



Literary Text Selection in Russian Émigré Education, 1918–1939: A Corpus-Based Study of School Anthologies

Mariia Kolenova

Lomonosov Moscow State University, Russian Federation

Abstract

The selection of literary texts for schooling offers a historical perspective on the relationship between language education and cultural continuity in diaspora communities. This study examines the coexistence of classics, folklore and contemporary literature in Russian-language educational publications produced abroad between 1918 and 1939. It draws on Daria Strizhkova's open dataset of the Russian émigré school literary canon [1], comprising 3,005 records of works and excerpts across 36 bibliographic units. The study provides a secondary quantitative analysis, distinguishing the original dataset from the author's analytical procedures. Author representation is measured through record counts, distinct title strings and presence across publications. To assess the influence of serial publication, the 36 bibliographic units are additionally consolidated into 20 series or projects. The results reveal a stable classical core: Pushkin appears in 23 bibliographic units, Tolstoy in 22, and Nekrasov and Turgenev in 20 each. However, the repertoire also regularly includes later writers, notably Balmont, Blok, Bunin and Sologub; their presence remains substantial after consolidation. A conservative classification identifies 273 anonymous tales, songs and byliny, distributed across 23 bibliographic units and 15 consolidated projects. These findings suggest selective cultural continuity: educational reading preserved a classical foundation while incorporating folklore and selected contemporary writing. The study contributes a reproducible approach to analysing literary selection in diaspora education and provides historical material for discussions of cultural content in language teaching. Its conclusions concern the printed repertoire; the corpus alone cannot establish classroom practices, learners' responses or language-learning outcomes.

Keywords: Russian émigré education; literary canon; school anthologies; cultural continuity; corpus analysis.

1. Introduction

Following the 1917 Revolution, Russian schools abroad operated under differing national regulations and material conditions, yet Russian language, literature and history remained central to the educational component known as Rossika [7, 8]. Unlike curricula, anthologies and readers document which authors and texts were actually selected for pupils.

Research on émigré schooling has largely examined institutions, educational ideas, periodicals and individual textbooks. Daria Strizhkova's machine-readable dataset permits a complementary approach: identifying recurring selections across 3,005 records rather than a handful of illustrative publications [1, 2]. The question is not simply whether schools “preserved culture”, but which texts persisted, which changed and how widely selections recurred. Reading in Russian could also provide access to written language in settings where another language often dominated public life. Heritage-language education offers a useful comparison, not a retrospective label for interwar pupils [12–14].

This article examines Russian-language educational publications of the first wave of emigration (1918–1939). It asks how classical literature, folklore and later writing were selected and combined in the émigré school literary canon. “Contemporary” is not treated as a single political or geographical category: émigré, Soviet-Russian and transnational trajectories are distinguished.

The analysis measures authors' presence at several levels, tests the effects of multi-part series, identifies a reproducible minimum folklore layer, records significant absences and compares the classical core with pre-revolutionary anthologies. It also considers the relevance—and limits—of heritage-language pedagogy.



The dates follow the dataset and end before the Second World War [1]. Its geographical coverage should not be mistaken for a complete map of émigré education: the Pushkin House announcement lists six countries, whereas the analysed spreadsheet also contains Prague publications [2].

The hypothesis is selective cultural continuity: an inherited classical core persisted but was reordered, folklore remained widespread, and later authors were selected without a simple émigré/Soviet divide. The persistence of these patterns after consolidating serial publications is tested separately.

The article's contribution is a secondary, multi-level analysis of an existing dataset, explicitly separating Strizhkova's metadata from the present aggregation, operational classification and historical interpretation.

2. Historiography and Research Perspective

Studies of pedagogical periodicals place émigré schooling within projects of linguistic and cultural continuity. Terenia analyses debates over school organisation and Rossika [7]; Gladysheva and Sedova examine the journal *Russian School Abroad* and its educational context [8, 9]. These studies establish why reading materials mattered but do not measure their distribution across textbooks.

