Russian Journal of Linguistics
ISSN 2687-0088 (Print), ISSN 2686-8024 (Online)

2026 Vol. 30 No. 2 352–374
http://journals.rudn.ru/linguistics

https://doi.org/10.22363/2687-0088-47384
EDN: CKPXVR

Research article / Научная статья

Language ecology of minority languages of Sakha (Yakutia) in the mirror of the All-Russian Census 2020–2021

Gulnara F. GABDRAKHMANOVA[1] ✉ and Nikolai B. VAKHTIN[2]

[1]Sh. Marjani Institute of History, Academy of Sciences of the Republic of Tatarstan, Kazan, Russian Federation
[2]European University at St. Petersburg, St. Petersburg, Russian Federation
✉medi54375@mail.ru

Abstract
This study applies the framework of the ecology of language, defined by Einar Haugen as the study of interactions between a language and its environment, to evaluate the reliability of population census data for sociolinguistic analysis. Focusing on the 2020–2021 Russian Census, the paper investigates its usefulness in describing language situations among indigenous minorities in the north-east of the Sakha Republic (Yakutia). The study adopts a qualitative and critical analytical approach. First, general features of the census are assessed, including coverage completeness and the clarity and adequacy of terminology used in ethnic and linguistic categories. Second, the broader sociolinguistic context—particularly language endangerment in north-eastern Russia—is examined. Finally, selected census data are analyzed for ten indigenous languages: Aleut, Aliutor, Chukchi, Dolgan, Even, Evenki, Eskimo, Itel'men, Koryak, and Yukaghir. The analysis reveals inconsistencies in census outcomes, including unexpected increases and decreases in reported speaker populations across several languages. These fluctuations do not consistently align with established trends in language vitality or decline, raising questions about data reliability. The findings suggest that census responses are shaped more by speaker attitudes, identity, and self-perception than by actual language use or proficiency. Consequently, while the census provides valuable insights into sociocultural identification, it is less effective as a precise instrument for measuring linguistic vitality. This underscores the need for complementary methodologies in assessing minority language situations.

Keywords: *Russian 2020/21 Population Census, indigenous languages, Yakutia (Sakha) Republic, language as identity marker*

For citation:
Gabdrakhmanova, Gulnara F. & Nikolai B. Vakhtin. 2026. Language ecology of minority languages of Sakha (Yakutia) in the mirror of the All-Russian Census 2020-21. *Russian Journal of Linguistics* 30 (2). 352–374. https://doi.org/10.22363/2687-0088-47384 EDN: CKPXVR

© Gulnara F. Gabdrakhmanova, Nikolai B. Vakhtin, 2026
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Лингвистическая экология миноритарных языков Республики Саха (Якутия) в зеркале Всероссийской переписи населения 2020–2021 годов

Гульнара Фаатовна ГАБДРАХМАНОВА[1] ✉,
Николай Борисович ВАХТИН[2]

[1]Институт истории имени Ш. Марджани, Академия наук Республики Татарстан, Казань, Российская Федерация
[2]Европейский университет в Санкт-Петербурге, Санкт-Петербург, Российская Федерация
✉medi54375@mail.ru

Аннотация
Данное исследование использует концепцию экологии языка (Haugen 1971, 1972) как изучение взаимодействия языка с его средой для оценки надежности данных переписей населения в социолингвистическом анализе. Сосредотачиваясь на Российской переписи 2020–2021 г., статья исследует ее применимость для описания языковой ситуации среди коренных малочисленных народов на северо-востоке Республики Саха (Якутия). В исследовании используется качественный и критико-аналитический подход. Сначала рассматриваются общие характеристики переписи, включая полноту охвата населения, а также ясность и адекватность терминологии, применяемой в этнических и языковых категориях. Далее анализируется более широкий социолингвистический контекст, в частности проблема исчезновения языков на северо-востоке России. Наконец, проводится анализ отдельных данных переписи по десяти языкам коренных народов: алеутскому, алюторскому, чукотскому, долганскому, эвенскому, эвенкийскому, эскимосскому, ительменскому, корякскому и юкагирскому. Анализ выявляет несоответствия в результатах переписи, включая неожиданные увеличения и сокращения численности носителей ряда языков. Эти колебания не всегда согласуются с установленными тенденциями жизнеспособности или угасания языков, что ставит под сомнение надежность данных переписи. Полученные результаты показывают, что ответы респондентов в большей степени отражают языковые установки, идентичность и самоощущение, чем реальное использование языка или уровень владения им. Таким образом, несмотря на ценность переписи для изучения социокультурной идентификации, она менее эффективна как точный инструмент измерения жизнеспособности языков, что подчеркивает необходимость использования дополнительных методологических подходов при изучении миноритарных языков.
Ключевые слова: *Всероссийская перепись населения 2020/21 года, языки коренных народов, Республика Якутия (Саха), язык как маркер идентичности*

Для цитирования:
Gabdrakhmanova G.F., Vakhtin N.B. Language ecology of minority languages of Sakha (Yakutia) in the mirror of the All-Russian Census 2020-21. *Russian Journal of Linguistics.* 2026. Vol. 30. № 2. P. 352–374. https://doi.org/10.22363/2687-0088-47384 EDN: CKPXVR

1. Introduction

In the mid-nineteenth century, many countries around the world began conducting population censuses in ways that closely resemble modern practices. It was during this period that censuses were first defined as large-scale state

registration exercises covering all citizens of a country or its individual territories. In Russia, the first steps towards conducting censuses in this modern form were taken in the 1860s. At that time, one-day censuses were carried out in major cities such as St Petersburg, Moscow, and Kharkov. On 28 January 1897, the first scientifically organised population census was conducted across the Russian Empire (with the exception of Finland). For the first time, the population of the entire country was counted on a universal basis. However, the census questionnaire did not include a question on ethnic (then referred to as "tribal") affiliation (Neudachin 1905: x). Such information could only be inferred from responses to questions on native language and religion, and no guidance was provided on how the term "native language" should be interpreted.

The inclusion of the question "your native language" in the questionnaire was an innovation for characterizing the country's citizens. The question about the native language in one form or another was required to be asked by census takers in all subsequent (Soviet) censuses, beginning in 1926; in the instructions for the 1926 census, the definition was as follows: "The native language is recognized as the one that the respondent has the best command of or usually speaks" (Vsesoyuznaya... 1928: 5). Beginning with the 1926 census,[1] census takers were also required to ask questions about ethnicity (*национальность*)[2] while post-Soviet censuses, due to the abolition of recording ethnicity in passports, became the only all-Russian statistical source of information about the ethnic and linguistic affiliation of citizens. Therefore, the census data was important for government agencies responsible for planning and implementing ethnic and language policy at the national and regional levels as well as for researchers — ethnographers, sociologists and (socio)linguists (Sotsial'noe i natsional'noe 1972, Etnografiia perepisi-2002, Sotsial'noe neravenstvo 2002) and many others.

The scientific study of censuses has been approached from a variety of perspectives: examining their instrumental function and their role in nation-building (Anderson 2001: chapter 10); considering them as a form of power capital (Tishkov 2003: 11); and analysing them as a means of constructing social groups (Krasnopolskaya & Solodova 2016: 69) or shaping ethnic identities through census procedures (Varshaver 2022: 201, Tishkov 2023). Attention has also been given to the subjective influences on census data, including the role of census takers (Farakhutdinov & Khairullina 2022) and respondents' own understandings of ethnicity (Filippova & Guerin-Paz 2015).

[1] It was preceded by two less successful and far from complete attempts: the 1920 census, which yielded few practical results due to wartime, territorial instability and population mobility, and the 1923 urban census, the main goal of which was to determine the social, rather than national or linguistic, composition of the country's population (see about them, as well as about the first Russian census of 1897, Vorobyov 1957: 3–16; for a comparison of language data from Russian censuses, see: Vakhtin 2001: 43–87).

[2] There is always confusion about translation of the Russian word *национальность* because its meaning is not the same as that of the English *nationality*; here we translate it as *ethnicity*.

