Kia kaua te reo e rite ki te moa, ka ngaro:

Do not let the language suffer the same fate as the moa

Tessa Barrett-Walker[1,2], Michael J. Plank[2,3], Rachael Ka'ai-Mahuta[3,4,6], Daniel Hikuroa[3,5,6], Alex James[2,3]

1. Aotahi - School of Māori and Indigenous Studies, University of Canterbury, Christchurch, New Zealand.
2. School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand.
3. Te Pūnaha Matatini, New Zealand Centre of Research Excellence.
4. Te Whare o Rongomaurikura Centre for Language Revitalisation, Auckland University of Technology, Auckland, New Zealand.
5. Department of Māori Studies, University of Auckland, Auckland.
6. Ngā Pae o te Māramatanga, New Zealand's Māori Centre of Research Excellence.

Abstract

More than a third of the world's languages are currently classified as endangered and more than half are expected to go extinct by 2100. Strategies aimed at revitalising endangered languages have been implemented in numerous countries, with varying degrees of success. Here, we develop a new model regarding language transmission by dividing the population into defined proficiency categories and dynamically quantifying transition rates between categories. The model can predict changes in proficiency levels over time and, ultimately, whether a given endangered language is on a long-term trajectory towards extinction or recovery. We calibrate the model using data from Wales and show that the model predicts that the Welsh language will thrive in the long term. We then apply the model to te reo Māori, the Indigenous language of New Zealand, as a case study. Initial conditions for this model are estimated using New Zealand census data. We modify the model to describe a country, such as New Zealand, where the endangered language is associated with a particular subpopulation representing the Indigenous People. We conclude that, with current learning rates, te reo Māori is on a pathway towards extinction, but identify strategies that could help restore it to an upward trajectory.

Keywords: dynamical model; endangered language; Indigenous language; language learning; language transmission; language revitalisation.

Introduction

It is estimated that the world has approximately 7000 languages, of which over 2800 are endangered (Alliance for Linguistic Diversity, n.d.). Between 50% and 90% of the world's languages are expected to be extinct by 2100 (Krauss 1992; Austin and Sellabank, 2011). Many languages have become endangered as a result of colonisation, associated with a long history of oppression of Indigenous Peoples and cultural subjugation, in a pattern repeated around the world since the early 17th century, including Indigenous languages in North America, South America, Africa and Australia (Krauss 1992; McConvell and Thieberger, 2010; Moseley, 2010). Te reo Māori, the language of the Indigenous People of Aotearoa New Zealand, is one example that follows the pattern. The number of speakers of te reo Māori was systematically reduced in the colonial period via numerous means including prohibition of the use of te reo Māori in schools, stigmatisation of Māori culture, severance of intergenerational transmission, the Tohunga Suppression Act of 1907, and government relocation policy (Waitangi Tribunal, 1986; New Zealand Government, 2010; Ka'ai-Mahuta, 2011). Te reo Māori is intrinsically linked to all that Māori culture encompasses (Moorfield, 2006; Henare, 2011), and suffered greatly to an apparent verge of extinction in the 1960s and 70s (Moorfield, 2006; New Zealand Government, 2010). There is a widespread sense that to lose one's language is to lose one's culture and identity (e.g. McCarthy, 1997; Glynn, 1998; Ka'ai-Mahuta, 2011).

More recently, some countries have begun to recognise the effects of their history of cultural oppression on language endangerment. Language revitalisation strategies have been implemented, with varying degrees of success, for endangered languages around the world. These include such measures as: affording official language status (Waitangi Tribunal, 1986); integration into school curricula (Welsh Government, 2017); bilingual or language-immersion schooling (Ka'ai-Mahuta, 2011; Welsh Government, 2017; Te Puni Kōkiri, 2018); mentor-apprentice programs (Jenni et al., 2017); provision or support of language media (Timms, 2013); strategies to re-establish and maintain intergenerational transmission (Forrest, 2017; Te Mātāwai, 2018); and strategies to promote engagement with an endangered language (Te Puni Kōkiri, 2018). In 2018, the New Zealand Government set a national target that, by 2040, one million New Zealanders (approximately 18% of the population) will be able to speak at least basic te reo Māori, and it will be the primary language for 150,000 Māori (Te Puni Kōkiri, 2018). The ultimate success or failure of such strategies in revitalising an endangered language will depend on their effectiveness in reversing the centuries-long trend of decline in the number of users of that language.

Several modelling studies have explored how the grammar and syntax of languages have evolved over time (Kroch, 1989; Niyogi and Berwick, 1997; Nowak et al., 2002; Wang and Minett, 2005). Existing modelling of trends in the number of speakers of a language commonly derives from a competitive perspective, an approach popularised by Abrams and Strogatz (2003). This model assumes that two languages compete with each other in a population of monolingual speakers, where increases in the number of speakers and perceived status of one language increase its "attractiveness". The assumption is that, ultimately, one language will dominate and push the other to the point of extinction. Extensions to this model consider a wider range of societal conditions and parameters including bilingualism, intergenerational transmission and conversations heard (Wang and Minett, 2005; Minett and Wang, 2008; Fernando et al., 2010). Modelling based on these works has been applied to te reo Māori by Ruckstuhl and Wright (2012), where the amount of te reo Māori heard, family contribution and community contribution were recognised as influential factors. These models

all consider the relationship between a dominant language and a minority language, rather than the dynamics of the number and proficiency of users of the minority language itself and how revitalisation strategies may influence the language's survival. It is certainly true that language extinction is typically caused by other languages (Ka'ai-Mahuta, 2011; Rewi and Rewi, 2015) and one-way bilingualism, whereby speakers of an endangered language typically also speak the majority language but not vice versa, can be a threat (Fitzpatrick, 1987; Beardsmore, 2003). Nevertheless, most endangered languages will only be able to survive and flourish by coexisting with other languages via bilingual or multilingual speakers and multiple language domains (Fishman, 1965; Fishman, 1971; Spolsky 2004), rather than outcompeting them (McConvell and Thieberger, 2010). This requires that languages are accorded equal socioeconomic status (Fitzpatrick, 1987; Fishman, 2001; Minett and Wang, 2008).

With any language, there are varying degrees of proficiency of language use. The Common European Framework of Reference for Languages (CEFR) (Council of Europe, n.d.) categorise language proficiency into three broad user levels: basic, independent and proficient. Basic users are able to communicate, with assistance and in simple terms, matters related to their immediate self, actions and environment. Independent users can understand and communicate experiences, explanations and opinion. Proficient users can communicate fluently, express themselves in all circumstances and understand complex or implicit meanings within text and conversation (Council of Europe, n.d.).

