The Position of the Komi-Yaz'va Idiom in the Komi Languages Based on Phonetic, Lexical, and Morphological Features: Evidence from the LingvoDoc Platform

O.N. Bazhenova, ORCID: 0000-0002-5666-2260 <bazhenova-olga2011@mail.ru>

"Linguistic platforms" laboratory,
Ivannikov Institute for System Programming of the Russian Academy of Sciences,
25, Alexander Solzhenitsyn st., Moscow, 109004, Russia.

Abstract. The present article is devoted to determining the place of the Komi-Yaz'va idiom in the Komi dialect continuum. To this end, 8 dictionaries were analyzed using the programs of the linguistic platform LingvoDoc (search for cognates of languages / dialects, analysis of cognates of languages / dialects, glottochronology of languages / dialects, degree of morphological closeness between dialects / languages, and composite distance between languages). The dictionaries analyzed were Komi-Yaz'va, Upper Kama, Upper Sysola, Mysovsky and Kudymkarsky dialects. The comparative analysis of these dictionaries revealed that, in terms of phonetic, lexical, and morphological features, the Komi-Yaz'va idiom is significantly divergent from other Komi dialects. The degree of similarity with these dialects ranges from 86 to 88%, indicating that the Komi-Yaz'va is now considered a distinct language that has retained Proto-Komi language's features.

Keywords: Komi-Yaz'va; Komi-Zyrian; Komi-Permyak; audio dictionaries; morphological dictionaries; glottochronology; classification of languages and dialects; LingvoDoc.

Место коми-язьвинского в классификации языков коми на основании фонетических, лексических и морфологических особенностей с помощью программ платформы ЛингвоДок

О.Н. Баженова, ORCID: 0000-0002-5666-2260 <bazhenova-olga2011@mail.ru>

Институт системного программирования им. В.П. Иванникова РАН,
Лаборатория 16.2 "Лингвистические платформы",
Россия, 109004, г. Москва, ул. А. Солженицына, д. 25.

Аннотация. Статья посвящена определению места коми-язьвинского идиома в коми диалектном континууме. Для этой цели на лингвистической платформе LingvoDoc было обработано 8 коми словарей (аудиословари коми-язьвинского, верхнекамского, верхнесысольского, мысовского и кудымкарского диалектов, а также морфологические словари коми-язьвинского, коми-зырянского и коми-пермяцкого идиомов). Благодаря применению инструментов «поиск когнатов языков/диалектов», «анализ когнатов языков/диалектов», «глоттохронология языков/диалектов», «степень морфологической близости между диалектами/языками», «суммарный коэффициент различий между языками» было выявлено, что по фонетическим, лексическим и морфологическим признакам коми-язьвинский говор значительно отличается от других коми диалектов. Степень сходства с этими диалектами составляет от 86 до 88%, что свидетельствует о том, что коми-язьвинский в настоящее время можно рассматривать как отдельный язык, сохранивший черты пракоми языка.

Ключевые слова: коми-язьвинский; коми-зырянский; коми-пермяцкий; аудиословари; морфологические словари; глоттохронология; классификация языков и диалектов; LingvoDoc.

1. Introduction

The Komi-Yaz'va idiom is spoken by an ethnolocal group residing in the villages of Upper Yaz'va, Bychina, Parshakova, Arefina, Antipina, Van'kova, Talavol, and other settlements of the Krasnovishersky District in the north-eastern part of Perm' Krai (Russian Federation). A significant population of Russians living along the Vishera and Kama rivers creates a geographical and linguistic separation between the Komi-Yaz'va and the rest of the Komi dialect continuum. From the All-Union census of 1926 to the All-Russian census of 2010, the population of Komi-Yaz'vinians decreased by more than sevenfold (from 3,163 to 436) [1]. According to the observations of R. Bartens, at the turn of the millennium, only a few dozen people of the older generation spoke Komi-Yaz'va [2]. This information is confirmed by the expedition data of R. V. Gaidamashko and Ju. A. Shkuratok, who classified the Komi-Yaz'va idiom as a language on the verge of extinction in terms of viability on the UNESCO scale [1].

2. Background of the problem

The question of the status of the Komi-Yaz'va idiom has been ambiguously resolved throughout its nearly 140-year history of study. Researchers consider it as a dialect [3-6] or as a distinct language [7-10]. Finnish linguist R. Bartens uses the term "Jaźvan Komi päämurre" 'the Jaźva's Komi main dialect', along with Komi-Zyrian and Komi-Permyak [2].