In world practice, censuses have long been considered the most reliable source of information on the composition, distribution and occupation of the population. This was explained by the comprehensiveness of the survey, standard methods and tools (Kitchin 2017: 113). However, in the modern world, the attitude towards the quality of census data is ambiguous (see on possible approaches to ensuring the quality of censuses (Gabdrakhmanova & Alos-i-Font 2024)[3]. The quality of census data is influenced by such factors as the number of people who did not answer a particular question, the validity of the answers, data bias by region, urban-rural area, age of respondents, similarity in the distribution of answers obtained by different methods (Myagkov & Zhuravleva 2011: 33–34). The quality of data on the issue of language and ethnicity is also influenced by such objective factors as the state of these languages and the stability of ethnic self-determination.

The aim of this study is to assess the quality of the results of the Russian Population Census 2020–2021 (hereinafter RPC-2020). The focus is less on the accuracy of the statistics themselves and more on the extent to which they reflect linguistic and ethnic characteristics, both across the Russian Federation as a whole and, in particular, with regard to the languages of the indigenous peoples of the Sakha Republic (Yakutia).

Accurate information about the linguistic situation is particularly important for minority languages with decreasing number of speakers. This applies to almost all the languages of the indigenous minorities of Northern Siberia and the Far East, and especially to those of the Sakha Republic (Yakutia). Most of these languages are characterised by what is commonly termed "language shift" (Vakhtin 2001, Spolsky 2004, 2009, Grenoble 2021), particularly among urban youth (Ignatyeva 2022: 178). Language shift is a situation in when speakers gradually switch to another language, often losing proficiency in their original one. Without reliable data on the state of these languages — above all, accurate figures for the number of active and passive speakers — no language policy aimed at their preservation can be effective. In such circumstances, efforts to slow down, halt, or reverse the process of language shift are unlikely to succeed.

As one of the many authors assessing the prospects of the 2020 census prior to its implementation observed, "Population censuses are the only source that allows us to reliably estimate the range and number of peoples of the Russian Federation, determine the number of speakers of various languages, and clarify the educational, social, and demographic characteristics of ethnic groups" (Vorontsov 2020: 94). In this paper, we argue that this claim is, at best, questionable.

Our analysis follows the approach of "ecology of language" proposed by Einar Haugen (Haugen 1971, 1972: 356–339). There are several approaches to how "ecology of language" can be understood (e.g. Troshina 2020: 9–10, Kravchenko 2025); in this paper we understand "ecology of language" as a relationship between languages and their environments: how languages interact with each other, how

[3] Some scholars even claim that the days of the population census as we know it are numbered (Thorvaldsen 2017).

they are used by all types of speakers, and how their well-being is affected by social and political factors. We limited our analysis to the languages of the northeast — the languages of the Sakha Republic (Yakutia) and adjacent regions, namely: Aleut, Alyutor, Dolgan, Itelmen, Kerek, Koryak, Chukchi, Evenki, Even, Yupik (Eskimo) and Yukaghir. One can expect that the situation with other languages of the indigenous peoples of the North will be similar.

In the first part of the article, the results of the RPC-2020 are assessed against several of the criteria outlined above, drawing primarily on responses to the questions "What is your ethnicity (*национальность*)?", "What is your native language?", "Do you speak Russian?", "What other languages do you speak?", and "Which languages do you use in everyday life?" (for details, see Gabdrakhmanova & Alos-i-Font 2024). The second part provides an overview of the current sociolinguistic situation of the languages of the indigenous peoples of north-east Asia, while the third part analyses the RPC-2020 results in relation to these languages.

2. Analysis of the 2020 Russian Population Census data quality indicators

2.1. Population coverage by the Census

According to surveys conducted prior to the RPC-2020, 85% of respondents stated that they intended to take part in it, while 8% of Russian citizens reported that they did not plan to participate [4]. The actual participation rate was estimated at 57%, falling to 48% among those aged 18–24 (Farakhutdinov & Khairullina 2022: 69). The number of individuals who did not provide information about themselves is unusually high. This can be attributed primarily to the absence of visits from census enumerators, poor organisation of the census takers' work, and the recruitment of students with little experience of engaging with diverse groups. Additional factors include public indifference towards the census, driven by declining trust in the authorities and inadequate information provision, as well as the broader epidemiological situation associated with COVID-19 (ibid: 72–73).

Experts also link the low participation rate in RPC-2020 to the fact that, under the Federal Law *"On the Russian Population Census"* (25 January 2002, No. 8-FZ), the census is defined as a voluntary exercise. As a result, there is no legal liability for non-participation. Twenty years ago, residual expectations from the Soviet period — when participation in censuses was effectively mandatory — helped to ensure 93% population coverage in the 2002 census (Polyan 2004). This effect has since dissipated. Amendments to the 2010 Census Law [5] allowed population data to be supplemented with administrative records, a move criticised

[4] Russian Population Census 2021. VTsIOM. URL: https://wciom.ru/analytical-reviews/analiticheskii-obzor/vserossiiskaja-perepis-naselenija-2021 (accessed: 23.06.2025).

[5] Federal Law of November 28, 2009 No. 293-FZ "On Amendments to the Federal Law "On the Russian Population Census"; Federal Law of July 27, 2010 No. 204-FZ "On Amendments to Certain Legislative Acts of the Russian Federation".

by experts due to the difficulty of verifying data completeness (Zakharov & Vishnevsky 2010: 14). In cases where individuals were absent or refused to respond, it became sufficient to record basic information such as age and sex from registration databases. This increased the risk of both double counting and the imputation of missing data by enumerators (Andreev 2012).

The 2020 RPC was also the first census in Russia to introduce a fundamentally new method of data collection, allowing citizens to complete census forms online via the Unified Portal of State Services (Gosuslugi.ru). Although 41% of respondents indicated prior to the census that they intended to use this option[6], only 29% of actual participants ultimately did so.

With regard to ethnicity and native language, there is considerable regional variation in population coverage in RPC-2020. This is partly explained by the proportion of individuals not residing at their registered address and by non-response to questions on ethnicity[7] (for example, ethnicity is not recorded in three million electronic entries where language information is present)[8].

2.2. Overall response rate and non-response

Table 1 presents the dynamics of missing responses to questions on ethnicity and language across the last three Russian censuses. With the exception of the ethnicity question, the number of non-respondents is not explicitly reported in published census data. Instead, negative responses are combined with non-response, as no data are available to distinguish between these categories. This issue is particularly evident in responses to the question on proficiency in Russian. According to the 2002 census, 97.6% of the population spoke Russian; in 2010, the figure was 96.2%, while in RPC-2020 it fell to 91.3%. Taken at face value, this would imply that the number of people who do not speak Russian increased by 2.4 times compared with 2010. However, this figure is clearly unreliable, as the proportion of non-respondents significantly exceeds the proportion of citizens who genuinely lack proficiency in Russian (Gabdrakhmanova & Alos-i-Font 2024: 33).

In 2002, individuals “undecided about their ethnicity” were concentrated mainly in Moscow, St Petersburg, and the Moscow region (1 million out of 1.5 million). By 2010, similar patterns had emerged in additional ten regions (each with over 100,000 such cases) (Bogoyavlensky 2012: 5), and by 2020 this extended to 16 constituent entities of the Russian Federation[9]. Experts suggest that

[6] Russian Population Census-2021. VTsIOM. URL: https://wciom.ru/analytical-reviews/analiticheskii-obzor/vserossiiskaja-perepis-naselenija-2021 (accessed: 06/23/2025).
[7] Population Census 2020 // Demoscope Weekly. November 30 — December 14, 2021. No. 923-924. URL: http://www.demoscope.ru/weekly/2021/0923/perep01.php#2 (accessed: 06/23/2025).
[8] Kazakova D. When we saw the loss of 600 thousand Tatars, we were also surprised: Rosstat suggests looking for them in “non-answers”. Business online. URL: https://www.business-gazeta.ru/article/584255 (accessed: 23.06.2025).
[9] In Russian, this is “subjects of the federation”; the term refers to constituent entities of the Russian federation, the highest level of political division.

this increase reflects a growing reluctance among citizens to self-identify in terms of ethnicity or language, as these are increasingly viewed as private matters (ibid.: 5).