In this study, we explore factors affecting the survival of a minority language by developing a dynamic model which divides the population by level of language proficiency based on the CEFR categories. The approach has some similarities with the ZePA model of language revitalisation, which classifies people into one of three states: zero, passive and active (Higgins and Rewi, 2014; Rewi and Rewi, 2015). In addition to language proficiency, these states encompass people's values, attitudes and behaviour. We take a simplified approach of defining the model categories solely on an individual's level of proficiency with the language. However, unlike the ZePA model, we attempt to quantify transitions between categories as a result of language acquisition and basic demographic processes. We do this via a system of differential equations that describe changes in the numbers of people in each category. The model is a coarse-scale model for processes affecting the number of people with varying levels of language proficiency at a population level. The model neglects variations among individuals in the myriad of factors that can affect learning outcomes and in domains of language use. Nor does it attempt to account for historical processes and events. Instead, the aim is to predict whether the language is on trajectory towards extinction or recovery, and to explore how different national interventions may affect the language's trajectory.

We parameterise the model using data from the recent resurgence of the Welsh language. Significant development in bilingual and Welsh-medium education and the presence of the language throughout the public and private sectors have positively contributed to an increase in the number of Welsh speakers (Welsh Government, 2015). The Welsh Government's series of Welsh language use in Wales surveys (Jones, 2012; Welsh Government, 2015; Welsh Language Board, 2008) contain categorical proficiency data and the three available surveys provide a useful set of longitudinal data over an extended time period. The context for the model in te reo Māori in New Zealand is different from that in Wales, partly because New Zealand contains relatively well defined Māori and non-Māori subpopulations. We recalibrate the model for te reo Māori using the limited volume of census data available. We use the model to assess the current trajectory of the language and to estimate the rates of learning that would be needed, relative to the situation in Wales, to achieve language revitalisation.

Informed by the model, we evaluate the likelihood of meeting government targets by 2040, and compare the potential strategies for promoting revitalisation.

Model development

We divide the population into three categories according to their level of language proficiency: basic, independent and proficient. In reality, language proficiency does not fall neatly into discrete categories, but is a continuous and multifaceted spectrum. However, in most cases, data on language proficiency are not sufficient to calibrate anything other than a very simple model. Hence, we limit the model to three proficiency categories. Therefore, although the CEFR definition of the basic category is those with some level of capability with language, we define this category to include those with no capability at all. We consider a system of differential equations for the proportion of the population in the basic $B(t)$, independent $I(t)$, and proficient $P(t)$ categories, where $B(t) + I(t) + P(t) = 1$. We assume that the total population size is fixed by assigning each category identical per capita birth and death rates r. Individuals transition between categories as a result of language acquisition. We describe the model for this process below (see below).

We assume that once a proficiency level is reached it cannot be lost, i.e. language skills do not deteriorate with time. We are confident this assumption holds for those with te reo Māori as their first language and, given the drivers for te reo Māori acquisition as a second language, is reasonable for that group as well. Although there might be some minor proficiency loss, it would not result in movement between proficiency categories.

Intergenerational transmission. We assume that children of individuals in the basic and independent categories are born into the basic category. A fraction α of the children of proficient users are born into the proficient category, the remaining fraction $1 - \alpha$ are born into the basic category. This assumption is an approximate model for intergenerational language transmission that assumes these children acquire language proficiency from birth. This simplifies the model because it avoids the need to explicitly model a parent-infant learning route separately from other learning routes. The parameter α is the intergenerational transmission rate, which is a quantity that is commonly estimated in studies of endangered languages (e.g. Alliance for Linguistic Diversity, n.d.; Welsh Government, 2015; King and Cunningham, 2017).

Language acquisition. Individuals can progress from one proficiency level to the next by acquiring language competency. This usually depends on direct interaction with proficient language users, either as formal language teachers or as informal social contacts. To model this learning route, we assume that individuals in either the basic or independent category progress to the next category at a per capita rate that is proportional to the size of the proficient category, i.e. the number of people capable of acting as teachers. This results in the system of differential equations:

$$
\begin{array}{ccccc}
\text{rate of change} & = & \text{births} & \text{learning} & \text{deaths} \\
\dot{B} & = & r(B + I + (1-\alpha)P) & -\beta_{BI}BP & -rB, \\
\dot{I} & = & & +\beta_{BI}BP - \beta_{IP}IP & -rI, \\
\dot{P} & = & \alpha rP & +\beta_{IP}IP & -rP,
\end{array}
$$

where β_{BI} and β_{IP} are learning rate parameters representing the rate of transition from the basic to the independent and from the independent to the proficient category respectively. These parameters

will be affected by a number of important variables, including the quality of teachers, the motivation and commitment of learners, access to educational facilities and programmes, and the relative roles of different formal and informal learning routes. We do not attempt to account for these factors individually, but instead take a coarse-scale approach of exploring the effect of variations in the learning rate parameters β_{BI} and β_{IP}.

As the total population size is assumed to be fixed, the system of equations above can be reduced to two equations for I and P:

$$\dot{I} = \beta_{BI}(1 - I - P)P - \beta_{IP}IP - rI,$$
$$\dot{P} = \alpha rP + \beta_{IP}IP - rP.$$

Depending on parameter values, this dynamical system has either one or three steady states. The zero steady state $(I, P) = (0,0)$ exists and is stable for all parameter values. When the system is in this steady state, all individuals are in the basic category and we interpret this as extinction or near-extinction of the language. The non-zero steady states (I^*, P^*) are given by $I^* = (1 - \alpha)r/\beta_{IP}$ and P^* is the solution of the quadratic equation

$$\beta_{BI}P^{*2} - \left(\beta_{BI}\left(1 - \frac{(1-\alpha)r}{\beta_{IP}}\right) - (1-\alpha)r\right)P^* + \frac{(1-\alpha)r^2}{\beta_{IP}} = 0.$$

Hence, the system has a saddle-node bifurcation: below the bifurcation the only steady state is the zero steady state and the language is guaranteed to go extinct from any initial condition; above the bifurcation there is a stable steady state with $I^*, P^* > 0$ in addition to the stable zero steady state, and the eventual fate of the language depends on the initial conditions. Figure 2A shows the dynamics of the model in the phase plane in the case where the positive steady state exists. The stable manifold of the saddle steady state (red line in Fig. 2A) divides the phase space into two regions: initial conditions below the red line will result in extinction of the language; initial conditions above the red line will result in long-term survival of the language. Figure 2B shows the saddle-node bifurcation dividing the (β_{BI}, β_{IP}) parameter space into: (i) a region where the language will go extinct regardless of initial conditions and; (ii) a region where survival of the language is possible depending on initial conditions.

