

Language shift, bilingualism and the future of Britain's Celtic languages

Anne Kandler^{1,*}, Roman Unger² and James Steele¹

¹*AHRC Centre for the Evolution of Cultural Diversity, Institute of Archaeology, University College London, 31–34 Gordon Square, London WC1H 0PY, UK*

²*Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz, Germany*

'Language shift' is the process whereby members of a community in which more than one language is spoken abandon their original vernacular language in favour of another. The historical shifts to English by Celtic language speakers of Britain and Ireland are particularly well-studied examples for which good census data exist for the most recent 100–120 years in many areas where Celtic languages were once the prevailing vernaculars. We model the dynamics of language shift as a competition process in which the numbers of speakers of each language (both monolingual and bilingual) vary as a function both of internal recruitment (as the net outcome of birth, death, immigration and emigration rates of native speakers), and of gains and losses owing to language shift. We examine two models: a basic model in which bilingualism is simply the transitional state for households moving between alternative monolingual states, and a diglossia model in which there is an additional demand for the endangered language as the preferred medium of communication in some restricted sociolinguistic domain, superimposed on the basic shift dynamics. Fitting our models to census data, we successfully reproduce the demographic trajectories of both languages over the past century. We estimate the rates of recruitment of new Scottish Gaelic speakers that would be required each year (for instance, through school education) to counteract the 'natural wastage' as households with one or more Gaelic speakers fail to transmit the language to the next generation informally, for different rates of loss during informal intergenerational transmission.

Keywords: language competition; Celtic; Gaelic; Welsh; reaction–diffusion; intergenerational transmission

1. INTRODUCTION

'Language shift' is the process whereby members of a community in which more than one language is spoken abandon their original vernacular language in favour of another. Membership of a community defined by its language selectively facilitates and inhibits interaction, enables entry into social contracts and cooperative exchange and gives access to a reservoir of accumulated and linguistically encoded knowledge. In cases of language contact, therefore, people are inevitably confronted with difficult choices about which language they wish or need to speak. The major driver of language shift is the decision to abandon a more local or less prestigious language, typically because the target of the shift is a language seen as more modern, useful or giving access to greater social mobility and economic opportunities (McMahon 1994; Mufwene 2001; Brenzinger 2006). In the modern era, nation states, globalization and selective migration (Boyd & Richerson 2009) have been potent forces of language standardization and

of minority language endangerment or extinction. The expected scale of global loss of contemporary linguistic diversity over the next 50–100 years is immense (Krauss 1992; Nettle & Romaine 1999).

The basis of the phylogenetic explanation in historical linguistics is that human populations have in the past undergone expansions, with the mechanism of expansion being local population increase, fissioning and spatial relocation of some fraction of that population. Subsequent divergence from a common linguistic root is driven by the natural tendency of languages to diversify under the combined effects of inherited mutation and isolation by distance, with the diversification accelerated by physical barriers to interaction and by effective population size-related sampling effects (drift). If the fissioning is kin-structured, with sub-populations splitting off who already share idiosyncratic linguistic features by virtue of membership of the same part of the larger interaction network (for instance, because of kinship ties), then these effects will be accelerated (Croft 2003).

There is a substantial body of recent scientific literature on the large-scale correlations between genetic and linguistic variation, much of it influenced by the integrative approach of Cavalli-Sforza and his collaborators who see the two systems as coevolving as a result of population expansion and splitting, geographical isolation and parental transmission (the

* Author for correspondence (a.kandler@ucl.ac.uk).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rstb.2010.0051> or via <http://rstb.royalsocietypublishing.org>.

One contribution of 14 to a Theme Issue 'Cultural and linguistic diversity: evolutionary approaches'.

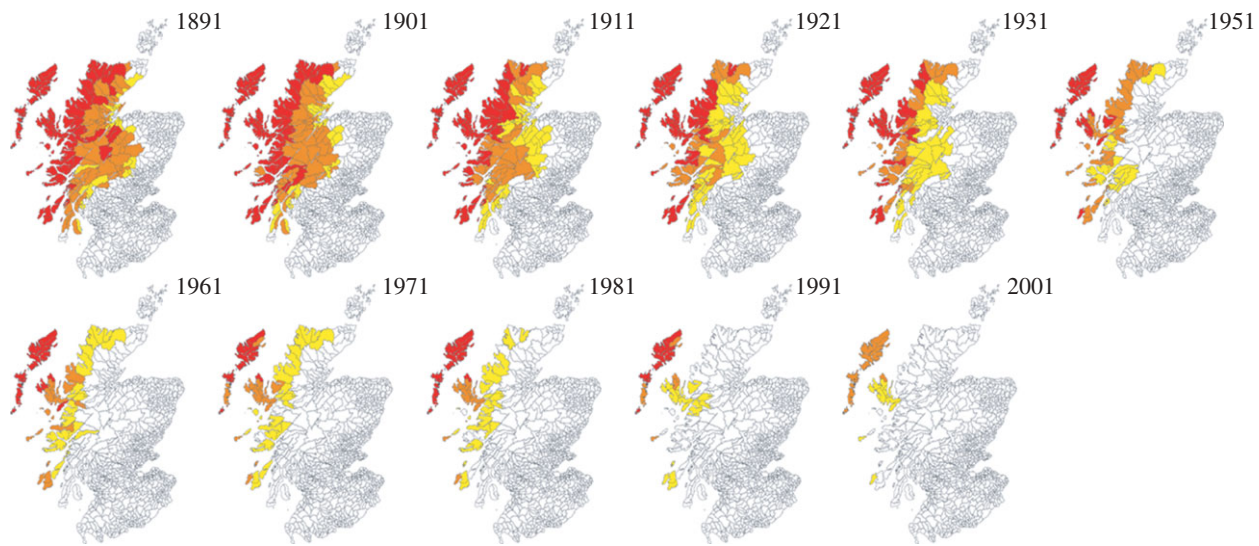


Figure 1. Percentages of Gaelic speakers (mono- and bilingual) in Scotland in successive census years, 1891–2001. Data for civil parishes: 1891–1971 from Withers (1984, pp. 227–234); 1981 from Withers (1988, p. 40); 1991–2001 from General Register Office for Scotland (2005, table 3). Red, 75–100% Gaelic speaking; orange, 50–74.9% Gaelic speaking; yellow, 25–49.9% Gaelic speaking; white, less than 25% Gaelic speaking.

latter being the sole mechanism of genetic inheritance and, they would argue, the predominant mechanism of linguistic inheritance in small-scale societies; e.g. Cavalli-Sforza *et al.* 1988, 1992). In prehistoric archaeology, such demographic interpretations of cultural macroevolution are familiar from the much-debated farming/language dispersal hypothesis for the spatial spread and diversification of languages such as those of the Bantu, Austronesian or Indo-European groups (Diamond & Bellwood 2003).

