diffusion-fundamentals.org

The Open-Access Journal for the Basic Principles of Diffusion Theory, Experiment and Application

Diffusive and non-diffusive language shift. Why are language borders moving and language islands shrinking?

Katharina Prochazka[1], Gero Vogl[2]*

[1] Institute for Slavistics, University of Vienna, 1090 Vienna, Austria
[2] Faculty of Physics, University of Vienna, 1090 Vienna, Austria
*Corresponding author: gero.vogl@univie.ac.at

Keywords: language shift, language borders, language islands, diffusion, Hungary

(Received: 2022/11/15, Accepted: 2022/11/16, Published online: 2022/12/03)

Gewidmet Jörg Kärger
mit Dank für seinen unermüdlichen Einsatz für die Diffusion
weit über deren ursprünglichen Rahmen hinaus.

Abstract

We compare two situations of language shift: When a language border exists, as e.g. between Slovenian and German in Carinthia, Austria before World War I or between Gaelic and English in Northwestern Scotland, then close to that border language shift is promoted by diffusive contact between different settlements. Such contact, however, was of no significant influence on language shift from German to Hungarian in Southern Hungary where no language border exists. There language shift rather occurred intrinsically, the linguistic composition of the settlements playing a dominant role, the shrinking of the minority language being minor in settlements with dominating minority language and increasing the smaller the fraction of that language was. The two different paradigms covering situations with ample data from censuses reveal two different language shift scenarios.

1. Introduction

Many of today's languages are in fast retreat losing their importance for communication aside of local communication. The cultural heritage connected with their use is threatened to be lost. An unrivaled collection of empirical data is available from the censuses of the Austrian-Hungarian Empire and its successor states. We use these data for studying language shift in time and space and claim that the model scenarios may help understanding the dynamics behind language shift in present scenarios and to stop or at least mitigate language retreat.

Since several years language has been in the focus of physicists in the context of language as a complex system which is changing over time (see e.g. the reviews [1-4]). These changes have been modeled as language change (changes within one language's system, e.g. changes in pronunciation or word use) [5-8], but also as language shift [2-4, 9-20] meaning that people and societies change their language use by shifting from one language to another [21]. The majority of the investigations until now have been performed on abstract scenarios. Abstract physical models should be compared to empirical data to check the validity.

Data with sufficient spatiotemporal resolution, however, are generally rare, but extremely well-resolved information on population details is available from the 1880 till 1910 censuses of the Austrian-Hungarian Empire (Fig.1 left) and after the collapse of the Empire in the Republics of Hungary and Austria. These data offer themselves for extended studies on the development of the eleven different languages listed in the Austro-Hungarian censuses.

Here we present a model for language shift from German to Hungarian in Southern Hungary and compare the underlying mechanism with a different process of language shift, i.e. the motion of a language border, as described earlier in literature [16, 17].

2. Historical Background

Until the end of the seventeenth century large parts of Hungary were under Turkish administration. After the failed attempt of the Osmanic Empire to conquer Vienna in 1683 („Second Turkish Siege of Vienna“) the Austrian troops, soon led by Prince Eugene of Savoy, drove out the Turks in a few years from all occupied Hungary. Before and during the battles with the Turks, the population had been markedly reduced. When the Turks left, the region, now in the Austrian Habsburg Empire, was very quickly re-populated starting around 1720 [22, 23]. A major fraction of the new settlers were Germans, a considerable part from Southwestern Germany, the region of Schwaben. Because of this history the Baranya and Tolna counties in Southern Hungary are sometimes called “Schwäbische Türkei” (Swabian Turkey). The villages settled by Germans were dispersed as language islands with settlements where the majority spoke other languages, predominantly Hungarian.

The German speaking society was a typical immigration society, widely self-contained, the economic exchange with other settlements and thus also language exchange being limited.

From the end of the 19th century on, these settlers – like all other non-Hungarian ethnic groups – were pushed to adopt the Hungarian language and culture under the government’s assimilation policy of “magyarisation” (hungarisation) in order to consolidate the national identity of the state by prioritizing the Hungarian language over other languages [22-24].

By understanding how language shift happened in that scenario which may be regarded as typical for immigration societies we hope to learn how to deal with similar present situations.

