

# Modelling the dynamics of language death

Wellington J. Leite da Silva <sup>1</sup>

<sup>1</sup>Escola de Matemática Aplicada da FGV (FGV/EMAp), Brasil

November 2021

# Table of contents

- 1 Introdução
- 2 Metodologia
- 3 Resultados
- 4 Conclusão

## Mandan language (pressão social)

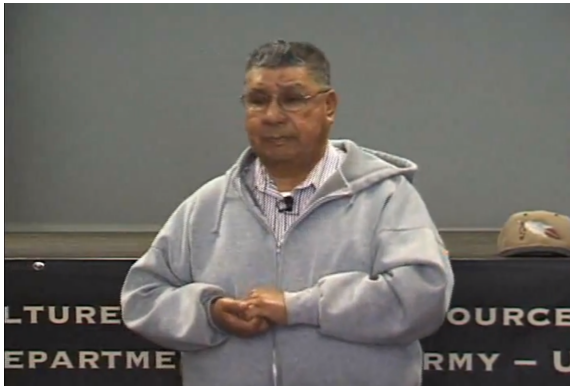


Figure: Edwin Benson, o último falante do Mandan (1931 - 2016)

# Cacaopera language (violência)

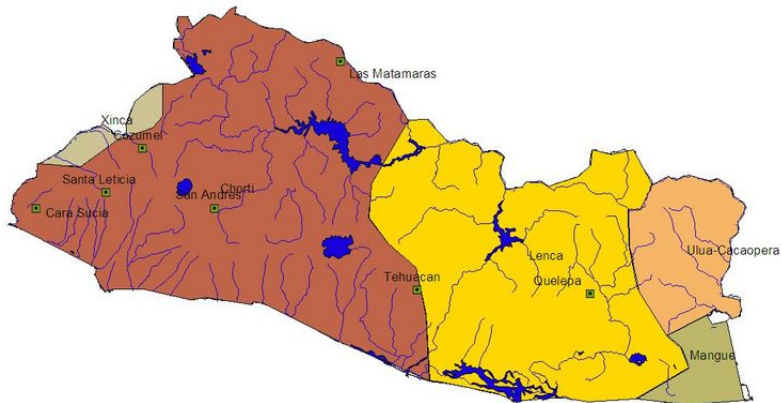


Figure: El Salvador (local dos indígenas falantes de Cacaopera)

# A extinção de idiomas

- Mais de **7.000** línguas faladas ao redor do mundo, mas cerca de 1/3 delas têm menos de **1.000 falantes**.
- Segundo UNESCO mais de 40% de todas as línguas estão em perigo de extinção.
- No Brasil, há 190 línguas ameaçadas de extinção

# Ideia do trabalho

Published: 21 August 2003

Linguistics

## Modelling the dynamics of language death

[Daniel M. Abrams](#) ✉ & [Steven H. Strogatz](#)

[Nature](#) **424**, 900 (2003) | [Cite this article](#)

10k Accesses | 271 Citations | 41 Altmetric | [Metrics](#)

### Abstract

Thousands of the world's languages are vanishing at an alarming rate, with 90% of them being expected to disappear with the current generation<sup>1</sup>. Here we develop a simple model of language competition that explains historical data on the decline of Welsh, Scottish Gaelic, Quechua (the most common surviving indigenous language in the Americas) and other endangered languages. A linguistic parameter that quantifies the threat of language extinction can be derived from the model and may be useful in the design and evaluation of language-preservation programmes.

Download PDF



Sections

Figures

References

Abstract

Main

References

Author information

Ethics declarations

Rights and permissions

About this article

Further reading

Comments

Figure: Artigo inicial<sup>1</sup>

<sup>1</sup><https://www.nature.com/articles/424900a#Bib1>

# O que é tratado neste trabalho?

Figure 1: The dynamics of language death.

