Linguistics

Modelling the dynamics of language death

Thousands of the world’s languages are vanishing at an alarming rate, with 90% of them being expected to disappear with the current generation. Here we develop a simple model of language competition that explains historical data on the decline of Welsh, Scottish Gaelic, Quechua (the most common surviving indigenous languages in the Americas) and other endangered languages. A linguistic parameter that quantifies the threat of language extinction can be derived from the model and may be useful in the design and evaluation of language-preservation programmes.

Previous models of language dynamics have focused on the transmission and evolution of syntax, grammar or other structural properties of a language itself. In contrast, the model we describe here idealizes languages as fixed, and as competing with each other for speakers. For simplicity, we also assume a highly connected population, with no spatial or social structure, in which all speakers are monolingual.

Consider a system of two competing languages, X and Y, in which the attractiveness of a language changes with both its number of speakers and its perceived status—a parameter that reflects the social or economic opportunities afforded to its speakers. Suppose an individual converts from Y to X with a probability per unit of time, of $P_{yx}(x,s)$, where $x$ is the fraction of the population speaking X, and $0 \leq x \leq 1$ is a measure of X’s relative status. A minimal model for language change is therefore

$$\frac{dx}{dt} = yP_{yx}(x,s) - xP_{xy}(x,s)$$

where $y = 1 - x$ is the complementary fraction of the population speaking Y at time t. By symmetry, interchanging languages should yield the same transition probability as a swap in the fraction of speakers and relative status; thus $P_{yx}(x,s) = P_{xy}(1-x,1-s)$. We also assume that no one will adopt a language that has no speakers ($P_{yx}(0,s) = 0$) or no status ($P_{xy}(x,0) = 0$), and that $P_{yx}$ is smooth and monotonically increasing in both arguments.

These mild assumptions simply that equation (1) generally has three fixed points. Of these, only $x = 0$ and $x = 1$ are stable. The model therefore predicts that two languages cannot coexist stably — one will eventually drive the other to extinction.

To test our model, we collected data on the number of speakers of endangered languages in 42 regions of Peru, Scotland, Wales, Bolivia, Ireland and Alsace-Lorraine, four instances of which are shown in Fig. 1. We fit the model’s solutions to the data, assuming transition functions of the forms $P_{yx}(x,s) = a x^s$ and $P_{xy}(x,s) = c(1-x)^s(1-y)$.

Unexpectedly, the exponent was found to be roughly constant across cultures, with $a = 0.31 \pm 0.05$ (mean ± standard deviation; further details are available from the authors).

Of the remaining parameters, status, s, is the most relevant linguistically; it could serve as a useful measure of the threat to a given language. Quechua, for example, still has many speakers in Huanuco, Peru, but its low status is driving a rapid shift to Spanish, which leads to an unfortunate situation in which a child cannot communicate with his or her grandparents.

Contrary to the model’s stark prediction, bilingual societies do, in fact, exist. But the histories of countries where two languages coexist today generally involve split populations that lived without significant interaction, effectively in separate, monolingual societies. Only recently have these communities begun to mix, allowing language competition to begin.

So what can be done to prevent the rapid disintegration of our world’s linguistic heritage? The example of Quebec French demonstrates that language decline can be slowed by strategies such as policy-making, education and advertising, in essence increasing an endangered language’s status. An extension to equation (1) that incorporates such control on $s$ through active feedback does indeed show stabilization of a bilingual fixed point.

Daniel M. Abrams, Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, New York 14853, USA
e-mail: dma32@cornell.edu

Competing financial interests: declared none.

Figure 1 The dynamics of language death. Symbols show the proportions of speakers over time for: a, Scottish Gaelic in Sutherland, Scotland; b, Quechua in Huanuco, Peru; c, Welsh in Monmouthshire, Wales; d, Welsh in all of Wales, from historical data (blue) and a single modern census (red). Fitted curves show solutions of the model in equation (1) with parameters $c$, $s$, $a$ and $s(x)$ estimated by least absolute-values regression. Where possible, data were obtained from several population censuses collected over a long timespan; otherwise, a single recent census with age-structured data was used (although errors are introduced, the size of which are reflected in the differing fits in d). Using the fraction of Catholic masses offered in Quechua in Peru as an indicator, we reconstructed an approximate history of the language’s decline.