However, accepting a role for the dispersal of its speakers in the initial spread of these major linguistic groupings does not preclude contact-induced language change and recruitment into the speaker population by language shift, either at the time of initial spread or subsequently. In fact, Campbell (2006, p. 2) suggests that empirically, in terms of the likelihood of finding complete gene–language congruence in language contact situations, ‘All of the following are attested (‘no’ here means ‘little or no’):

- (1) no linguistic admixture—no genetic admixture
- (2) no linguistic admixture—genetic admixture
- (3) linguistic admixture—no genetic admixture
- (4) linguistic admixture—genetic admixture

where much work in language–gene correlation has tended to privilege (1) [...], linguists expect (1) least, with (4) perhaps the most common.’

In this paper, we focus on the social processes underlying Campbell’s scenarios (2) and (4) to frame the following questions: when does branch pruning on a linguistic phylogeny (language death) reflect local population extinction, and when does it reflect a purely cultural extinction process with the descendants of its speakers simply transferring to a different branch of the language tree (language shift)? Thomason (2001; cf. McMahon & McMahon 2005, pp. 78–79) has suggested that the effects of language contact can be arranged on a continuum from

contact-induced language change (which may involve just non-basic vocabulary elements, or basic vocabulary and structural features, depending on the level of contact and of bilingual interaction), to extreme language mixture (involving pidgins, creoles and mixed languages; cf. Mufwene 2008), to language death, with people abandoning one language outright and shifting to adopt another. Tree-building methods attempt to reconstruct the aspects of similarity and divergence that are due to conservative transmission with mutation-based modification. However, the phylogenetic approach ignores the important role of selective cultural migration (or shifting between branches) in determining the extinction rates of different branches of such trees.

In this paper, we will describe our recent work on language competition and language shift using the example of the recent history of Britain’s major Celtic languages. We emphasize the extreme lack of congruence between genetic and linguistic trees that results from language shift, and stress that the frequent shift of individuals between branches of a linguistic tree is not only a contemporary phenomenon (for discussion, see Steele & Kandler 2010).

The historical shifts to English by Celtic language speakers of Britain and Ireland are particularly well-studied examples of language competition for which good census data exist for the most recent 100–120 years in many areas where Celtic languages were once the prevailing vernaculars (see figure 1 for a visualization of the Scottish Gaelic census data). Some of the earliest fieldwork on language death was done in communities where Scottish Gaelic was endangered or dying out (MacKinnon 1977; Dorian 1981). The last monolingual speakers of Cornish died in the late seventeenth century, although their language survived locally among Cornish–English bilinguals until the end of the nineteenth century. On the Isle of Man, the last native speaker bilingual in Manx died in the 1970s. Following the extinction of these informal

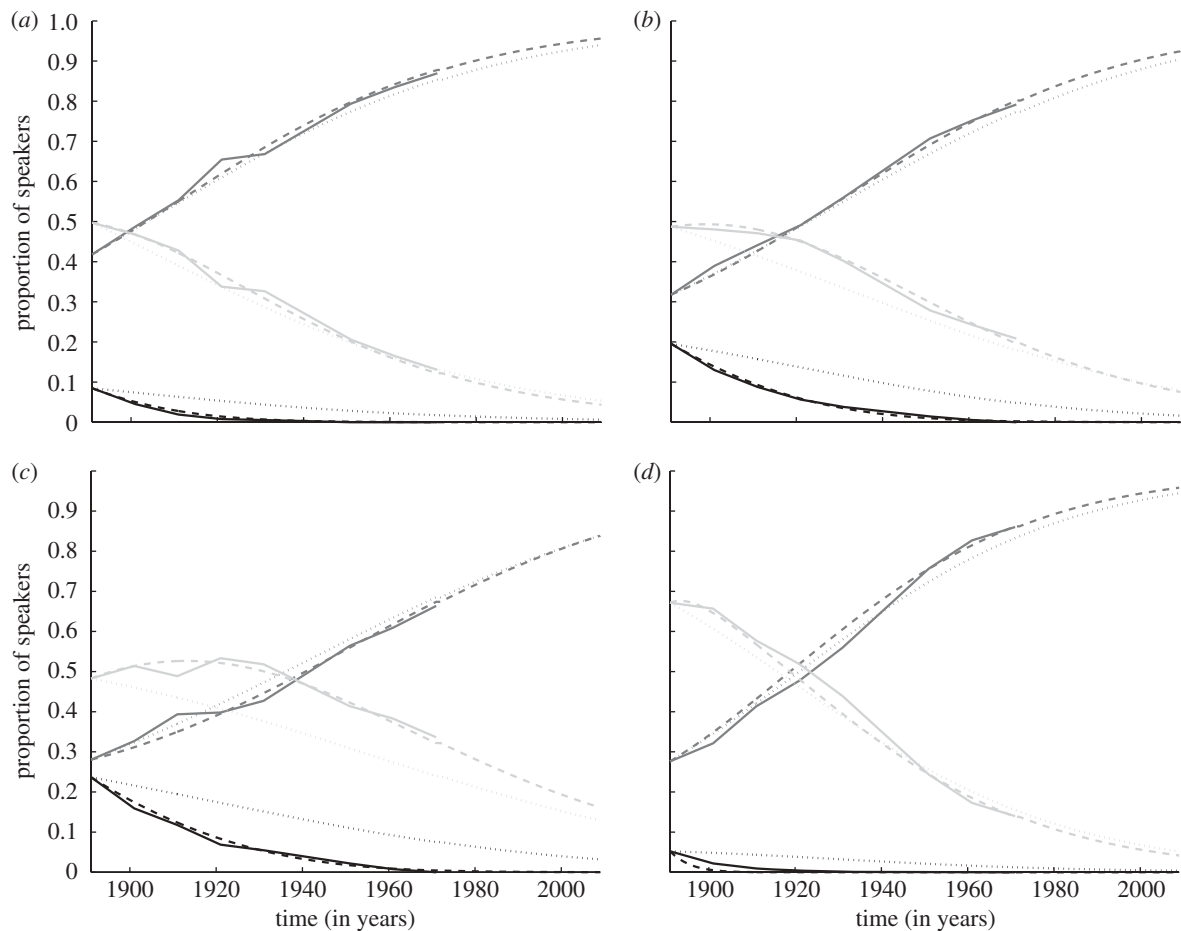


Figure 2. Frequencies of the three sub-populations in the four Scottish Highland counties for the time period 1891–2010. Empirical data (solid lines) and predictions of model (5.1) under the assumptions $c_{31} = c_{32}$ and $c_{13} = c_{12}$ (dotted lines) and $c_{31} \neq c_{32}$ and $c_{13} \neq c_{12}$ (dashed lines) of the frequencies of Gaelic (black), bilingual (light grey) and English (grey) speakers in (a) Argyll, (b) Inverness, (c) Ross and Cromarty and (d) Sutherland over time.