Fig. 1. Left: Map of the Austro-Hungarian Empire in 1880. Purple: Austrian part, light orange, brown: Hungarian part, green: Bosnia-Herzegovina which was occupied by Austria-Hungary in 1878 and later annexed. The area of investigation is shown in darker orange-brown (counties Baranya and Tolna). Right: Map of Austria and Hungary in today’s Europe. The area of investigation is again shown in darker orange-brown.

3. Language shift in the settlements of Southern Hungary

3.1 Data

We decided to study the language shift from German to Hungarian in Southern Hungary, in particular in the counties Baranya and Tolna, the central counties of Swabian Turkey (Fig. 1).

All through the Austrian-Hungarian "double monarchy" from 1880 on censuses asked for people's language. In the Austrian part the vernacular (everyday) language was asked for, in Hungary the censuses asked for the mother tongue, and in both cases there was no possibility to indicate bilingualism. The censuses were continued in Hungary and Austria as well after the collapse of the Austrian-Hungarian Empire.

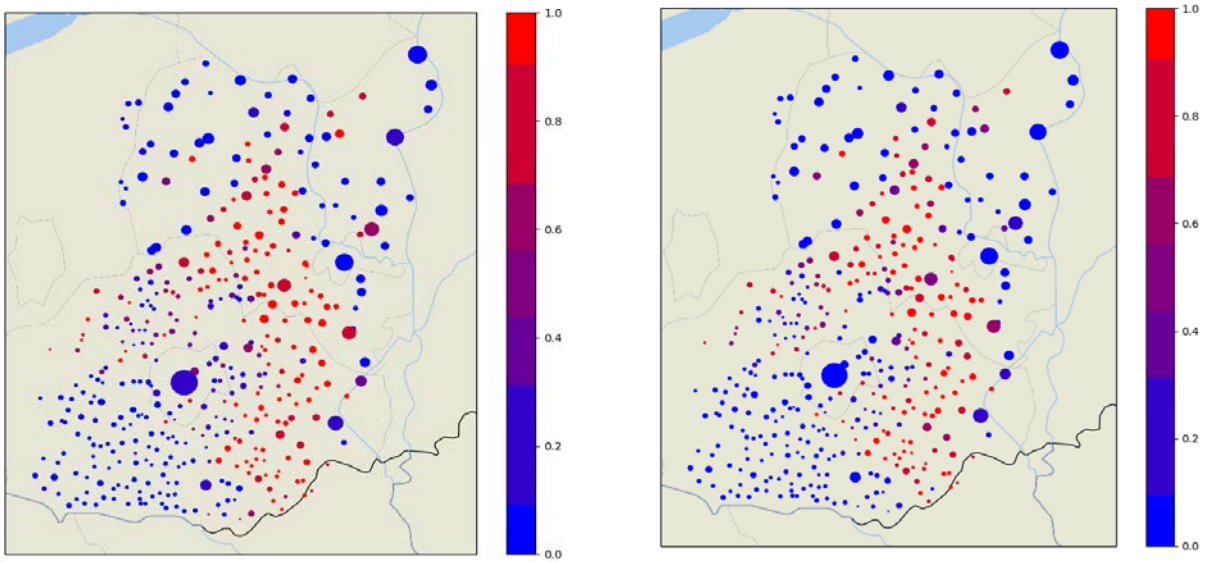

Fig. 2. Fraction of German speakers in the counties Baranya and Tolna according to census results. The dark brown area in Fig. 1 indicates the geographic position of the region. Each circle symbolizes a settlement, circle size proportional to number of inhabitants. Colors from red to blue indicate fraction of German speakers (from 100 % down to 0 %). Left: Census results for 1880. Right: Census results for 1930. Rivers and lakes are shown in light blue. The black line at the bottom is the border of Hungary to Yugoslavia after 1918. Though at first glance there is no striking difference between 1880 and 1930, the detailed data permitted to model language shift.

Using the censuses, we first digitised the language data for each settlement from the ten-yearly Hungarian censuses between 1880 and 1930 [26]. Fig. 2 gives data for 1880 and 1930. An extension to later times does not appear meaningful since the data taken in 1941 (already during World War II) appear problematic, and after World War II a major part of the German speaking population was forced to leave Hungary.