The isolation from other Komi dialects played a pivotal role in the formation and consolidation of distinctive linguistic characteristics within the Komi-Yaz'va idiom, particularly within the domain of phonetics. V. I. Lytkin identifies four primary distinctions in the vocalism system of the Komi-Yaz'va dialect:

1) the unique vowel phonemes not found in other Komi dialects (ɵ, ÿ);
2) distinctive features in the pronunciation of certain phonemes, such as the labial pronunciation of non-rear vowels;
3) numerous specific regularities in the domain of vowels, including the absence of labialized vowels in suffixal syllables;
4) a distinctive system of stress (the dependence of the quality of vowels on the place they occupy in relation to stress) [5].

The aforementioned peculiarities of the vocalism of the Komi-Yaz'va idiom led V. I. Lytkin to conclude that this dialect stands further from the Komi-Permyak and Komi-Zyrian literary languages than the latter from each other [5].

Nowadays, the linguistic platform LingvoDoc has possibilities for the calculation of the proximity between languages and dialects based on their phonetic, lexical, and morphological characteristics. A study conducted by Yu. V. Normanskaya in 2020 examined the phonetic proximity of the Komi-Yaz'va idiom to the Komi-Permyak and Komi-Zyrian dialects. The research revealed that the Komi-Yaz'va idiom is significantly different from the modern dialects of the Komi-Zyrian and Komi-Permyak languages: "Consequently, Komi-Yaz'va is currently classified as a distinct language, exhibiting a closer resemblance to the old Komi-Permyak dictionaries recorded by P. S. Pallas and N. Ovchinnikov in the 18th century" [11].

With regard to the lexicon and morphology of the Komi-Yaz'va idiom, V. I. Lytkin asserts that it exhibits significant morphological and lexical affinities with Komi-Permyak dialects, and to a lesser extent with the Komi-Zyrian dialects [5].

3. Materials and methods

The objective of the present study is to use the tools of the LingvoDoc platform to ascertain the position of the Komi-Yaz'va idiom in the classification of Komi languages based on phonetic, lexical, and morphological features.

In the initial phase of the study, we created morphological dictionaries for the Komi-Yaz'va, the literary Komi-Permyak, and the literary Komi-Zyrian languages. These dictionaries were based on V. I. Lytkin's monograph [5], and the manual compiled by A. S. Lobanova [8]; for the Komi-Permyak language we used the textbooks [12], [13], and G. A. Nekrasova's monograph [14]. The morphological dictionary of the Komi-Zyrian language was compiled on the basis of the monograph [15].

Subsequently, the morphological dictionaries were uploaded to the LingvoDoc platform and connected via the Cognates field by etymological links (Fig. 1).

Fig. 1. A fragment of the morphological dictionary of the literary Komi-Zyrian language on lingvodoc.ispras.ru.

Then, the "Degree of Morphological Closeness of Languages" option was used to compare all etymologically related affixes. The percentage of matches between the lists of idioms was subsequently calculated. This process yielded a summary table that contained all related and unique affixes (Fig. 2).

4. Results

The results of the automatic calculation of the degree of morphological closeness of the Komi-Yaz'va idiom to the Komi-Zyrian and Komi-Permyak literary languages revealed, that in cases where several languages have an affix with the same meaning in the system of word and form formation, the percentage of coincidence between the above idioms is 100% (Fig. 2).

LingvoDoc presents all related morphological indicators in a table, in which affixes are arranged in alphabetical order (Fig. 3).

	1. Морфологический словарь коми-язьвинского идиома	2. Морфологический словарь литературного коми-зырянского языка	3. Морфологический словарь литературного коми-пермяцкого языка
1	n/a	-0.00 (100%)	-0.00 (100%)
2	-0.00 (100%)	n/a	-0.00 (100%)
3	-0.00 (100%)	-0.00 (100%)	n/a

Fig. 2. Percentage of matches among etymologically related affixes in the morphological dictionaries analyzed.