within-household transmission pathways, Cornish and Manx are now subjects of local revival efforts to bring these languages back into the community via schools, print and broadcast media, the arts and traditional community events. In Scotland and Wales, the original Gaelic- and Welsh-speaking populations were more numerous and the pattern of decline has been more influenced by local geographical factors. During the twentieth century, Welsh remained widely spoken, and even in 1961 it was still possible to traverse Wales from north to south without leaving a parish in which 80 per cent or more of the residents spoke Welsh (Aitchison & Carter 1985). This is despite long-term pressures for Anglicization owing to interventions such as the Act of Union of 1536, which incorporated Wales into the realm of the English monarch and included a stipulation that ‘no Person or Persons that use the Welsh Speech or Language shall have or enjoy any Manor Office or Fees within the Realm of England, Wales or Other the King’s Dominion’ (Bowen 1908), and much later, promotion of the use of English in schools to eradicate Welsh from the industrial heartlands after rural–urban migration had created self-contained Welsh-speaking communities in the coalfields (Commissioners of Inquiry into the State of Education in Wales 1847). However, in the last 50 years, monolingual Welsh speakers declined towards extinction (in 1981 there were 21 283 Welsh

monolinguals recorded in the official census, 0.8% of the total population), and a vigorous programme of Welsh language revitalization since the 1970s has been targeted at creating the conditions for stable bilingualism¹ (Jones 1993).

In Scotland, by late mediaeval times, Gaelic was the main language of the Highlands and western islands, with Scots (descended from the Old Northumbrian dialect of Old English) and English prevailing in the Lowlands. This division appears to have been reinforced by a contrast between these two regions in their social structure, marriage and migration patterns (with the clan system predominating in the Highlands): the subsequent breakdown of the geographical ‘niche’ for Scottish Gaelic is closely linked to the political and economic dominance of actors to the south, and their interference with the Highlands’ political and economic systems. Drastic demographic changes (the eighteenth–nineteenth century ‘Highland clearances’) and the associated establishment of English as the language of education and advancement were associated with increasing rates of Gaelic-to-English language shift (Murdoch 1996). The late stages of this shift process can be reconstructed from census records. Figure 2 (solid lines) shows the change in the proportions of monolingual English and Gaelic speakers and bilinguals for the counties of Argyll, Inverness, Ross and Cromarty

and Sutherland during the time period 1891–1971. These four counties are seen as the ‘core land’ of the Gaelic language (*Gaidhealtachd*): in 1891, 73 per cent of all Scotland’s Gaelic speakers were located among the 8 per cent of Scotland’s population that lived in these ‘Highland Counties’, covering the mainland Highlands and the Western Isles. By 2001, economic adversity in Highland areas, the ‘pull’ factor of economic opportunity in urban, industrial areas and Gaelic revivalism in the Lowlands have produced a substantial Gaelic presence in the Lowlands, with only 52 per cent of all Gaelic speakers resident in the wider *Gaidhealtachd* (where only 6.5% of Scotland’s population now live), and 48 per cent residing in the rest of Scotland (figure 1). The absolute numbers of Gaelic speakers in Scotland have however declined through this period, from about 250 000 in the 1891 census of Scotland to about 65 000 in the most recent (2001) census. Of these, the majority were always bilingual in Gaelic and English, with the last census record of Gaelic monolinguals finding fewer than 1000 still alive in 1961. Recent revitalization efforts have included the establishing of Gaelic-medium pre-school and primary school units (MacKinnon 1993) and the development of Gaelic-medium broadcasting (Murdoch 1996). In 2005, the Gaelic Language (Scotland) Act was passed by the Scottish Parliament, providing a planning framework for a number of additional shift-reversal measures, while Comhairle nan Eilean Siar, the Western Isles Council, has adopted Gaelic as its primary language.

2. MATHEMATICAL MODELS OF LANGUAGE SHIFT

(a) *Basic model*

We model the dynamics of language shift as a competition process in which the numbers of speakers of each language vary as a function both of internal recruitment (as the net outcome of birth, death, immigration and emigration rates of native speakers), and of gains and losses owing to language shift. Mathematical work on language shift dynamics² has been stimulated by Abrams & Strogatz (2003), who proposed a simple two-language competition model in which the outcome (extinction of one or other language) is determined by the strength of innate attraction to the higher status language and by the initial conditions (with preferential attachment—the nonlinear effect of initial concentrations on shift rates—capable of driving the higher status language to extinction when its speakers are rare). Our own basic model is very different. In addition to the status-related shift term, we model the changing sizes of speaker sub-populations as the balance of births and deaths, and of immigration and emigration, and we model a bilingual transition state.³ There is no process of preferential attachment—absolute rates of shift are a simple linear function of sub-population sizes. In our basic model of the shift process, the variables u_1 , u_3 and u_2 represent the sizes of the two monolingual and the bilingual sub-populations and the parameter c_{ij} represents the strength of the innate attraction of language i to speakers currently situated in

sub-population j (for a graphical representation of the shift process, see electronic supplementary material, figure S2a). Each sub-population also recruits internally by reproduction, spatial dispersal and long-distance migration, which is modelled as a reaction–diffusion process with logistic growth to a carrying capacity K and (in spatially explicit formulations) with diffusion of speakers between adjacent locations (§5). This model of language shift leads inevitably to the extinction of one or other monolingual sub-population, followed by the extinction of the language itself in the bilingual community. In the absence of the bilingual transition state, extinction would always be the fate of the lower status language; however, including the bilingual transition state fundamentally changes the dynamics. The less attractive or lower status language can now prevail, provided that its speakers have an initial numerical advantage that outweighs their language’s intrinsic status disadvantage. In formal terms, and if overall population size is stable, this outcome requires that there are initially few enough monolinguals in the high-status language, and therewith enough pressure on them to become bilingual, for it to always hold that $c_{12}u_2 < c_{31}u_3$ (where u_1 defines the frequency of the sub-population speaking the high-status language). These dynamics are analysed in more detail in the electronic supplementary material.

(b) *Diglossia model*

Many advocates of the preservation of endangered languages as living languages have promoted strategies in which the objective is stable societal bilingualism, by creating or preserving essential social domains (perhaps quite prestigious domains, such as political fora) in which the endangered language is the preferred or only acceptable medium of communication. Although such reversal strategies all require some measure of planned intervention to revive demand for skill in the endangered language, language planners typically cite as precedent the apparent stability of language coexistence in cases of diglossia (e.g. Fishman 1991). Diglossia, in the strict sense, refers to situations where the mother tongue of the community is used in everyday (low status) settings, but another language (or another form of the vernacular language) is used in certain high-status domains typically involving religious ceremonies, or written transactions in societies with low levels of literacy (Ferguson 1959; Hudson 2002). Language coexistence is possible in such diglossic situations because the demand for the high-status language is specific to social context. Language shift, in the sense of our basic model, relates instead to situations where the high-status language is associated with entire social identities that are seen as desirable and worthy of emulation. Such situations are not compatible with stable language coexistence, because the languages are competing as the medium of communication in all social contexts.