Model calibration and predictions for Welsh

The model has three language-related parameters: the intergeneration transmission rate α and the two learning rate parameters β_{BI} and β_{IP}. The birth/death rate parameter r is purely demographic and not language-related; we set this as $r = 1/70\,\mathrm{yr}^{-1}$, which corresponds to an average life expectancy of 70 years.

We use data on Welsh language users to estimate the parameters α, β_{BI} and β_{IP} for that context. The 1991 Welsh census found that 18.7% of the population were able to speak Welsh and the 1992 Welsh Language Survey found that 61% of these individuals were fluent. Interpreting the fluent speakers as being in the proficient category, the non-fluent speakers as being in the independent category, and the non-speakers as being in the basic category, this corresponds to a system state of $(B, I\ P) = (0.81, 0.07, 0.11)$ in 1991. The 2004-06 Welsh language use survey found that 20.5% of the population aged three and over were able to speak Welsh and 12% were able to speak Welsh fluently (Welsh

Language Board, 2008). This corresponds to a system state of $(B, I\ P) = (0.795, 0.085, 0.12)$ in 2005. Finally the Welsh Government's report on Welsh language use in Wales 2013-15 (2015) categorised 24% of the population as Welsh speakers and 11% as fluent, giving $(B, I\ P) = (0.76, 0.13, 0.11)$ in 2014.

The intergenerational transmission rate has been estimated for several languages including Welsh. The report on Welsh language use in Wales 2013-15 (Welsh Government, 2015) defined the transmission rate as the proportion of children aged 3 to 4 who could speak Welsh in households where one or both parents were speakers, estimated to be 45% and 82% respectively. Assuming homogeneous mixing between different categories and 11% of the population in the proficient category, this corresponds to an average probability of intergenerational transmission of $\alpha = 0.47$.

The learning rate parameters β_{BI} and β_{IP} represent the per capita rate at which individuals in one proficiency category progress to the next category when taught by a proficient language user. Individuals can progress form the basic to the independent category in a variety of ways, ranging from mainstream (i.e. non-immersion) schooling, social interactions with language speakers, hearing the language on the radio and television, supermarket and road signs, and enrolment in evening classes. The transition from the independent to the proficient category is more likely come about through tertiary level study, immersion schooling or family environment. Even with available data, estimating these parameters directly is difficult, so we estimate them by fitting the model to the data on (B, I, P) at the three time points 1991, 2005 and 2014.

We use the 1991 data to define the initial condition (see Table 1) and we then use the 2005 and 2014 data to fit the model using least squares to give $\beta_{BI} = 0.0510\ \mathrm{yr}^{-1}$,$\beta_{IP} = 0.0969\ \mathrm{yr}^{-1}$ (Figure 3A). These fitted parameter values have a real-world interpretation: if almost all of the population were proficient, then 5.1% of the basic category would progress to independent per year, and 9.7% of the independent category would progress to proficient per year; when less than 100% of the population is proficient, these transition rates are reduced pro rata.

The model predicts that the revitalisation efforts will be successful and, in the long term, Wales will have over 70% proficient Welsh language users. Figure 2A shows that, at the estimated learning rates, almost any initial condition with more than 6% proficient speakers would lead to language revitalisation. The state of the language in 1991 (blue star Fig. 2A) was sufficient to put it on a trajectory towards successful revitalisation.

Model calibration for te reo Māori

We now consider how the model can be recalibrated for te reo Māori. In countries such as Wales, where an endangered language exists but is not associated with a particular subpopulation, it is reasonable to treat the population as homogeneously mixed with uniform intergenerational transmission and learning rates. However, a more nuanced approach is needed for countries, such as New Zealand where an endangered language is associated with the Indigenous People and a dominant language is typically associated with colonisation. Within New Zealand, Māori are a well-defined population group making up approximately 15% of the total population (StatsNZ, 2013c), with much higher rates of reo Māori use than in the population as a whole. In this section, we use data from the

2013 New Zealand census (StatsNZ, 2013a, 2013b) to estimate model parameter values and initial conditions for the population as a whole and for the Māori and non-Māori subpopulations

Data from the 2013 survey on Māori wellbeing, Te Kupenga (StatsNZ, 2013b) show that, in 2013, 5% of Māori were able to speak te reo very well, 5.6% were able to speak well, 12% were able to speak fairly well, 32.1% were not able to speak very well and 45.3% were able to speak no more than a few words or phrases. We interpret the 10.6% of those able to speak very well or well as the proficient category, the 12% of those able to speak fairly well as the independent category, and the remaining 77.4% as the basic category. Of the population as a whole, 3.7% were able to speak te reo Māori (StatsNZ, 2013a). We estimate that 1% are in the proficient category, 2.7% in the independent category, and the remaining 96.3% in the basic category. For the non-Māori subpopulation, it is estimated that only 0.6% can speak te reo Māori (StatsNZ, 2013a). We err on the side of generosity and assume these are all in the proficient category and that an additional 1% of the non-Māori subpopulation are in the independent category. Table 1 shows the model initial conditions for the overall population and the Māori and non-Māori subpopulations.

King and Cunningham (2017) inferred from the 2013 census data that the intergenerational transmission rate for te reo Māori is 43.6%, close to the Welsh rate. This result encompasses Māori and non-Māori speakers of te reo.

Data for estimating the learning rate parameters from changes in proficiency levels over time are not readily available for te reo Māori. We therefore derive estimates for these parameters from the estimated Welsh values by comparing rates of participation in language learning in English-medium schools and in language-immersion schools across the two countries. Although progression to the next level of language proficiency may also occur by other means, comparing rates of participation at school-level can be used to gain an approximate estimate for learning rates in one country relative to those in another country. Although proficiency levels may differ across countries, second language learning in English-medium schools can, broadly speaking, be regarded as progression from the basic to the independent category; immersion or language-medium education can be regarded as progression from the independent to the proficient category. Table 2 shows language learning participation rates in English-medium and language immersion schools in Wales and in New Zealand. Participation rates in New Zealand are substantially lower than in Wales, suggesting that the learning rate parameters for te reo Māori are lower than for Welsh. Table 2 shows estimates for the learning rate parameters for te reo Māori, derived by reducing the Welsh estimates pro rata by the relevant school participation rate.

Model predictions for te reo Māori

We first apply the model to New Zealand's population as a whole, and then to the Māori and non-Māori populations as two distinct, closed subpopulations. We assume that these three groups have the same intergenerational transmission rate α, but differ in their learning rates and initial conditions (see Table 1). To explore the potential effect of language revitalisation strategies, we present model results across a range of learning rates, using the values estimated in Table 2 as benchmarks.

