languages

MDPI

Article

The Application of Geospatial Analysis Methods for the Reconstruction of Lithuanian–Slavic Ethnolinguistic Boundaries in Southeastern Lithuania

Aidas Gudaitis

Centre of Geolinguistics, The Institute of the Lithuanian Language, 10308 Vilnius, Lithuania; aidas.gudaitis@lki.lt or aidas.gudaitis@gmail.com

Abstract: (1) Background: The article addresses the issue of geospatial dynamics of Lithuanian–Slavic ethnolinguistic boundaries in Southeastern Lithuania (SEL) that were influenced by long-term Lithuanian–Slavic linguistic competition. The aim of the study was to reconstruct the Lithuanian–Slavic ethnolinguistic boundaries and reveal the intensive contact zones in the late 19th century based on published data. Additionally, the study aimed to assess the geospatial changes in the ethnolinguistic situation in the research area during the period 1890–2021. (2) Methods: The ESRI ArcGIS technology geoprocessing tools were applied for boundary reconstruction and geospatial change detection. Cartographic materials, statistical data, and national census information were utilized in the process. (3) Results: The gained results provided a better understanding of Lithuanian–Slavic ethnolinguistic dynamics over space and time in the research area. The study reveals that the ethnolinguistic boundary in the Vilnius–Trakai urbanized area shifted in favor of the Lithuanian language, suggesting its potential influence on the metropolitan suburbs in the future. However, insufficient social infrastructure and weak economic development in rural settlements have led to a negative migration balance, a low birth rate, and rapid population aging. These challenges might have a negative effect on the future survival of the Lithuanian language in the multilingual rural area of SEL, especially considering the recent geopolitical realia in the region. (4) Conclusions: The study anticipates an increase in the influence of the Lithuanian language in the Vilnius–Trakai metropolitan area at the expense of further decline in the rural Lithuanian-speaking population in the next decade.

Keywords: geolinguistics; GIS; geospatial analysis; Lithuanian–Slavic ethnolinguistic boundary; Southeastern Lithuania

1. Introduction

In the last decade, the global development of geoinformation systems[1] (GIS) and location intelligence[2] (LI) technologies has gone beyond their traditional application fields, such as geodesy, remote sensing, transport, spatial planning, or public security, and is actively expanding into new areas, such as financial management, sociology, epidemiology, or statistics (Ripley 2005; Krieger 2003; Waters 2017; etc.).

Geolinguistics, which studies the geographical distribution of linguistic phenomena and their alternation trends that can be affected by political, economic, cultural, or natural factors, is not an exception. The application of GIS technology as well as the fundamentals of geography and cartography in geolinguistics has enabled the integration of both well-developed and innovative issues within the research field (Hoch and Hayes 2010, p. 24; see also: Lameli et al. 2010; Ayad and Luthin 2009). These circumstances provided the possibility to create digital data models that can help to analyze the spatial and temporal dynamics of linguistic phenomena as well as predict their future scenarios.

Among various geolinguistic research areas, GIS also provides substantial benefits for the research of multilingual environments. In this field, GIS technology can be applied to identify and map language contact zones and estimate language influences (McGuirk 2004).

When analyzing the processes of ethnolinguistic dynamics in Lithuania, there is an obvious lack of quantitative research based on statistical data. A more unprejudiced analysis of the ethnolinguistic and ethnographic processes of the Lithuanian population would allow for a more critical assessment of the tendencies of dynamics in both the ethnic identity and spoken languages in the region.

The first attempts at applying GIS-based geospatial processing techniques in geolinguistic research in Lithuania started a few years ago (Čepaitienė and Gudaitis 2021, 2022). The insights gained from previous research encourage us to continue using these methods as well as applying them in the analysis of the dynamics of ethnolinguistic processes in the multilingual environment of Southeastern Lithuania (SEL).

The territorial distribution of the Lithuanian language and its historical dynamics have been and continue to be extensively researched (Anonim 1898; Rozwadowski 1904; Chominski 1933, 2010; Gaučas 1993, 2004; Čekmonas and Grumadienė 1993; Kurila 2005a, 2005b; Zinkevičius 1998, 2006; Liseichykau 2017; etc.). Studies show that during the period between the early 16th and ate 20th centuries, the area of the Lithuanian language almost halved from ~103,000 sq. km to ~59,000 sq. km due to unfavorable socio-cultural conditions (Figure 1, red arrows show the historical contraction directions of the Lithuanian language area).
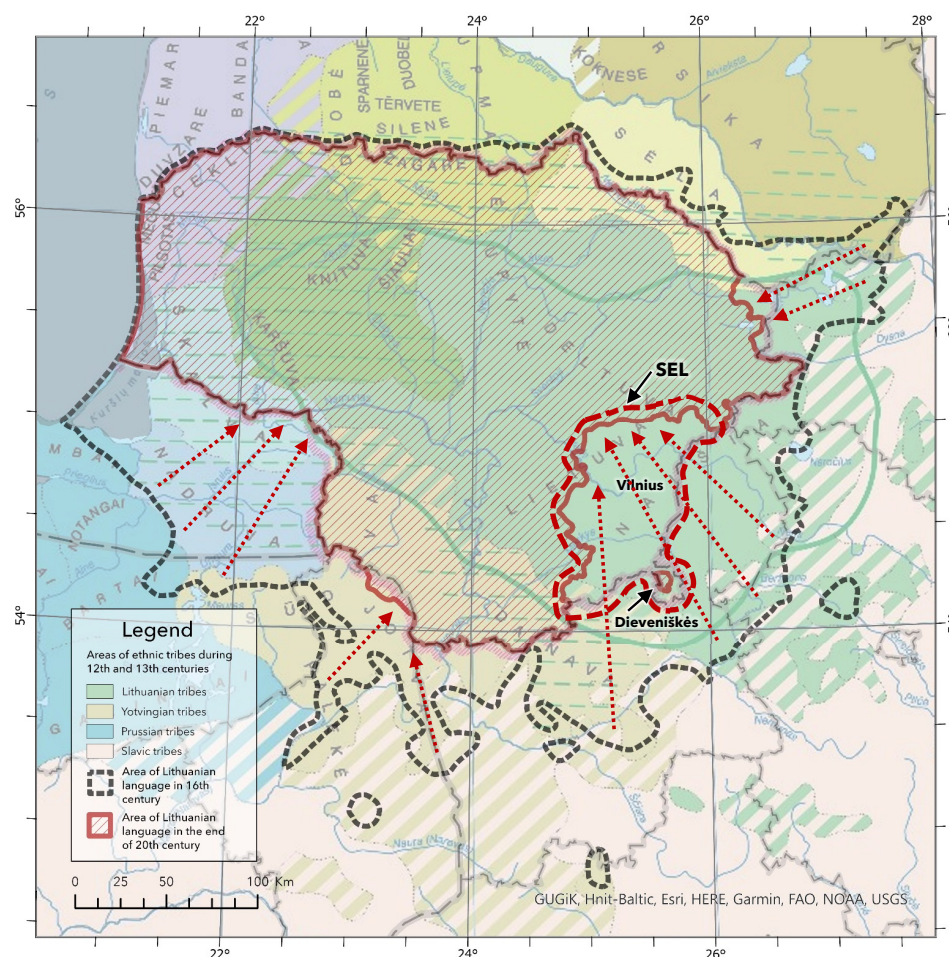

Figure 1. The dynamics of the Lithuanian language spread during the early 16th–late 20th centuries. Basemap—territories of the Baltic and Slavic tribes in the late 12th–early 13th centuries (Gudavičius 2008; the map by Science and Encyclopedia Publishing Center), the 16th-century boundary of the Lithuanian language according to Z. Zinkevičius (black dashed line, Zinkevičius 2006, p. 106), the late 20th-century ethnolinguistic boundary of the Lithuanian language according to V. Čekmonas, P. Gaučas, and L. Grumadienė (red stenciled area, Čekmonas and Grumadienė 1993). Red arrows show contraction directions.

The assimilation processes of Lithuanian speakers became particularly acute from the beginning of the 19th century and stabilized only at the end of the 20th century when Lithuania restored its independence. The intensive assimilation processes significantly affected the southeastern part of the Lithuanian-speaking area through the influence of Slavic languages. An illustration of such assimilation is the Dieveniškės case (see Figures 1 and 2), where, according to statistical data (Anonim 1898; State Data Agency of Lithuania 2021a, 2021b, 2021c, 2021d), the Lithuanian-speaking population decreased more than 13 times (from 7108 to 523 inhabitants) during the period 1890–2021.
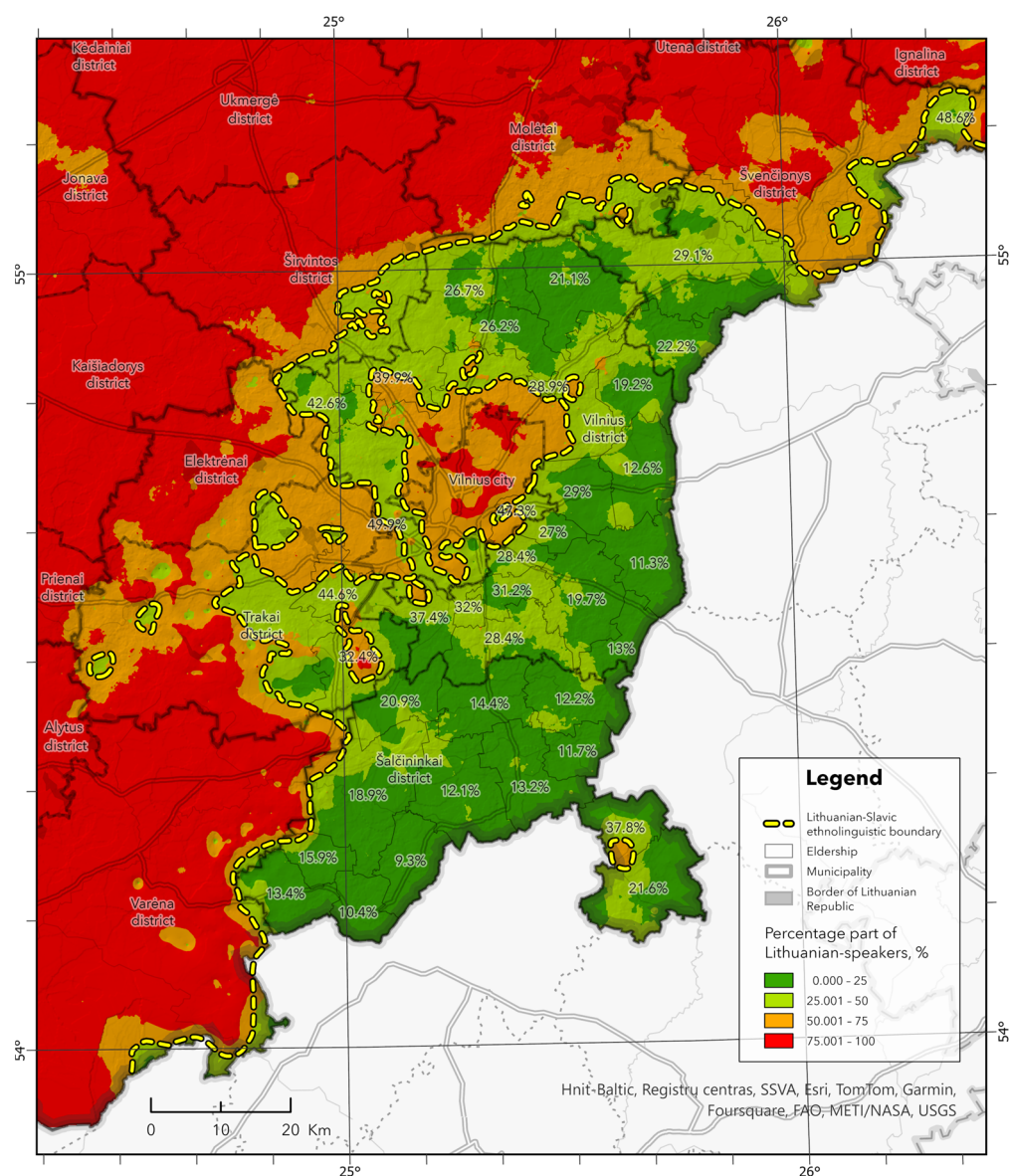