To consider the effects of the creation and maintenance of segregated and complementary sociolinguistic domains, in each of which both languages are differentially preferred as the medium of communication, we have examined a second model in which bilingualism

Table 1. Fitted shift coefficients for the basic model with $c_{31} = c_{32}$ and $c_{13} = c_{12}$, respectively, $c_{13} \neq c_{32}$ and $c_{31} \neq c_{12}$.

	Wales	Scottish Highlands	Argyll	Inverness	Ross and Cromarty	Sutherland
shift from Celtic to bilingual and/or to monolingual English (c_{13})	0.025	0.03	0.03	0.035	0.03	0.035
shift from English to bilingual and/or to monolingual Celtic (c_{31})	0.005	0	0	0	0	0
shift from Celtic-only to bilingual (c_{13})	0.06	0.07	0.115	0.1	0.12	0.075
shift from bilingual to English-only (c_{12})	0.02	0.025	0.03	0.03	0.025	0.035
shift from English-only to bilingual (c_{31})	0	0	0.005	0	0.005	0
shift from bilingual to Celtic-only (c_{32})	0	0	0	0	0.005	0

is no longer simply the transitional state for households moving between alternative monolingual states. Superimposed on the basic shift dynamics, there is an additional demand for the endangered language as the preferred medium of communication in some restricted sociolinguistic domain, and this demand persists regardless of the numbers of speakers of the endangered language until that number becomes very small (at which point the demand ceases; for a graphical representation of the shift process, see electronic supplementary material, figure S2*b*). This additional dynamics creates a steady reverse flow of monolingual speakers in the dominant language who enter or re-enter the bilingual sub-population. Because this second model allows for demand for both languages, each in its own preferred domain, bilingualism is now a stable final state; we now find that a wider range of extinction and coexistence states is possible, depending on the strength of the various in- and out-fluxes between the three sub-populations. We can now model—for any given case of well-advanced language shift—the rates of acquisition of skills in the endangered language that would be required from monolingual households fluent in the dominant language for shift reversal to take off. These dynamics are also analysed in more detail in the electronic supplementary material.

3. RESULTS

Using our basic model, we have estimated the strengths of the competitive advantage driving language shift from Scottish Gaelic to English in Highland Scotland (1891–2001), and from Welsh to English in Wales (1901–2001). We fitted⁴ the model to official census data (see electronic supplementary material, S1 ‘Data’ for more details) for these time periods. Historical census data on language use will include some ‘noise’ owing to inaccurate answers (for instance, owing to the perceived social status implications of self-classification into a particular category), and to changes in the phrasing of the questions in successive censuses. To avoid over-fitting (where the model fits the noise in the data as well as the significant trends), we initially reduced the model’s degrees of freedom by assuming the parameter constellation $c_{31} = c_{32}$ and $c_{13} = c_{12}$. The results are shown in table 1 and figures 2 and 3 (dotted lines). Our basic model captures well the general dynamics of the language shift process (the decrease in

the Gaelic and Welsh monolingual and bilingual sub-populations and the increase in the English monolingual sub-population). Table 1 (top two rows) gives the estimated values for the shift coefficients. These show that while the Celtic monolingual sub-populations were not able to attract a significant number of English speakers or bilinguals (cf. $c_{31} = c_{32} = 0-0.005$), the shift from the Celtic monolingual to the bilingual sub-populations and from the bilingual to the English monolingual sub-populations happened at high rates owing to the competitive advantage of the dominant language (cf. $c_{13} = c_{12} = 0.025-0.035$). Additionally, the competitive advantage for English speakers in Highland Scotland was greater than in Wales.

However, the fitted curves in figures 2 and 3 also suggest that the parameter constellations given in table 1 generally overestimate the Celtic monolingual sub-population and slightly underestimate the bilingual sub-population. Therefore, we also fitted the basic model with constellations in which $c_{31} \neq c_{32}$ and $c_{13} \neq c_{12}$ (so that, for example, the balance of competitive advantage driving the shift from monolingual Welsh to bilingualism can be different from that driving the shift from bilingualism to monolingual English). The results are illustrated by the dashed lines in figures 2 and 3 and by the values for the competition coefficients in table 1 (bottom four rows). The fit is improved,⁵ and table 1 (bottom four rows) shows that the key to improvement in fit lies in the increase in the shift parameter from Celtic-only to bilingual (c_{13}). All other coefficients stay roughly constant. Celtic monolinguals were more affected by the status difference between English and Celtic than were bilinguals. This implies that the priority was to learn the high-status language and not to abandon the Celtic language. Bilinguals tended to stay bilingual longer than Gaelic speakers stayed monolingual. We also found that the fit of the basic shift model is most sensitive to changes in the coefficient c_{12} , implying that small changes in the rate of shift from bilingualism to monolingual English may result in significantly changed competition dynamics.

Figure 3 highlights a further deviation of the basic model’s predictions from the census data for the very recent period in Wales. In the above results, we have assumed constant shift coefficients over time (i.e. that the ‘environment’ for language competition does not change within the considered time period). However, political, social and/or economic changes can

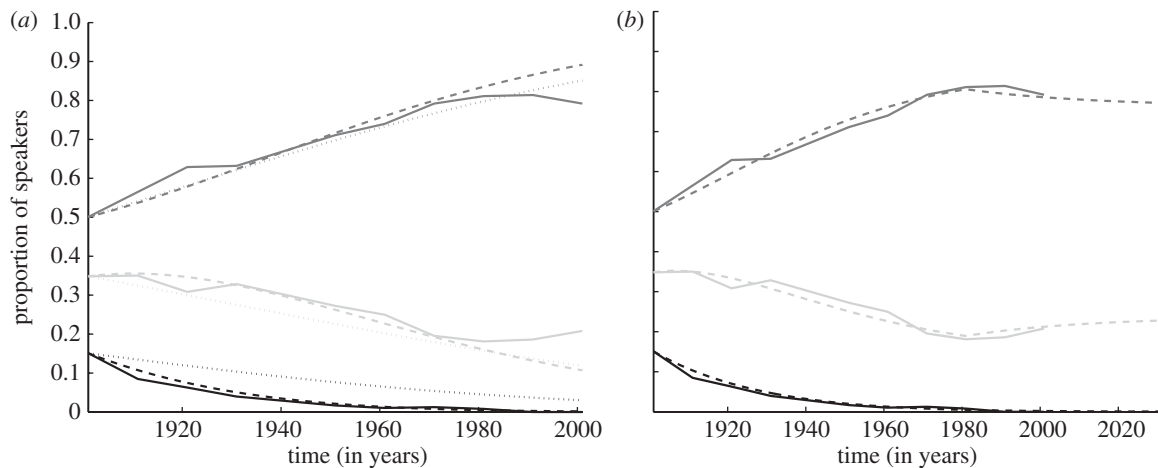


Figure 3. Empirical and projected frequencies of the three sub-populations in Wales for the time period 1901–2001. Empirical data (solid lines) and predictions of model (5.1) under the assumptions $c_{31} = c_{32}$ and $c_{13} = c_{12}$ (dotted lines) and $c_{31} \neq c_{32}$ and $c_{13} \neq c_{12}$ (dashed lines) of the frequencies of Welsh (black), bilingual (light grey) and English (grey) speakers. (a) Prediction of model (5.1) with parameters given in table 1 (bottom four rows) and (b) prediction of model (5.2) with the same c -values and $w_1 = 0.005$ and $w_3 = 0$ for the time period 1901–1971 and $w_1 = 0.01$ and $w_3 = 0$ for the time period 1971–2001.