Figure 4 shows the long-term, steady-state proficient population for the overall New Zealand population and the Māori and non-Māori subpopulations, for a range of learning rate parameters. The

Welsh learning rates are marked with a green star; the learning rates estimated for the New Zealand population are marked with yellow stars.

The model predicts that, when the New Zealand population is considered as a whole (Figure 4A), the language is on a path towards extinction at current learning rates. Even if learning rates can be improved to the Welsh values, the language is still on a downward trajectory; learning rates would need to be significantly higher than in Wales to enable long-term revitalisation of the language in the New Zealand population as a whole.

When the Māori subpopulation is considered separately (Figure 4B), the model predicts that, with current learning rates, the language is on a downward trajectory, but language revitalisation will be successful in the long-term if similar learning rates to those in Wales can be achieved. This is because the current proportion of Māori in the independent and proficient categories is higher, so the initial condition is more favourable than in the population as whole. Model predictions for the non-Māori subpopulation (Figure 4C) are similar to those for the population as a whole (Figure 4A) because non-Māori make up the majority of the total population.

The prediction that the language is on a downward trajectory in the Māori subpopulation is consistent with census data showing a decline on the proportion of Māori under the age of 24 who are able to speak the language from 21% in 2001 to 16% in 2013 (StatsNZ, 2013a; Te Puni Kōkiri, 2018). However, data from the survey of Māori wellbeing show an increase in the proportion of Māori under the age of 24 able to speak more than a few words or phrases from 42% in 2001 to 55% in 2013 (StatsNZ, 2013b). These differences could be due to different respondent populations or different categorisations of self-assessed proficiency levels. The survey of Māori wellbeing also showed that the biggest increases in the proportion able to speak the language were in the 25-34 and 35-44 age brackets (StatsNZ, 2013b). This suggests that adult learning is a significant pathway to language acquisition and therefore the true learning rate parameters may be higher than the estimates in Table 2, which were derived from statistics on school-level learning.

The calibrated model can be used to evaluate whether the language's current trajectory will meet stated targets. In 2018, the New Zealand government set two targets: (1) by 2040, 150,000 Māori will speak te reo Māori as a primary language; and (2) by 2040, one million New Zealanders will be able to speak at least basic te reo Māori (Te Puni Kōkiri, 2018). Target (1) translates into model variables as 14% of Māori in the proficient category. Achieving target (1) would require increasing the learning rates in the Māori subpopulation to a level approximately 50% higher than the Welsh learning rates. For target (2), "at least basic" means to being able to speak more than a few words or phrases (Te Puni Kōkiri, 2018). This does not map as easily onto model variables, as we defined the independent category as being able to speak "fairly well", which is a higher level of proficiency. If we assume that, of those with "at least basic" competency, half are in the basic category and half are in the independent or proficient categories, target (2) translates to 8.7% of the whole population in the independent or proficient categories. However, even with this relatively generous interpretation, the model predicts that achieving this target would require the learning rates for the population as a whole to be approximately 5 times higher than the Welsh rates. See Supplementary Information for details of target calculations and model predictions.

Model extension to two interacting population groups

It is likely that reality is somewhere between the scenario shown in Fig. 4A, which assumes a single population with fully homogeneous intermixing, and that in Fig. 4BC, which assumes two completely separate subpopulations with strictly homophilous mixing (no multi-ethnicity families). In addition, learning rates in each of the subpopulations are likely to differ significantly, given the influence of cultural exposure and access to language domains on language acquisition and development. Capacity for intergenerational transmission and demographic parameters such as birth and death rates may also differ between ethnicities. The degree of intermixing, as well any variation in transmission and learning rate parameters, will affect model predictions. Addressing this variation would not necessarily lead the model towards a competition approach, but would allow for a more realistic reflection of the development of multicultural and multi-language coexistence.

We propose a more realistic model that splits the total population into two subpopulations, each having three proficiency levels. In the New Zealand example, we denote the proportion of the Māori subpopulation in the three proficiency categories as B_M, I_M, P_M and the proportion of the non-Māori subpopulation as B_X, I_X, P_X. For simplicity, we assume that both subpopulations are of fixed size and that all individuals belong to the same subpopulation as their parents (i.e. homophilous mixing). However, we allow for some degree of mixing via proficient users of one subpopulation acting as teachers for either subpopulation. The degree of mixing of teachers is represented by an additional model parameter γ. This parameter takes a value between $\gamma = 0$, meaning that there is no mixing and learners only interact with teachers from their own subpopulation, and $\gamma = 1$, meaning there is fully homogeneous mixing of teachers across both subpopulations. This gives a set of four differential equation

$$\dot{I_M} = \beta_{BIM} B_M\big((1-\gamma)P_M + \gamma(MP_M + XP_X)\big) - \beta_{IPM} I_M\big((1-\gamma)P_M + \gamma(MP_M + XP_X)\big) - rI_M,$$
$$\dot{P_M} = \alpha r P_M + \beta_{IPM} I_M\big((1-\gamma)P_M + \gamma(MP_M + XP_X)\big) - rP_M,$$
$$\dot{I_X} = \beta_{BIX} B_X\big((1-\gamma)P_X + \gamma(XP_X + MP_M)\big) - \beta_{IPX} I_X\big((1-\gamma)P_X + \gamma(XP_X + MP_M)\big) - rI_X,$$
$$\dot{P_X} = \alpha r P_X + \beta_{IPX} I_X\big((1-\gamma)P_X + \gamma(XP_X + MP_M)\big) - rP_X,$$

where M and X are the proportion of the total population in the Māori and non-Māori subpopulations respectively. We assume that $M = 0.15$ and $X = 0.85$ and use the same initial conditions and intergenerational transmission rate as previously. Figure 5 shows the effect of mixing of teachers between the two subpopulations, as measured by the parameter γ, under the Welsh learning rates. As seen previously in Fig. 4BC, when there is no mixing of teachers ($\gamma = 0$, completely independent subpopulations), the model predicts that te reo Māori will thrive among Māori but die out in the non-Māori subpopulation. However, even a relatively small degree of mixing of teachers drastically reduces the predicted long-term number of proficient users among Māori, without producing any significant benefit for non-Māori. This is because the Māori subpopulation has a strong pool of teachers, due its relatively high proportion of proficient users. However, if these teachers are spread across the much larger total population, the pool of Māori teachers becomes diluted and is not able to sustain the language on an upward trajectory in either subpopulation.

Discussion

Prevalent themes among language revitalisation strategies include increasing knowledge, skill and proficiency across the population and creating the conditions that support the sustenance of intergenerational transmission (Timms, 2006; Fernando et al., 2010; Te Puni Kōkiri, 2018; Welsh Government, 2017). The strategies ultimately aspire to develop bilingual or multilingual populations where languages can coexist rather than compete (May et al., 2004; Rewi and Rewi, 2015). We have developed a dynamic model that considers a population in terms of levels of proficiency with an endangered language. This simple model can offer mathematical validation of language revitalisation strategies and aligns well with their key objectives and cultivation of bilingualism.