Overall, we obtained data for 425 settlements in the counties of Baranya and Tolna. For simplicity, we use only information on the number of German and Hungarian speakers, disregarding other languages. While there are speakers of other languages listed in the censuses, their number is rather small compared to the number of German speakers: In only 53 of the 425 total settlements (12.4%) there are more than 100 speakers of other languages in 1880. In the sample chosen for analysis (50 settlements) there are more than 100 speakers of other languages in 12 settlements (25.5%). The total population of each settlement is therefore calculated by adding the numbers of German and Hungarian speakers.

The linguistic mixture of the region originated directly from the resettlement policy of the Habsburg Empire from about 1720 on. Language shift started after the magyarisation politics began towards the end of the 19th century and was unidirectional from German to politically intended Hungarian. For modelling language shift over time we therefore assume that the fraction of the German minority language in the population can only decrease. Additionally, we assume that the process of language shift is the same in all settlements, i.e. it takes the same shape but with different initial conditions (different initial linguistic compositions) based on the fraction of German speakers

"implanted" in the resettlement process. In default of information we have to neglect migration processes.

3.2 Model

A first inspection of the development of the fraction of German speakers in the settlements from 1880 till 1930 indicated that magyarisation occurred the faster the higher the fraction of Hungarian speakers in the settlement was. This let us suppose that it was not so much the diffusive contact between neighboring settlements which promoted linguistic magyarisation than rather intrinsic factors, i.e. the linguistic composition of the settlements.

The data scattering is very notable. Since relative scattering is the larger the smaller the settlement we restricted the detailed study to 50 larger settlements with a non-negligible fraction of German speakers. This choice resulted in a minimum settlement size of 1200 inhabitants with at least 100 German speakers. Even in this reduced sample considerable scattering remained.

As indicated above, a first inspection of the development of the fraction of German speakers in the settlements from 1880 till 1930 suggested that magyarisation occurred the faster the lower the fraction of German speakers or the higher the fraction of Hungarian speakers in the settlement was. We describe this supposed relation by setting the retreat of German language as a negative exponential in analogy to phenomena as the radioactive decay. Our hypothesis to be tested in the following is that this retreat will be the stronger the smaller the fraction $p_G(t)=n_G(t)/N(t)$ of German speakers in a settlement with $N(t)$ inhabitants composed of $n_G(t)$ German and $n_H(t)= N(t)-n_G(t)$ Hungarian speakers was (remember that we omit speakers of other languages). To model that situation we assume that in each settlement after one year's time the fraction $p_G(t + 1)$ of German speakers will be proportional to that fraction in the preceding year $p_G(t)$ weighted by a factor $\exp[-m\cdot(1-p_G(t)]$ which takes account of the fraction of German speakers in the settlement in the way that the lower the fraction $p_G(t)$ the faster is its decrease.

The probable fraction of German speakers after one year can then be written

$$p_G(t+1) = p_G(t)\cdot\exp\left[-m\cdot(1-p_G(t))\right] \qquad (1)$$

where $p_G(t)= n_G(t)/N(t)$ and m is the annual magyarisation rate averaged over all settlements in the limit of a negligible fraction of German speakers.

3.3 Search for diffusional interaction

We consider an additional impact on language shift through interactions $F_G(\mathbf{r}, t)$ with German speakers and $F_H(\mathbf{r}, t)$ with Hungarian speakers in the neighborhood settlements in a similar way as applied to the shift of Slovenian to German in Carinthia [17]. For each settlement at position $\mathbf{r}$ and time t we obtain Eq. 2 for the probable fraction $p_G(\mathbf{r}, t + 1)$ to speak German in the settlement located at position $\mathbf{r}$ at time $t + 1$ by assuming the probable fraction of German speakers $p_G(\mathbf{r}, t + 1)$ to be normalized to the total number of speakers and the total interaction $F_G(\mathbf{r}, t)$ and $F_H(\mathbf{r}, t)$ in that settlement.

$$p_G(\mathbf{r}+\mathbf{1},t) = \frac{n_G(\mathbf{r},t)\cdot\exp\left[-m\cdot(1-p_G(\mathbf{r},t))\right]+F_G(\mathbf{r},t)}{N(\mathbf{r},t)+F_G(\mathbf{r},t)+F_H(\mathbf{r},t)}. \qquad (2)$$

Each interaction term $F_G(\mathbf{r}, t)$ is a sum over the contributions of German speakers in all other settlements surrounding the initial settlement at position $\mathbf{r}$ and in an analogous way this applies to $F_H(\mathbf{r}, t)$ for Hungarian speakers.