Dimianenko's study of children's-book criticism demonstrates that reviewers explicitly debated suitability, audience and values [6]. Petkova and Barkovskaia show how P. M. Bitsilli's textbooks embodied individual decisions about literary selection [4, 5]. Such work illuminates particular editorial projects rather than corpus-wide patterns.

Vdovin and Leibov analysed 108 Russian Imperial anthologies and readers (1805–1912), showing that a small group of authors dominated textual inclusions [3]. Their later dataset provides a historical reference point, although it is not a statistically matched control [10].

In 46 anthologies dated 1843–1904, Vdovin also found that texts by recently published authors rose from 4% around 1860 to 35% by 1900: anthologies could mediate between conservative curricula and recent literature even before 1917 [11].

Heritage-language scholarship distinguishes learners with familial or community exposure to Russian from foreign-language learners and highlights literacy and cultural content [12, 13]. Kulibina similarly discusses literary texts in teaching bilingual children abroad [14]. This is an analytical parallel, not an assumption about interwar pupils' linguistic biographies.

Strizhkova's study of Russian modernism in the interwar émigré school canon examines authors and works in 27 educational books [15]. The present analysis instead tests series effects, operationalises folklore identification, compares pre-revolutionary practice and explores the pedagogical implications of the printed repertoire.

3. Data and Methods

The source is Daria Strizhkova's open dataset, *School Literary Canon of Emigration, 1918–1939* (version 1.0; DOI: 10.31860/openlit-2025.7-B020), hosted by the Pushkin House repository [1]. Its *bibliography.tab* contains 36 records and *contents.tab* contains 3,005 records; the sources were reportedly examined *de visu* [2].

Strizhkova defined the corpus to include works and excerpts in Russian-language textbooks, anthologies, readers and collections used for émigré educational reading [1, 2]. The observational unit is an entry for a work or excerpt within a publication, not a unique literary work.

The source metadata include author, text, genre, publication details and *biblio_id*, among other fields. Original values were retained; aggregation, sensitivity checks and additional classifications belong to the present study.

Four author-level measures were used: (1) number of records; (2) distinct text-field strings; (3) distinct *biblio_id* values; and (4) consolidated series/projects. Title normalisation standardised case, spacing, ë/e, quotation marks and dashes; it did not merge semantically equivalent titles or excerpts. Accordingly, distinct strings are not necessarily distinct works.

To avoid treating parts of a single publication series as independent books, 36 *biblio_id* values were conservatively consolidated into 20 projects. Only explicitly identified parts of the same named series were merged: Arronet (two), Bitsilli (two), Gudkov (five), Osmolovsky (five), Spevak (three), Paegle (three), and



Tikhonitsky, Andreeva and Maksimovich (three). Other units remained separate. This grouping is an analytical sensitivity test, not source metadata.

A record counted as folklore only if author = “б/а” (anonymous) and genre was “сказка” (tale), “песня” (song) or “былина” (bylina). Of 739 anonymous records, 273 met this rule. The other 466—including 369 without a genre label—and authored literary tales were excluded pending manual verification. This deliberately yields a lower bound.

The “later literary layer” comprises selected authors active around the turn of the twentieth century and its early decades, rather than an exhaustive category whose overall share can be calculated. Writers were distinguished by post-1917 trajectories: abroad, predominantly in Soviet Russia, or transnational. “Absent” means no author-field entry among these 3,005 records, not absence from émigré schooling generally.

For historical comparison, the study uses findings on 108 anthologies (1805–1912) and their published dataset [3, 10], together with research on 46 anthologies (1843–1904) [11]. Given differences in chronology, size and coding, only broad rank and structural comparisons are made.

All émigré counts were recalculated from the 3,005 records. Unreproducible indicative rankings in the working file were excluded. Genre is missing in 583 rows, and genre and kind_of_literature sometimes disagree; such uncertainties limit inference about compilers’ intentions.