The model predicts that there are two possible long-term outcomes for an endangered language, depending on the average rate at which individuals progress from one proficiency level to the next. If the learning rates are too low, the language will be on a trajectory towards extinction; if the learning rates are sufficiently high, the language will be on a trajectory to revitalisation. The threshold value of the learning rates needed to ensure successful revitalisation depends on the number of current language users via the model's initial condition. We used data from the 2013 New Zealand census to determine these initial conditions. These values are approximate because the data come from a self-assessment of proficiency (Te Puni Kōkiri, 2018), which does not map perfectly to the proficiency categories in the model. Better quality data, or additional time points, would allow the threshold values for the learning rates to be determined more accurately. However, one advantage of the model is that the qualitative predicted outcome (extinction or revitalisation) is not sensitive to the exact quantitative values of parameters and initial conditions, provided these are not close to the threshold level.

We estimated the learning rate parameters in the model using a combination of data on trends in the number of speakers of Welsh, and the relative rates of participation in language learning at different school levels. With the estimated learning rates for te reo Māori, the model predicts that the language is currently on a downward trajectory within the Māori population and will not meet government targets by 2040 without a significant increase in learning rates to levels above those achieved in Wales.

We have used the model to explore language revitalisation strategies aimed at increasing the learning rate parameters of the model. These parameters represent the average rate at which individuals progress from one proficiency level to the next, when taught by a proficient language user. The education system is recognised as one of the most powerful levers available to government policymakers for the acquisition of an endangered language (Te Puni Kōkiri, 2018). Statistics broken down by age brackets (StatsNZ, 2013b) suggest that adult learning of te reo Māori is relatively strong and school-age learning lags behind. This suggests that strategies targeting learning at schools are likely to have the highest potential benefits. Possible government actions include the integration of the language into all early-childhood education; development of the quantity and quality of teachers; investment in language-immersion education as a crucial avenue for the development of proficiency.

The model shows that if proficient teachers, who are predominantly Māori, are spread across the whole population, this is detrimental to the language trajectory in the population as a whole, because the limited pool of teachers is spread too thinly. This suggests that, where capacity is limited either by resource constraints or by the number of teachers, learning among Māori should be prioritised initially. This is broadly consistent with government strategy (Te Puni Kōkiri, 2018; Te Mātāwai, 2018). If and when the language is determined to be on a healthy trajectory towards revitalisation among

the Māori population, resources and teachers can be diverted to promote learning across the whole population.

Another potential route to success would be a significant increase in the intergenerational transmission rate, which is the other key parameter of the model. However, in reality, this is likely to be a slowly varying parameter because an increase requires increased intergenerational participation and continuation of language use, and so progress may only be evident when considered over a time period which spans multiple generations (Forrest, 2017).

The model neglects a range of factors that are likely to be important for the rates of language use and language learning. This is appropriate given the limited data available, and allows the model to focus on comparing alternative strategies or interventions. The model has the potential to provide quantitative numerical targets for proficiency levels, which may be used to inform language revitalisation policy and implementation, and evaluate its effectiveness. Furthermore, it is likely that the social norms and values in which such strategies take place are at least as important as the strategies themselves. Attitudes towards te reo Māori among younger generations are shifting, and the government has specifically identified the importance of New Zealanders recognising the value of te reo Māori as a key element of the national identity (Te Puni Kōkiri, 2018). These factors may also be affected by changing demographics. For example, in New Zealand, the relative size of the Māori population is predicted to grow from 15% to 18% over the next 20 years (StatsNZ, 2013c).

Acknowledgements

The whakataukī (Māori proverb) in the title of this paper was adapted from He inoi kia mau te reo Māori (a prayer for the Māori language) by Tā Kīngi Matutaera Īhaka. The authors are grateful to Mary Boyce (Kaihautū Ako Māori, University of Canterbury) and Dr Abby Suszko (Kaiārahi Māori, University of Canterbury) for useful discussions about this research project. MJP and AJ were partly supported by Te Pūnaha Matatini.

References

1. Abrams, D.M. and Strogatz, S.H. (2003). Linguistics: modelling the dynamics of language death. Nature 424(6951), 900.
2. Alliance for Linguistic Diversity. (n.d.). About the endangered languages project. Retrieved from http://endangeredlanguages.com/about/
3. Austin, P.K. and Sellabank, J. (2011). The Cambridge Handbook of Endangered Languages. Cambridge University Press.
4. Beardsmore, H. B. (2003). Who's afraid of bilingualism? In: Dewaele, J.-M., Housen, A. and Wei, L. (Eds.) Bilingualism: Beyond Basic Principles, Multilingual Matters vol. 123, pp 10-27.
5. Council of Europe (n.d.). Common European framework of reference for languages: common reference levels. Retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale
6. Education Counts (2018). Māori language in education. Retrieved from https://www.educationcounts.govt.nz/statistics/maori-education/maori-in-schooling/6040
7. Fernando, C., Valijärvi, R. L. and Goldstein, R. A. (2010). A model of the mechanisms of language extinction and revitalization strategies to save endangered languages. Human Biology 82, 47-75.
8. Fishman, J. (1965). Who speaks what language to whom and when? La Linguistique, 1(2), 67-88.
9. Fishman, J. (1971). Sociolinguistic perspective on the study of bilingualism. In: Fishman, J. A., Cooper, R.L. and Newman, R.M. (Eds.) Bilingualism in the Barrio (Language Science Monographs Vol. 7), pp. 557-582. Bloomington: Indiana University Publications.
10. Fishman, J.A. (2001) Can threatened languages be saved? Reversing language shift, revisited: a 21st century perspective. Multilingual Matters, vol. 116.
11. Fitzpatrick, F. (1987). The open door: the Bradford bilingual project. Multilingual Matters vol. 26.
12. Forrest, W. (2017). The intergenerational transmission of Australian Indigenous languages: why language maintenance programmes should be family-focused. Ethnic and Racial Studies 41, 303-323.
13. Glynn, T. (1998) Bicultural challenges for educational professionals in Aotearoa. Waikato Journal of Education 4, 3-16.
14. Henare, M. (2001). Tapu, mana, mauri, hau, wairua: A Māori philosophy of vitalism and cosmos. In: Grim, J.A. (ed.) Indigenous traditions and ecology: the interbeing of cosmology and community, pp 197-221. Harvard University Press.
15. Higgins, R. and Rewi, P. (2014). ZePA - right-shifting: reorientation towards normalisation. In: Higgins, R., Rewi, P. and Olsen-Reeder, V. (Eds.) The Value of the Māori Language - Te Hua o te Reo Māori, pp. 7-32. Wellington: Huia.
16. Jenni, B., Anisman, A., McIvor, O., Jacobs P. (2017). An exploration of the effects of mentor-apprentice programs on mentors' and apprentices' wellbeing. International Journal of Indigenous Health 12, 25-42.
17. Jones, H. (n.d.). A survey of the Welsh language: the 1992 Welsh social survey. Retrieved from http://statiaith.com/gwefan/pdfs/arolwgcymdeithasolcymru1992/Erthygl%20Stats%20News.PDF