Demographers and statisticians have long recognised that certain questions are inherently sensitive and unlikely to elicit candid responses. In the Soviet census of 1937, for example, questions on religion and social status were considered problematic. As Elena Semenova notes, "The inclusion of questions on religion and social group membership in the 1937 census contradicted classical statistical principles, since such questions, affecting personal interests, often produced unreliable answers" (Semenova 1994: 16). Today, questions on ethnicity and native language appear to have acquired a similar sensitivity.

More broadly, "the accuracy of enumeration is unattainable where public trust is lacking. Given widespread scepticism towards the census, even the threat of criminal penalties for non-participation cannot ensure reliable results. Many individuals will simply avoid enumerators. […] Leading statisticians as early as the 1920s advocated excluding sensitive questions from census forms, particularly those affecting respondents' personal interests" (Zhiromskaya & Kornilov 2020: 36).

As Table 1 shows, the share of missing information was growing between the 2002, the 2020, and the 2020 censuses.

Table 1. **The share of missing information on national and language issues in the Russian Population Census of 2002, 2010 and 2020, people, %[10]**

	2002		2010		2021	
	Number	**%**	**Number**	**%**	**Number**	**%**
The entire population	145 166 731	100	142 856 536	100	147 182 123	100
Those who did not specify…						
ethnicity	1 460 751	1	5 629 429	3.9	16 594 759	11.3
proficiency in Russian	2 593 399	1.8	5 361 643	3.8	12 905 881	8.8
knowledge of languages other than Russian	1 420 544	1	4 544 001	3.2	12 056 452	8.2
native language	—[11]		4 544 578	3.2	16 638 532	11.3

The results of the RPC-2020 appear to be much less reliable than in previous censuses, and are difficult to compare with them due to insufficient consideration of various types of "non-responses". Their share was 16.5% (answers: "no ethnicity", "ethnicity not specified", "refusal") and is comparable to the share

[10] Russian Population Census 2002. Volume 4. Ethnic Composition and Language Proficiency, Citizenship. Rosstat. Available at: http://www.perepis2002.ru/index.html?id=17 (accessed: 23.06.2025). Russian Population Census 2010. Volume 4. Ethnic Composition and Language Proficiency, Citizenship. Rosstat. Available at: https://rosstat.gov.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm (accessed: 23.06.2025). Results of the 2020 Russian Population Census. Volume 5. Ethnic Composition and Language Proficiency. Rosstat. Available at: https://rosstat.gov.ru/vpn/2020/Volume5_National_composition_and_ownership_of_languages (accessed: 23.06.2025).

[11] The question was not asked.

of non-responses to questions about knowledge of languages (Gabdrakhmanova & Alos-i-Font 2024: 33).

2.3. Data shift by region and by urban-rural parameter

Let us consider this issue using the census statistics by Republic (constituent entities of the Russian Federation) (Table 2).

Table 2. **The share of missing information on ethnic and language issues of the 2020 RPC for the Republics of the Russian Federation, %[12]**

	Not specified…					
	ethnicity		native language		language proficiency	
	urban	rural	urban	rural	urban	rural
For the Russian Federation as a whole	14.0	3.3	14.2	2.8	10.3	1.8
Kabardino-Balkarian Republic	3.6	1.4	6.1	0.7	0.8	1.1
Karachay-Cherkess Republic	2.1	0.9	3.5	0.8	1.4	0.9
Republic of Adygea	14.8	5.6	14.8	3.8	12.4	3.0
Altai Republic	11.3	3.9	11.5	3.8	10.5	3.1
Republic of Bashkortostan	2.0	0.9	2.9	0.6	0.8	0.5
Republic of Buryatia	10.6	1.9	10.5	1.8	8.8	1.3
Republic of Dagestan	2.0	1.1	15.1	1.3	7.7	3.1
Republic of Ingushetia	4.6	2.4	5.8	3.1	6.4	6.2
Republic of Kalmykia	8.3	1.5	9.8	1.3	7.4	1.3
Republic of Karelia	13.9	2.2	13.2	1.9	12.0	0.5
Komi Republic	**28.2**	**2.9**	**27.4**	**2.4**	**25.6**	**1.6**
Republic of Crimea	11.8	4.4	10.9	3.7	7.0	2.2
Republic of Mari El	12.3	2.5	13.1	3.1	10.0	1.5
Republic of Mordovia	5.5	2.0	6.0	1.1	3.3	0.4
Republic of Tatarstan	3.1	0.8	3.6	0.6	2.5	0.6
Republic of Tuva	10.1	1.6	9.7	0.8	9.8	1.6
Republic of Sakha (Yakutia)	21.8	0.7	22.6	0.8	20.4	2.3
Republic of North Ossetia-Alania	9.2	0.5	11.6	0.8	8.2	0.8
Republic of Khakassia	26.1	3.2	25.9	3.1	24.0	1.5
Udmurt Republic	20.8	2.2	20.4	1.9	19.3	1.1
Chechen Republic	0	0	1.7	0.9	1.4	1.5
Chuvash Republic	13.6	2.0	13.6	2.1	9.9	1.3

The most striking are figures from Komi Republic where the share of people with missing information on their ethnicity was 22.5%. In five republics it exceeded 10%, in seven — from 6.2 to 9.5%, in 15 — less than 4.6%. On average, the share of people who did not indicate their ethnicity in the census forms was twice as low in Komi (5,7%) as in the country as a whole. The share of "non-responses" for native language was 6.6% (in the Russian Federation — 11.3%), proficiency in languages other than Russian — 5.2% (8.2%). For comparison: in the Ulyanovsk

[12] Results of the 2020 VPN. Volume 5. National composition and language proficiency. Rosstat. URL: https://rosstat.gov.ru/vpn/2020/Volume5_Nacionalnyj_sostav_i_vladenie_yazykami (date accessed: 23.06.2025).

region, 15.9% of the population do not have data on their ethnicity, 15.5% on their native language, 12% on their language proficiency; in Moscow, the corresponding figures are 22.7, 23.8, 15.3%. The statistics on ethnicity and language among city dwellers is thus less reliable than among rural dwellers. The distribution of "non-responses" on language proficiency by age gives a different picture for the four age groups (Gabdrakhmanova & Alos-i-Font 2024: 33). For the age groups 10–19 years and over 60 years, the share of non-responders is lower than the national average (6.9% in both cohorts versus 8.2% on average for the population). On the contrary, for the age cohort 20-39 years, the situation is the opposite (9.5% of non-responders). The general trend is common for all regions, but there are also some differences: for example, in Moscow, where the share of non-responders is significantly higher than the national average, the differences by age group are not so manifest[13].

2.4. Validity of answers

The increase in the number of "non-responses" and their uneven distribution do not explain all the differences in the results of the RPC-2020 compared to the 2010 RPC. In 2010, 15.6% of the population in Moscow who indicated language proficiency claimed to know English (Borisovich 2017: 32). In 2020, this figure was 5.8%, although the number of people who indicated language proficiency in the two censuses remained virtually unchanged (around 11 million). Has the population begun to understand differently what proficiency in English means? Is the population less willing to admit knowledge of English? Or were different criteria used during the 2020 census to determine the number of "those who indicated language proficiency"? The lack of answers to these questions makes it impossible to compare the dynamics of multilingualism between the two most recent censuses. In addition, in 2010 respondents in Moscow who answered questions about knowledge of languages claimed to know an average of 1.23 languages per person; in 2020, this figure was 1.1. It seems highly unlikely that Muscovites have become less knowledgeable of languages, especially English, over the past 11 years (Gabdrakhmanova & Alos-i-Font 2024: 35).

2.5. Question-answer uncertainty

In Russian censuses, ethnicity and native language are recorded based on the respondent's answers. Ambiguity in the meaning of the terms may have resulted in uncertainty when providing the answers. Some of the RPC-2020 participants perceived the question about ethnicity as a question about their identification with an ethnic group, while others perceived it in the national sense (206,081 people

[13] Results of the 2020 VPN. Volume 5. Table 12. Language proficiency by population of different age groups. Rosstat. URL: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Frosstat.gov.ru%2Fstorage%2Fmediabank%2FTom5_tab12_VPN-2020.xlsx&wdOrigin=BROWSELINK (accessed: 23.06.2025).

claimed their ethnicity was "Russian citizens", another 652,387 people indicated it as "RF"). Some experts explain the emergence of supra-ethnic self-identification through the decreasing significance of ethnic identity[14], while others emphasize that national identity must be included as an option into the census form due to its growing significance in contemporary Russia (Tishkov 2023: 202).