1. Морфологический словарь коми-язьвинского идиома	2. Морфологический словарь литературного коми-пермяцкого языка	3. Морфологический словарь литературного коми-зырянского языка
-ӟик; -ӟиг (COMP)	-жык (COMP)	-джык (COMP)
-ө (ACC)	-ӧ (ACC)	-ӧ (ACC)
-ө (ILL)	-ӧ (ILL)	-ӧ (ILL)
-ө (IMP.2PL)	-ӧ (IMP.2PL)	-ӧй (IMP.2PL)
-ө (POSS.1SG)	-ӧ (POSS.1SG)	-ӧй (POSS.1SG)
-ө (PRS.3SG)	-ӧ (PRS.3SG)	-ӧ (PRS.3SG)
-өӟ' (TERM)	-ӧдз (TERM)	-ӧдз (TERM)
-өд (ORD)	-ӧт (ORD)	-ӧд (ORD)
-өм (PTCP.PST)	-ӧм (PTCP.PST)	-ӧма (PTCP.PST)
-өмөн (CVB.SIM)	-ӧмӧн (CVB.SIM)	-ӧмӧн (CVB.SIM)
-өма (PST2)	-ӧм (PST2)	-ӧма (PST2)
-өн (INE)	-ын (INE)	-ын (INE)
-өнөс (PRS.3PL)	-ӧны (PRS.3PL)	-ӧны (PRS.3PL)
-өс (ACC)	-ӧс (ACC)	-ӧс (ACC)
-өсь (PL (прил, числ))	-ӧсь (PL (прил, числ))	-ӧсь (PL (прил, числ))
-өт (PROL)	-ӧт (PROL)	-ӧд; -ті (PROL)
-өшт (ATTEN)	-ышт (ATTEN)	-ышт (ATTEN)
-a (FUT.1SG)	-a (FUT.1SG)	-a (FUT.1SG)
-a (ILL)	-a (ILL)	-a (ILL)
-a (INE)	-a (INE)	-a (INE)
-a (PRS.1SG)	-a (PRS.1SG)	-a (PRS.1SG)

Fig.3. Fragment of a comparative morphological table of Komi idioms, automatically generated on the LingvoDoc platform.

4.1 Phonetics

As the comparative table on lingvodoc.ispras.ru shows, Komi-Yaz'va is more similar to Komi-Permyak (13 positions) than to Komi-Zyrian (1 position) in the phonetic arrangement of affixes.

In contrast to the phonetic forms observed in Komi-Zyrian, Komi-Yaz'va and Komi-Permyak exhibit distinctive characteristics, particularly the presence of phonetic forms with a deaf consonant:

KY. *-атѳ*, KP. *-атö*, KZ. *-ад* 'PRS / FUT.2PL',
KY. *-итѳ*, KP. *-итö*, KZ. *-ид* 'PST.2PL',
KY. *-кѳт*, KP. *-кöт*, KZ. *-кöд* 'COM',
KY. *-нит*, KP. *-ныт*, KZ. *-ныд* 'POSS.2PL',
KY. *-ит*, KP. *-ыт*, KZ. *-ыд* 'POSS.2SG',
KY. *-к*, KP. *-ик*, KZ. *-иг* 'CVB.SIM',
KY. *-ѳт*, KP. *-öт*, KZ. *-öд* 'ORD',
KY. *-ѳт*, KP. *-öт*, KZ. *-öд* 'PROL'.

4.2 Morphology

While the Komi-Zyrian and Komi-Permyak dialects have witnessed the emergence of a substantial group of approximate local cases, resulting from the combination of the suffix of the approximative (*-лань*) with suffixes of other local cases, the Komi-Yaz'va dialect exhibits an such cases are absent in its texts. A similar trend of reduced usage of the approximative and substitution by afterwords has been observed in the Vym' and Izhma dialects of the Komi-Zyrian language. According to G. A. Nekrasova's study, the approximative is considered to be a highly infrequent component of the modern Udmurt literary language, with its semantics typically conveyed through adverbs and afterwords [14].

Contrary to the Komi-Permyak dialects, Komi-Yaz'va and Komi-Zyrian languages do not exhibit morphologized postpositions with *-выв-* and *-дін-* bases. The increase in the number of cases resulting from the agglutination of afterwords with the base *-дін-* 'at, near' is also present in certain northern dialects of the Udmurt language [16, 17]. As demonstrated in G. A. Nekrasova's research, the surface-local and approximate-local cases of Komi-Permyak are the result of an innovation that occurred between the late 19th and mid-20th century [18].

The Komi-Permyak language is distinguished by the presence of the comparative (*-ся*), which is also recorded in the southern dialects of the Komi-Zyrian language. In Finno-Ugric studies, the comparative is considered to be a southern Komi innovation [14]. In the Komi-Yaz'va idiom, there is no separate case indicator for expressing comparison; most likely, this meaning is conveyed by means of the elative (as in Komi-Zyrian) and with afterword *кин'а*: *сотн'а кин'а уна̌жик мѳнти* '(I) paid more than a hundred' [5].

In the Komi-Permyak language there is a typologically uncommon altiterminative case (*-ви*) employed to denote the vertical boundary. G. A. Nekrasova posits that the suffix *-ви* emerged as a consequence of the apocopy of two subsequent words: *виын* 'at' and *виö* 'to' [14].