4. Results

4.1. The Classical Core: Frequency and Distribution

Record counts reveal a classical core: Pushkin (193), Krylov (148), Tolstoy (126), Lermontov (108) and Nekrasov (102). Yet distribution across publications matters as well: Pushkin appears in 23 of 36 biblio_id values, Tolstoy in 22, and Nekrasov and Turgenev in 20 each (Table 1).

Table 1. Author representation in the original dataset and after series consolidation

Author	Records	Distinct text values	biblio_id	Series/projects (of 20)
A. S. Pushkin	193	73	23	14
I. A. Krylov	148	64	18	10
L. N. Tolstoy	126	81	22	13
M. Yu. Lermontov	108	39	18	13
N. A. Nekrasov	102	28	20	12
I. S. Turgenev	63	30	20	14
K. D. Balmont	54	30	19	12
N. V. Gogol	54	16	15	11
A. A. Blok	32	17	16	9
I. A. Bunin	19	15	14	11
F. K. Sologub	11	9	10	9

After series consolidation, Pushkin and Turgenev each occur in 14 of 20 projects, Tolstoy and Lermontov in 13, Nekrasov in 12 and Gogol in 11. Turgenev’s 63 records versus Pushkin’s 193 illustrate why frequency and breadth of distribution measure different dimensions of canonical presence.

4.2. The Folklore Layer

The minimum folklore layer comprises 273 records (9.1%): 117 songs, 112 tales and 44 byliny. They occur in 23 of 36 bibliographic units and 15 of 20 projects (Table 2), indicating distribution beyond a few anthologies.

Table 2. Conservative estimate of the folklore layer (author = ‘б/а’; genre ∈ {‘песня’, ‘сказка’, ‘былина’})

Folklore form	Records	Share of 3,005 (%)	biblio_id	Series/projects
---------------	---------	--------------------	-----------	-----------------



Song	117	3.9	—	—
Tale	112	3.7	—	—
Bylina	44	1.5	—	—
Total	273	9.1	23	15

Some folklore texts recur: “Ilya Muromets” appears 20 times, while “Za rekoiu, za bystroiu”, “Ne shumi, mati, zelenaiia dubrovushka”, “Morozko” and “Volga Sviatoslavovich i Mikula Selianinovich” appear five times each. Recurrence alone does not establish why compilers chose them.

The 9.1% figure is a lower bound: 466 anonymous records failed the formal filter, including 369 without genre labels and others marked as poems, stories or different genres. Including them without attribution would compromise reproducibility.

4.3. The Later Literary Layer: Emigration, Soviet Russia and Significant Absences

Later authors span different post-revolutionary trajectories (Table 3). Émigré-associated writers include Balmont (54 records; 19 bibliographic units; 12 projects), Bunin (19; 14; 11), Sasha Chorny (15; 10; 8), Shmelev (11; 5; 5) and Remizov (7; 6; 5). Writers predominantly based in Russia include Blok (32; 16; 9), Briusov (16; 10; 7), Sologub (11; 10; 9), Gorodetsky (6; 5; 5) and Esenin (6; 4; 3). Gorky (16; 12; 9) requires separate treatment because of his transnational trajectory.

Table 3. Representation of selected later authors by post-revolutionary trajectory

Group	Author	Records	biblio_id	Series/projects
Russian diaspora	K. D. Balmont	54	19	12
Russian diaspora	I. A. Bunin	19	14	11
Russian diaspora	Sasha Chorny	15	10	8
Russian diaspora	I. S. Shmelev	11	5	5
Russian diaspora	A. M. Remizov	7	6	5
Authors remaining in Russia after 1917	A. A. Blok	32	16	9
Authors remaining in Russia after 1917	V. Ya. Bryusov	16	10	7
Authors remaining in Russia after 1917	F. K. Sologub	11	10	9
Authors remaining in Russia after 1917	S. M. Gorodetsky	6	5	5
Authors remaining in Russia after 1917	S. A. Yesenin	6	4	3
Complex trajectory	M. Gorky	16	12	9

Thus “contemporary” cannot mean exclusively émigré: Balmont and Blok, or Bunin and Briusov, could coexist in educational publications. Residence outside Soviet Russia was not necessary for inclusion, while residence within it did not entail exclusion.