It appears that the question about the native language was not clear enough for some of the RPC-2020 respondents. The existing ambivalence of this term in public perception is shown by the results of a sociological survey in 2019 by the N.N. Miklouho-Maclay Institute of Ethnology and Anthropology of the Russian Academy of Sciences, conducted in different regions of the country. According to the survey, a fifth of the respondents found it difficult to explain the meaning of the term "native language", while the opinions of the rest were distributed between "language of constant use", "language of childhood, family, relatives", "ethnic language" (Stepanov 2019: 143). At first glance, this problem can be solved by census questions on proficiency in Russian and other languages of the peoples of Russia and foreign languages. However, when answering these questions, the respondents are free to decide what level of knowledge is necessary for proficiency in the language. The question on the use of languages introduced in the RPC-2020 is of the same nature. Is passive use of language sufficient? Or is it necessary to speak it on a daily basis? Is a school teacher of a foreign language considered a daily user of a language if he or she uses it exclusively in the teaching process? The lack of clarification of the census question makes it difficult to interpret the answers (see below for more details).

2.6. Data processing

The RPC-2020 recorded 1,658 names of nationalities in Russia. The column "other nationalities" included the answers of 109,891 respondents. Among the ethnonyms, there are responses of children from mixed families (Russian-Greeks, Sakhalyars[15]), "new" ethnic communities (Rusich, Soviet, Old Believer, Orlovtsy, Zemlyanin[16]), different names of one ethnic community (Bashkirts — Bashkirs — Bashkort — Bashkurt), obsolete self-names of ethnic and tribal communities (Yuits). The information background created by lobby groups before and during the RPC-2020 prompted people, under pressure, to revise their self-identification or

[14] Kolebakina-Usmanova E. "Identified themselves as Russians": Rosstat explained the "loss" of 590 thousand Tatars // Business online. URL: https://www.business-gazeta.ru/article/583782 (accessed: 23.06.2025).

[15] *Sakhaliar* (Rus. *сахаляр*) — colloquial name widely used in the Sakha Republic for children of mixed descent (mixed marriages: Yakut / Russian, other non-Yakut ethnicities / Yakut, etc).

[16] *Rusich* (*русич*) is a mythical name of a Slavic tribe used by some right-wing groups as an alleged ethnonym for the ancestors of today's Russians. *Orlovtsy* are those who live in the *Orlov Region* (*Orlovskaya Oblast'*). *Zemlyanin* means Earthling.

switch to a more neutral status[17]: this is what happened, for example, in the Astrakhan region where some Tatars were "counted" as Nogais[18].

The analysis of the data is complicated by the option of indicating a second ethnicity, while only the first answer is taken into account in the published results. The exception was the indigenous peoples of the Russian Federation, for whom "one should take into account all persons who indicated the corresponding ethnicity as any (first or subsequent) answer to the question on ethnicity".[19] The desire of some experts to identify multiple ethnic and linguistic identities of Russians, which was embedded already in the 2002 Russian Population Census and fully used in the RPC-2020 (Etnografiya… 2003, Stepanov 2001), is assessed by some researchers as an unsuccessful attempt to use an expensive census for scientific experimentation (Polyan 2004). At the same time, according to some estimates, the number of those in Russia who can indicate dual ethnicity and two or more native languages during the census ranges from 7 to 20 million people (Stepanov 2019: 153).

The list of nationalities of Russia used in processing the RPC-2020 includes 194 items and practically repeats the tools of the 2002 and 2010 censuses. In the list of ethnicities used in the 2002 Russian Population Census, independent subgroups were identified in some ethnic groups, but this principle was not used consistently (Iskhakov 2002); this selective approach was retained in the RPC-2020 which recorded 372 languages, including 155 languages of the peoples of Russia. This list also includes non-existent languages: European, Asian, Central Asian, that of the CIS, Northern, Finno-Ugric, Canadian, Swiss, Monegasque, Senegalese, Cameroonian, Polynesian, etc. A programming language was recorded, the speakers of which believe that they can speak it (apparently due to the peculiarities of the census instructions[20]). The number of "new" languages increased from 10 in 2010 to 45 in 2020. The largest new "Dagestani" language (19 thousand responses) is a mystery for specialists[21]. Some languages are combined into one: Catalan and Spanish, Languedoc and French, Gascony and Basque. RPC-2020 recorded people who can speak (not just read!) Latin, Ancient Greek, Ancient Egyptian, Avestan, Polovtsian, Old Church Slavonic, and the Etruscan language that still remains undeciphered. Such absurdities are less common in data on native languages, although Gabonese, Latin, and Old Church Slavonic are also recorded there.

[17] Kolebakina-Usmanova E. "Identified themselves as Russians": Rosstat explained the "loss" of 590 thousand Tatars. Business online. URL: https://www.business-gazeta.ru/article/583782 (accessed: 23.06.2025).

[18] "The enemy was the COVID pandemic": how VKT explained the "loss" of 600 thousand Tatars. Business online. URL: https://www.business-gazeta.ru/article/579311 (accessed: 23.06.2025).

[19] Russian Population Census 2020. Methodological notes. Rosstat. URL: https://rosstat.gov.ru/storage/mediabank/Tom5_Metodolog_VPN-2020.docx (accessed: 23.06.2025).

[20] Russian Population Census 2020. Methodological notes. Rosstat. URL: https://rosstat.gov.ru/storage/mediabank/Tom5_Metodolog_VPN-2020.docx (accessed: 23.06.2025).

[21] Kazakova D. When we saw the loss of 600 thousand Tatars, we were also surprised: Rosstat suggests looking for them in "non-answers". Business online. URL: https://www.business-gazeta.ru/article/584255 (accessed: 23.06.2025).

To conclude this section, a number of inaccuracies and contradictions are found in the resulting statistics of the RPC-2020 regarding the national and linguistic composition of the population of the country. The census turned out to be significantly less reliable than previous Russian censuses, and therefore its results must be handled with caution in scientific and analytical work, as well as for implementation of state ethnic and language policy.

3. "Minority languages" and language shift

For the languages of the indigenous peoples of the North of the Russian Federation (hereinafter referred to as the Languages of the Peoples of the North, LPN), the problem of "language shift" is extremely acute: the transition from one language to another is common not at the level of individual speakers, but at the level of the entire language community, which, as is known, inevitably leads to the complete loss of language, or "language death" (Grenoble 2021). The speed of the process is very fast: if at the end of the 20th century most of the LPN languages still had a chance to survive (see Vakhtin 2001 for details), today the situation is completely different: for all LPN languages, intergenerational transmission has been interrupted, the language is actively used by the oldest generation at best, but even if they use it, it is in a simplified language with strong Russian interference.

The authors of the project "Languages of Russia"[22] distinguish ten levels of language preservation ("vitality statuses", in their terminology): from "extinct" (1A) to "prosperous" (4B). Of this set, only the first five levels are relevant for LPN, see Tables 3 and 4.

