

Modelling language dynamics in Carinthia, Austria

K. Prochazka* and G. Vogl

University of Vienna, Faculty of Physics, Boltzmanngasse 5, 1090 Vienna, Austria
*katharina.prochazka@univie.ac.at

Background

Use of the minority language Slovenian in Carinthia, Austria has been steadily declining over the past century. Despite supportive measures, language shift (speakers abandoning use of one language for another) is taking place, reducing cultural and intellectual diversity.

One way of monitoring this language shift on a large scale is using methods from the natural sciences where dealing with big sets of data is common. Language shift has previously been described by computer simulations and equations based on models of diffusion from solid state physics.^[1–3]

Most of these simulations use differential equations and look only at the proportion of speakers in the whole population with no spatial dynamics. We present a different approach: **a microscopic model to follow language shift over time and space** in Carinthia based on data from the Austrian census.

[1] C. Schulze, D. Stauffer, S. Wichmann. Commun Comput Phys 3, 271–294 (2008).

[2] A. Kandler. Hum Biol 81, 181–210 (2009).

[3] N. Isern, J. Fort. J R Soc Interface 11 (2010).

The data from the Austrian census

The Austrian census was **the primary method for collecting data on language use** in Austria between 1880 and 2001. Data is **available every ten years between 1880 and 1910 and 1951 to 2001**, plus for some years in between the two periods (with varying data quality).

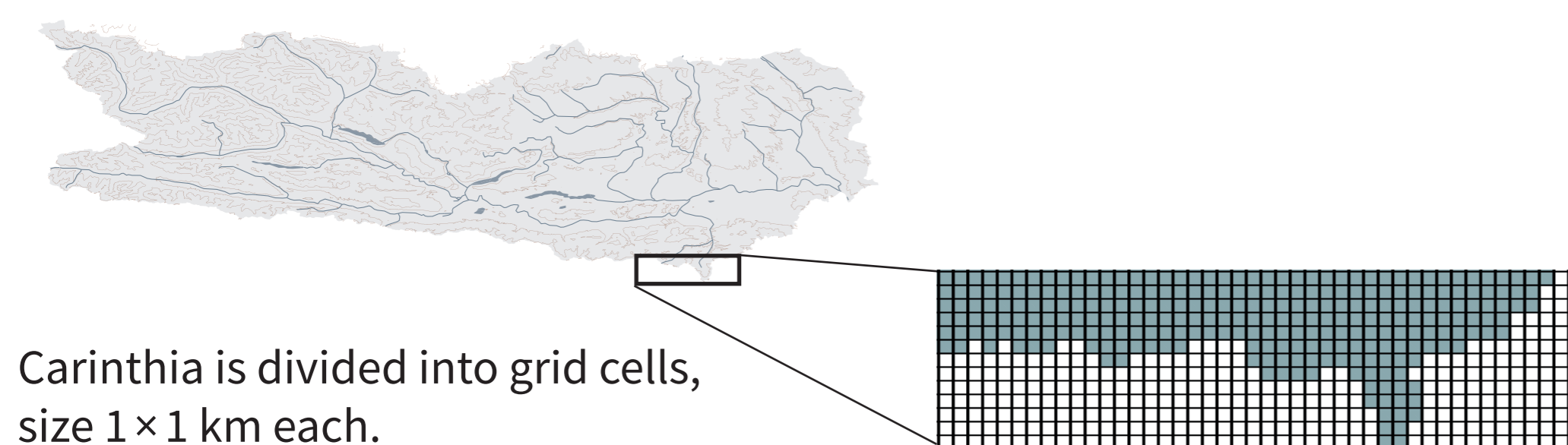
A page from the **1890 census** with information about the “vernacular language” (*Umgangssprache*) in each village (Slovenian or German).

Kärnten, Koroško. — Völkermarkt, Velikovec.

91

Bezirke, Gemeinden, Ortschaften (Okraj, občine, kraji)	Häuser (hiše)	Anwesende Bevölkerung (Pričujoče prebivalstvo)							Umspr. d. einh. Bevölk. (Obč. jez. domač. prebiv.)			
		männlich (moško)	weiblich (žensko)	zusammen (vkupe)	Confession (Po veri)				deutsche (nemški)	sloven. (slo- venski)	andere (drugi)	
					kathol. (kato- liški)	evangel. (evan- gelski)	israelit. (izraelci)	andere (drugi)				
Mitterdorf, Ober-, Sred- nja ves (Dorf, vas) . .	18	56	68	124	124	.	.	.	17	107	.	
Mitterdorf, Unter-, Spod- nja ves (Dorf, vas) . .	17	51	62	113	113	.	.	.	1	112	.	
Sch. (š.)												
Nikolai, St., Miklavž Sv., Dorf mit E.-H. (vas s p. h.) Tiefenbach . . .	23	63	70	133	133	.	.	.	5	128	.	
Cur. [St. Nikolai in Go- rentschach] (kur.)												
Radigund, St., Radigunda Sv. (Dorf, vas)	11	35	44	79	79	.	.	.	2	77	.	