Figure 2. Ethnolinguistic situation in SEL according to the 2021 census (the percentage values of Lithuanian-speaking population are shown for the elderships with less than 50% Lithuanian speakers).

The SEL region (Figures 1 and 2) is currently characterized by a heterogeneous multilingual environment with varying linguistic attitudes and national identity assessments among the local population (Gaučas 1992, pp. 141–43; Zinkevičius 2006, pp. 105–18). Nowadays, four languages, Lithuanian, Polish, Belorussian, and Russian as well as their dialects, are actively used for communication in this area.

The term "Southeastern Lithuania (SEL)" used in this article refers to a region discussed by Lithuanian scholars for its historical, socio-economic, and socio-cultural significance as being an important multilingual and multicultural area. The SEL area does not have a formal administrative designation and current falls within the County of Vilnius. It also forms a part of the ethnographic region of Dzūkija. During the period of the Russian Empire, the area belonged to the Vilna Governorate, with Vilna (Vilnius) City as its administrative center. In the interwar years (1921–1939), this region was part of the Republic of Poland and was commonly referred to as Wileńszczyzna by Polish scholars and media, a term that persists in Polish literature.

The status of other Lithuanian-speaking enclaves within the neighboring Republic of Belarus remains uncertain. According to the latest available public survey by the National Statistical Committee of the Republic of Belarus (2019), around 1300 individuals of Lithuanian nationality, constituting approximately 25% of the Lithuanian population in Belarus, reside in the neighboring Ostrovetsky, Voronovsky, and Lida districts (National Statistical Committee of the Republic of Belarus 2019). However, these individuals represent less than 1% of each district's total population. Only 327 individuals in these districts reported the Lithuanian language as native. Based on prior research on Lithuanian-speaking communities in Gervėčiai, Pelesa, and Rodūnia conducted in the late 20th century (Čekmonas and Grumadienė 1993; Zinkevičius 1998, 2006), the Lithuanian language was considered vibrant within these communities (Figure 10). Unfortunately, geopolitical factors have restricted Lithuanian linguistic researchers' access to these communities since 2014, preventing a more precise assessment of their current linguistic situation.

The aim of this study is to reconstruct the ethnolinguistic boundaries and reveal contact zones between Lithuanian and Slavic languages of the SEL based on the data published in the late 19th and early 20th centuries as well as assess their geospatial change over the last 130 years (1890–2021) by applying GIS-based geoprocessing techniques. This study focuses on methodological issues, presenting one of the technological ways to reconstruct ethnolinguistic boundaries. The geospatial processing methods and principles of spatiotemporal change analysis were applied for the first time in the research area. The techniques used can provide new technological opportunities to deepen the understanding of ethnolinguistic processes and assess future trends in their development.

2. Materials and Methods

2.1. Study Area

The study area, where, according to the 2021 census data, more than 50% of the population indicated their mother tongue language as Slavic (Polish, Russian, Belarusian, or their variants), covers around 7% of the territory of the Republic of Lithuania (~4370 sq. km) and, excluding the main urban zones, has around 154,000 inhabitants (State Data Agency of Lithuania 2021a, 2021b, 2021c, 2021d). Figure 2 represents the study area with 38 elderships[3] where the percentage of Lithuanian-speaking population is below 50%.

2.2. Study Material and Context

In talking about the selection of a period beyond the latter half of the 19th century for ethnolinguistic boundary reconstruction, the main factor lies in the shortage of precise geolinguistic and cartographic data. Quantitative research methods require accurate data, but no precise linguistic investigations or linguistic maps within the designated research area were available prior to the latter half of the 19th century.

One of the first notable linguistic attempts from that period is the Lithuanian language distribution map "Karte des littauischen Sprachgebiets" (Map of the Lithuanian language area), compiled by Friedrich Kurschat in 1876 (Kurschat 1876; Mikulėnienė 2018, pp. 69–72). The demarcation of the approximated southeastern boundary of the Lithuanian language was based on limited scientific research data at that time (Figure 3).

Figure 3. The map of the area of the Lithuanian language in the second half of the 19th century by Friedrich Kurschat (Kurschat 1876; Mikulėnienė 2018, pp. 69–72).

For the reconstruction of the historical ethnolinguistic boundaries of the late 19th century in the research area, the manuscript map "Mapa języka litewskiego w gubernii Wileńskiej" (Map of the Lithuanian language in Vilnius Governorate) published by Jan Michal Rozwadowski was used (Rozwadowski 1901, 1904; Figure 4). In terms of cartography, the map is sufficiently precise. The scale of the map is 1:600,000. It includes over 600 accurately geolocated and marked settlements. The settlements on the map are classified into six classes according to the percentage of Lithuanian-speaking inhabitants (percentage classes of Lithuanian-speaking population: 75–100%, 50–75%, 25–50%, 10–25%, 1–10%, none—0%). Most of the identified settlements were the centers of districts (Pol. *gmina*) or communes (Pol. *włość*) of the Vilnius Governorate.

It should be noted that Rozwadowski's map was based on the survey data on the ethnic and linguistic composition of the population of the Vilnius Governorate collected through the initiative and strong support of Count Adam Alfred Gustaw Broel-Plater (1836–1909)[4] in 1890. The survey data were summarized and published in 1898 by an individual who called himself Anonym (Anonim 1898, pp. 3–72). The significance of Anonym's work lies in the fact that, for the first time in the region's history, population survey data were collected and aggregated not according to ethnicity but the main spoken language. This source played an important role in the research for assessing the statistically valid ethnolinguistic situation in the late 19th century as well as adjusting the Lithuanian–Slavic language boundaries and main contact zones in the research area.

The main drawback of the Plater–Anonym–Rozwadowski survey is that it does not reflect the population's multilingualism, even though a significant proportion of the population in the region was bilingual or even trilingual at the time (Vidugiris 1993, p. 121; Zinkevičius 2006, pp. 105–19; Čekmonas 2017, pp. 115–26). It is likely that this method's weakness might have led to an underestimation of the number of Lithuanian speakers in the region (Gaučas 1993, p. 65).

The obtained ethnolinguistic boundary reconstruction results of the late 19th century were compared with geostatistical data from the Lithuanian censuses of the early 21st century (State Data Agency of Lithuania 2021a, 2021b, 2021c, 2021d) as well as other sources.

The methods of the reconstruction of the ethnolinguistic boundaries and the analysis of the spatial data are described below.

Figure 4. Ethnolinguistic manuscript map "Mapa języka litewskiego w gubernii Wileńskiej" by J. M. Rozwadowski (1901) used for ethnolinguistic Lithuanian–Slavic boundary reconstruction.

2.3. Research Methods

For the ethnolinguistic boundary reconstruction, the contact zone revelation, and their spatiotemporal change analysis, the ESRI ArcGIS platform application ArcGIS Pro 3.2.0[5] and its geospatial data editing and analysis tools were applied.

The reconstruction and analysis process itself was carried out in six stages (Figure 5).

Below, on the example of the works published by Anonym in 1898 and Rozwadowski in 1901, the stages of the process used are elaborated:
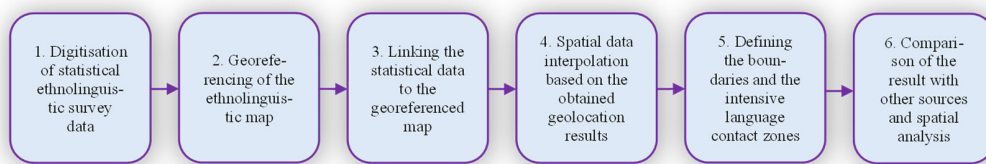
Figure 5. Stages in the process of reconstructing ethnolinguistic boundaries.

1. Digitization of statistical ethnolinguistic survey data. At this stage, the digitized text (Figure 6) of Anonym's 1898 article "Obszar jezyka litewskiego w gubernii Wilenskiej" (The area of the Lithuanian language in the Vilnius governorate), in which he provided detailed statistics on the main language of the population, was extracted and transferred to ArcGIS Pro for further geoprocessing.

