts (bottom) for various values of mutation rate μ and assortative mixing strengths $\sigma = 0$ (red), $\sigma = 0.5$ (green), and $\sigma = 1$ (blue). Means ($D_{0.1}$) or medians ($S_{0.1}$) over 50 runs.

§4 Results & Conclusion

We ran a batch of computer simulations on the model for various combinations of values of σ and μ. The other parameters were $N = 100$, $K = 10$, $C = 3$; each simulation lasted for 10^5 iterations, and a total of 50 runs were conducted with each parameter setting.

Figure 3 shows that there exists a region of the parameter space in which the language community satisfies two features simultaneously: (1) a high dominance time, and (2) an ability to traverse from dominance by one variant to dominance by another. This behaviour occurs when $0.005 \lesssim \mu \lesssim 0.01$ and $\sigma \gtrsim 0.5$. Figure 4 suggests that shifts from one dominant variant to another are well-behaved for these values of μ and σ.

We conclude that social network (demographic) effects are sufficient to drive well-behaved language change even when adoption of linguistic variants is neutral, i.e. frequency-based and not biased by differential fitness values.

Future research should tackle these questions by more analytic means. The limiting case $\sigma = 1$ lends itself particularly well to such treatment.

I wish to thank Professor David Denison, Dr George Walkden, Dr Tobias Galla, and Dr Mark Muldoon for discussions which have been instrumental in directing the course of this research. I gratefully acknowledge financial support from the School of Arts, Languages and Cultures, University of Manchester.