lead to a change in the sociolinguistic environment and consequently to a change in the competition dynamics. Figure 3a shows that the basic model with time-independent shift coefficients captures well the dynamics of Welsh–English language competition until about 1971, but not the change in the competition dynamics that is observable more recently. During the last 40 years, Welsh language-planning initiatives and legislation have led to several maintenance interventions that were able to alter the shift dynamics. The decline in the bilingual sub-population appears to have been reduced or halted, leading to a stable coexistence condition. This new situation must be explained using our diglossia model, since it is inconsistent with the basic model (in which bilingualism is assumed to be a transitional state and not a final stable state). Figure 3b shows fitted curves for the diglossia model in relation to the Welsh census data. Here, time-dependent coefficients are crucial to capture the change in the competition dynamics. The fitted values of the same shift coefficients as just discussed in the basic model for the two time periods before and after 1970 show that the language-planning initiatives resulted in an increased ‘force’ for English speakers to learn Welsh (w_1 , being a measure of the strength of this force for the time period 1971–2001), which means sociolinguistic domains in which Welsh is the advantageous language have been created and supported. Figure 3b also projects ahead the fate of Welsh–English bilingualism if the ‘environment’ stays the same, and indicates that Welsh is then preserved in the bilingual sub-population. However, at present, this is due to the maintenance activities of the planners creating an influx of English monolinguals into the bilingual sub-population that balances the continuing ‘organic’ loss of bilingual households to English monolingualism owing to low levels of intergenerational transmission of Welsh within the home.

How might the experiences of language planners intervening to limit the shift from Welsh to English be used to ‘save’ the Gaelic language in the Scottish Highlands? We applied the diglossia model to the

Gaelic–English situation and asked how strong an intervention would need to be (in other words, how many English monolinguals have to learn Gaelic per year) in order to alter the shift dynamics. We note that the number of Gaelic monolinguals is now effectively zero, so that the term w_3u_3 does not play a role in the competition dynamics: therefore, we set $w_3 = 0$. We obtain that $w_1 = 0.0035$ is sufficient to stabilize the bilingual population at its current level (cf. figure 4). This implies that roughly 860 English speakers have to become bilingual every year (based on a Highland population of about 315 000 individuals). However, the coexistence between the bilingual and the English-speaking sub-populations depends in this case entirely on the planners’ initiatives and on legislation. Intervention strategies may prove much more successful if the rate of intergenerational transmission of the bilingual strategy could be increased as well. Thus, for example, the number of English monolinguals required to learn Gaelic each year could drop down to roughly 440 if the rate of intergenerational transmission of Gaelic at home could be increased (c_{12} from 0.025 to 0.0125). This means that beside the 440 new recruits to bilingualism, roughly 340 more children who live in bilingual households would have to be raised in both languages to stabilize the bilingual population at the current level. These numbers indicate that an increase in the rate of intergenerational transmission is a highly effective language maintenance strategy, although one that is also harder to achieve in practice.

4. DISCUSSION

The current linguistic ‘extinction crisis’ is expected to decimate global cultural diversity. As outlined in this paper in the Gaelic–English example, most of the recent language extinction events are caused by language shift rather than by the extinction of the population speaking this language. This inevitably results in an increasing divergence between the

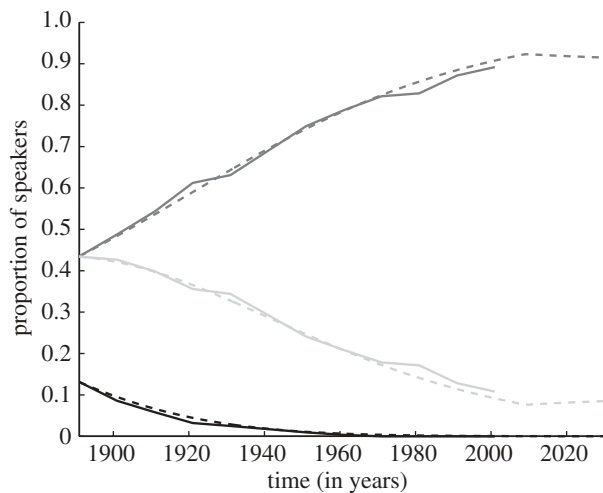


Figure 4. Empirical and projected frequencies of the three sub-populations in the Scottish Highlands for the time period 1901–2030 with assumed intervention after 2009. Empirical data (solid lines) and predictions of model (5.1) until 2009 and model (5.2) after 2009 (dashed lines) of the frequencies of Gaelic (black), bilingual (light grey) and English (grey) speakers. Parameter values for model (5.1) are given in table 1 (bottom rows); after 2009, a diglossia model with the same c -values and $w_1 = 0.0035$ and $w_3 = 0$ is assumed.

transmission histories represented in genetic and in linguistic trees. What provokes shift is not cultural selection acting on grammatical or prosodic potential, but people shifting between two competing languages because of their associated social ecologies. There may of course be some associated variation in expressive potential relating to those ecologies (for example, in terms of specialized vocabulary); in the case of the Gaelic-speaking fishing communities of East Sutherland, the death of whose language was studied closely by Dorian, problems arose when their niche was irrevocably altered:

Now I should stress here that fisherfolk Gaelic was not lexically impoverished. The trouble was, that like any other strictly local speech form deeply associated with a traditional lifestyle, the richness of the lexicon was chiefly connected with their own specialized way of life. There wasn't much connected with the sea or with boats that they didn't have a word for, and they had a lot of weather terms that reflected the importance of decisions about whether to put to sea or not. When I acquired the dialect I learned the names of more varieties of seaweed than I had ever known existed, the names for parts of a rabbit snare, and the term for an egg that emerged from the hen without an exterior shell. But, not surprisingly, there were no local words for the parts of a car or for the national health service

(Dorian 2006, p. 7).

The solution used by Gaelic speakers was to adopt the English words as loanwords; what drives the shift process is not the available specialized lexicon, but the wider contrast in social and economic potential that participation in one or other linguistic community opens up.