Our approach: A microscopic model



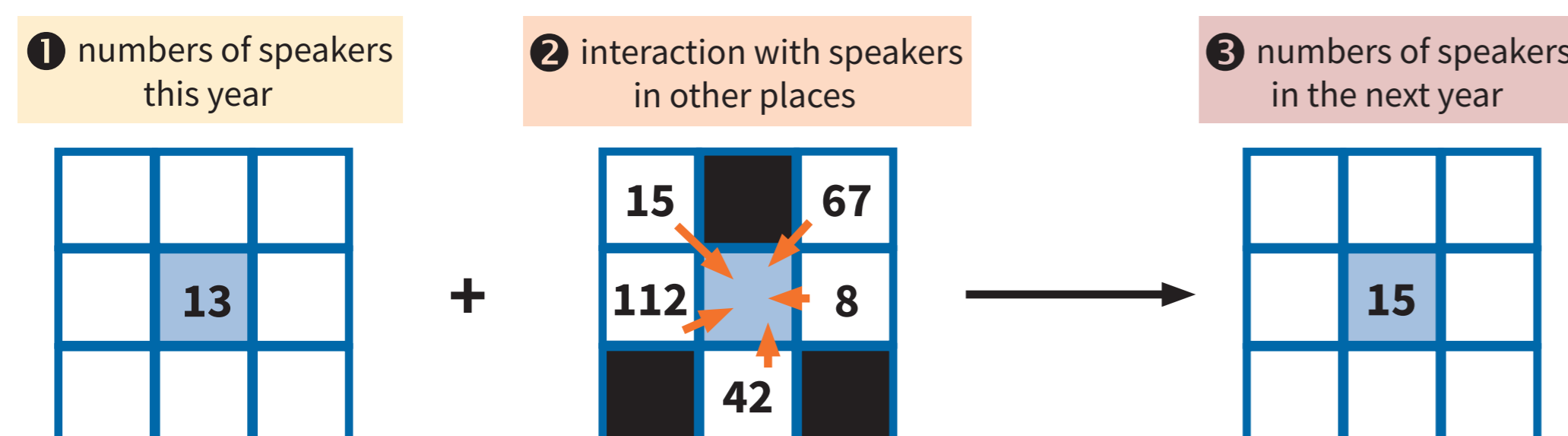
Carinthia is divided into grid cells, size 1×1 km each.

To get the number of speakers of a language α in the next year, each grid cell (■) is updated according to the mathematical rule:

$$n_{\alpha}(\mathbf{r}, t+1) = n_{\text{total}}(\mathbf{r}, t+1) \cdot \frac{n_{\alpha}(\mathbf{r}, t) + F_{\alpha}(\mathbf{r}, t)}{n_{\text{S}}(\mathbf{r}, t) + F_{\text{S}}(\mathbf{r}, t) + n_{\text{G}}(\mathbf{r}, t) + F_{\text{G}}(\mathbf{r}, t)}$$

what this means: the number of speakers in the next year ③ is proportional to:

- ① the number of current speakers and
- ② the interaction with speakers in other places



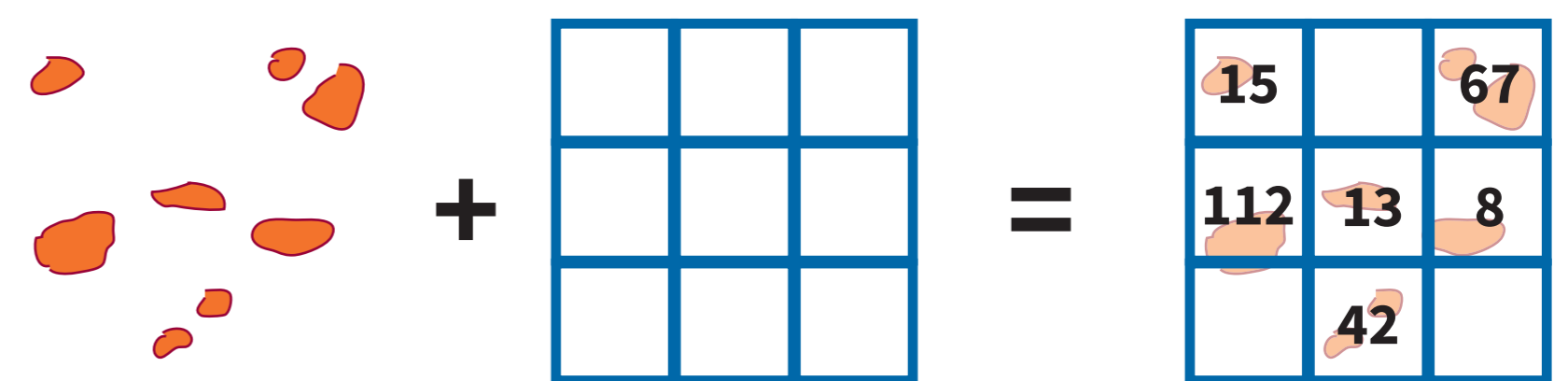
Challenges

Creating new data is impossible, thus we **need to get the most out of the historical data available**.