Table 3. **Vitality statuses (reference)**

Label	Status name	Description
1A	disappeared	The last active speakers died in the 20th century.
1B	asleep	The last active speakers died in the 21st century
1C	falling asleep	Regular communication does not occur
2A	interrupted	Intergenerational transmission has been interrupted throughout the entire territory
2B	intermittent	Intergenerational transmission is maintained within a small area, which is an insignificant part of the entire community range.
3A	localized	Persisting in a limited range; intergenerational transmission persists
. . .		
4B	prosperous	A majority language with stable intergenerational transmission and provided with a developed language infrastructure
(R)	*index*	*Only data on the language within Russia are taken into account*
-	*index*	*Possibly should be classified into the previous group*
+	*index*	*There are actions in relation to language that allow it to be classified into the following group*

[22] This project supported by Russian Science Foundation and led by Yu. Koryakov aims at compiling a complete list of the languages of the Russian Federation; see: https://jazykirf.iling-ran.ru/

Table 4. **Indigenous languages in Yakutia by degree of vitality**

Language	Status	Comment
Yakut	3B	Given for comparison
Minor languages of the Republic of Sakha (Yakutia) and Adjacent Territories		
Dolgan	2B	
Chukchi	2B-	
Koryak	2A	
Northern Yukaghir	2A	
Alutor	2A-	
Naukan Eskimo	2A-	
Chaplino Eskimo	2A(R)-	
Itelmen	1B	
Southern Yukaghir	1B+	
Aleut	1B(R)	The last speaker died in 2021
Kerek	1B	By 2005, there were no speakers left

Formally, the languages of the indigenous minorities of the North are taught in schools in the Republic of Sakha (Yakutia). However, pupils often lack motivation to study these languages, most teachers do not speak them fluently, and lessons frequently amount to little more than the rote memorisation of vocabulary (see Table 5). In this situation, the functions of the disappearing languages change radically. As sociological surveys show, for a considerable number of the Russian population, the term "native language" has acquired the meaning of the language of ethnicity, and not the language they spoke in childhood (Drobizheva & Ryzhova 2015: 17).

There is a notable relationship between processes of ethnic revival and language shift, which renders the concept of "language proficiency" particularly ambiguous (Vakhtin 2010: 415). A paradox emerges: on the one hand, the community requires its language as a key marker of ethnic identity; on the other, only a small proportion of its members are actually able to speak it. In such circumstances, communities adopt a compensatory strategy: individuals claim to have a good command of the language, even though they do not use it in everyday communication. As a result, "language proficiency" proves not merely to be a matter of degree (i.e. speaking better or worse), but also one in which assertions that individuals or their fellow villagers "know their native language" may in fact signify the opposite — that they do not speak it at all. In this way, the language loses its communicative function and comes to be used primarily as a marker of ethnic identity.

One of the authors of the present article once called this extreme state of language "a demo version": members of the community demonstrate to the researcher their alleged ability to speak the language, while in reality the language is no longer used for communication and exists only in its symbolic function. This is, of course, an extreme case, but to some extent it is present everywhere where a language is classified as endangered, that is, where a situation of language shift can be identified (Vakhtin 2006). These processes are also recorded by other researchers; cf. the description of the situation with the Yupik (Eskimo) language

of the village of Novoye Chaplino (south-eastern Chukotka), where there are practically no fluent native speakers of the language left, but the residents nevertheless constantly insert individual Eskimo words and common phrases into their Russian speech, thereby demonstrating their commitment to the language. "Bringing Yupik words… into a conversation allows local people not only to express their attachments but, more importantly, to feel connected, both with each other and with past experiences, to feel that they "belong" (Mourgunova-Schwalbe 2022: 195).

Table 5. **Number of schools and schoolchildren where LPNs were taught in 2024 in Republic of Sakha (Yakutia)[23]**

	Number of schools		Number of schoolchildren	Of these, type of education programme			Comments
				primary	general	secondary	
	2023	2024					
Dolgan	1	0	6	0	4	2	for 2023 only
Chukchi	4	3	41	14	25	2	
Even	7	21	530	206	263	58	
Evenki	9	25	358	120	226	12	
Yukaghir	3	2	52	18	26	8	

It is obvious that in a situation of language shift the concept of "language proficiency" is blurred; the last section of this article is devoted to describing the consequences of this for the adequacy of census results.

4. Russian Population Census 2020 and minor languages of the Republic of Sakha (Yakutia)

The documentation of languages is at the heart of the modern approach to linguistic diversity. "Only a thorough situational analysis of languages and specific linguistic ecosystems can guarantee the development of policies that can successfully protect linguistic diversity".[24] In this regard, the question of the source of accurate and reliable data on languages and language situations becomes important. Are the data we have on the languages of the Russian Federation in general and on the languages of the indigenous peoples of northeast Asia in particular reliable enough? And if not, from what sources can reliable data be obtained?

In the 2020 Census data release,[25] all information on the minority languages of the North-East is concentrated in Volume 5 "National Composition and Language

[23] Portal of the indigenous peoples of the North "Ilken": https://ilken.ru/2023/11/28/sostoyanie-obrazovaniya-kmns-v-yakutii/ (accessed: 06/23/2025).
[24] See: World Atlas of Languages — https://en.wal.unesco.org/world-atlas-languages (accessed: 23.06.2025).
[25] See: https://rosstat.gov.ru/folder/56580 (accessed: 23.06.2025)

Proficiency", in four tables[26]. Table 6 summarizes the data from these four tables for the languages listed above.

Table 6. **Summary data for the four tables of the RPC-2020**

Census 2020/2021	Population of the Russian Federation by native language[27]	Indigenous peoples of the North by native language[28]	Indigenous peoples of the North by use[29]	Indigenous peoples of the North by language command[30]
1	*2*	*3*	*4*	*5*
Aleut	136	135	91	109
Alutor	168	182	164	174
Dolgan[31]	-	298	74	404
Itelmen	808	783	356	489
Koryak	4164	4123	2024	2549
Chukchi	8526	8619	1615	2485
Even	7208	6956	3724	4954
Evenki	8724	8559	3311	5488
Eskimo	816[32]	810	90	168
Yukaghir	516	512	235	330

The discrepancies between the data obtained in response to different questions for the same language are striking. Of course, some discrepancies, like those between columns 1 and 2 (the population of the Russian Federation by native language and indigenous peoples by native language) may be regarded as insignificant (for Aleut 136 and 135, i.e. less than 1%, for Koryak 4164 and 4123, i.e. exactly 1%, etc.); these discrepancies can be explained by recording, calculation and other types of errors.

The situation is different with other columns of the table where the discrepancies are systematic. In all cases, the number of those who named a language as their native language (*column 3*) is significantly, sometimes several times, greater than the number of those who "use" this language (*column 4*) — from

[26] *Table 6*: Population by Native Language; *Table 18*: Proficiency in Languages of Indigenous Peoples of the Russian Federation; *Table 19*: Use of Languages by Indigenous Peoples of the Russian Federation; and *Table 20*: Population of Indigenous Peoples of the Russian Federation by Native Language. Information on language proficiency and use in everyday life was obtained from the answers to the questions "What other languages (besides Russian) do you speak?" and "Which of them do you use in everyday life?" Information on native language was obtained from the answers to the question "Your native language".

[27] VPN-2020, Volume 5, table. 6: Population by native language.

[28] VPN-2020, Volume 5, Table 20. Population of indigenous peoples of the Russian Federation by native language. See below for details

[29] In Russian, *ispolzovanie*, "using"; VPN-2020, Volume 5, Table 19. Use of languages by indigenous peoples of the Russian Federation

[30] In Russian, *vladenie* can mean "possession" or "command"; VPN-2020, Volume 5, Table 18: Proficiency in languages of the indigenous minorities of the North.

[31] Only Dolgans of the Republic of Sakha (Yakutia) are shown.

[32] Plus one person indicated Yuitsky, another name for the Eskimo language.

33% Aleut to 90% Eskimo. The number of those who "command" the language (*column 5*) is always somewhat higher than the number of those who "use" it (from 87% Eskimo to 6% Alutor).

How can such a discrepancy be explained? Why is the number of respondents who "command" the language is significantly higher than that of who "use" it, and at the same time both figures are significantly lower than the number of those who call this language their native language? Who are these approximately 300 Itelmens who called Itelmen their native language, despite the fact that they do not "command" it or "use" it? The answer to this question is more or less clear: for respondents, the phrase "our native language" does not mean the language they have known since childhood (that is, this is not their "mother tongue"): this is their "language of ethnic self-identification", or the titular language, that is, the language whose name matches the name of the ethnicity (for more details, see: Belikov 1997, 1999, Belikov & Krysin 2002).

It is harder to understand the discrepancies between columns 4 and 5. Who are these approximately 500 Koryaks and approximately 750 Chukchi who have "command" of their titular languages but do not "use" them? Do these figures conceal a situation in which people know their titular language and could speak it, but have no one to do so with? Or does "command" here again mean the feeling that the titular language belongs to a person regardless of whether they actually speak it or not? These and other questions cannot be clarified without special research.