No author-field entries were found for Tsvetaeva or Nabokov/Sirin, nor for Mayakovsky, Akhmatova, Mandelstam, Pasternak or Khlebnikov. These significant absences apply only to this file and suggest selection more complex than a political-geographical divide.

Repeated titles also suggest selectivity within authors’ oeuvres: Balmont’s “K zime” occurs five times and “Osen” and “Za gribami” four times each; Blok’s “Verbochki” appears eight times and “Zaichik” four. Titles alone, however, cannot establish thematic or pedagogical motives.



4.4. The Pre-Revolutionary Reference Point

Vdovin and Leibov's 1805–1912 corpus comprises 108 anthologies and 10,473 textual inclusions; its five leading authors account for 43% [3, 10]. Their counts—Pushkin 1,577, Krylov 1,126, Zhukovsky 862, Lermontov 563 and Dmitriev 377—provide a historical rank reference, not directly comparable absolute totals (Table 4).

Table 4. The pre-revolutionary high-frequency core and the presence of the same authors in the émigré corpus

Author	1805–1912 occurrences [3]	1918–1939 records	biblio_id	Series/projects
A. S. Pushkin	1577	193	23	14
I. A. Krylov	1126	148	18	10
V. A. Zhukovsky	862	41	15	11
M. Yu. Lermontov	563	108	18	13
I. I. Dmitriev	377	7	3	3

Pushkin, Krylov and Lermontov remain prominent in the émigré corpus, supporting continuity. Yet Zhukovsky falls behind the leaders and Dmitriev has only seven records in three bibliographic units, while Tolstoy, Nekrasov and Turgenev occupy central positions. The broad 1805–1912 reference period cannot establish whether these shifts occurred before or after 1917.

The inclusion of Balmont, Blok and Bunin also extends an earlier anthology practice: recent literature had entered readers well before the Revolution [11]. What changed was the historical and linguistic setting, not necessarily the principle of including recent authors.

4.5. Corpus Geography and the Effect of Serial Publication

The corpus is geographically uneven: 20 of 36 bibliographic units were published in Riga, three each in Yuriev, Paris and Prague, two in Belgrade, and one each in Berlin, Reval, Tallinn, Chaville and Sofia. Riga's multi-part series further inflate its apparent weight; consolidation reduces but does not eliminate the imbalance.

The Pushkin House announcement names six countries, but the spreadsheet also includes three Prague units dated 1923–1924 [2]. Geographical counts therefore describe this file, not the full network of émigré schools.

5. Discussion

The pre-revolutionary comparison qualifies “continuity”: some nineteenth-century authors retained strong positions, but the internal hierarchy changed. The broad chronology of the reference corpus prevents assigning those changes specifically to the post-1917 period [3, 10, 11].

Folklore was also widely distributed, although its formal minimum count cannot be equated with its total cultural or pedagogical importance. In particular, anonymous entries without qualifying genre labels were not automatically treated as folklore.

The later repertoire resists an émigré-versus-Soviet binary. Age suitability, genre, topic, availability, editorial habit and wider cultural context may all have mattered; their relative influence cannot be established from record counts alone.

Heritage-language research foregrounds literacy, academic registers and cultural content when a family or community language coexists with a socially dominant language [12, 13]. Literary texts may support such teaching [12, 14]. Interwar émigré schools were historically different from present-day programmes, but their anthologies provided printed Russian-language reading material.

The three layers potentially served complementary functions: classics could offer established literary models, folklore recurring genres and formulae, and later authors texts closer to pupils' own period. These are interpretative possibilities, not measured language-learning outcomes.



The corpus identifies what was printed and repeated, not pupils' experience, classroom use or identity formation. Periodical research provides separate evidence of educational aims [6–9]; the present claim of cultural continuity concerns selection and reproduction of texts.