Wykaz włości i gmin.						
Włości litewskie		I.	II.	III.	IV.	V.
I. Gielwany.						
1 Gm. Gielwany	1921	—	12	1609	300	—
2 „ Bogusławiszki	1244	25	97	922	200	—
3 „ Powicina	594	—	52	542	—	—
4 „ Soboliszki	1325	25	58	1242	—	—
5 „ Kiejżany	1101	126	—	975	—	—
6 „ Czabiszki	1066	—	46	908	112	—
7 „ Gielwanki	1052	100	29	923	—	—
Razem	8303	276	294	7121	612	—

Figure 6. Excerpt of statistical table from Anonym's 1898 article "Obszar jezyka litewskiego w gubernii Wilenskiej" (Gielwany districts, Pol. gminy). Population survey classification by main speaking language: I. Russians, II. Belorussians-Poles, III. Lithuanians, IV. Jews, V. Others.

2. Georeferencing of the ethnolinguistic map. The high-resolution image of Rozwadowski's map (width: 7649 px, height: 8310 px, 380 dpi) was georeferenced using the ArcGIS Pro imagery editing tool and bound to the LKS-94 coordinate system[6]. The location binding was performed by attaching the image to at least 360 known geographical locations.

In the georeferenced map, 638 historical settlements were identified. The average percentage values of Lithuanian speakers that were indicated in the map legend were assigned to identified settlements. For example, if a settlement was categorized as 75–100% of Lithuanian speakers, an average value of 87.5% was assigned to it. The assigned percentage values were used later for the data interpolation process that is described in step 4.

3. Linking the statistical data to the georeferenced map. At this stage, the statistical data from Anonym's article was linked to the geolocated settlements of Rozwadowski's map. The linking was performed according to the main settlement attributes (address name, district (Pol. *gmina*), commune (Pol. *włość*)). A total of 464 settlements from 473 listed in Anonym's article were linked with the identified locations from Rozwadowski's map. These linked settlement geolocations were supplemented with the values of the number of inhabitants per linguistic group and their percentages, which were later used in the spatial analysis (Figure 7).
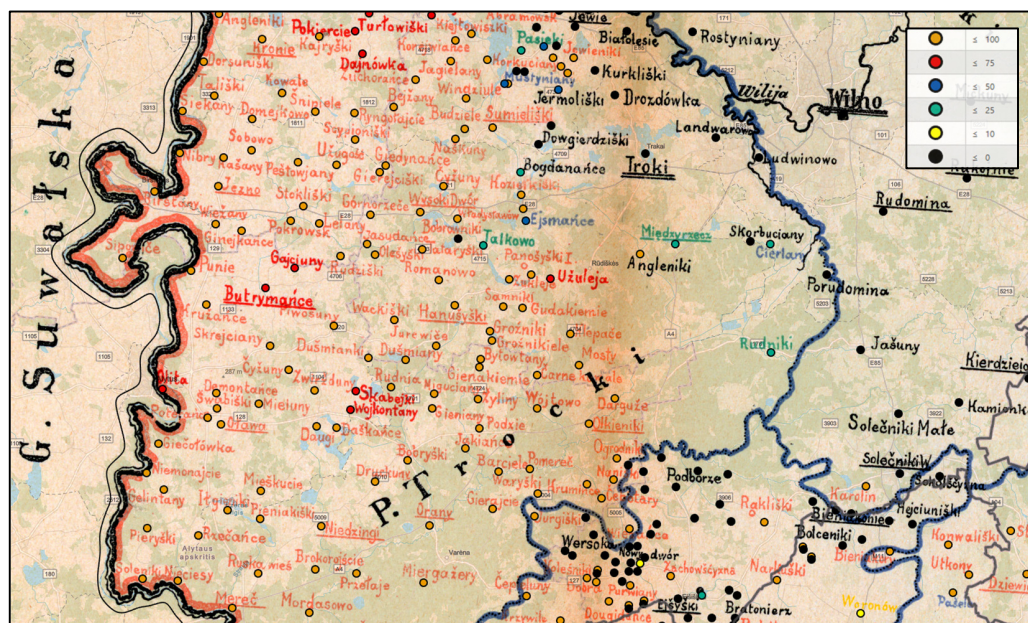
Figure 7. Fragment of Rozwadowski's map with the localized settlements (colors of the dots indicate the distribution of Lithuanian speakers according to their percentage share in a settlement).

For each of the linked settlements, the percentage of the Lithuanian and Slavic-speaking (Russian + Polish-Belarussian) population from the total Lithuanian and Slavic-speaking population (Lithuanian + Russian + Polish-Belarussian) of the settlement was calculated. Because the estimation of the ethnolinguistic boundary was based on the Lithuanian–Slavic percentage ratio, the calculations followed the Rozwadowski rule and did not include data from other language groups, such as Jews, Germans, Latvians, Tatars, etc. Otherwise, the percentage ratio would either be constantly to the disadvantage of the Lithuanian language or it would require additional data manipulation (Rozwadowski 1904, p. 91).

For the geolocated settlements from Rozwadowski's map, which were not linked to the statistical data from Anonym's article (156 settlements), the average percentage values of the Lithuanian-speaking population according to Rozwadowski's map classification were assigned.

After linking the statistical data to the localized settlements, the localization refining task was performed. In this phase, the georeferenced detailed military topographic maps of German, Russian, Lithuanian, and Polish origin dating from 1914 to 1942, which were accessible online in the geoportal of the Wroblewski Library of the Lithuanian Academy of Sciences (Wroblewski Library of the Lithuanian Academy of Sciences 2019), were used. Of particular significance was the 1914–1921 German military topographic map at a scale of 1:100,000. The historical locations of settlements were primarily adjusted by referring to this map.

4. Spatial data interpolation based on the geolocation results. At this stage, interpolation[7] of the distribution of the percentage of the Lithuanian-speaking population of localized settlements in the territory of the Vilnius Governorate was carried out.

The Empirical Bayesian Kriging[8] (EBK) interpolation was chosen for this research as it is the most versatile and most conservative method with respect to the data, with the least distorting impact on results.

Due to being able to perform geospatial comparisons between the results of different time periods, all the geospatial analysis results in this study were generated within the administrative boundaries of the former Vilnius Governorate of the period 1843–1915[9] (Figure 8).
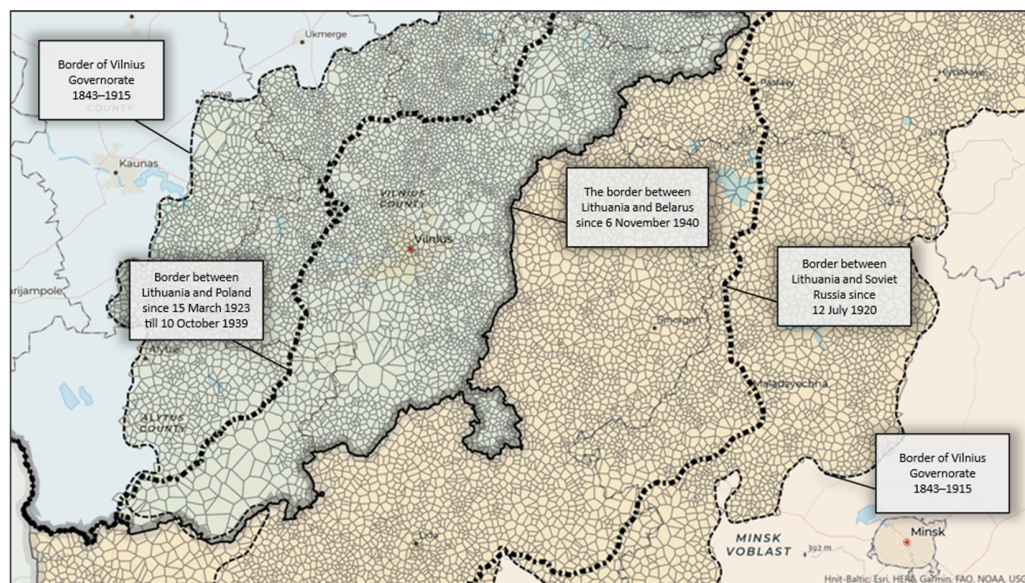
Figure 8. The fragment of the administrative boundaries of the Vilnius Governorate (1843–1915) and other historic boundaries that were important for the research.

5. Defining the ethnolinguistic boundaries and the intensive language contact zones. The contour geoprocessing tool[10] was used to extract the ethnolinguistic Lithuanian–Slavic boundaries. The boundaries were extracted by deriving an isoline based on the value of 50% of the percentage of the Lithuanian-speaking population in the interpolated raster surface. The isolines of 25% and 75% of the percentage of the Lithuanian-speaking population were derived with the purpose of delineating the intensive Lithuanian–Slavic language contact zones.

6. Comparison and spatial analysis of the results with other sources. Various spatial analysis methods were used to compare the results with studies by other authors as well as the latest public census data. In order to compare the reconstructed ethnolinguistic situation of 1890 with the data of the 2021 public census, the ArcGIS Pro change detection raster statistics tool[11] was used, that allowed to detect statistically significant changes between two raster datasets. In this research, the pixel value change evaluation method was applied. After computing the difference between two raster datasets, the tool generated a new raster dataset containing the change information. The tool can be combined with other tools for a more complex workflow.

The reconstruction and analysis yielded a number of results which are summarized in the section below.

3. Results

3.1. The Boundary Reconstruction and Ethnolinguistic Situation Analysis

The statistical analysis of the digitized and processed Anonym's article data shows (Figure 9) that at the end of the 19th century, the population of the Vilnius Governorate was close to 1 million (998,611 inhabitants). According to the main spoken language, the population of the governorate was distributed, as follows: Polish-Belarussian: 688,915 (69.0%), Lithuanian: 232,790 (23.3%), Yiddish: 62,005 (6.2%), Russian: 13,998 (1.4%), and other: 1403 (0.1%).