4.3 Glottochronology

In order to facilitate the analysis of differences in the basic vocabulary of languages and dialects, the LingvoDoc platform provides a tool "Glottochronological analysis of dialects / languages". It can be used for any set of languages with dictionaries containing more than 50 words from M. Swadesh's word list [19]. This tool is based on a formula developed by S. A. Starostin [20], which has been incorporated into the LingvoDoc program for calculating the proximity of languages and dialects. The algorithm first removes borrowings from a 100-word list of basic vocabulary, then connects related words by etymological links and calculates the percentage of matches between lists of idioms [11].

The glottochronological analysis was conducted using materials from five Komi audio dictionaries: the Komi-Yaz'va idiom, the Upper Kama dialect, the Upper Sysola dialect of the Komi-Zyrian language, the Mysovsky and the Kudymkar dialects of the Komi-Permyak language. Finally, a summary table was generated, which displays the percentage of common lexicon in the idioms under consideration (Table 1).

Table 1. Percentage of shared basic vocabulary among Komi idioms..

	KY.	Mys.	Kud.	UK.	US.
KY.	---	95%	96%	90%	91%
Mys.	95%	---	98%	92%	95%
Kud.	96%	98%	---	96%	95%
UK.	90%	92%	96%	---	92%
US.	91%	95%	95%	92%	---

The Komi-Yaz'va idiom exhibits the highest degree of lexical similarity with the dialects of **the Komi-Permyak language**, reaching 95–96% (in bold in Table I). Conversely, it exhibits the lowest degree of similarity with the Upper Kama dialect, at 90%, which is due to the fact that the latter contains some words recorded also in the Komi-Zyrian dialects, e.g.: 'white': UK., US., LL. *jed͡ʒud* < ComP. **jeǯ-* 'unripe, white, green' [21]; KY. *t͡ɕótɕɛkʌm*, Mys. *t͡ɕótɕɛkəm*, Kud. *t͡ɕótɕɛkom* < ComK. **čọčkem* [21]; 'mountain': UK., LL. *t͡ɕoj* < ComK. **čọj* [21]; KY. *kerʌs*, Mys., Kud., US. *kerəs* < *ʞер-* 'cut down (wood)' [21].

Despite the fact that the majority of lexemes from the 100-word list in the Komi-Yaz'va idiom have correspondences in Komi-Permyak dialects (95–96%), it also contains word forms that are not recorded in any other dictionaries on Komi dialects, but have parallels in the Udmurtian language: KY. *døl* 'all', Udm. *долаʞ* 'all' [22]; KY. *dʌlót* 'rolled road', Udm. **дол** *вольыт сюрес* 'rolled road' [22]. The verb *vérpʌ* 'to say' used in the Komi-Yaz'va idiom is also recorded in the old Permian monuments (*verpθ*) and in the Izhma dialect of the Komi-Zyrian language (*вӧйпны*) [21]. In other Komi dialects, this word has been supplanted by the verb *шуны* 'to say'.

4.4 Composite distance between languages

The calculation of phonetic, morphological, and lexical proximity between the Komi idioms was performed by "Get composite distance between languages" tool at LingvoDoc. The final matrix (Table 2) illustrates the extent of Komi-Yaz'va deviation from other Komi dialects, with a range of **86% to 88%** similarity. Notably, the Upper Kama dialect, which is also isolated, exhibits the highest degree of similarity with the Upper Sysola dialect of the Komi-Zyrian language, reaching 95% similarity.

Table 2. Matrix of inter-dialectal similarity among Komi dialects in terms of phonetic, lexical, and morphological features.

	KY.	UK.	Mys.	Kud.	US.
KY.	---	87%	88%	86%	88%
UK.	87%	hh---	93%	94%	95%
Mys.	88%	93%	---	97%	95%
Kud.	86%	94%	97%	---	95%
US.	88%	95%	95%	95%	---

5. Conclusions

After processing 8 dictionaries of the Komi language on LingvoDoc, it was determined that the Komi-Yaz'va idiom exhibits significant differences from other Komi dialects in terms of phonetic, lexical, and morphological peculiarities (86–88% percent of the similarity). Concurrently, the percentage of

similarity between the other Komi dialects under consideration ranges from 93% to 97%. These figures permit the conclusion that Komi-Yaz'va can be classified as a distinct language that has preserved archaic features of the Proto-Komi language.