18. Jones, H. M. (2012). A statistical overview of the Welsh language. Cardiff: Welsh Language Board.
19. Ka'ai-Mahuta, R. (2011). The impact of colonisation on te reo Māori: a critical review of the State education system. Te Kaharoa 4(1), 195-225.
20. King, J. and Cunningham, U. (2017). Tamariki and fanau: child speakers of Māori and Samoan in Aotearoa/New Zealand. Te Reo 60, 29.
21. Krauss, M. (1992). The world's languages in crisis. Language 68, 4-10.
22. Kroch, A.S. (1991). Reflexes of grammar in patterns of language change. Language Variation and Change 1, 199-244.
23. May, S., Hill, R. and Tiakiwai, S. (2004). Bilingual/immersion education: indicators of good practice. Report to the Ministry of Education. New Zealand: the Ministry of Education. Retrieved from: https://hdl.handle.net/10289/4342
24. McCarthy, M. (1997). Raising a Māori child under a new right state. In: Te Whāiti, P., McCarthy, M. and Durie, A. (Eds.), Mai i Rangiātea Māori Wellbeing and Development, pp.25-38. Auckland: Auckland University Press with Bridget Williams Books.
25. McConvell, P. and Thieberger N. (2010). State of Indigenous languages in Australia – 2001. Australia State of the Environment Second Technical Paper Series (Natural and Cultural Heritage), Department of the Environment and Heritage, Canberra. Retrieved from https://minerva-access.unimelb.edu.au/bitstream/handle/11343/34091/66504_00001559_01_indigenous-languages.pdf?sequence=1
26. Minett, J.W. and Wang, W.S.-Y. (2008). Modelling endangered languages: the effects of bilingualism and social structure. Lingua 118, 19-45.
27. Moorfield, J.C. (2006). Teaching and learning an Indigenous language through its narratives: Māori in Aotearoa/New Zealand. Junctures: The Journal for Thematic Dialogue 6, 107-116.
28. Moseley, C. (Ed.) (2010). Atlas of the world's languages in danger, 3rd edition. Paris: UNESCO Publishing.
29. New Zealand Government (2010). The Māori language: selected events 1800-2010. Parliamentary Library Research Paper 10/08. Retrieved from https://www.parliament.nz/resource/en-NZ/00PLSocRP09071/7fbd5da67d74336d2e486c3c881a2f919f662b92
30. Nowak, M. A., Komarova, N. L. and Niyogi, P. (2002). Computational and evolutionary aspects of language. Nature 417, 611–617.
31. Niyogi, P. and Berwick, R. C. (1997). Evolutionary consequences of language learning. Linguistics and Philosophy 20, 697–719.
32. Rewi, T. and Rewi, P. (2015). The ZePA model of Māori language revitalization: key considerations for empowering Indigenous language educators, students, and communities. In: Reyhner, J., Martin, J., Lockard, L. and Gilbert, W.S. (Eds.) Honoring Our Elders: Culturally Appropriate Approaches for Teaching Indigenous Students, pp. 136-153. Flagstaff, AZ: Northern Arizona University.
33. Ruckstuhl, K. and Wright, J. (2012). Māori language modelling. In Proceedings of the International Indigenous Development Research Conference 2012, pp. 122-128.
34. Spolsky, B. (2004). Language Policy. Cambridge: Cambridge University Press.

35. StatsNZ (2013a). Māori language speakers. Retrieved from http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/nz-social-indicators/Home/Culture%20and%20identity/maori-lang-speakers.aspx
36. StatsNZ (2013b). Te Kupenga 2013: a survey of Māori well-being. Retrieved from http://archive.stats.govt.nz/browse_for_stats/people_and_communities/maori/TeKupenga_HOTP13/Commentary.aspx
37. StatsNZ (2013c). National ethnic population projections, by age and sex, 2013(base)-2038 update. Retrieved from http://nzdotstat.stats.govt.nz/wbos/Index.aspx?DataSetCode=TABLECODE7560&_ga=2.126766537.530629275.1563142782-1973661841.1532044666&_gac=1.216187298.1561435728.EAIaIQobChMIo9Sjjt-D4wIVkg4rCh2bxAA3EAAYASAAEgIzk_D_BwE#
38. Te Mātāwai (2018). Maihi Māori 2017-2040. Retrieved from https://www.tematawai.maori.nz/maihi-maori-english
39. Te Puni Kōkiri (2018). Maihi Karauna, the Crown's Strategy for Māori Language Revitalisation, 2019–2023. ISBN: 978-0-9951038-6-3. Retrieved from https://www.tpk.govt.nz/docs/tpk-maihi-karauna-en-2018.pdf
40. Timms, C.E. (2013). Indigenous language revitalisation in Aotearoa New Zealand and Alba Scotland (Thesis, Doctor of Philosophy). University of Otago. Retrieved from http://hdl.handle.net/10523/3851
41. Waitangi Tribunal. (1986). Report of the Waitangi Tribunal on the te reo Māori claim (Report no. Wai 11). Retrieved from https://forms.justice.govt.nz/search/Documents/WT/wt_DOC_68482156/Report%20on%20the%20Te%20Reo%20Maori%20Claim%20W.pdf
42. Wang, W. S. and Minett, J. W. (2005). The invasion of language: emergence, change and death. Trends in Ecology and Evolution 20(5), 263-269.
43. Welsh Government (2015). Welsh language use in Wales, 2013-15. Retrieved from https://gov.wales/docs/statistics/2016/160301-welsh-language-use-in-wales-2013-15-en.pdf
44. Welsh Government (2017). Cymraeg 2050. Retrieved from https://gov.wales/docs/dcells/publications/170711-welsh-language-strategy-eng.pdf
45. Welsh Language Board (2008). The Welsh language use surveys of 2004-06. Retrieved from http://doc.ukdataservice.ac.uk/doc/7477/mrdoc/pdf/7477_methodology_in_english.pdf

	Basic	Independent	Proficient
Welsh, 1991	81.0%	7.0%	11.0%
Te reo Māori (overall), 2013	96.3%	2.7%	1.0%
Te reo Māori (Māori subpopuation), 2013	77.4%	12.0%	10.6%
Te reo Māori (non-Māori subpopulation), 2013	98.4%	1.0%	0.6%

Table 1. Initial conditions for the model applied to various populations, specifying the proportion of the population in the three proficiency categories.