It should be noted that the data provided by Count Adam Plater and processed by Anonym do not contain data for Vilnius City as well as some urban centers and large cities (Anonim 1898, p. 10). Anonym's article refers to the 1892 Governorate's Yearbook, in which a population of 1 million 336 thousand was recorded. Thus, Anonym, appealing to other empirical data from that period, estimated the total Lithuanian-speaking population

of the Vilnius Governorate at around 272 thousand inhabitants (about 20% of the total population of the Vilnius Governorate).
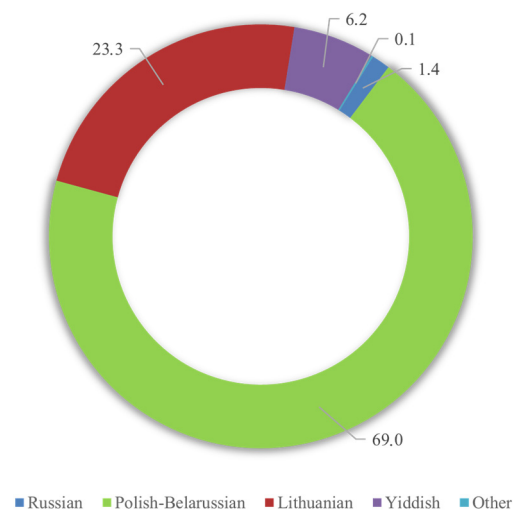

Figure 9. Statistics of the processed Plater–Anonim data (1898). Percentage (%) of the population of the Vilnius Governorate by main spoken language in 1890 (n = 998,611 inhabitants).

After the spatial data editing and refinement stages, the geospatial data interpolation process was carried out by applying the EBK interpolation method. The geospatial estimation of the Lithuanian–Slavic language distribution in the study area was based on the percentage of the Lithuanian-speaking population (Figure 10).
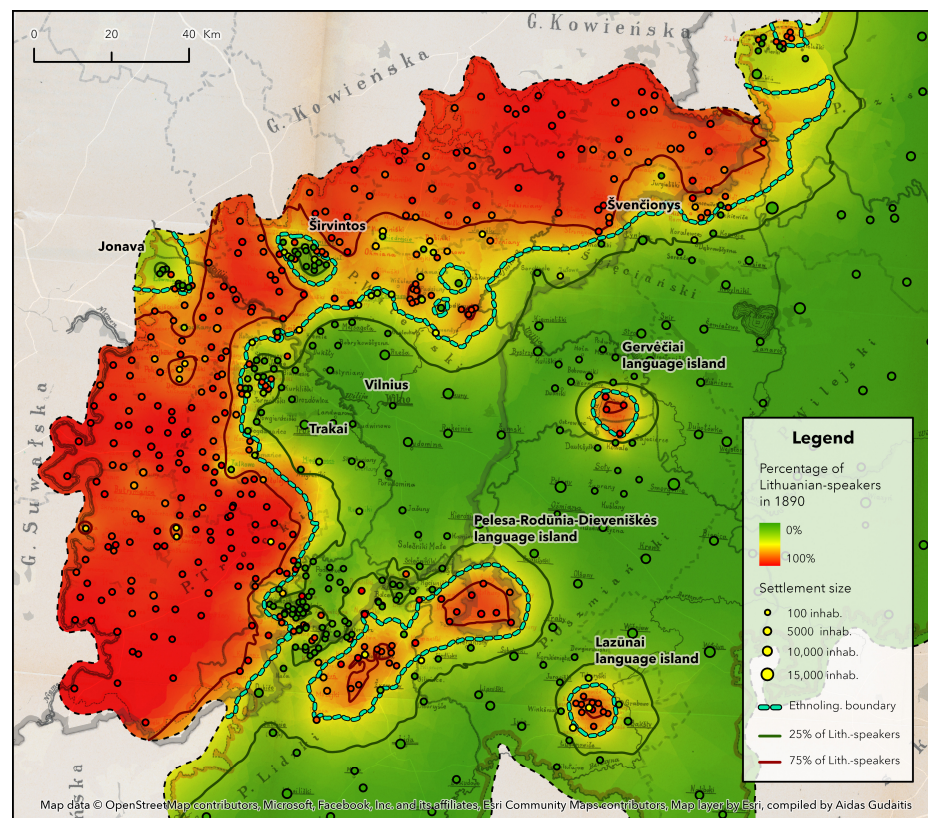

Figure 10. Map of the reconstructed Lithuanian–Slavic ethnolinguistic boundaries and language distribution in the Vilnius Governorate at the end of the 19th century.

The obtained ethnolinguistic reconstruction results revealed three main Lithuanian-speaker isolated zones: Lazūnai ([lɐ[1]ˈzuːnɐɪ], Blr. *Лаздуны*, Pol. *Łazduny*), Gervėčiai ([gʲɛrʲ[1]ˈʋʲeːʧʲɛɪ], Blr. *Гервяты*, Pol. *Gierwiaty*) as well as Pelesa-Rodūnia-Dieveniškės ([pʲɛlʲɛˈsɐ | [1]ˈroːduːnʲɛ | dʲiɛ[2]ˈʋʲæːnʲɪʃkʲeːs], Blr. *Пеляса-Радунь-Дзевянішкі*, Pol. *Pielasa-Raduń-Dziewieniszki*) zone, which, according to the Plater–Anonym–Rozwadowski data at the end of 19th century, had already been isolated from the main Lithuanian language array and became a linguistic island. The latter had also already started to split into smaller isolated parts, where Lithuanian speakers still constituted the majority (>75%).

On the other hand, two Slavic-speaking dominated linguistic islands in the vicinities of the Jonava ([jɔːnɐˈʋɐ]) and Širvintos ([ʃɪrʲʋʲɪn[2]ˈtoːs]) administrative centers in the north-western part of the research area (to the north-west from Vilnius) were identified.

It should be mentioned that the term "linguistic island" used in this article in the context of Lithuanian and other languages is based on the definition proposed by Kardelis (Kardelis 2014, p. 5; 2015; see also Metzler Lexikon Sprache 2005, p. 619).

The obtained reconstruction results were compared with the map "Mapa rozsiedlenia ludności polskiej na obszarze dawnej Rzeczypospolitej i krajów ościennych" (Map of the distribution of the Polish population in the former Polish-Lithuanian Commonwealth and neighboring countries), created by cartographer Józef Gruenberg in 1913 (Gruenberg 1913), which reflected the distribution of the Polish population according to their language and dialects (Figure 11).
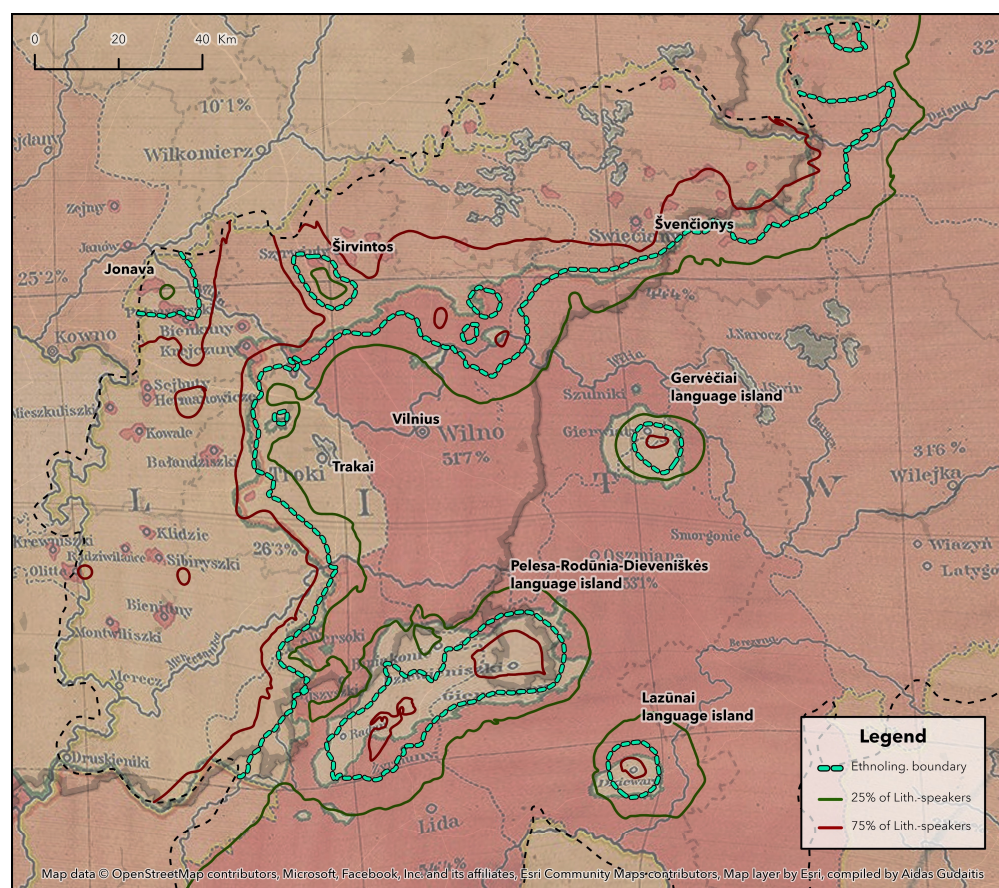

Figure 11. Comparison of the reconstructed Lithuanian–Slavic ethnolinguistic boundaries of the late 19th century with the ethnolinguistic boundaries extracted by Józef Gruenberg in the 1913 map "Mapa rozsiedlenia ludności polskiej na obszarze dawnej Rzeczypospolitej i krajów ościennych" (Map of the distribution of the Polish population in the former Polish-Lithuanian Commonwealth and neighboring countries).

Józef Gruenberg's map clearly shows a similarity to the Plater–Anonym–Rozwadowski ethnolinguistic studies. According to Gruenberg, the Pelesa–Rodunia–Dieveniškės linguistic island has a larger coverage in the northern part. This assessment of the situation is in line with the works of authors from later periods (Zinkevičius 1993, pp. 18–19; 2006, p. 107; Gaučas 1993, pp. 68–69).