Abbreviations

- ComK.   – Common Komi
- ComP.   – Common Permian
- Kud.   – Kudymkar dialect of the Komi-Permyak language
- KY   – the Komi-Yaz'va idiom
- LL   – Luza-Letka dialect of the Komi-Zyrian language
- Mys.   – Mysovsky dialect of the Komi-Permyak language
- PK   – Komi-Permyak dialects
- Udm.   – Udmurtian
- UK   – the Upper Kama dialect
- US   – the Upper Sysola dialect of the Komi-Zyrian language

References

[1]. R. V. Gaydamashko, Yu. A. Shkuratok, "Modern state of the Komi-Yaz'va language", Yearbook of Finno-Ugric Studies, no. 4, pp. 576–590, 2019.

[2]. R. Bartens, "Permiläisten kielten rakenne ja kehitys", Helsinki, 2000.

[3]. A. Genetz, "Ost-permische Sprachstudien". Helsingfors, 1897.

[4]. M. Kövesi, "A permi nyelvek ősi képzői", Budapest, 1965.

[5]. V. I. Lytkin, "The Komi-Yaz'va dialect", Moscow, 1961.

[6]. Fundamentals of Finno-Ugric linguistics: Mari, Perm and Ugrian languages, Moscow, 1976.

[7]. V. Kel'makov, "Experience of creating a written language for the Komi-Yazvinians (A. L. Parshakova, Komi-Yazvin Primer. Educational edition, Perm Book publ. house, 2003. 135 p.)", Linguistica Uralica, vol. XL, no. 2, pp. 135–147, 2004.

[8]. Materials on the study of the case system of the Komi-Yaz'va language. Compiled by A.S. Lobanova, Perm', 2005.

[9]. V. V. Napol'skikh, "Introduction to historical Uralistics", Izhevsk, 1997.

[10]. G. N. Chagin, "Komi-Yazvin Permyaks are ancient people of the Northern Urals", Krasnovishersk; Solikamsk, 2002.

[11]. Yu. V. Normanskaya, "Is the Komi-Yaz'va a dialect of Komi-Permyak or a separate language?", Yearbook of Finno-Ugric Studies, vol. 14, no. 4, pp. 628–641, 2020.

[12]. Komi-Permyak language. Introduction, phonetics, lexicon and morphology: textbook for universities, Kudymkar, 1962.

[13]. E. A. Tsypanov, "Komi-Permyak literary language", Syktyvkar, 1999.

[14]. G. A. Nekrasova, "The category of noun case in Permian languages", Syktyvkar, 2023.

[15]. Modern Komi language. Morphology, edited by G.V. Fedyuneva. Syktyvkar, 2000.

[16]. S. A. Maksimov, "About secondary spatial cases in the Udmurt language", Problems of Udmurt and Finno-Ugric philology: Interuniversity collection of scientific works, Izhevsk, pp. 193–208, 1999.

[17]. N. M. Lyukina, "Phonetic and morphological peculiarities of the language of Lekminsky and Yundinsky Besermyans", Izhevsk, 2016.

[18]. G. A. Nekrasova, "Changes in the case system of the Komi-Permyak language of the XIX-XX centuries (based on written sources)", Yearbook of Finno-Ugric Studies, vol. 10, no. 3, pp. 7–16, 2016.

[19]. A. Kassian, G. Starostin, A. Dybo and V. Chernov, "The Swadesh wordlist. An attempt at semantic specification", Journal of Language Relationship, no. 4, pp. 46–89, 2010.

[20]. S. A. Starostin, "Comparative-historical linguistics and lexicostatistics", Linguistic reconstruction and ancient history of the East, P. 1, pp. 3–39, 1989.

[21]. V. I. Lytkin and E. S. Gulyaev, "Brief etymological dictionary of the Komi language", Syktyvkar, 1999.

[22]. Udmurt-Russian dictionary, edited by L.E. Kirillova. Izhevsk, 2008.

Информация об авторах / Information about authors

Ольга Николаевна БАЖЕНОВА – кандидат филологических наук, младший научный сотрудник Лаборатории 16.2 «Лингвистические платформы» Института системного программирования РАН, научный сотрудник сектора языка Института языка, литературы и истории ФИЦ Коми научного центра УрО РАН. Сфера научных интересов: коми диалекты, компаративистика, история коми-зырянского литературного языка, стилистика коми языка.

Olga Nikolajevna BAZHENOVA – Cand. Sci. (Philology), junior researcher of Laboratory 16.2 "Linguistic Platforms" of the Institute of System Programming of the Russian Academy of Sciences, researcher of the Language Department of the Institute of Language, Literature and History of the Komi Scientific Center of the Ural Branch of the Russian Academy of Sciences. Research interests: Komi dialects, comparative linguistics, history of the Komi-Zyrian literary language, stylistics of the Komi language.