The contributions of all other settlements positioned at $\mathbf{r}_j$ surrounding the initial settlement at position $\mathbf{r}$ are modeled by two-dimensional Gaussian functions identical to distributions describing the diffusion of particles in physics or chemistry:

$$F_G(\mathbf{r},t) = \sum_{r_j \neq r} \frac{n_G(\mathbf{r_j},t)}{4\pi D_G \cdot t} \cdot \exp\left(-\frac{\left|\mathbf{r}-\mathbf{r_j}\right|^2}{4D_G \cdot t} \right) \quad (3)$$

where D_G is the diffusivity of German language, that is, a measure for its spread. The diffusivity can also be seen as measure for the region of influence of a language. $F_H(\mathbf{r},t)$ is conceived in an analogous way, with D_H the diffusivity of Hungarian language.

To take account of weighted contributions from neighboring cells (here neighboring settlements) is a method known as cellular automaton or agent-based modeling.

3.4 Evaluation Procedure

Simulations were performed using GNU Octave 4.0.0. The data from the first census in each period (1880 till 1920) were set as the initial state from which the number of speakers in each settlement changes assuming a linear population development between censuses. To evaluate the goodness of fit between simulated data and census data, we used ordinary least squares to minimize the squared sum of errors. Optimization was done using the Nelder–Mead method [27].

4. Results

In Fig. 3 we plot the result for $p_G(1930)$ divided by $p_G(1880)$, i.e. the fraction of German speakers in 1930 relative to the initial fraction in 1880, vs. the initial fraction in 1880. Even though data scattering obliterates the graph it is evident that the description of $p_G(1930)/p_G(1880)$ vs. $p_G(1880)$ by an exponential is an adequate hypothesis.

From the best fit of the model to the data we receive the average magyarisation rate m for the time period from1880 till 1930 as 0.0192 ± 0.0010 per year. The mean deviation of the best simulation of the German population of each of the fifty settlements in our sample from the census data is 74 persons. The RMS error is 186. This should be seen in relation to the average size of the settlements in 1880, 3,433 persons, and the average number of German speakers in the settlements, 1,633 persons.

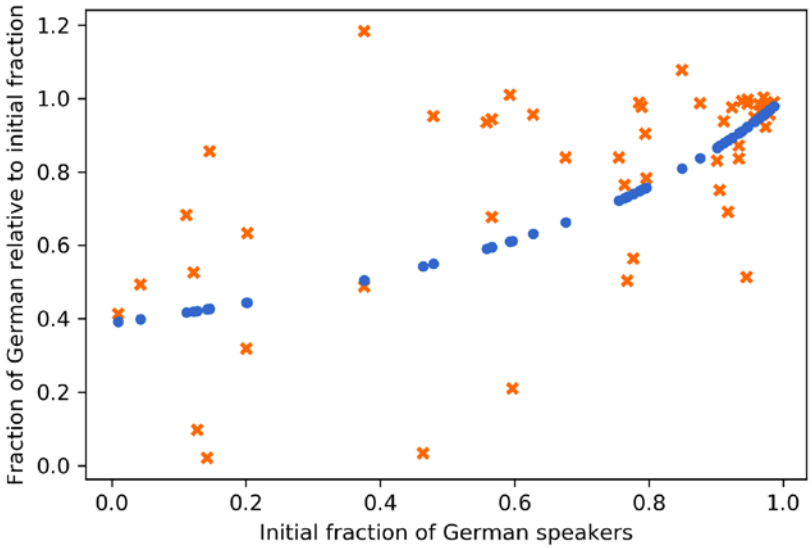

Fig. 3. Fraction of German speakers in 1930 relative to initial fraction in 1880 vs. initial fraction in 1880. Orange crosses are the census data, blue dots the results of the best fit with the model of Eq. (1).