The results of the reconstructed boundaries were also compared with the manuscript ethnolinguistic map "Die Sprachen im Wilnageibiet" (The languages in the Vilnius region) of 1942 (Figure 12), created by linguists of the Institute of the Lithuanian Language (Vidugiris 1993, p. 128, Mikulėnienė 2022, p. 128, Mikulėnienė 2022, pp. 78–79). The map was based on the 1942 public census data[12] (Kairiūkštytė 1992, pp. 116–23).
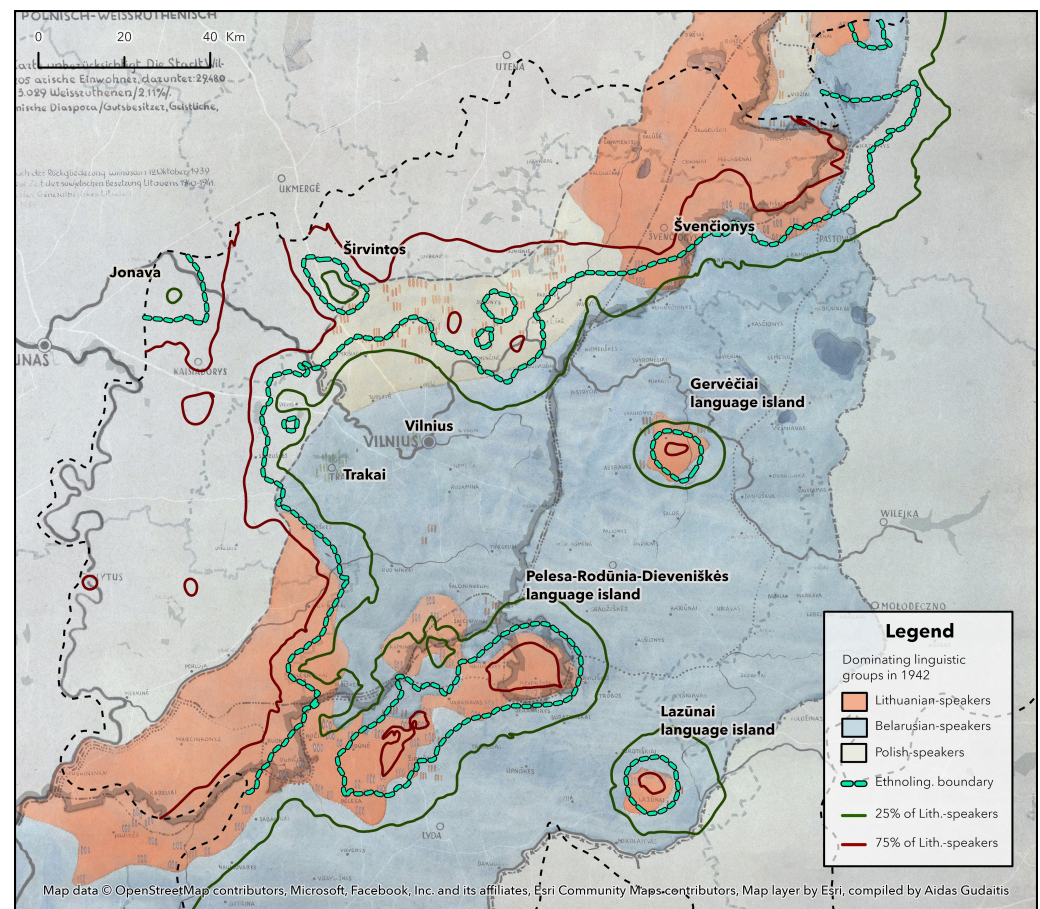

Figure 12. Comparison of the reconstructed 19th-century ethnolinguistic boundaries with the data of the 1942 ethnolinguistic map "Die Sprachen im Wilnageibiet" (by K. Alminas-Alminauskis and G. Šlapelytė-Sirutienė (Alminas-Alminauskis and Šlepelytė-Sirutienė 1942), the Institute of the Lithuanian Language).

It is important to note that in this map, the Polish language enclave (marked in yellow color) was distinguished for the first time in SEL. This Polish-speaking zone was also confirmed by other scholars in later research (Zinkevičius 1993, pp. 22–23; 2006, pp. 110–11; Čekmonas and Grumadienė 1993, pp. 132–36; Gaučas 1993, pp. 76–77).

In comparing the obtained 1890 reconstruction results with the ethnolinguistic map of 1942, a significant overlapping between the broad intensive Lithuanian–Slavic contact zone of 1890 and the delineated Polish language zone of 1942 can be observed. This reinforces the assumption that it was not until the first half of the 20th century (during 1921–1939) that the language code changed irreversibly in the area from Lithuanian to Polish due to the intensive Polonization of the Lithuanian-speaking population.

3.2. The Comparative Geospatial Analysis of the Ethnolinguistic Data of Late 19th and Early 21st Centuries

The results of the ethnolinguistic situation reconstruction in the late 19th century were compared with the contemporary demographic situation. For this purpose, the latest open geospatial data from the 2021 public census of Lithuania were used (State Data Agency of Lithuania 2021a, 2021b, 2021c, 2021d)[13].

Although the detailed geospatial data on the population distribution by the mother tongue was not officially published by the State Data Agency of Lithuania, a comparison of the mother tongue and ethnicity data on a municipal level revealed a statistically significant correlation between Lithuanian ethnicity and the mother tongue ($r = 0.99986$, $p < 0.001$) (State Data Agency of Lithuania 2021b, 2021c). According to the public census of 2021, the citizens with declared Lithuanian ethnicity made up 84.61% of the total population ($n = 2,810,761$), while citizens who declared the Lithuanian language as their mother tongue made up 85.33%. It is important to mention that 98.67% of the population who declared Lithuanian ethnicity reported the Lithuanian language as their mother tongue, while from the remaining 15.4% of Lithuanian citizens ($n = 432,600$) who declared another ethnicity, 52 thousand (12.01%) indicated Lithuanian as their mother tongue. Thus, it could be stated that ethnicity data could be used to estimate the population distribution by the mother tongue in 2021.

By comparing the ethnolinguistic boundaries of 1890 with those of 2021 on the map, we can visually identify the main areas with the greatest ethnolinguistic shifts in Lithuanian–Slavic language use over the last 130-year period (Figure 13).
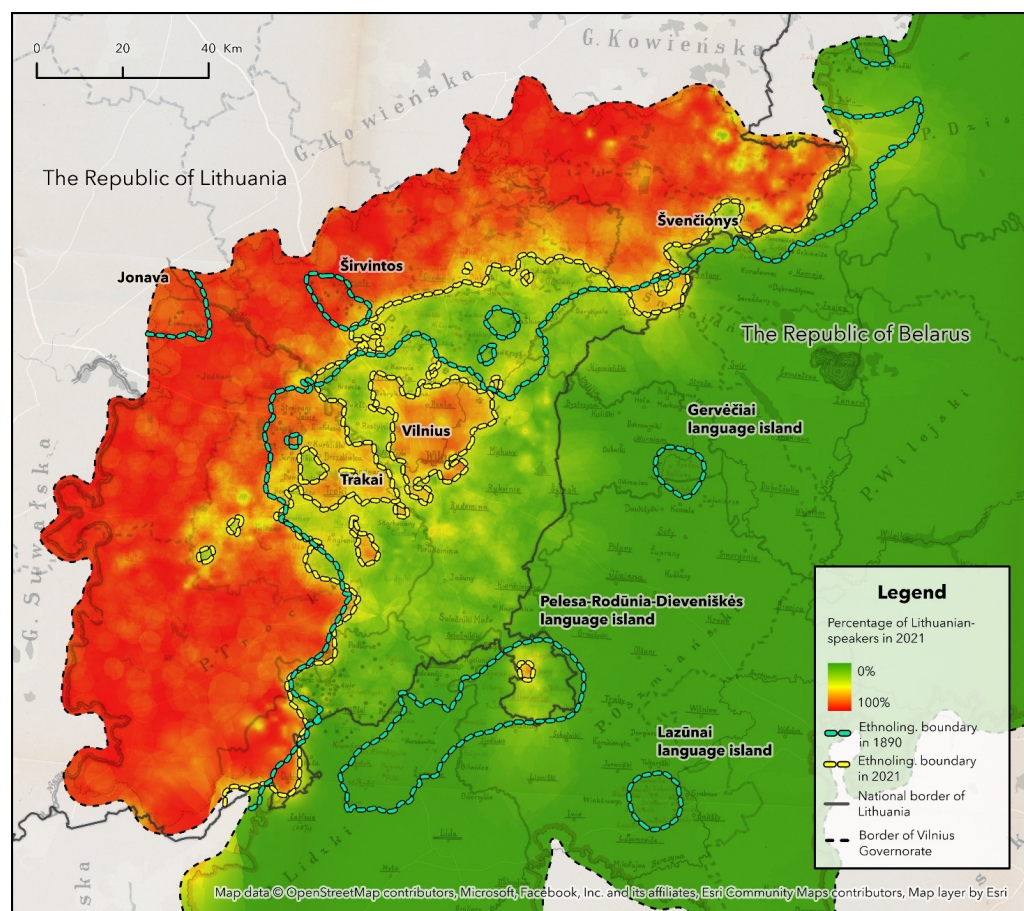

Figure 13. Comparison of the Lithuanian–Slavic ethnolinguistic boundaries of 1890 and 2021.

A more efficient way to quantify the results of two spatial datasets from different time periods is to apply the change detection raster geospatial analysis method (Figure 14).
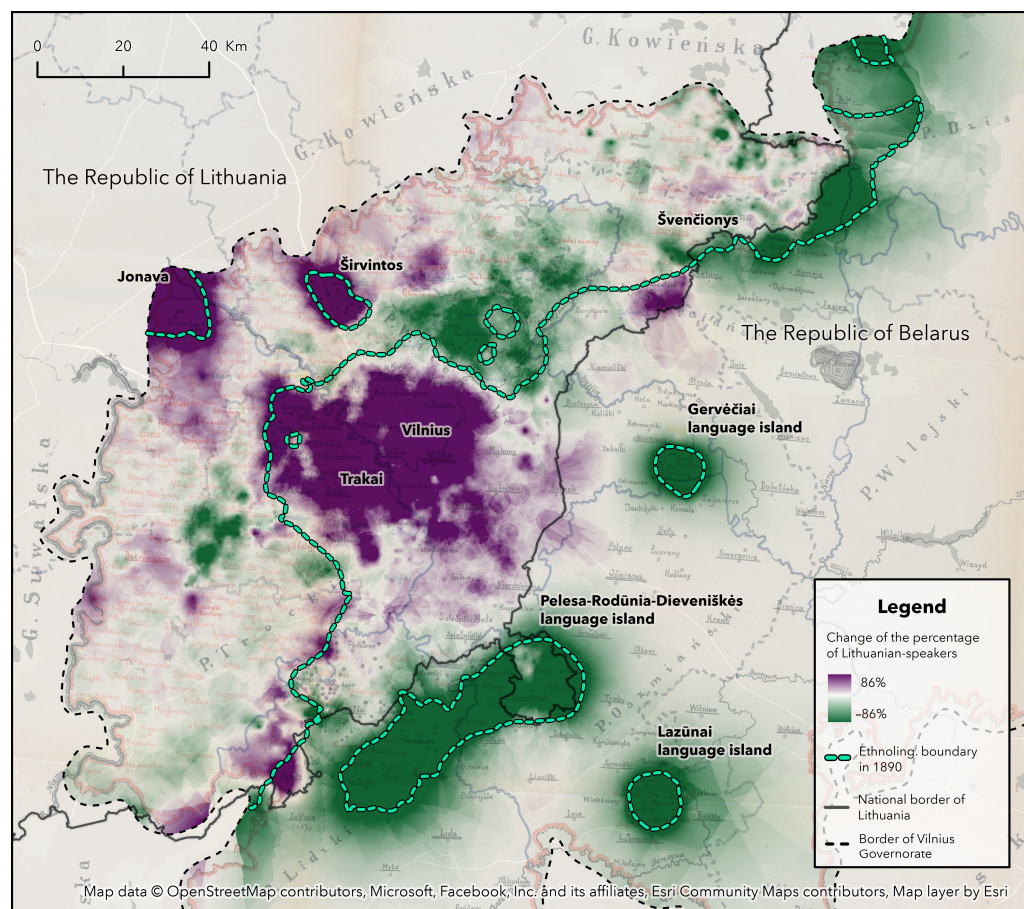