	Language learning in English-medium school	Language immersion school	Estimated learning rates (yr^{-1})	
			β_{BI}	β_{IP}
Welsh	78%	22%	0.0510	0.0969
Te reo Māori (Māori)	31.3%	9.6%	0.0205	0.0423
Te reo Māori (overall)	21.1%	2.5%	0.0138	0.0016

Table 2. School participation rates and estimated learning rates for Welsh and te reo Māori. Columns 1 and 2 shows rates of participation in language learning in English-medium school and in language-immersion schools. Columns 3 and 4 show estimated learning rate parameters for the model. Estimated learning rates for Wales are obtain by fitting the model to data of changes in proficiency levels over time; estimated learning rates for New Zealand are derived from the Welsh estimates using the relative school participation rates. References for participation rate data: New Zealand – Education Counts (2018); Wales – Welsh Government (2017).

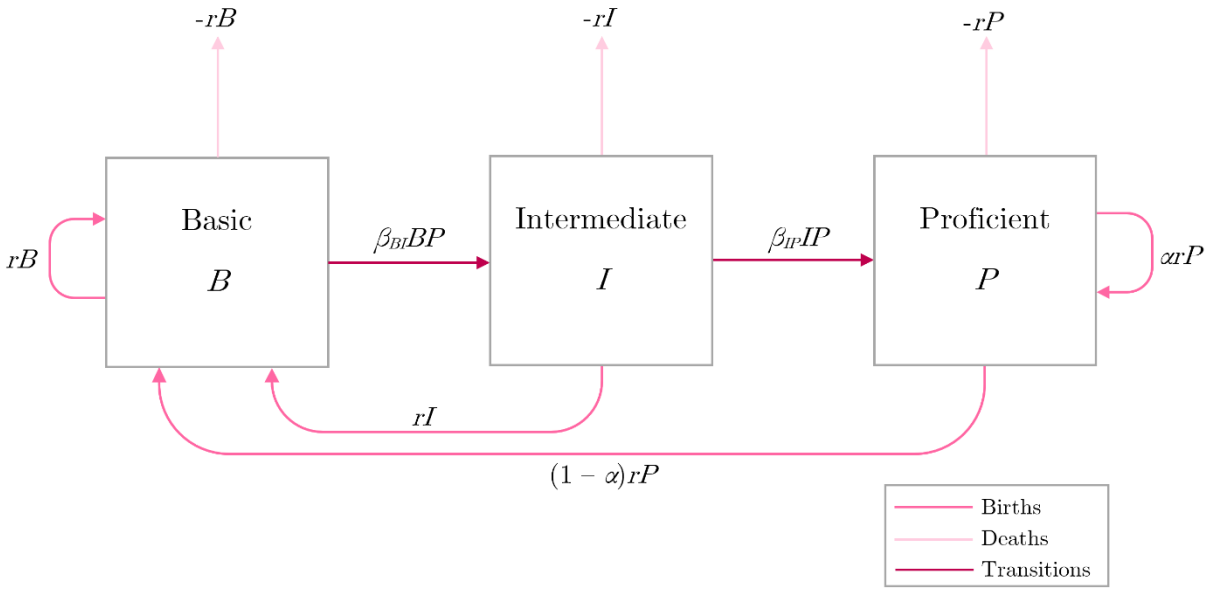

Figure 1: Transitions between categories occurring as a result of births, deaths and learning from proficient users.

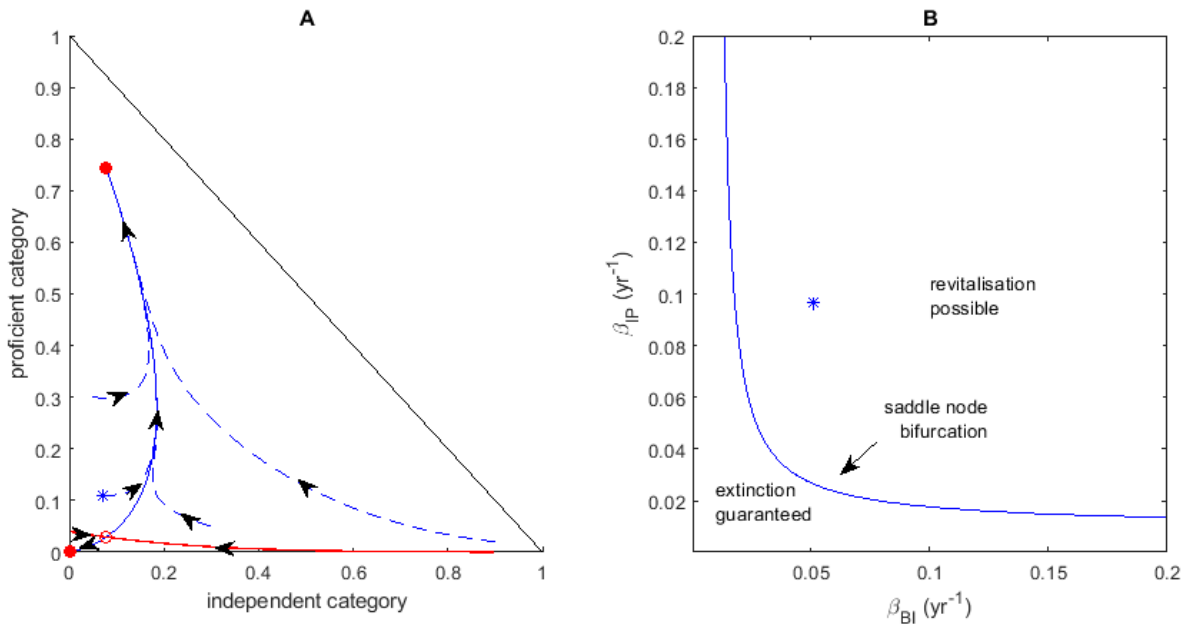

Figure 2. Structure of the dynamical system for the proportion of the population in the independent category (I) and the proficient category (P). (A) Trajectories (dashed blue curves) and steady states (red circles) in the (I, P) phase plane for learning rate parameters $\beta_{BI} = 0.051$ yr^{-1} and $\beta_{IP} = 0.0969$ yr^{-1}. For these values of β_{BI} and β_{IP}, there are two positive steady states as well as the steady state at $(I, P) = (0,0)$. The steady states at $(I, P) = (0.08,0.74)$ and $(I, P) = (0,0)$ are locally stable (filled red circles) and the steady state at $(I, P) = (0.08,0.03)$ is a locally unstable saddle node (open red circle). The stable manifold of the saddle node (red curve) separates the phase space into two basins of attraction: initial conditions below the red curve tend towards the zero steady state (language extinction); initial conditions above the bold curve tend to the positive stable steady state (language revitalisation). The blue star shows the 1991 initial conditions and predicted trajectory for Welsh. (B) Two-parameter bifurcation diagram showing behaviour of the model in (β_{BI}, β_{IP}) parameter space. There is a saddle-node bifurcation (solid curve) separating parameter combinations for which extinction is guaranteed from those for which language revitalisation is possible. Blue star shows estimated learning rates for Welsh. Other parameter values: $\alpha = 0.47$ and $r = 1/70$ yr^{-1}.