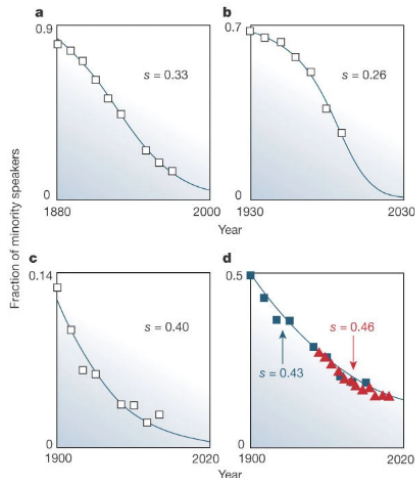


Figure: Gráficos do artigo inicial

## Premissas para o modelo

Com intuito de modelar uma função  $(x(t))$ , que representa a porcentagem da população que fala um certo idioma pelo tempo, tomamos as seguintes premissas:

- Vamos considerar um sistema com 2 línguas concorrentes Y e X (onde só é possível falar uma das línguas).
- Vamos considerar também a existência de um parâmetro  $s$  ( $0 \leq s \leq 1$ ) que chamamos de status da língua.
- Sendo  $P_{yx}(x(t), s)$  a probabilidade de um falante de Y ir para X no tempo  $t$ , onde  $s$  é o status de X em relação a Y.

Assim uma EDO que modela o sistema pode ser dada por:

$$x'(t) = (1 - x(t))P_{yx}(x(t), s) - x(t)P_{xy}(x(t), s)$$



# O modelo

Também podemos assumir que:

- Ninguém adotará uma linguagem que não tenha falantes ( $P_{yx}(0, s) = 0$ ).
- Ou nenhuma língua com status zero ( $P_{yx}(x(t), 0) = 0$ ).
- Vamos propor que  $P_{yx}$  é da seguinte forma  $P_{yx} = cs(x(t))^a$  e  $P_{xy} = c(1 - s)(1 - x(t))^a$

Dessa forma o modelo fica assim:

$$x'(t) = (1 - x(t))cs(x(t))^a + x(t)c(1 - s)(1 - x(t))^a$$

Onde  $c$ ,  $s$ ,  $a$  e  $x(0)$ , são parâmetros do modelo.

# Proposta de mudança

$$x'(t) = (1 - x(t))cs(x(t))^a + x(t)c(1 - s)(1 - x(t))^a$$

Estimação dos parâmetros:

- (Artigo inicial) Regressão Linear
- (Este trabalho) Inferência Bayesiana

# Aplicação do modelo

Para testar o modelo escolhemos 2 línguas: Gaelic e Welsh, então precisamos estimar os parâmetros do modelo, para isso:

- Usamos **pymc3**.
- Fixamos  $x(0)$  como o primeiro valor das observações.
- Sugerimos distribuições a priori para cada um dos outros parâmetros, que foram sendo refinadas, até encontrar bons valores para cada parâmetro.
- Em seguida plotamos o gráfico da  $x$  com os "melhores" parâmetros encontrados, junto com as observações.