We have not considered here the reasons why a phylogenetic model might explain the historical evolution of languages in terms of their basic vocabularies; rather, we have shown that language shift (seen as selective migration between branches of a language tree) is another significant force in cultural evolution, one which may also—in some circumstances—serve as a mechanism of cultural selection acting on alternative systems of economic practices and social norms. Language planners are active in many situations attempting to reverse or modify this shift process, while academic linguists increase their effort to record details of representative samples of these endangered languages (most of which have no or minimal written corpora) before they disappear. With the English–Gaelic and the English–Welsh case studies, we analysed two different scenarios. While the 2001 census showed that the decline in numbers of Scottish Gaelic speakers had not yet been halted, census data for Wales in the same year showed that Welsh seemed to be being maintained at stable levels in a bilingual sub-population. Analysis of our diglossia model has shown that the key language-planning issues for maintenance of an endangered language are (i) to create or support social domains in which the endangered language is the preferred or only acceptable medium of communication and (ii) to increase the rate of intergenerational transmission of the endangered language. Other important dimensions of language maintenance are the creation of economic incentives (e.g. jobs created to implement language-planning-related initiatives and which themselves require skills in the endangered language), and the establishment of corpora of written texts in the endangered language as a cultural archive and as a medium of continuing cultural self-expression. Without stabilizing a sustainable level of intergenerational transmission, language planners will have to rely on constant interventions in formal public domains (e.g. in the school curriculum) to counter the continuing outflux from bilingualism by individual households. An indication of one cause of this background outflux from Gaelic-speaking bilingualism can be found in the 2001 Scottish census data (General Register Office for Scotland 2005): 70 per cent of children aged 3–15 years speak Gaelic in households in which a married or co-habiting couple both speak Gaelic, while the percentages are only 18 per cent if the male partner alone speaks Gaelic, and 27 per cent if the female partner alone speaks Gaelic. This is the current reality of intergenerational transmission in an environment where languages compete with very unequal external advantages. The success of current planning interventions in reversing language shift and preserving Welsh and Gaelic as living languages will be assessed when results are available from the next Welsh and Scottish censuses in 2011.

5. MODEL AND METHODS

(a) *Basic model*

We examine the dynamics of language shift as a spatially dependent competitive process using the

reaction–diffusion system:

$$\left. \begin{aligned} \frac{\partial u_1}{\partial t} &= d_1 \Delta u_1 + a_1 u_1 \left(1 - \frac{u_1}{K - (u_2 - u_3)} \right) \\ &\quad - c_{31} u_3 u_1 + c_{12} u_2 u_1 \\ \frac{\partial u_2}{\partial t} &= d_2 \Delta u_2 + a_2 u_2 \left(1 - \frac{u_2}{K - (u_1 - u_3)} \right) \\ &\quad + (c_{13} + c_{31}) u_1 u_3 - (c_{12} u_1 + c_{32} u_3) u_2 \\ \frac{\partial u_3}{\partial t} &= d_3 \Delta u_3 + a_3 u_3 \left(1 - \frac{u_3}{K - (u_1 - u_2)} \right) \\ &\quad - c_{13} u_1 u_3 + c_{32} u_2 u_3, \end{aligned} \right\} \quad (5.1)$$

with the boundary conditions $\partial u_i / \partial n = 0$, $x \in \partial D$, $i = 1, 2, 3$, where $\partial / \partial n$ is the outer normal derivation. The time- and space-dependent variables u_1 and u_3 stand for the frequencies of monolingual speakers of Language A and Language B, respectively, whereas u_2 describes the frequency of bilingual speakers of both languages. The terms $\partial u_i / \partial t$, $i = 1, 2, 3$, indicate the rate of change in these frequencies over time. The terms on the right-hand side of the equations in system (5.1) describe the changes in the frequency of speakers in each of the three sub-populations u_1 , u_2 and u_3 . The components $a_i u_i (1 - u_i / (K - (u_j + u_k)))$ define the internal reproductive rates, which represent coupled biological and cultural reproduction within each sub-population. This is usually modelled (as shown here) as a logistic process with intrinsic rate of increase a_i . The variable K stands for the carrying capacity of the environment and defines an upper limit to the size of the whole population regardless of the languages spoken, which imposes the condition $u_1 + u_2 + u_3 \leq K$ for any time t (i.e. we assume that our human sub-populations must compete for a common resource base). For a detailed analysis of the relevance of this self-limiting term, see Kandler & Steele (2008). The mobility of speakers of each sub-population in space within the modelled region is modelled by the diffusion terms $d_i \Delta u_i$. The language shift dynamics is modelled in system (5.1) by the frequency-dependent conversion term $c_{ij} u_i u_j$. The coefficients c_{13} and c_{31} represent the likelihood of language shift causing speakers to become bilingual based on the differential prestige or attractiveness of the two competing languages. Following Minnett & Wang (2008), we assume $c_{31} = \tilde{c}_{31} s$ and $c_{13} = \tilde{c}_{13} (1 - s)$, where the variable s describes the social status differences between the two languages on a scale from 0 to 1. The higher the status of a language, the higher is the likelihood of being the preferred target of shifting. The coefficients \tilde{c}_{13} and \tilde{c}_{31} model the likelihood that monolinguals will respond to these status differences by learning the other language. Language shift cannot happen by passing directly from being monolingual in one language to being monolingual in the other language, but must involve a bilingual transition state. The bilingual sub-population therefore recruits from both monolingual sub-populations at a rate $(c_{13} + c_{31}) u_1 u_3$. In turn, bilinguals shift to being monolingual in one or other

language at rates $c_{12} u_1 u_2$ (representing the loss to monolingualism in Language A) and $c_{32} u_3 u_2$ (representing the loss to monolingualism in Language B). The coefficients c_{12} and c_{32} represent the likelihood of bilingual speakers then becoming monolingual in each of the two languages. In real life, this transition back to monolingualism happens when bilingual parents choose to raise their children monolingually, or when speakers reared as bilinguals in bilingual households abandon one of their languages during their lifetime. We define the overall balance of competitive advantage to speaking each language on the base of the conversion rates: for example, fluency in Language A can be assumed to be more advantageous if it holds that $c_{31} < c_{13}$ and $c_{12} > c_{32}$. This implies that when the monolingual sub-populations are compared, monolinguals of Language A are less likely to become bilingual, and bilinguals are more likely to shift to speaking only Language A.