Figure 14. Estimation of the ethnolinguistic Lithuanian–Slavic change in SEL over the period of 1890–2021 (purple—the change in favor of the Lithuanian speakers, green—the change in favor of the Slavic speakers).

The analysis of the Lithuanian–Slavic ethnolinguistic change in the study area within the period 1890–2021 clearly shows the disappearance of the Lithuanian language linguistic islands and margins in the territory of Belarus as well as the disappearance of the former Lithuanian-speaking area to the North from Vilnius City. On the other hand, we can observe the disappearance of northwestern Slavic-speaking linguistic islands in the vicinity of the Jonava and Širvintos cities as well as the intensive Lithuanization of the Vilnius–Trakai urbanized area.

By analyzing the ethnolinguistic data from a slightly different perspective, other aspects of the ethnolinguistic change in the region can be observed. Figure 15 represents the same processed data of 1890, but in this case, a heatmap of the Lithuanian-speaking population density per sq. km at the end of the 19th century was generated by applying the Kernel density[14] geospatial analysis method.
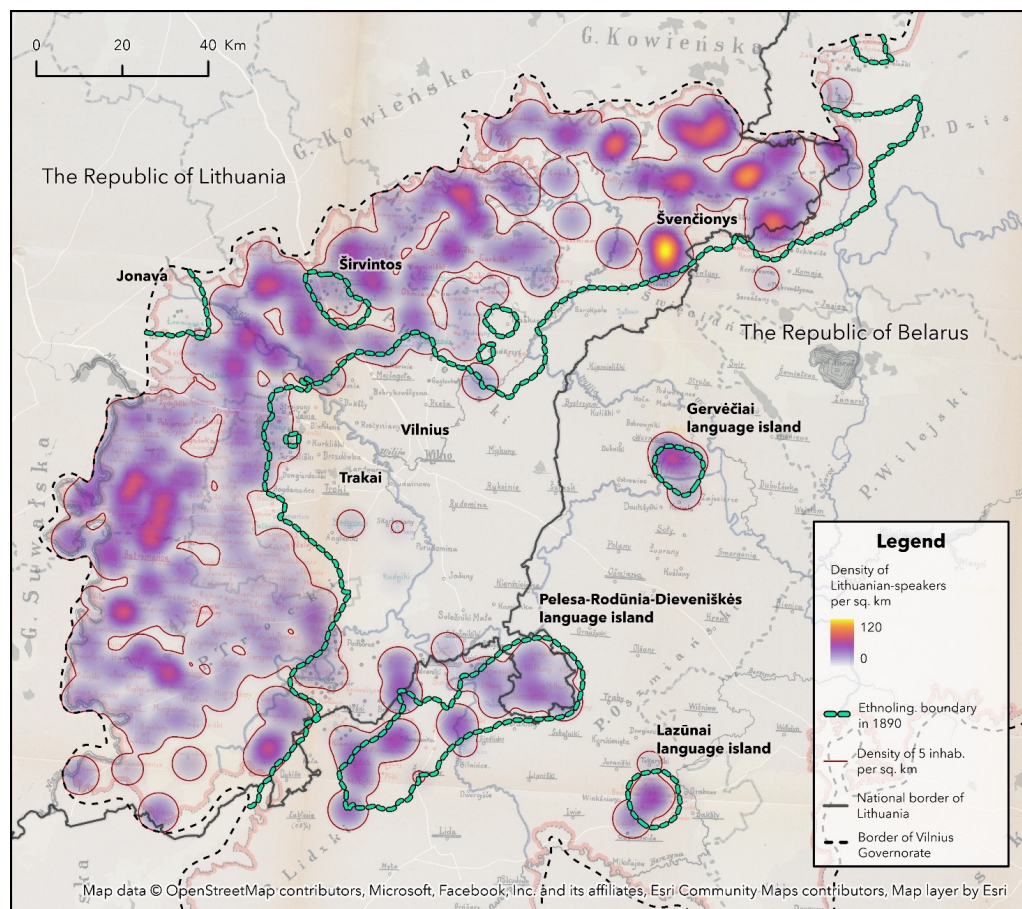
Figure 15. Density of Lithuanian-speaking population per sq. km in the late 19th century. The red outline—the area where the density of the Lithuanian-speaking population was at least five inhabitants per sq. km.

As we can see from the analysis results, at the end of the 19th century, the Lithuanian-speaking population was mainly concentrated to the north and the west from the ethnolinguistic boundary, bypassing the cities of Vilnius and Trakai. We can also observe that Lithuanian speakers were quite homogeneously distributed throughout the whole area of their living range, besides the zone of more intense concentration in the northeast near Švenčionys (around 120 inhabitants per sq. km).

For comparison purposes, the density results of 1890 and 2021 were aggregated into 100 sq. km (10 × 10 km) grids. The application of the change detection tool for the comparison of the Lithuanian-speaking population density in 1890 and 2021 revealed that even though the Lithuanian-speaking population has increased threefold (from 232,790 to 639,437) in the research area over 130 years, an intensive decline in Lithuanian speakers was observed not only in the islands in the territory of Belarus but in the entire rural area (Figure 16). The maximum increase in the population density was more than 200,000 inhabitants per 100 sq. km (in Vilnius city), while the decrease was more than 2700.

The total increase in the Lithuanian-speaking population in the Vilnius–Trakai area was more than 540,000, while in the rest of the study area, the Lithuanian-speaking population decreased by over 173,000 over the last 130 years.
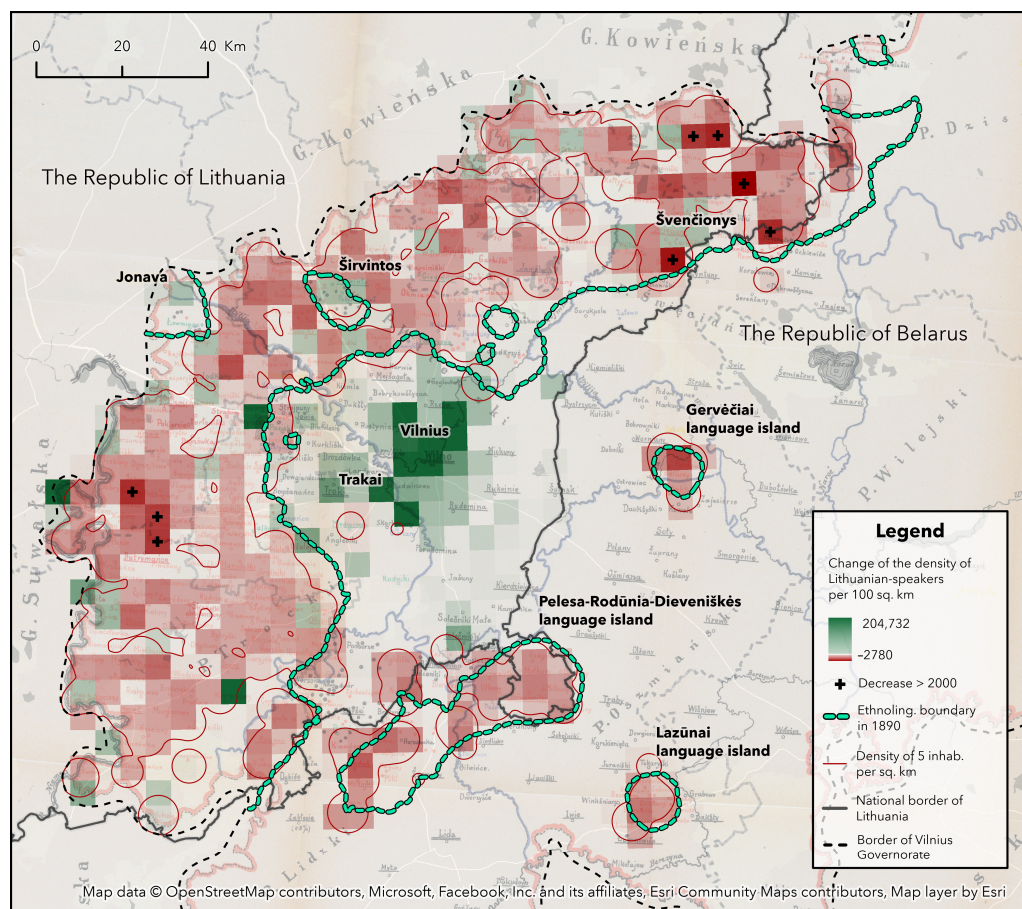
Figure 16. Changes in the density of Lithuanian-speaking population per 100 sq. km during the period 1890–2021.

4. Discussion

It may be presumed that the rapid decline in the number of Lithuanian speakers in the country's rural areas over the last 130 years was primarily influenced by the intensive emigration of the younger generation to urban areas. It is likely that the decline was also influenced by other socio-economic factors: the low quality of social infrastructure, weak economic development and job supply, a negative birth–death ratio, increase in ethnically mixed marriages, a change in ethnic self-awareness and the linguistic code affected by the decreasing prestige of the Lithuanian language.