Literature [23-25] tells us that magyarisation pressure was less strong before the turn to the twentieth century but increased later on. For comparison Fig. 4 shows a plot for the earlier period.

Indeed a fit of the data yields a 23 percent lower annual magyarisation rate m=0.0148 ± 0.008 from 1880 till 1910 compared to m=0.0192 ± 0.010 for the full period from 1880 till 1930.

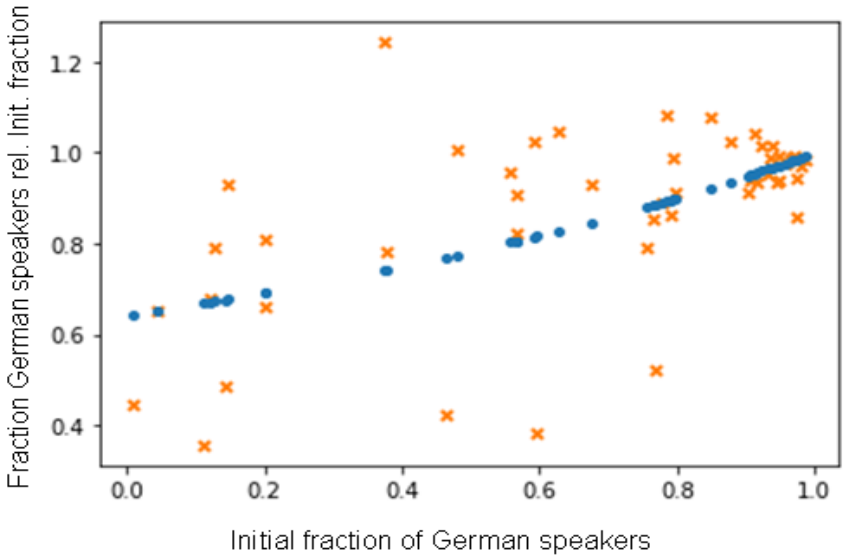

Fig. 4. Fraction of German speakers in 1910 relative to initial fraction in 1880 vs. initial fraction in 1880. Orange crosses are the census data, blue dots the results of the model calculated with Equ. (1).

Here the mean deviation of the best simulation of the German population of each of the fifty settlements in our sample from the census data is 49 persons. The RMS error is 146.

Our attempts to determine an additional influence of language contact with surrounding settlements by varying the diffusivities D_G and D_H was unsuccessful. There was no relevant improvement of the model fit by considering diffusive contact.

5. Conclusions and Outlook

Our central result is: the linguistic magyarisation proceeds intrinsically, which means that it is due to contact within the settlement, be it by interactions with other persons in the settlement or by the government's magyarisation policy. Magyarisation proceeded the faster the higher the fraction of Hungarian speakers in the settlement was. The mean annual magyarisation factor m for the period from 1880 till 1910 was 0.0148 increasing for the full period till 1930 by 23 percent to 0.0192.

In contrast to previous studies [3, 16, 17] we cannot trace down an effect of diffusive contact with other settlements on language shift.

5.1 Comparison with language shift under different preconditions (Southern Hungary vs. Southern Carinthia and Northwestern Scotland)

In the present study of Southern Hungary the languages were Hungarian and German. The object of a previous study has been Southern Carinthia, Austria, where the languages were German and Slovenian [17]. Kandler et al. [3, 16] have dealt with a related scenario in Northwestern Scotland.

There are two different preconditions regarding the language situations in Southern Hungary and in Southern Carinthia or Northwestern Scotland, i.e. our present study and the previous ones [3, 16, 17].

Firstly geographically (for a comparison see Fig. 5): Southern Hungary is different from Southern Carinthia (and from Northwestern Scotland) in that is predominantly flat country side, large settlements well separated from each other, the average distance between settlements being significantly larger in Southern Hungary than in Southern Carinthia.