It is interesting to compare the data on the languages of the indigenous minorities of the North in the 2020 census with the data of the two previous censuses — 2002 and 2010 (Table 7). If a comparison of the 2002–2010 data shows a steady *decrease* in the number of people who called a specific LPN their native language, the 2020 census demonstrates a paradoxical *increase* in the number of those who called these languages their native language. This, of course, does not mean that the language shift has stopped and the number of speakers of indigenous languages is growing rapidly. The increase in the number of those who called a language their native language indicates a change in the meaning of the phrase "native language".

It should be noted, finally, that the 2020 census materials contain interesting data on the relationship between language and ethnic identity. Tables 8 and 9 provide data on the ethnicity of people who speak a particular LPN [33]: the second column shows numbers of those whose ethnic identity matches the name of the language, the third column shows the numbers of all the others who indicated proficiency in this language.

The lists of respondents of different ethnicities shown in the third column who reported their knowledge of the LPNs that are not their title languages require a comment. Among them, there are obvious errors, random or frivolous answers (three people of Vod (northwest European part of the Russian Federation) ethnicity

[33] VPN-2020, Volume 5, tables 18 and 19.

who speak the Aleut language; one Nenets, who speaks several languages of the indigenous peoples of the North; etc.). There are isolated proficiencies (one Koryak who speaks Aleut, one Chukchi who speaks Itelmen). Finally, there are different names of ethnicity (some call themselves Alyutors, some — Koryaks, the same is true for the pairs Itelmen-Kamchadals, Evens-Orochi-Uilta). If we exclude these cases, one sees that the languages fall into two groups (Table 9).

Table 7. **Comparison of data on the indigenous peoples of the North by native language in three censuses**

	Census of the Russian Federation 2002: Indigenous minorities by native language[34]	**Census of the Russian Federation 2010: Indigenous minorities by native language[35]**	**Census of the Russian Federation 2021/22: Indigenous minorities by native language[36]**
Aleut	175	45	135
Alutor	40	25	182
Itelmen	385	82	783
Dolgan (RF)	4865	1054	5346
Koryak	3019	1665	4123
Chukchi	7742	5095	8619
Even	7168	5656	6956
Evenki	7584	4802	8559
Eskimo	410	508[37]	810
Yukaghir	604	370	512

Table 8. **Proficiency in the languages of the indigenous minorities of the North**

Language	**Number of "title" speakers**	**Speakers of the language of other ethnicities**
Aleut	105 Aleuts	3 Vod, 1 Koryak
Alutor	11 Alutorians	163 Koryaks
Itelmen	455 Itelmens	25 Kamchadals, 8 Koryaks, 1 Chukchi
Kerek	1 Kerek	
Koryak	2205 Koryaks	202 Chukchi, 59 Evens, 39 Itelmens, 22 Kamchadals, 17 Alyutors, 4 Evenks, 1 Nenets
Chukchi	2308 Chukchi	71 Evens, 39 Yukaghirs, 20 Eskimos, 19 Koryaks, 15 Chuvans, 5 Nanais, 4 Alutors, 1 Nenets, 1 Ulchi, 1 Evenki
Evenki	5394 Evenks	74 Evens, 6 Kets, 3 Dolgans, 3 Selkups, 1 Nanai, 2 Nenets, 2 Negidals, 1 Udege, 1 Chukchi, 1 Enets
Even	4645 Evens	118 Evenkis, 72 Yukagirs, 63 Chukchi, 35 Koryaks, 7 Kamchadals, 3 Orochis, 3 Nanais, 2 Nenets, 2 Itelmens, 1 Vepsian, 1 Dolgan, 1 Selkup, 1 Uilta
Eskimo	160 Eskimos	7 Chukchi, 1 Even
Yukaghir	300 Yukaghirs	24 Evens, 5 Chukchis, 1 Nenets

[34] https://ru.wikipedia.org/wiki/Languages_of_Russia_in_order_of_the_number_of_speakers.
[35] Ibid.
[36] VPN-2020, Volume 5, Table 20. Population of indigenous peoples of the Russian Federation by native language.
[37] Another 4 people indicated *the Yuit* language, and another 5 indicated *the Sireniki language*.

Table 9. **Proficiency in the languages of indigenous minorities of the North by representatives of "their own" and "foreign" ethnicities**

The language is spoken mainly by "our own people"		
Aleut	105 Aleuts	
Itelmen	455 Itelmens + 25 Kamchadals	8 Koryaks
Eskimo	160 Eskimos	7 Chukchi
Yukaghir	300 Yukaghirs	24 Evens, 5 Chukchi
The language is spoken by representatives of several ethnicities		
Dolgan	4701 Dolgans	20 Nganasans, 13 Nenets, 25 Evenkis, 2 Evens
Koryak	2368 Koryaks + 28 Alutors	202 Chukchi, 59 Evens, 39 Itelmens, 22 Kamchadals, 4 Evenkis
Chukchi	2308 Chukchi + 15 Chuvans	71 Evens, 39 Yukaghirs, 20 Eskimos, 19 Koryaks, 5 Nanais, 4 Alutors
Evenki	5394 Eveniks	74 Evens, 6 Kets, 3 Dolgans, 3 Selkups
Even	4645 Evens + 3 Orcs	118 Evenkis, 72 Yukaghirs, 63 Chukchi, 35 Koryaks, 7 Kamchadals, 3 Nanais

It should be noted that the upper part of the table contains the languages of groups that were traditionally sedentary or semi-sedentary hunters (maritime or land-based), while the lower part contains the languages of reindeer-herding groups. Thus, the Chukchi language is reported to be spoken by representatives of all groups except the Aleuts, with whom the Chukchi never had contact; the Even language is reported to be spoken by everyone except the Aleuts and Eskimos, and so on. It is known that the languages of reindeer herders (for example, the Chukchi) served as a common language for the entire region: the Eskimos living in the eastern part of the Chukchi nomadic area got the word about, for example, the annual Anyui fairs on the western edge of the area through communication with the Chukchi (Bogoraz 1934: 53 ff., Menovshchikov 1969). It is interesting that the VPS-2020 confidently records the remnants of this multilingualism, when (former?) reindeer herding peoples turn out to be speakers of languages that are more or less spoken by all their neighbours, both (former?) nomads and (former?) sedentary hunters and fishermen, while the languages of the latter are known only to themselves.

It is not important whether eight Koryaks actually speak Itelmen or 39 Itelmens speak Koryak: what is important is that in response to the census taker's question, these people preferred to stress that they know not only their own language and Russian, but also the language of their neighbours.

5. Conclusion

This paper set out to assess the quality of the results of the Russian Population Census 2020–2021, both for Russia as a whole and, in particular, with regard to the languages of the indigenous peoples of the Sakha Republic (Yakutia).

The findings are as follows.

1) At the national level, the 2020 census was the first large-scale attempt to incorporate administrative (register-based) data into census inputs. Further critical analysis is required for each census variable, including those relating to ethnicity

and language. This is especially important given the sensitivity of ethnic and linguistic relations in a country characterised by a wide range of ethnic communities.

2) The census data on the languages of indigenous minorities in the Sakha Republic (Yakutia) and neighbouring areas allow for a number of interesting observations concerning the relationship between ethnic and linguistic self-identification, as well as the significance of historical language contact and multilingualism for present-day populations. However, the 2020 census data cannot be used uncritically. Taking at face value the reported numbers of people "commanding" or "using" a given language, or those identifying a particular indigenous language as their "mother tongue", carries a considerable risk of error in the formulation of practical language policy in the region. In other words, the census offers limited value for research in language ecology, language policy, or the revitalisation of endangered languages.

3) Fieldwork evidence indicates that, for instance, there are far fewer than 109 Aleut speakers, 168 Eskimo speakers, or 330 Yukaghir speakers if "speaking" is understood as the ability to communicate rather than merely recalling a small number of words.