Surprisingly, the Lithuanian linguistic islands in the territory of Belarus are not the areas with the highest population loss. The largest proportion of the Lithuanian-speaking population has been lost in the northeastern zone (five 100 sq. km squares with a loss of more than 2000 Lithuanian-speaking inhabitants) and the western zone (three 100 sq. km squares with a loss of more than 2000 Lithuanian-speaking inhabitants) of the research area (Figure 16, marked as black crosses).

It is also evident that in recent decades, the Vilnius–Trakai urbanized area, previously dominated by the Slavic-speaking population, has been intensively Lithuanized. Based on the analysis of the public survey data of 2011 and 2021, the Lithuanization process in the metropolitan area is predicted to expand further, affecting the Slavic-speaking surroundings (State Data Agency of Lithuania 2011a, 2011b, 2021d).

When considering the competition between the Lithuanian language and other languages in the study area, as well as assessing the status dynamics of Lithuanian as a communicative language in the region, it is evident that Lithuanian has retained its significance over time. Census data from the last three decades indicate that its position in the SEL region has strengthened (see Figures 13 and 14). This trend is likely to have been influ-

enced by the restoration of the country's independence, the designation of Lithuanian as the official state language, and the strengthened education of the Lithuanian language in the region.

The analysis reveals a significant increase in the prevalence of Lithuanian across the urbanized areas of the SEL in recent decades. This growth is largely attributed to the rapid economic expansion of Vilnius City and its suburbs, which has stimulated the internal migration of Lithuanian-speaking residents from across the country.

In Belarus, however, there has been a significant decline in the use of Lithuanian. According to the Belorussian public census of 2019, out of approximately 1300 Lithuanians residing in the Ostrovetsky, Voronovsky, and Lida districts, only 327 individuals (25%) considered Lithuanian their mother tongue, with as few as 119 people (9%) using it as their primary language of communication. The predominant languages among Lithuanians in these districts are Belarusian (26.5%) and Russian (68.1%). These figures suggest that Lithuanian may be pushed to extinction amongst the Lithuanian-speaking communities of Belarus in the foreseeable future.

It should also be noted that the recent geopolitical developments are creating language dynamics within the SEL that are not favorable for Lithuanian. Increased migration from former, primarily Russian-speaking Soviet Union countries to the Vilnius city area, has been recorded over the last few years. As a result, the competition between Lithuanian and Slavic languages in the SEL region remains a relevant and ongoing concern.

To more accurately assess language balance trends in the SEL region and anticipate potential future scenarios, consistent monitoring of the vitality of the Lithuanian language through geostatistical analysis is necessary.

5. Conclusions

To summarize the results of the study, the following conclusions can be drawn.

The chosen methods for reconstructing the ethnolinguistic boundaries of SEL, using GIS technology and geospatial analysis tools, have proven to be effective. The results obtained can be used for a more comprehensive examination of the ethnolinguistic boundary shifts that have occurred in the region as well as a deeper exploration of their causes and consequences.

The data presented by Plater–Anonym–Rozwadowski aligns with the findings of later studies. It is therefore possible to conclude that, despite certain methodological limitations (data collected only about the primary language, not considering bilingualism or multilingualism), the early research is an accurate reflection of the ethnolinguistic situation within the SEL region at the end of the 19th century.

Even though the Lithuanian-speaking population in the study area has increased by almost threefold (from 233,000 to 639,000) over the last 130 years and the ethnolinguistic boundary at the Vilnius–Trakai agglomeration has shifted in favor of the Lithuanian language, the Lithuanian-speaking population has been intensively declining in the entire rural area of SEL.

As the Vilnius urbanized area expands, it is likely that the influence of the Lithuanian language in the metropolitan surroundings will continue to increase. However, this will occur at the expense of more distant rural settlements with poorer social infrastructure and weaker economic development which lead to a negative migration balance, a low birth rate, and intensive population ageing. Economic development and investment in social infrastructure in the region could help to stabilize the situation and even reverse demographic losses.

The ethnolinguistic study of the SEL area will be continued by analyzing other aspects of ethnolinguistic processes to further clarify the causes of the changes observed and assess future trends.

Notes

[1] Geoinformation system (GIS). It is an information technology platform used to describe and characterize the Earth and other surfaces to visualise and analyse geographically (spatially) coordinated information. GIS provides the tools to collect, store, and analyse spatial and geographic data and their attributes. Further: https://www.earthdata.nasa.gov/learn/backgrounders/gis (accessed on 6 November 2024).GIS technology was first used in 1963. Roger Tomlinson, in Canada, developed a vector-based GIS framework and a geodatabase, which at that time made it possible to store and analyse large amounts of spatial land use data. The technical solution enabled the Canadian government to efficiently implement a national land management programme (Waters 2017).

[2] Location intelligence (LI) is the insight gained from visualizing and analyzing geospatial data. More about LI: https://www.esri.com/en-us/location-intelligence/overview (accessed on 6 November 2024).

[3] Eldership (Lith. seniūnijà) is the smallest structural territorial subdivision of a municipal administration in Lithuania, a branch of a municipality's executive authority having the rights of a legal person, which operates in a given part of the municipality's territory. Further: https://www.vle.lt/straipsnis/seniunija-2/ (accessed on 6 November 2024).

[4] Further: https://lad.lt/broelalfred/ (accessed on 6 November 2024).

[5] Further: https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview (accessed on 6 November 2024).

[6] LKS-94 is a national realization of ETRS89 and is coincident to WGS 84 within 1 metre. This transformation has an accuracy equal to the coincidence figure. Further: https://epsg.io/3346 (accessed on 6 November 2024).

[7] Interpolation—a method of geospatial forecasting in which the intermediate value of a variable is approximated from known values. In geospatial analysis, various statistical interpolation methods (e.g., Inverse Distance Weighting (IDW), Spline, Kriging, etc.) are used, depending on the specifics of the research (See also: Ripley 2005, chap. 4).

[8] Further: https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm (accessed on 6 November 2024).

[9] More about Vilna Governorate: https://www.vle.lt/straipsnis/vilniaus-gubernija/ (accessed on 6 November 2024).

[10] Further: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/contour.htm (accessed on 6 November 2024).

[11] Further: https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/change-detection-in-arcgis-pro.htm (accessed on 6 November 2024).https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/pixel-value-change-detection.htm (accessed on 6 November 2024).

[12] The map does not consider the ratio of languages in cities.

[13] Description: https://osp.stat.gov.lt/gyventoju-ir-bustu-surasymai1 (accessed on 6 November 2024).Open data: https://open-data-ls-osp-sdg.hub.arcgis.com/datasets/80272a0f7a7e4905bc379f0ae921e5b7_1/about (accessed on 6 November 2024).

[14] Further: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-kernel-density-works.htm (accessed on 6 November 2024).

References

Alminas-Alminauskis, K., and G. Šlepelytė-Sirutienė. 1942. *Map. Die Sprachen im Wilnagebiet [Languages in the Vilnius Region]*. Vilnius: The Institute of the Lithuanian Language.

Anonim. 1898. *Opszar Języka Litewskiego w Gubernii Wilenskiej [Lithuanian Language Area in Vilnius Governorate]*. Krakow: Nakładem Akademii Umiejętności. Available online: https://archive.org/details/podernia-obszar-jezyka-litewskiego...-1898/page/8/mode/2up (accessed on 6 November 2024).

Ayad, Yasser, and Herb Luthin. 2009. Mapping Dialect: GIS Applications in Studying Dialect of Western Pennsylvania. Paper presented at the 2009 ESRI International User Conference Proceedings, San Diego, CA, USA, July 13–17; pp. 1–26. Available online: http://proceedings.esri.com/library/userconf/proc09/uc/papers/pap_1686.pdf (accessed on 6 November 2024).

Chominski, Olgierd. 2010. *Obszar Językowy Litewski w Panstwie Polskim 1927–1933 [The Lithuanian Language Area in the Polish State 1927–1933]*. Edited by L. Bednarczuk. Warszawa: Acta Baltico-Slavica, 34 SOW. pp. 197–261.

Chominski, Olgierd, ed. 1933. *Mapa Obszaru Językowego Litewskiego w Rzeczypospolitej Polskiej 1928–1933 [Map of the Lithuanian Linguistic Area in the Republic of Poland 1928–1933]*. Warszawa: Zakł. Graf. Straszewiczów.

Čekmonas, Valerijus. 2017. *Nad Etniczną i Językową Mapą Polaków Litewskich—O Teraźniejszości i Przyszłości [On the Present and Past Considering the Ethnic and Linguistic Map of Lithuania's Poles]*. Vilnius: The Institute of the Lithuanian Language, pp. 115–26, 867–69, 968–70.

Čekmonas, Valerijus, and Laima Grumadienė. 1993. *Kalbų Paplitimas Rytų Lietuvoje [Language Distribution in Eastern Lithuania]*. Edited by Lietuvos Rytai, Kazimieras Garšva and Laima Grumadienė. Vilnius: State Publishing Centre, pp. 132–36.

Čepaitienė, Agnė, and Aidas Gudaitis. 2021. *Rytų Aukštaičių Panevėžiškių Diferencinio Požymio Kaita: Geoerdvinis Aspektas [Change of Phonetic Differential Feature in the Eastern Aukštaitian Subdialect of Panevėžys: The Geospatial Approach]*. Vilnius: Acta Linguistica Lithuanica LXXXIV, pp. 149–72.

Čepaitienė, Agnė, and Aidas Gudaitis. 2022. Sociokultūrinių tinklų tyrimo modelio tarmės kaitai nustatyti koncepcija ir realizacija [Concept and realization of a socio-cultural network research model for the analysis of a dialectal shift]. *Baltistica* 57: 291–327.

Gaučas, Petras. 1992. Etnolingvistinė situacija ir bendrojo lavinimo mokyklos Pietryčių Lietuvoje 1945–1991 metais [Ethnolinguistic situation and general education schools in Southeastern Lithuania in 1945–1991]. In *Rytų Lietuva. Istorija, Kultūra, Kalba [Eastern Lithuania. History, Culture, Language]*. Edited by Vacys Milius, Kazimieras Garšva, Nijolė Laurinkienė and Mindaugas Tamošiūnas. Vilnius: Science Publishing House, pp. 141–50.