(b) *Diglossia model*

To model the effects of planned interventions on stable societal bilingualism, we generalize the basic language shift model (5.1) by incorporating a simplified concept of (extended) diglossia. While in the majority of social domains the shift mechanisms of the basic model apply, diglossia pertains to some restricted social domain in which the balance of competitive advantage differs from that which drives the main shift process. We now assume that the language that tends to lose its speakers in the majority of social domains can nonetheless be the preferred language in a more restrictive domain or set of domains. We therefore generalize the basic model (5.1) by allowing for the possibility that in such domains language use is determined by an alternative set of social norms or prescriptions. This assumption results in a change of the shift dynamics. In our basic model, the reason for monolinguals becoming bilingual is simply that it is a required transition state on the way to being monolingual in the other language. In the diglossia model, we now also allow people to become bilingual as the preferred ‘end state’. If monolinguals of the disadvantaged language want to participate in domains where the advantaged language is required (such as higher education or ‘global’ businesses), they need to learn that second language. This is modelled by the term $w_3 u_3$ where w_3 measures the demand for participation in these domains. However, as long as the low-status language is still used, there is also the possibility that the low-status language is the required language in some domains (such as small ‘local’ businesses or service encounters). Such domains might also be created by political interventions (e.g. legislation that requires use of the endangered local language in a specific set of contexts). This dynamics is incorporated into model (5.2) by the term $w_1 r(u_2) u_1$. Again, w_1 models the demand of participation of monolinguals of the high-status language in these domains and the function $r(u_2)$ controls for the frequent use of the

low-status language.⁶ These considerations lead to our second model:

$$\left. \begin{aligned} \frac{\partial u_1}{\partial t} &= d_1 \Delta u_1 + a_1 u_1 \left(1 - \frac{u_1}{K - (u_2 - u_3)} \right) \\ &\quad - w_1 r(u_2) u_1 - c_{31} u_3 u_1 + c_{12} u_2 u_1 \\ \frac{\partial u_2}{\partial t} &= d_2 \Delta u_2 + a_2 u_2 \left(1 - \frac{u_2}{K - (u_1 - u_3)} \right) + w_3 u_3 \\ &\quad + w_1 r(u_2) u_1 + (c_{13} + c_{31}) u_1 u_3 \\ &\quad - (c_{12} u_1 + c_{32} u_3) u_2 \\ \frac{\partial u_3}{\partial t} &= d_3 \Delta u_3 + a_3 u_3 \left(1 - \frac{u_3}{K - (u_1 - u_2)} \right) \\ &\quad - w_3 u_3 - c_{13} u_1 u_3 + c_{32} u_2 u_3. \end{aligned} \right\} (5.2)$$

Both systems of partial differential equations are implemented in C++ and solved numerically using the finite-element method.

(c) Data

Data for Scottish Gaelic speakers are from the decennial census of Scotland (see electronic supplementary material). The first census to enumerate Gaelic speakers was that of 1881, but only from 1891 were data gathered separately on numbers of Gaelic monolinguals and Gaelic–English bilinguals (in all cases, among those aged 3 years or older). After 1961, no data were collected on the incidence of Gaelic monolinguals, as these were assumed by that time to be approaching extinction. From 1891 until 1971, the census enumerations were collated and analysed on the basis of the old county divisions (the Highland counties of the *Gaidhealtachd* included Argyll, Inverness, Ross and Cromarty and Sutherland). From 1981 onwards, these counties were subsumed into new administrative units. The new Highland region includes most of Inverness, the majority of Ross and Cromarty, Sutherland and a small portion of Argyll; it also includes Caithness and Nairn, and a small portion (5%) of Moray. The remaining portions of Inverness and Ross and Cromarty make up the new Western Isles region, while the remainder of Argyll and Bute is included in the new Strathclyde region. To document trends in Gaelic speaking in the wider *Gaidhealtachd* as a single area, we have therefore also collated data from the pre-1981 counties of Argyll, Bute, Caithness, Inverness, Nairn, Ross and Cromarty and Sutherland, and compared this with collated data from the most recent three censuses for the administrative regions of Highlands and Western Isles, and for the unitary council area of Argyll and Bute within the modern Strathclyde administrative region. In the most recent 2001 census, enumeration was extended to include those who stated that they could understand Gaelic but not speak it; we have excluded these instances in order to retain comparability with the earlier records (which enumerate only those with Gaelic-speaking skills).

Data for Welsh speakers are from the decennial census of England and Wales. The first census to

enumerate Welsh speakers (monolinguals and Welsh–English bilinguals) was that of 1891, but there was some dissatisfaction with the phrasing of the language question and with the definition of an age cut-off for young children. From 1901, the enumeration was limited to those aged 3 years or older. After 1981, no data were collected on the incidence of Welsh monolinguals, as these were assumed by that time to be approaching extinction. In the most recent 2001 census, enumeration was extended to include those who stated that they could understand Welsh but not speak it; we have excluded these instances in order to retain comparability with the earlier records (which enumerate only those with Welsh-speaking skills). The model-fitting procedure is described in the electronic supplementary material, S3.

We would like to thank Peter Austin and April McMahon for discussions of this work during its development. We also thank our colleagues at the AHRC Centre for the Evolution of Cultural Diversity (www.cecd.ucl.ac.uk) for creating a hospitable interdisciplinary environment in which to develop models of cultural dynamics. This work was funded by an AHRC Phase Two Research Centre grant, and by a Leverhulme Early Career Fellowship for A.K.

ENDNOTES

¹In this context, Jones studied two Welsh dialects and pointed out that ‘dialect death in Wales may involve the divesting of regional features and an approximation to a commonly accepted uniform variety that is being proliferated throughout the speech community’ (Jones 1998, p. 2).

²See Patriarca & Leppänen (2004), Mira & Paredes (2005), Stauffer & Schulze (2005), Pinasco & Romanelli (2006), Castelló *et al.* (2007), Kandler & Steele (2008), Schulze *et al.* (2008), Minett & Wang (2008), Kandler (2009) and Patriarca & Heinsalu (2009).

³See also Baggs & Freedman (1990, 1993) and Wyburn & Hayward (2008, 2009).

⁴We estimate the growth and diffusion parameters a_i and d_i from demographic data. In order to determine the shift rates c_{ij} , we calculate the best fit (in a quadratic sense) of model (5.1) to the empirical census data, using the pre-estimated parameters a_i and d and leaving the competition terms free to vary (see electronic supplementary material, S3, ‘Model fitting’, for further information).

⁵Improvement is quantified in terms of a smaller quadratic distance between the model outcome and the empirical data.

⁶The function $r(u_2)$ is assumed to be 1 if u_2 is sufficiently large but tends to zero if the frequency of the bilingual population becomes too small (e.g. $r(u_2)$ can be modelled as a step function).

REFERENCES

- Abrams, D. M. & Strogatz, S. H. 2003 Modelling the dynamics of language death. *Nature* **424**, 900. (doi:10.1038/424900a)
- Aitchison, J. & Carter, H. 1985 *The Welsh language, 1961–1981: an interpretative atlas*. Cardiff, UK: University of Wales Press.
- Baggs, I. & Freedman, H. I. 1990 A mathematical model for the dynamics of interactions between a unilingual and a bilingual population: persistence versus extinction. *Math. Sociol.* **16**, 51–75. (doi:10.1080/0022250X.1990.9990078)
- Baggs, I. & Freedman, H. I. 1993 Can the speakers of a dominated language survive as unilinguals? A mathematical model of bilingualism. *Math. Comput. Model.* **18**, 9–18. (doi:10.1016/0895-7177(93)90122-F)
- Bowen, I. 1908 *The statutes of Wales*. London, UK: T. Fisher.