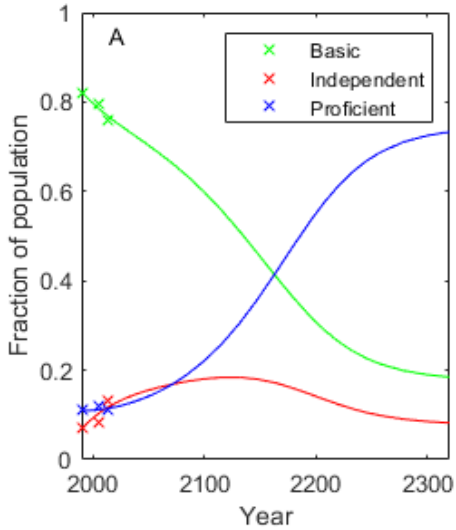

Figure 3: The model predicts long-term success for Welsh language revitalisation. The predicted model time series fitted to data for the Welsh language, with initial conditions taken from the first data point: $B = 0.81$, $I = 0.07$, $P = 0.11$ in 1991. Parameter values: $r = 1/70\ \mathrm{yr}^{-1}$, $\alpha = 0.47$, $\beta_{BI} = 0.051\ \mathrm{yr}^{-1}$, $\beta_{IP} = 0.0969\ \mathrm{yr}^{-1}$.

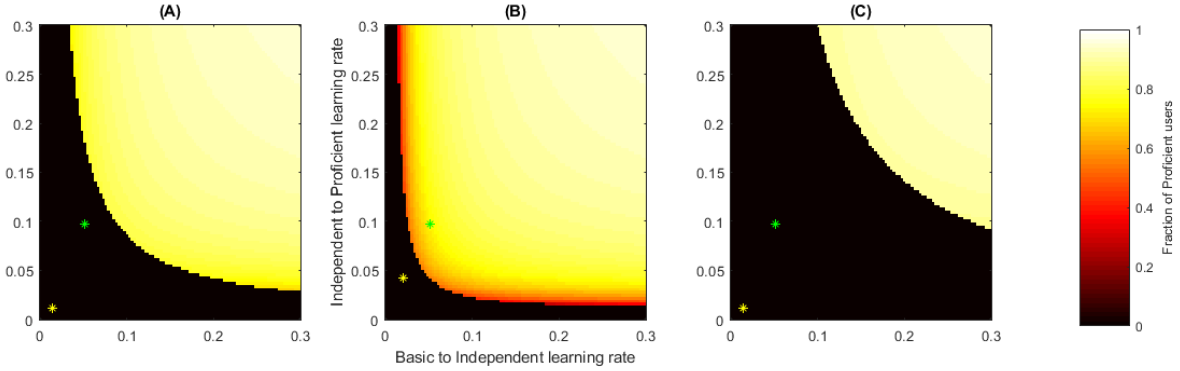

Figure 4: Successful revitalisation of te reo Māori is possible if the learning rate parameters are high enough. Long-term language outcomes for: (A) the New Zealand population as a whole; (B) the Māori subpopulation as a separate group; (B) the non-Māori subpopulation as a separate group. Colours show the long-term proportion of the population in the proficient category starting from initial conditions estimated for the relevant population (Table 1) and for a range of learning rate parameters β_{BI} and β_{IP}. Black regions correspond to learning rate parameters that lead to long-term extinction of the language; coloured regions correspond to long-term success of the language. Green stars indicate the estimated learning rate parameters for Welsh; yellow stars indicate the estimate learning rate parameters for te reo Māori in the relevant group.

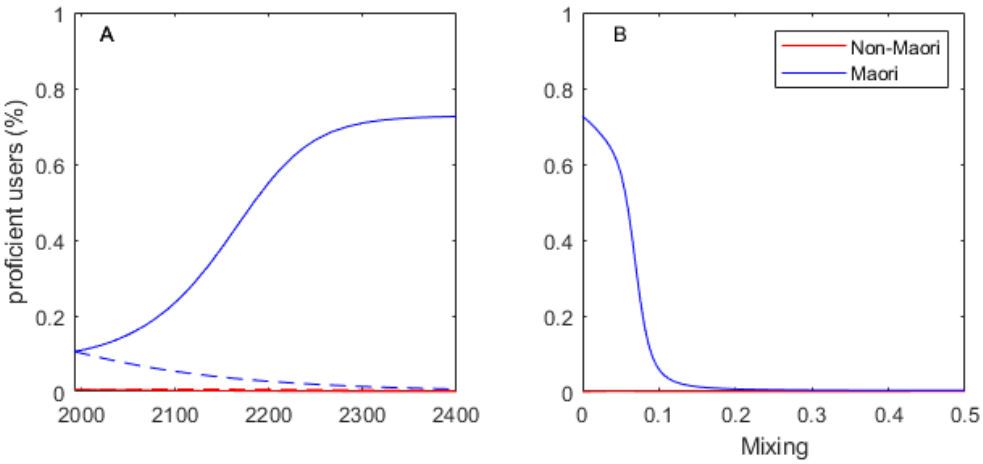

Figure 5: Overuse of Māori teachers in the non-Māori population could have a detrimental effect on long-term language revitalisation. (A) With no mixing of teachers between the populations ($\gamma = 0$, solid curves) te reo Māori is predicted to thrive in the Māori subpopulation (blue) under Welsh learning rates, but decline to extinction in the non-Māori subpopulation (red). With fully homogeneous mixing of teachers ($\gamma = 1.0$, dashed curves), te reo Māori is predicted to go extinct in both subpopulations. (B) Even a small degree of mixing of teachers between populations ($\gamma > 0$) drastically reduces the long-term number of proficient users in the Māori subpopulation (blue), without having any significant beneficial effect for the non-Māori subpopulation (red).