In contrast, Southern Carinthia is a geographically heavily structured region with many of its settlements in the narrow valleys and small basins betweens hills and mountains. Thus often a settlement where the majority spoke German was in close neighborhood to another settlement with Slovenian language majority, particularly close to the language border. Diffusive contact between the settlements had a significant influence on language shift [17]. The situation in Northwestern Scotland is similar [3, 16].

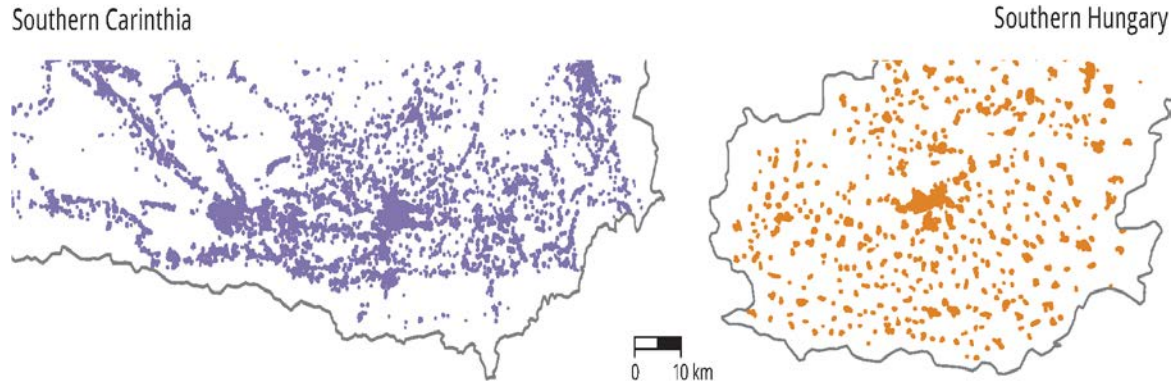

Fig. 5. Settlement structure of parts of Southern Carinthia and Southern Hungary on the same scale. Color shows populated areas. All settlements are shown regardless of the languages used there.

Secondly historically: In Hungary German settlers were given land by the Hungarian nobility or the crown in widely devastated country often founding new settlements quasi as language islands.

In Carinthia, different preconditions prevailed: the immigration of a population speaking a Slavic language is timed to about 600 AD, the immigration of German speakers from about 800 AD on. Both immigrations may have been “diffusion processes”. By the time a language border between German and Slovenian developed which slowly moved southwards. Thus these two languages coexist since about 1.200 years and their peaceful interaction has a long history, only interrupted by World Wars I and II and the time in between.

5.2 Interpretation of the consequences of different preconditions

Historical conditions in Southern Carinthia and in Northwestern Scotland created a moving language border, and close to that border language shift was significantly promoted by diffusive contact between different settlements. We have shown that such contact was of no significant influence on language shift in Southern Hungary where no language border exists. In Southern Hungary language shift rather occurred intrinsically, the prevalent linguistic composition of the settlements playing a dominant role.

Furthermore, in Southern Hungary linguistic homogeneity of population groups “implanted” in a short time period forming rather large settlements obviously promoted linguistic inertia: In settlements with German prevailing, language shift proceeded retarded in comparison with linguistically mixed settlements.

We keep this difference as a paradigm for two very different situations. There may be interesting analogies to present scenarios. Language border motion will become rare, though, and in the majority of cases language shift will rather manifest through the shrinking or finally dissolution of language islands. These may be the remnants of former widely spread language regions as presently the Slovenian language islands in Southern Carinthia [17] or modern quasi-language islands as formed by “implanted” immigrants in cities of the Western world. For the latter scenarios data over a longer time period still hardly exist, though, i.e. why data collection is necessary and should be performed in order to enable serious treatments and predictions.

ACKNOWLEDGEMENTS. The team of the Research Center of Hungarian Germans (Forschungszentrum der Ungarndeutschen) at Eötvös Loránd University, Budapest, is gratefully acknowledged for their support: E. Knipf-Komlósi, K. Manherz, M. Müller, M. Erb. We thank T. Tölgyesi (University of Vienna) for help in translating the Hungarian census data, and M. Glauninger (Austrian Academy of Sciences/University of Vienna), S. Puchegger and H. Peterlik (both University of Vienna) for many helpful discussions and input. Geographical data for Figs. 1 and 2 were obtained from the MPIDR Population History GIS Collection (http://censusmosaic.org/data/historical-gis-files), OpenDEM (www.opendem.info), KAGIS Kärnten Atlas (https://gis.ktn.gv.at) and OpenStreetMap (www.openstreetmap.org). K. Prochazka was supported by a uni:docs fellowship from the University of Vienna. We thank San Yu Lai from University Hongkong for his contributions during a research internship at our university.
Calculations by GNU octave and python.