4) A key practical implication for language maintenance and revitalisation is that census data alone are insufficient and cannot provide a reliable picture of language use or endangerment. Detailed, localised information is required — concerning specific languages, in particular villages, and among distinct communities. Such data can only be obtained through regular fieldwork conducted in close collaboration with local researchers and their networks. Encouragingly, colleagues in the Sakha Republic have begun this work in 2025 as part of the RSE project "Languages and Cultures of the Peoples of the North and the Arctic of the Russian Federation".

Acknowledgements

The work of Nikolai Vakhtin was supported by RSF, project "Languages and Cultures of the Peoples of the North and the Arctic of the Russian Federation: Comprehensive socio-humanitarian research (on the basis of big data)" (Agreement No 25-78-30006 of 22.05.2025).

CRediT authorship contribution statement

Gulnara F. Gabdrakhmanova: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing — Original Draft. **Nikolai B. Vakhtin:** Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing — Original Draft, Writing — Review & Editing, Funding Acquisition.

References

Anderson, Benedict. 2001. *Voobrazhaemye Soobshchestva. Razmyshleniya ob Istokakh i Rasprostranenii Natsionalizma* (Imagined communities. Reflections on the Origin and Spread of Nationalism). Moscow: KANON-press-C, Kuchkovo pole. (In Russ.).

Andreev, Evgenii M. 2012. O tochnosti rezul'tatov rossiiskikh perepisei naseleniya i stepeni doveriya k raznym istochnikam (On the accuracy of the results of Russian population censuses and the degree of trust in different sources of information). *Voprosy statistiki* 11. 21–35. (In Russ.).

Arutyunyan, Yurii V. (ed.). 1972. *Sotsial'noe i natsional'noe. Opyt ehtnosotsiologicheskikh issledovanii po materialam Tatarskoi ASSR* (Social and National: Essay in Sociological Research in Tatar ASSR). Moscow: Nauka. (In Russ.).

Belikov, Vladimir I. & Leonid P. Krysin. 2002. *Sotsiolingvistika*. (Sociolinguistics). Moscow: Rossiiskii Gosudarstvennyi Gumanitarnyi Universitet Publ. (In Russ.).

Belikov, Vladimir I. 1997. Nadezhnost' sovetskikh ehtnodemograficheskikh dannykh (Nadezhnost' sovetskih etnodemograficheskih dannyh (Reliability of Soviet ethnodemographic data). *Malye yazyki Evrazii: Sotsiolingvisticheskii aspekt*. Moscow: Moscow State University Publ.). (In Russ.).

Belikov, Vladimir I. 1999. Metodicheskie neudachi v sotsiolingvisticheskikh oprosakh. (Methodological failures in sociolinguistic surveys). *Tipologiya i teoriya yazyka: Ot opisaniya k ob''yasneniyu. K 60-letiyu A. E. Kibrika*. Moscow: Yazyki slavyanskikh kul'tur Publ.). (In Russ.).

Bogoraz, Vladimir G. 1934. *Chukchi. Chast' 1: Sotsial'naya organizatsiya*. (Chukchi. Part 1: Social organization). Leningrad: Institute of the Peoples of the North Publ. (In Russ.).

Bogoyavlenskij, Dmitrii. 2012. Perepis' 2010: Ehtnicheskii srez (Census 2010: An ethnic cross-section). *Demoskop Weekly*. 531–532. URL: www.demoscope.ru/weekly/2012/0531/demoscope531.pdf (accessed 20.05.2023). (In Russ.).

Drobizheva, Leokadiia M. (ed.). 2002 *Sotsial'noe neravenstvo ehtnicheskikh grupp: Predstavleniya i real'nost'*. (Social Inequality of Ethnic Groups: Perception and Reality). Moscow: Academia. (In Russ.).

Drobizheva, Leokadiia M. & Svetlana V. Ryzhova. 2015. Grazhdanskaya i ehtnicheskaya identichnost' i obraz zhelaemogo gosudarstva v Rossii (Civil and ethnic identity and the image of the desired state in Russia). *Polis. Politicheskie Issledovaniya* 5. 9–24. (In Russ.).

Farakhutdinov, Shamil F. & Nursafa G. Khairullina. 2022. Vserossiiskaya perepis' naseleniya 2020–2021 gg. glazami perepischika (Russian population census 2020–2021 through the eyes of a census taker). *Sotsiologicheskie Issledovaniya* 10. 66–73. (In Russ.).

Filippova, Elena, Dominique Arel & Catherine Gousseff (eds.). 2003. *Ehtnografiya perepisi-2002*. (Ethnography of the 2002 census). Moscow: Aviaizdat. (In Russ.).

Filippova, Elena I. & Franc Geren-Pas. 2015. Rasovye i ehtnicheskie kategorii v perepisyakh 2010–kh godov (Racial and ethnic categories in the 2010s censuses). *Ehtnograficheskoe obozrenie* 1. 5–25. (In Russ.).

Gabdrahmanova, Gulnara F. & Ektor Alos-i-Font. 2021. K voprosu o fiksatsii natsional'noi prinadlezhnosti i vladeniya yazykami vo Vserossiiskoi perepisi naseleniya 2020–2021 gg. (On the issue of recording ethnicity and language proficiency in the Russian population census 2020–2021). *Sotsiologicheskie Issledovaniya* 1. 28–39. (In Russ.).

Grenoble, Lenore. Language Shift. *Oxford Research Encyclopaedia of Linguistics*. Retrieved from https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-347 (accessed 30.03.2026)

Haugen, Einar. 1971. The ecology of language. *Linguistic Reporter* 13 (1). 19–26.

Haugen, Einar. 1972. *The Ecology of Language: Language Science and National Development*. Stanford University Press*.

Ignat'eva, Vanda B. 2022. Minoritarnye yazyki Severa v sotsiologicheskoi perspektive (Minority languages of the North in a sociological perspective). *Identichnost', yazyk i*

* Включён в перечень иностранных организаций, деятельность которых признана нежелательной на территории России.

kul'tura molodezhi korennykh narodov Severa: Itogi issledovaniya v Yakutii. Novosibirsk: Nauka. 163–184. (In Russ.).

Kitchin, Rob. 2017. Bol'shie dannye, novye ehpistemologii i smena paradigm (Big data, new epistemologies and paradigm shifts). *Sotsiologiya 4M* 44. 111–152. (Transl. from Eng.).

Koryakov, Yurii B. 2017. Yazyki Moskvy po dannym perepisi 2010 goda (Languages of Moscow according to the 2010 census). *Rodnoj yazyk* 7 (2). 24–44. (In Russ.).

Krasnopol'skaya Irina I. & Galina S. Solodova 2016. Perepis' naseleniya kak instrument stroitel'stva natsii (Population census as a tool for nation building). *Mir Rossii* 25 (1). 55–78. (In Russ.).

Kravchenko, Alexander V. 2025. Ecolinguistics: A paradigm shift. *Russian Journal of Linguistics* 29 (1). 17–36. https://doi.org/10.22363/2687-0088-39530

Menovshchikov, Georgii A. 1969. Lingvisticheskie kontakty aborigenov krainego severo-vostoka Sibiri (Linguistic contacts of the aborigines of the extreme north-east of Siberia). *Proiskhozhdenie aborigenov Sibiri i ikh yazykov. Materialy mezhvuzovskoi konferentsii 11–13 maya 1969 g*. Tomsk: Tomsk State University Publ. 27–30. (In Russ.).

Morgounova Schwalbe, Daria. 2022. In the language of their hearts: Emotions and language choice in child-parent interaction, insights from a Yupik village. *Études Inuit Studies* 45 (1–2). 171–205.

Myagkov, Aleksandr Yu. & Svetlana L. Zhuravleva. 2011. Ehksperimental'naya otsenka kachestva dannykh v telefonnom interv'yu (Experimental assessment of data quality in a telephone interview). *Sotsiologiya 4M* 32. 26–52. (In Russ.).

Neudachin, Vasilii. 1905. Kratkii obzor tsifrovykh dannykh po Yakutskoi oblasti (A brief overview of digital data on the Yakut region). *Pervaya Vseobshchaya perepis' naseleniya Rossiiskoi imperii 1897 g.* In Nikolai A. Troinitskii (ed.), *Vol 80: Yakutskaya oblast',* III–XIII. (In Russ.).