# Posterioris Gaelic

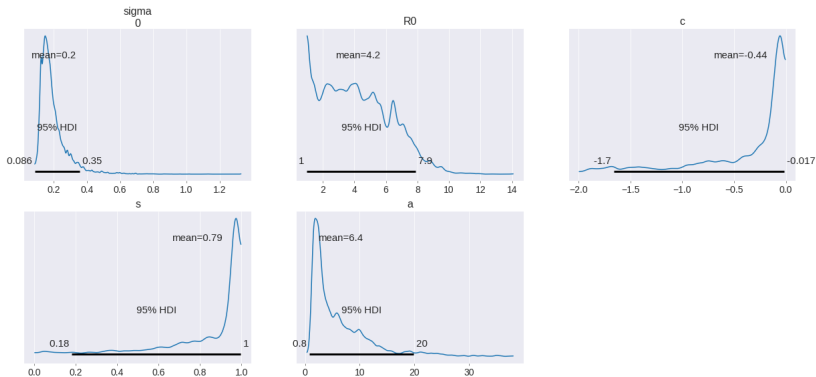


Figure: Posterioris dos parâmetro para gaelic

## Projeção da $x(t)$

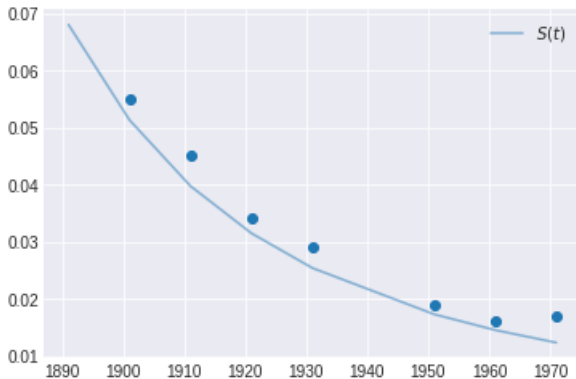


Figure: Modelagem gaelic

# Posterioris Welsh

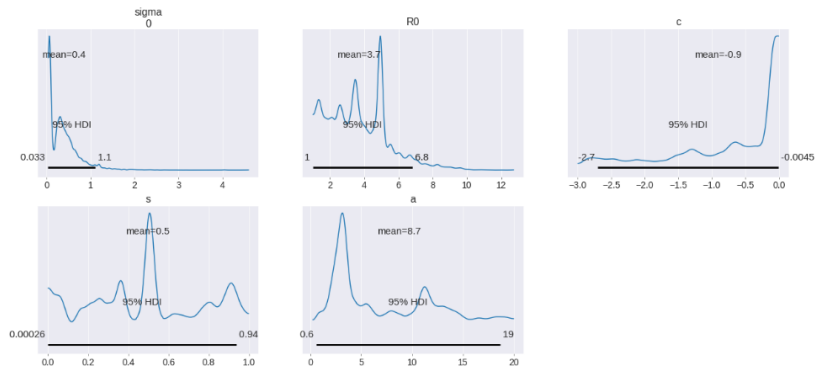


Figure: Posterioris dos parâmetro para welsh

## Projeção da $x(t)$

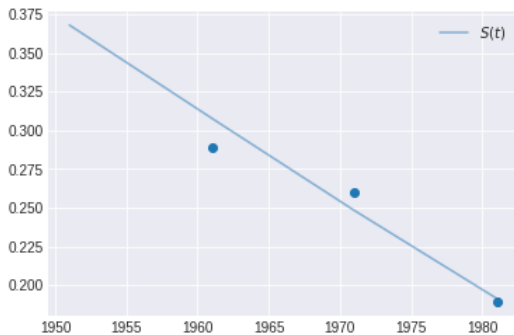


Figure: Modelagem welsh

## Para mais detalhes

### Jupyter-notebook com implementação de Galic

```
https://github.com/wellington36/  
Modelling-the-dynamics-of-language-death/blob/main/  
Implementa%C3%A7%C3%B5es/implementacao_gaelic.ipynb
```

### Jupyter-notebook com implementação de Welsh

```
https://github.com/wellington36/  
Modelling-the-dynamics-of-language-death/blob/main/  
Implementa%C3%A7%C3%B5es/Implementa%C3%A7%C3%A3o%20Welsh.  
ipynb
```



# Dificuldades durante o trabalho

- Falta de dados (e de referencias para os que tem).
- Vários aspectos podem levar um idioma a morte.
- Poucas pesquisas na área (principalmente no Brasil que é muito afetado).

# Ideias futuras

- Melhorar a modelagem através da estimação de parâmetros.
- Usar dados do artigo original para comparar as modelagens (dados dos parâmetros foram solicitados via e-mail ao autor).

# Thank you!

For contact:

[wellington.71319@gmail.com](mailto:wellington.71319@gmail.com)