Gaučas, Petras. 1993. *Lietuvių-Gudų Kalbų Paribio Etnolingvistinė Situacija 1875–1914 m. [The Ethnolinguistic Situation on The Lithuanian-Guddish Linguistic Periphery in 1875–1914]*. Edited by Lietuvos Rytai, Kazimieras Garšva and Laima Grumadienė. Vilnius: Valstybinis leidybos centras, pp. 42–100.

Gaučas, Petras. 2004. *Etnolingvistinė Rytų Lietuvos Gyventojų Raida: XVII a. Antroji Pusė—1939 m.: Istorinė-Geografinė Analizė [Ethnolinguistic Development of the Population of Eastern Lithuania: Second Half of the 17th Century—1939: A Historical-Geographical Analysis]*. Vilnius: Inforastras.

Gruenberg, Józef. 1913. Mapa Rozsiedlenia Ludności Polskiej na Obszarze Dawnej Rzeczypospolitej i Krajów Ościennych [Map of the Distribution of the Polish Population in the Former Polish-Lithuanian Commonwealth and Neighbouring Countries]. Lviv. Available online: https://collections.lib.uwm.edu/digital/collection/agdm/id/2701/ (accessed on 6 November 2024).

Gudavičius, Edvardas. 2008. Lietuvių Žemių Konfederacija [Lithuanian Land Confederation]. Visuotinė Lietuvių Enciklopedija [Universal Lithuanian Encyclopedia] T. XIII: Leo–Magazyn, Science and Encyclopedia Publishing Center. Available online: https://www.vle.lt/straipsnis/lietuviu-zemiu-konfederacija/ (accessed on 6 November 2024).

Hoch, Shawn, and James J. Hayes. 2010. Geolinguistics: The Incorporation of Geographic Information Systems and Science. *The Geographical Bulletin* 51: 23–36.

Kairiūkštytė, Nastazija. 1992. *Vilniaus Krašto Gyventojai. 1942 m. Gegužės 27 d. Surašymo Duomenimis [Population of Vilnius Region. According to the Census of 27 May 1942]*. Rytų Lietuva. Istorija, Kultūra, Kalba [Eastern Lithuania. History, Culture, Language]. Edited by Vacys Milius, Kazimieras Garšva, Nijolė Laurinkienė and Mindaugas Tamošiūnas. Vilnius: Science Publishing House, pp. 116–23.

Kardelis, Vytautas. 2014. Lietuvių Kalbos Salos: Probleminiai Aspektai [Islands of the Lithuanian Language: Problematic Aspects]. Lietuvių Kalba, 8, 1 pdf (18 pp.). Available online: https://www.lituanistika.lt/content/52931 (accessed on 6 November 2024).

Kardelis, Vytautas. 2015. Lietuvių Kalbos Salos: Tyrimo Perspektyvos [Islands of the Lithuanian Language: Research Perspectives]. Lietuvių Kalba, 9, 1 pdf (17 pp.). Available online: https://www.lituanistika.lt/content/61395 (accessed on 6 November 2024).

Krieger, Nancy. 2003. Place, Space, and Health: GIS and Epidemiology. *Epidemiology* 14: 384–85. [CrossRef] [PubMed]

Kurila, Laurynas. 2005a. Rytinė lietuvių genties riba IX–XII a. 1. archeologijos duomenys [The eastern boundary of the Lithuanian tribe in the 9th–12th centuries. 1. archaeological data]. *Lietuvos archeologija* 27: 59–84.

Kurila, Laurynas. 2005b. Rytinė lietuvių genties riba XI–XIV a. 2. rašytinių šaltinių duomenys [The eastern boundary of the Lithuanian tribe in the 11th–14th centuries. 2. data from written sources]. *Lietuvos archeologija* 28: 121–32.

Kurschat, Friedrich. 1876. *Grammatik der Littauischen Sprache: Mit einer Karte des Littauischen Sprachgebiets und einer Abhandlung über Littauische Volkspoesie nebst Musikbeilage von 25 Dainosmelodien, von Friedrich Kurschat*. Halle: Buchhandlung des Waisenhauses, S. 1–2.

Lameli, Alfred, Roland Kehrein, and Stefan Rabanus, eds. 2010. *Language and Space: An International Handbook of Linguistic Variation*. Vol. 2: Language Mapping. Berlin: Mouton de Gruyter.

Liseichykau, Dzianis. 2017. *Zasięg Osadnictwa Ruskiej Społeczności Etniczno-Wyznaniowej w Wielkim Księstwie Litewskim od Końca XVII do Połowy XVIII Wieku (w Świetle Materiałów Wizytacji Generalnych Cerkwi Unickich) [Extent of the Settlement Area of the Ruthenian Ethnic-Religious Community Within the Grand Duchy of Lithuania from the End of the Seventeenth to the Mid-Eighteenth Century (in the Light of General Visitation Documents of Uniate Churches)]*. Warsaw: Polskiej Akademii Nauk.

McGuirk, David G. 2004. An Ethnolinguistic Analysis of Hispanics in Miami-Dade County. Ph.D. dissertation, Florida International University, Miami, FL, USA.

Metzler Lexikon Sprache. 2005. *Metzler Lexikon Sprache*. Edited by Helmut Glück and Friederike Schmöe. Stuttgart: Weimar.

Mikulėnienė, Danguolė. 2018. *Lietuvių Tarmėtyra: Genezė, Raida ir Paradigminiai Lūžiai [Lithuanian Dialectology: Genesis, Development, and Paradigmatic Breakthroughs]*. Vilnius: Publishing Centre of Science and Encyclopaedias.

Mikulėnienė, Danguolė. 2022. *Lietuvių Kalbos Instituto Istorija. D. 2, Kalbotyros Aneksija [The History of the Institute of the Lithuanian Language. D. 2, Annexation of Linguistics]*. Vilnius: The Institute of the Lithuanian Language.

National Statistical Committee of the Republic of Belarus. 2019. The Results of Public Survey of the Republic of Belarus in 2019. Available online: https://census.belstat.gov.by/sections (accessed on 6 November 2024).

Ripley, Brian D. 2005. *Spatial Statistics*. Hoboken: John Wiley & Sons. 272p.

Rozwadowski, Jan Michał. 1901. *Mapa Języka Litewskego w Gubernii Wilenskiej [Map of the Lithuanian Language in Vilnius Governorate]. Materiały i Prace Komisii językowej Akademii Umiejętności w Krakowe*. Krakow: Nakładem Akademii Umiejętności. Krakow: Biblioteka Jagiellońska, Tomas Nenartovic. Available online: https://lithuanianmaps.com/images/1901_19_Rozwadowski_COMP15.jpg (accessed on 6 November 2024).

Rozwadowski, Jan Michał. 1904. *Objaśnienia do Mapy Języka Litewskego w Gubernii Wilenskiej [Explanations of the Map of the Lithuanian Language in Vilnius Governorate]. Materiały i Prace Komisii Językowej Akademii Umiejętności w Krakowe*. Krakow: Nakładem Akademii Umiejętności, pp. 89–94.

State Data Agency of Lithuania. 2011a. Geospatial Data on 1 × 1 km Grids from the 2011 Population and Housing Census of the Republic of Lithuania. Available online: https://open-data-sets-ls-osp-sdg.hub.arcgis.com/datasets/7fec4f1e8a73428283ea9e906aa89483_0/about (accessed on 6 November 2024).

State Data Agency of Lithuania. 2011b. Results of the 2011 Census of Population and Housing of the Republic of Lithuania. Available online: https://osp.stat.gov.lt/documents/10180/217110/Gyv_kalba_tikyba.pdf/1d9dac9a-3d45-4798-93f5-941fed00503f (accessed on 6 November 2024).

State Data Agency of Lithuania. 2021a. Population and Housing Census 2021—Population (GRID 1 km). Available online: https://open-data-sets-ls-osp-sdg.hub.arcgis.com/datasets/80272a0f7a7e4905bc379f0ae921e5b7_1/about (accessed on 6 November 2024).

State Data Agency of Lithuania. 2021b. Population by Ethnicity in Municipality. Available online: https://osp.stat.gov.lt/en/statistiniu-rodikliu-analize?hash=0078cd86-acd6-46a8-9843-623bdf998aba#/ (accessed on 6 November 2024).

State Data Agency of Lithuania. 2021c. Population by Mother Tongue in Municipality. Available online: https://osp.stat.gov.lt/en/statistiniu-rodikliu-analize?hash=03d78029-cb87-40e8-aff0-0b00b868015e#/ (accessed on 6 November 2024).

State Data Agency of Lithuania. 2021d. Population by Nationality, Mother Tongue and Religion (Metadata). Available online: https://osp.stat.gov.lt/documents/10180/5118910/Gyventojai+[EN]+490.html (accessed on 6 November 2024).

Vidugiris, Aloyzas. 1993. *Etnolingvistinė Pietryčių Lietuvos Padėtis XX a. Pirmojoje Pusėje [Ethnolinguistic Situation of South-Eastern Lithuania in the First Half of the 20th Century]*. Edited by Lietuvos Rytai, K. Garšva and L. Grumadienė. Vilnius: State Publishing Centre, pp. 115–31.

Waters, Nigel. 2017. *GIS: History, The International Encyclopedia of Geography*. Edited by Douglas Richardson, Noel Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu and Richard A. Marston. Hoboken: John Wiley & Sons, Ltd., pp. 2–12.

Wroblewski Library of the Lithuanian Academy of Sciences. 2019. Topographic Maps of Lithuania for 1914–1945. Available online: https://maps-lmavb.hub.arcgis.com/pages/topo (accessed on 6 November 2024).

Zinkevičius, Zigmas. 1993. *Pietryčių Lietuva nuo Seniausių Laikų iki Mūsų Dienų [Southeastern Lithuania from Ancient Times to the Present]*. Edited by Lietuvos Rytai, K. Garšva and L. Grumadienė. Vilnius: State Publishing Centre.

Zinkevičius, Zigmas. 1998. *The History of the Lithuanian Language*. Vilnius: Institute of Science and Encyclopedia Publishing.

Zinkevičius, Zigmas. 2006. *Lietuvių Tarmių Kilmė [The Origin of Lithuanian Dialects]*. Vilnius: Publishing House of the Institute of the Lithuanian Language.