6. Limitations of the Study

The 36 bibliographic units are not an exhaustive sample of émigré educational publishing; absent authors cannot be assumed absent from émigré schooling as a whole.

Consolidation into 20 projects is a sensitivity check, not a definitive count of textbooks; alternative grouping rules could alter individual values.

Metadata are incomplete: 583 genre fields are blank, and genre labels sometimes conflict. Normalised title strings may still refer to the same work, while records vary from short poems to excerpts from longer prose. The folklore filter deliberately excludes 466 other anonymous entries, some potentially folkloric. A fuller count requires manual attribution and a normalised inventory.

The pre-revolutionary and émigré corpora differ in coverage, coding and size; they support structural-historical comparison, not pooled statistical inference.

Riga is overrepresented, and published descriptions do not perfectly match the file's geography. Cross-city comparisons require bibliographic verification and normalisation.

Heritage-language theory is used comparatively: the dataset contains no pupil-level evidence about home language, proficiency or language use.

Finally, inclusion in an anthology is not equivalent to classroom emphasis. Prefaces, curricula, reviews and testimony would be needed to reconstruct teaching practice and reception.

7. Conclusion

The 3,005 records reveal a durable but reconfigured classical core. Pushkin occurs in 23 of 36 bibliographic units, Tolstoy in 22, and Nekrasov and Turgenev in 20; after consolidation, Pushkin and Turgenev occur in 14 of 20 projects.

Comparison with nineteenth-century anthologies supports continuity for Pushkin, Krylov and Lermontov, while showing a different hierarchy for Zhukovsky, Dmitriev, Tolstoy, Nekrasov and Turgenev. The timing of those changes remains unresolved.

A conservative, reproducible rule identifies 273 anonymous folklore records across 23 bibliographic units and 15 projects. The true folklore total may be higher.

The later layer includes both émigré writers and authors linked to Soviet Russia, while prominent figures from each group are absent. Selection therefore cannot be explained by a simple political-geographical filter.

Four observable principles emerge: retention of part of the pre-revolutionary core, reordering within that core, broad folklore inclusion and selective incorporation of later writing. Together they support a qualified model of selective cultural continuity.

Heritage-language education provides a comparative question about the potential role of anthologies in providing access to written Russian abroad; it does not demonstrate particular classroom practices or learning outcomes.

Further research should normalise titles semantically, attribute anonymous materials, code themes and compare anthology contents with prefaces, curricula, educational periodicals and classroom testimony.

REFERENCES

- [1] Strizhkova, D. (2026). Shkol'nyi literaturnyi kanon emigratsii 1918–1939 gg. [The school literary canon of emigration, 1918–1939]. Open Data Repository for Russian Literature and Folklore, Institute of Russian Literature (Pushkin House), Russian Academy of Sciences, version 1.0. <https://doi.org/10.31860/openlit-2025.7-B020>. <https://dataverse.pushdom.ru/dataset.xhtml?persistentId=doi:10.31860/openlit-2025.7-B020> (accessed 11 October 2026).
- [2] Pushkin House (13 March 2026). Novyi dataset [New dataset]. <https://pushkinskijdom.ru/2026/03/13/novyj-dataset-8/> (accessed 11 October 2026).