- Boyd, R. & Richerson, P. J. 2009 Voting with your feet: payoff-biased migration and the evolution of group beneficial behavior. *J. Theoret. Biol.* **257**, 331–339. (doi:10.1016/j.jtbi.2008.12.007)
- Brenzinger, M. 2006 Language maintenance and shift. In *The encyclopedia of language and linguistics* (ed. K. Brown), pp. 542–548, 2nd edn. In *Society and language*, vol. 6 (ed. R. Mesthrie). Oxford, UK: Elsevier.
- Campbell, L. 2006 Languages and genes in collaboration: some practical matters. In *Language and Genes: an Interdisciplinary Conf., Santa Barbara, 8–10 September 2006*. Berkeley, CA: University of California.
- Castelló, X., Loureiro, L., Eguiluz, V. M. & Miguel, M. 2007 The fate of bilingualism in a model of language competition. In *Advancing social simulation: the first World Congress* (eds S. Takahashi, D. Sallach & J. Rouchier), pp. 83–94. New York, NY: Springer.
- Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. & Mountain, J. 1988 Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc. Natl Acad. Sci. USA* **85**, 6002–6006. (doi:10.1073/pnas.85.16.6002)
- Cavalli-Sforza, L. L., Minch, E. & Mountain, J. 1992 Coevolution of genes and languages revisited. *Proc. Natl Acad. Sci. USA* **89**, 5620–5624. (doi:10.1073/pnas.89.12.5620)
- Commissioners of Inquiry into the State of Education in Wales 1847 *Reports of the Commissioners of Inquiry into the state of education in Wales*, vols 1–3. London, UK: William Clowes and Sons / HMSO.
- Croft, W. 2003 Social evolution and language change. See <http://www.unm.edu/~wcroft/Papers/SocLing.pdf>.
- Diamond, J. & Bellwood, P. 2003 Farmers and their languages: the first expansions. *Science* **300**, 597–603. (doi:10.1126/science.1078208)
- Dorian, N. C. 1981 *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia, PA: University of Pennsylvania Press.
- Dorian, N. 2006 Using a private-sphere language for a public-sphere purpose: some hard lessons from making a television documentary in a dying dialect. Paper presented at Bryn Mawr College, 16 March. See www.brynmawr.edu/emeritus/gather/Dorian.doc.
- Ferguson, C. A. 1959 Diglossia. *Word* **15**, 325–340.
- Fishman, J. A. 1991 *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages*. Clevedon, UK: Multilingual Matters Ltd.
- General Register Office for Scotland. 2005 *Scotland's Census 2001, Gaelic Report*. Edinburgh, UK: General Register Office for Scotland.
- Hudson, A. 2002 Outline of a theory of diglossia. *Int. J. Sociol. Lang.* **157**, 1–48.
- Jones, R. O. 1993 The sociolinguistics of Welsh. In *The Celtic languages* (ed. M. J. Ball), pp. 536–605. London, UK: Routledge.
- Jones, M. C. 1998 *Language obsolescence and revitalisation: linguistic change in two sociolinguistically contrasting Welsh communities*. Oxford, UK: Oxford University Press.
- Kandler, A. 2009 Demography and language competition. *Hum. Biol.* **81**, 181–210
- Kandler, A. & Steele, J. 2008 Ecological models of language competition. *Biol. Theor.* **3**, 164–173. (doi:10.1162/biot.2008.3.2.164)
- Krauss, M. 1992 The world's languages in crisis. *Language* **68**, 4–10.
- MacKinnon, K. 1977 *Language, education and social processes in a Gaelic community*. London, UK: Routledge and Kegan Paul.
- MacKinnon, K. 1993 Scottish Gaelic today: social history and contemporary status. In *The Celtic languages* (ed. M. J. Ball), pp. 491–535. London, UK: Routledge.
- McMahon, A. 1994 *Understanding language change*. Cambridge, UK: Cambridge University Press.
- McMahon, A. & McMahon, R. 2005 *Language classification by numbers*. Oxford, UK: Oxford University Press.
- Minett, J. W. & Wang, W. S.-Y. 2008 Modelling endangered languages: the effects of bilingualism and social structure. *Lingua* **118**, 1945.
- Mira, J. & Paredes, Á. 2005 Interlinguistic similarity and language death dynamics. *Europhys. Lett.* **69**, 1031–1034. (doi:10.1209/epl/i2004-10438-4)
- Mufwene, S. S. 2001 *The ecology of language evolution*. Cambridge, UK: Cambridge University Press.
- Mufwene, S. S. 2008 *Language evolution*. London, UK: Continuum.
- Murdoch, S. 1996 *Language politics in Scotland*. Aberdeen, UK: Aberdeen University Scots Leid Quorum.
- Nettle, D. & Romaine, S. 1999 *Vanishing voices: the extinction of the world's languages*. Oxford, UK: Oxford University Press.
- Patriarca, M. & Heinsalu, E. 2009 Influence of geography on language competition. *Physica A* **388**, 174–186. (doi:10.1016/j.physa.2008.09.034)
- Patriarca, M. & Leppänen, T. 2004 Modeling language competition. *Physica A* **338**, 296–299. (doi:10.1016/j.physa.2004.02.056)
- Pinasco, J. P. & Romanelli, L. 2006 Coexistence of language is possible. *Physica A* **361**, 355–360. (doi:10.1016/j.physa.2005.06.068)
- Schulze, C., Stauffer, D. & Wichmann, S. 2008 Birth, survival, and death of languages by Monte Carlo simulation. *Commun. Comput. Phys.* **3**, 271–294.
- Stauffer, D. & Schulze, C. 2005 Microscopic and macroscopic simulation of competition between languages. *Phys. Life Rev.* **2**, 89–116. (doi:10.1016/j.plrev.2005.03.001)
- Steele, J. & Kandler, A. 2010 Language trees ≠ gene trees. *Theor. Biosci.* **129**, 223–233. (doi:10.1007/s12064-010-0096-6)
- Thomason, S. G. 2001 *An introduction to language contact*. Edinburgh, UK: Edinburgh University Press.
- Withers, C. W. J. 1984 *Gaelic in Scotland, 1698–1981: the geographical history of a language*. Edinburgh, UK: John Donald.
- Withers, C. W. J. 1988 *Gaelic Scotland: the transformation of a cultural region*. London, UK: Routledge.
- Wyburn, J. & Hayward, J. 2008 The future of bilingualism: an application of the Baggs and Freedman model. *J. Math. Soc.* **32**, 267–284. (doi:10.1080/00222500802352634)
- Wyburn, J. & Hayward, J. 2009 OR and language planning: modeling the interaction between unilingual and bilingual populations. *J. Operat. Res. Soc.* **60**, 626–636. (doi:10.1057/palgrave.jors.2602600)