References

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

[2] C. Castellano, S. Fortunato, V. Loreto, Rev. Mod. Phys. 81 **(**2009) 591.

[3] A. Kandler, Hum. Biol. **(**2009) 181; A. Kandler, R. Unger, in: A. Bunde, J. Kärger, J. Caro, G. Vogl (Eds.), Diffusive Spreading in Nature, Technology and Society, Springer, Cham, 2018, pp. 351-373; 2nd ed. 2023.

[4] R.V. Solé, B. Corominas-Murtra, J. Fortuny, J. R. Soc. Interface 7 (2010) 1647.

[5] F. Ghanbarnejad, M. Gerlach, J.M. Miotto, E.G. Altmann, J. R. Soc. Interface 11 **(**2014) 1044.

[6] J. Burridge, Phys. Rev. X 7 (2017) 031008.

[7] Q. Feltgen, B. Fagard, J.-P. Nadal, R. Soc. Open Sci. 4 (2017) 170830.

[8] J. Burridge, R. Soc. Open Sci. 5 (2018) 171446.

[9] D.M. Abrams, S.H. Strogatz, Nature 424 **(**2003) 900.

[10] M. Patriarca, E. Heinsalu, Physica A 388 (2009) 174.

[11] L. Chapel, X. Castelló, C. Bernard, G. Deffuant, V.M. Eguíluz, S. Martin, M. San Miguel, PLOS ONE 5 (2010) 1.

[12] M. Patriarca , X. Castello, J.R. Uriarte, V.M. Eguiluz, M. San Miguel, Advances in Complex Systems 15 (2012) 1250048-1.

[13] C.E. Walters, Meccanica 49 (2014) 2189.

[14] N. Isern, J. Fort, J. R. Soc. Interface 11 (2014) 20140028.

[15] A. Carro, R. Toral, M. San Miguel, New J. Phys. 18 (2016) 113056.

[16] A. Kandler, R. Unger, J. Steele, Phil. Trans. R. Soc. B 365 (2010) 3855.

[17] K. Prochazka, G. Vogl, Proc. Nat. Acad. Sci. USA 114 (2017) 4365.

[18] T. Barrett-Walker, M.J. Plank, R. Ka'ai-Mahuta, D. Hikuroa, A. James, J. R. Soc. Interface 17 (2020) 20190526.

[19] J.R. Uriarte, J. Sperlich, PLoS ONE 16 (2021) e0252453.

[20] S. Mimar, M.M. Juane, J. Mira, J. Park, A. P. Muñuzuri , G. Ghoshal, P.R Res. 3 (2021) 023241.

[21] N.H. Hornberger, in: R.B. Kaplan (Ed.), The Oxford Handbook of Applied Linguistics, Oxford University Press, Oxford (2010), pp. 412–420.

[22] K. Manherz, Die Ungarndeutschen, Útmutató, Budapest 1998.

[23] G. Seewann, Geschichte der Deutschen in Ungarn. Band 2: 1860 bis 2006, Herder-Institut, Marburg 2012.

[24] L. Marácz, Jezikoslovlje 13 (2012) 269.

[25] J.C. Swanson, Tangible Belonging: Negotiating Germanness in Twentieth-Century Hungary, University of Pittsburgh, Pittsburgh 2017.

[26] Scans of census publications are available from the online collection HUNGARICANA: https://library.hungaricana.hu/en/collection/kozponti_statisztikai_hivatal_nepszamlalasi_digitalis_adattar. Digitized data used in this paper have been deposited on figshare (https://doi.org/10.6084/m9.figshare.6115634.v1).

[27] J. Nelder, R. Mead, Comput J 7 (1965) 308.