- [3] Vdovin, A. V., and Leibov, R. G. (2013). 'Khrestomatiinye teksty: russkaia poeziia i shkol'naia praktika XIX stoletii' [Anthology texts: Russian poetry and school practice in the nineteenth century]. In A. V. Vdovin and R. G. Leibov (eds), *Khrestomatiinye teksty: russkaia pedagogicheskaia praktika i literaturnyi kanon XIX veka*. Tartu: University of Tartu Press, pp. 7–34.
- [4] Petkova, G. T. (2014). "Kratkaia istoriia russkoi literatury" P. Bitsilli — uchebnoe posobie, lapsus, avtorskii kanon? [P. Bitsilli's Brief History of Russian Literature: textbook, lapsus or authorial canon?]. *Filologicheskii klass*, 1(35), 69–77.
- [5] Barkovskaia, N. V. (2017). 'Iz istorii russkoi emigrantskoi shkoly v Bolgarii: "Uchebnaia trilogiia" Petra Bitsilli' [From the history of Russian émigré schools in Bulgaria: Petr Bitsilli's textbook trilogy]. *Filologicheskii klass*, 3(49), p. 139 (book review). <https://doi.org/10.26710/fk17-03-21>.
- [6] Dimianenko, A. (2016). 'Kritika detskoj literatury russkogo zarubezh'ia v 1920-e gg. na stranitsakh zhurnala "Russkaia shkola za rubezhom"' [Criticism of Russian émigré children's literature in the 1920s in The Russian School Abroad]. *Detskie chteniia*, 10(2), 161–169.
- [7] Terenia, Yu. Yu. (2014). 'Problematika pedagogicheskikh zhurnalov russkogo zarubezh'ia (1920–1940 gg.)' [Themes in Russian émigré educational journals, 1920–1940]. *Istoriko-pedagogicheskii zhurnal*, 2, 180–191.
- [8] Gladysheva, S. N. (2022). 'Problemy obrazovaniia i vospitaniia emigrantskikh detei v zhurnale "Russkaia shkola za rubezhom"' [The education and upbringing of émigré children in The Russian School Abroad]. *Zhurnal filologicheskikh issledovanii*, 2, 71–75.
- [9] Sedova, E. E. (2009). 'Zhurnal "Russkaia shkola za rubezhom" v sotsiokul'turnom i obrazovatel'nom prostranstve Evropy' [The Russian School Abroad in Europe's sociocultural and educational space]. In *Russkaia shkola za rubezhom* (Prague, 1923–1931, nos 1–34): index of contents. St Petersburg: Sudarynia, pp. 7–30.
- [10] Vdovin, A. V., Leibov, R. G., and Kazakova, E. (2021). *Khrestomatii Rossiiskoi Imperii s 1805 po 1912 gg.* [Anthologies of the Russian Empire, 1805–1912]. Open Data Repository for Russian Literature and Folklore, Institute of Russian Literature (Pushkin House), Russian Academy of Sciences. <https://doi.org/10.31860/openlit-2021.9-B003>. <https://dataverse.pushdom.ru/dataset.xhtml?persistentId=doi:10.31860/openlit-2021.9-B003> (accessed 11 October 2026).
- [11] Vdovin, A. V. (2020). 'Sovremennaia russkaia literatura v khrestomatiakh 1843–1904 godov i literaturnyi kanon' [Contemporary Russian Literature in Reading Anthologies (1843–1904) and the Literary Canon]. *Quaestio Rossica*, 8(1), 85–101. <https://doi.org/10.15826/qv.2020.1.449>.
- [12] Kagan, O., and Dillon, K. (2001). 'A New Perspective on Teaching Russian: Focus on the Heritage Learner'. *The Slavic and East European Journal*, 45(3), 507–518. <https://doi.org/10.2307/3086367>.
- [13] Kisselev, O., Dubinina, I., and Polinsky, M. (2020). 'Form-Focused Instruction in the Heritage Language Classroom: Toward Research-Informed Heritage Language Pedagogy'. *Frontiers in Education*, 5, article 53. <https://doi.org/10.3389/feduc.2020.00053>.
- [14] Kulibina, N. V. (2017). 'The Role of Literary Text in Teaching Bilingual Children Russian Language'. *RUDN Journal of Language Education and Translingual Practices*, 14(3), 431–435. <https://doi.org/10.22363/2312-8011-2017-14-3-431-435>.
- [15] Strizhkova, D. A. (2025). 'Russkii modernizm v shkol'nom literaturnom kanone mezhvoennoi emigratsii' [Russian modernism in the school literary canon of interwar emigration], abstract of final qualification thesis. National Research University Higher School of Economics. <https://www.hse.ru/edu/vkr/1048200681> (accessed 12 October 2026).