Polyan, Pavel M. 2004. Podgotovka, provedenie i nekotorye kachestvennye itogi Vserossiiskoi perepisi naseleniya 2002 goda (Preparation, implementation and some qualitative results of the Russian population census of 2002). *Demoskop Weekly*. 155–156. (In Russ.). URL: http:.www.demoscope.ru/weekly/2004/0155/analit01.php (accessed: 20.04.2023).

Semenova, Elena A. 1994. Ehtnicheskaya informatsiya v materialakh perepisi naseleniya (istochnikovedcheskii aspekt) (Ethnic information in population census materials (sources study aspect)). *Ehtnokognitologiya. Vol. 1: Podkhody k izucheniyu ehtnicheskoi identifikatsii*. Moscow: Institute of Natural and Cultural Heritage Publ. (In Russ.).

Spolsky, Bernard. 2004. *Language Policy*. Cambridge: Cambridge University Press.

Spolsky, Bernard. 2009. *Language Management*. Cambridge: Cambridge University Press.

Stepanov, Valerii V. 2001. Rossiiskaya perepis' 2002 goda: puti izmereniya identichnosti bol'shikh i malykh grupp (Russian census of 2002: Ways of measuring the identity of large and small groups). *Issledovaniya po prikladnoi i neotlozhnoi ehtnologii* №.145. Moscow: Institute of Ethnology and Anthropology. (In Russ.).

Stepanov, Valerii V. 2019. Izmerenie kul'turnogo mnogoobraziya Rossii (Measuring the cultural diversity of Russia). In Marina Yu. Martynova & Valerii V. Stepanov (eds.), *Izmerenie kul'turnogo mnogoobraziya. Yazykovaya situatsiya, perepisi, polevaya ehtnostatistika*, 140–154. Moscow: Institute of Ethnology and Anthropology. (In Russ.).

Thorvaldsen, Gunnar. 2017. *Censuses and Census Takers: A Global History*. London: Routledge.

Tishkov, Valerii A. 2003. Perepisi naseleniya i konstruirovanie identichnostei. (Population censuses and construction of identities). *Na puti k perepisi*. Moscow: Aviaizdat. 9–38. (In Russ.).

Tishkov, Valerii A. 2023. O perepisyvanii narodov, ili dekonstruktsiya perepisei naseleniya (On the census of peoples, or deconstruction of population censuses). *Ehtnograficheskoe obozrenie* 4. 183–211. (In Russ.). https://doi.org/10.31857/S0869541523040085

Troshina, Natalia N. 2020. *Ehkologiya yazyka. Analiticheskii obzor* (Ecology of language. Analytical review) Moscow: INION. (In Russ.).

Vakhtin, Nikolai B. 2001. *Yazyki narodov Severa v KHKh veke: Ocherki yazykovogo sdviga* (Languages of the Peoples of the North in the Twentieth Century: Essays on Language Shift). St Petersburg: Dmitrij Bulanin Publ. (In Russ.).

Vakhtin, Nikolai B. 2010. Opisanie malykh yazykov Rossiiskoi Federatsii na sovremennom ehtape (Description of small languages of the Russian Federation at the present stage). *Adaptatsiya narodov i kul'tur k izmeneniyam prirodnoi sredy, sotsial'nym i tekhnogennym transformatsiyam*. Moscow: ROSSPEN. 413–417. (In Russ.).

Vakhtin, Nikolai. 2006. The Language of Markovo old settlers: From a communication system to a demo version. In Dieter Stern & Christian Voss (eds.), *Marginal linguistic identities: Studies in Slavic contact and borderline varieties*, 135–147. Wiesbaden: Harrassowitz

Varshaver, Evgenii A. 2022. V lovushke dvoinoi irrelevantnosti: (vos)proizvodstvo ehtnichnosti vo vzai¬modeistviyakh mezhdu perepischikami i perepisyvaemymi v khode Vserossiiskoi perepisi 2021 g. v Dagestane (In the trap of double irrelevance: (Re)production of ethnicity in ¬interactions between census takers and census takers during the 2021 Russian Census in Dagestan). *Monitoring obshchestvennogo mneniya: ehkonomicheskie i sotsial'nye peremeny* 4. 199–221. https://doi.org/10.14515/monitoring.2022.4.2150 (In Russ.).

Vorob'ev, Nikolai Ya. 1957. *Vsesoyuznaya perepis' naseleniya 1926 g* (All-Union Population Census of 1926). Moskva: Gosstatizdat. (In Russ.).

Vorontsov, Vladimir S. 2020. Perepisi naseleniya kak istochnik ehtnicheskikh i yazykovykh kharakteristik naseleniya (Population censuses as a source of ethnic and linguistic characteristics of the population). *Istoriko-kul'turnoe nasledie narodov Uralo-Povolzh'ya* 2 (9). 93–98. (In Russ.).

Zakharov, Sergei V. & Anatolii G. Vishnevskij. 2010. Chto znaet i chego ne znaet rossiiskaya demograficheskaya statistika (What Russian demographic statistics know and what they do not know). *Voprosy statistiki* 2. 7–17. (In Russ.).

Zhiromskaya, Valentina B. & Gennadii E. Kornilov. 2020. K istorii perepisnogo dela v Rossii: vseobshchie perepisi naseleniya KHKh veka (On the history of the census business in Russia: general population censuses of the twentieth century). *Instituty razvitiya chelovecheskogo potentsiala v usloviyakh sovremennykh vyzovov*. Vol. 1. Ekaterinburg. 28–42. (In Russ.).

Other sources

Vsesoyuznaya perepis' naseleniya. 17 dekabrya 1926 g. Kratkie svodki. Vypusk IV. Narodnost' i rodnoi yazyk naseleniya SSSR. (All-Union Population Census. December 17, 1926. Brief summaries. Issue IV. Ethnicity and native language of the population of the USSR) Moscow: Central Statistical Agency of the USSR. (In Russ.)

Vserossiiskaya perepis' naseleniya 2020 goda. (Russian Population Census 2020). Rosstat. URL: https://rosstat.gov.ru/vpn/2020 (retrieved 03.04.2026)

Article history:
Received: 29 October 2025
Accepted: 12 February 2026

Bionotes:

Gulnara F. GABDRAKHMANOVA is a Doctor of Sociology and Head of the Ethnological Research Department at the Sh. Marjani Institute of History, Academy of Sciences, Republic of Tatarstan. Her research interests include ethnosocial processes, identity, ethnic, linguistic and cultural policies, interethnic relations, integration of multiethnic societies, migration, and ethnosociological research methods.
e-mail: medi54375@mail.ru
https://orcid.org/0000-0003-0735-5181

Nikolai B. VAKHTIN is a professor at the Department of Anthropology, the European University at St. Petersburg, a leading researcher at the Institute of Linguistic Studies of the Russian Academy of Sciences, Doctor of Philology, and a Corresponding Member of the Russian Academy of Sciences. His research interests include sociolinguistics, particularly the sociolinguistic study of the languages of the peoples of the North, as well as the history of Arctic anthropology as a branch of social anthropology.
e-mail: vakhtin@eu.spb.ru
https://orcid.org/0000-0002-4011-2141

Сведения об авторах:

Гульнара Фаатовна ГАБДРАХМАНОВА — доктор социологических наук, заведующая отделом этнологических исследований Института истории им. Ш. Марджани Академии наук Республики Татарстан. Сфера научных интересов — этносоциальные процессы, идентичность, национальная и языково-культурная политика, межэтнические отношения, интеграция полиэтничного общества, миграция, методы этносоциологических исследований.
e-mail: medi54375@mail.ru
https://orcid.org/0000-0003-0735-5181

Николай Борисович ВАХТИН — профессор факультета антропологии Европейского университета в Санкт-Петербурге, главный научный сотрудник Института лингвистических исследований РАН, доктор филологических наук, член-корр. РАН. Сфера научных интересов включает социолингвистику, прежде всего социолингвистическое изучение языков народов Севера, а также историю североведения как отрасли социальной антропологии.
e-mail: vakhtin@eu.spb.ru
https://orcid.org/0000-0002-4011-2141