The sample available is very large and detailed, but due to the anonymity of the census **no additional information about individual persons** is available.

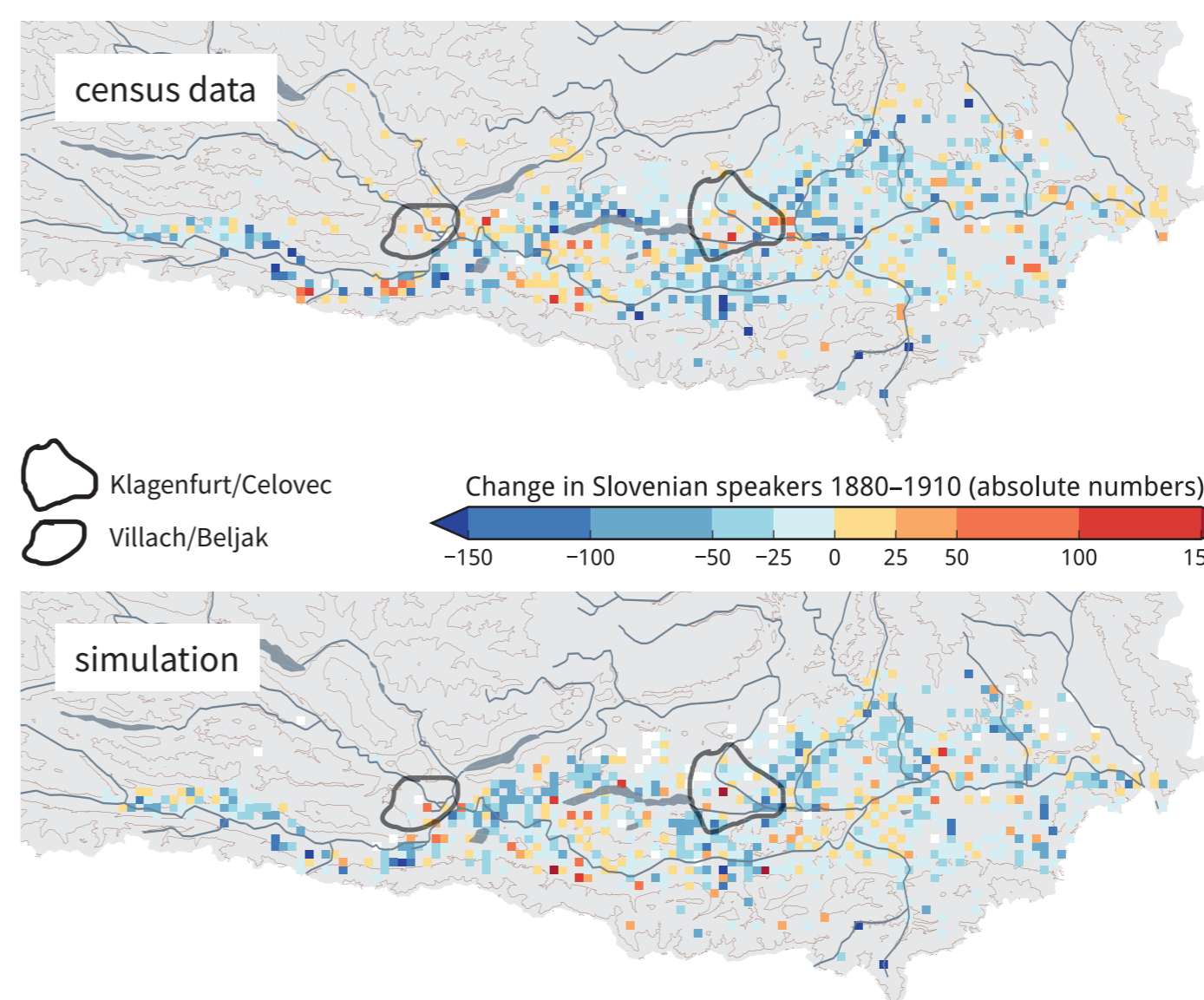
Ideally, the model **should only use parameters which (a) make sense linguistically and (b) can be empirically measured**. Testing the model with empirical data is only possible if parameters are used which correspond to the data.

Events influencing language use **which are not language shift as such** (e.g. wars) but visible in the data **cannot be captured by the model**.



Each village (●) is assigned to a grid cell based on its geographic coordinates and speaker numbers are attributed to the grid cells.

Some results: Carinthia 1880–1910



Graphs show absolute change in Slovenian speaker numbers (red = more, blue = fewer).

Simulated and real data agree well qualitatively.

All model parameters can be calculated directly from the census data, ensuring that the model is applicable even in situations where data on other factors influencing language use (e.g. perceived status of a language) is not available or even possible to obtain.