

Генерация именных словоформ южнолюдиковского диалекта

А. П. Родионова

*Институт языка литературы и истории
Карельского научного центра РАН,
г. Петрозаводск, Российская Федерация,
santrar@krc.karelia.ru*

Н. Б. Крижановская

*Институт прикладных математических исследований
Карельского научного центра РАН,
г. Петрозаводск, Российская Федерация,
nataly@krc.karelia.ru*

АННОТАЦИЯ

Введение. Людиковское наречие оказалось на периферии ревитализационных процессов: в России на нём пока не разработано ни одного академического словаря (по сравнению с ливвиковским наречием), что влияет на невысокие темпы наполняемости людиковского подкорпуса в ВепКар.

Центральной частью нашей работы являются разработанные правила генерации именных словоформ. Приводится обоснование выбора южнолюдиковского диалекта для создания правил. В подготовленных таблицах демонстрируются примеры парадигмы именного словоизменения людиковского наречия для одноосновных и двуосновных слов. Формализованы словоизменительные типы имён южнолюдиковского диалекта. Предложены правила выделения основ, которые нужны для последующей генерации именных словоформ святозерского диалекта.

Описана программа, являющаяся частью корпуса ВепКар, в которой запрограммированы эти правила. Программа позволяет ускорить генерацию словоформ в словарных статьях для последующей разметки текста.

Цель: проиллюстрировать правила для автоматической генерации словоформ, подготовленные при помощи списка основ именных частей речи людиковского наречия карельского языка.

Материалы исследования: леммы и словоформы из Открытого корпуса вепского и карельского языков (ВепКар).

Результаты и научная новизна. На основе изученных теоретических источников исследователям удалось выявить грамматические закономерности. В ходе проведённых экспериментов был сформирован список основ и псевдооснов именного словоизменения людиковского наречия карельского языка (южнолюдиковский диалект), разработана система правил генерации словоформ, подготовлена и апробирована соответствующая программа. Научная новизна исследования заключается в разработке системы единых правил автоматической генерации словоформ для людиковского наречия карельского языка, которая осуществляется впервые.

Ключевые слова: карельский язык, людиковское наречие, южнолюдиковский диалект, корпусная лингвистика, морфология, именное словоизменение, генерация словоформ

Благодарности: Исследование проведено в рамках выполнения государственного задания КарНЦ РАН. Родионова А. П. проводила исследования в рамках темы НИР № 124022000089-4 «Прибалтийско-финские языки Северо-Запада России в условиях цифровизации научных знаний»; Крижановская Н. Б. проводила исследования в рамках темы НИР № FMEN-2021-0015 «Математические модели и методы исследования и применения информационно-вычислительных систем и сетей».

Для цитирования: Родионова А. П., Крижановская Н. Б. Генерация именных словоформ южнолюдиковского диалекта // Вестник угроведения. 2024. Т. 14. № 3 (58). С. 476–488.

Generation of nominal word forms of the South Ludic dialect

A. P. Rodionova

*Institute of Linguistics, Literature and History,
Karelian Research Center of the Russian Academy of Sciences,
Petrozavodsk, Russian Federation,
santrar@krc.karelia.ru*

N. B. Krizhanovskaya

*Institute of Applied Mathematical Research,
Karelian Research Centre of the Russian Academy of Sciences,
Petrozavodsk, Russian Federation,
nataly@krc.karelia.ru*

ABSTRACT

Introduction: the Ludic dialect has practically found itself on the periphery of the revitalization processes. There are no academics dictionaries on this dialect in Russia (compared to the Livvi dialect), which affects the low rate of filling of the Ludic sub-corpus in the VepKar.

The central part of our work is the developed rules for generating nominal word forms. The rationale for choosing the South Ludic dialect for creating rules is given. The prepared tables demonstrate examples of the Ludic dialect's nominal inflection paradigm for single-stem and double-stem words. Inflectional types of names of the South Ludic dialect are formalized. Rules for identifying stems that are needed for the subsequent generation of nominal word forms of the dialect of Svyatozero are proposed.

A program is described, which is part of the VepKar corpus, in which these rules are programmed. The program allows us to speed up the generation of word forms in dictionary entries for subsequent text markup.

Objective: to illustrate the rules for the automatic generation of word forms, prepared using a list of stems of nominal parts of speech of the Ludic dialect of the Karelian language.

Research materials: lemmas and word forms from the Open Corpus of the Vepsian and Karelian languages (VepKar).

Results and novelty of the research: based on the studied theoretical sources, the researchers were able to identify grammatical patterns. In the course of the experiments carried out in the study, a list of the stems and pseudo-stems of the nominal inflection of the Ludic dialect of the Karelian language (Southern Ludic dialect) was formed; a system of rules for generating word forms was developed; a corresponding program was prepared and tested. The scientific novelty of the research lies in the development of a system of unified rules for the automatic generation of word forms for the Ludic dialect of the Karelian language, which is being implemented for the first time.

Key words: Karelian language, Ludic dialect, Southern Ludic dialect, corpus linguistics, morphology, nominal inflection, generation of word forms

Acknowledgments: the study is carried out under the state order of the Karelian Research Centre of the Russian Academy of Sciences. A. P. Rodionova conducted research within the framework of the research work topic "Baltic-Finnish languages of the North-West of Russia in the conditions of digitalization of scientific knowledge". State registration number: 124022000089-4. N. B. Krizhanovskaya conducted research within the framework of the research work topic "Mathematical models and methods for research and application of information computing systems and networks", state registration number: FMEN-2021-0015.

For citation: Rodionova A. P., Krizhanovskaya N. B. Generation of nominal word forms of the South Ludic dialect // Vestnik ugrovedenia = Bulletin of Ugric Studies. 2024; 14 (3/58): 476–488.

Введение

Корпусная лингвистика в Институте языка литературы и истории (ИЯЛИ) развивается с 2009 г. Благодаря тесной и плодотворной работе лингвистов с исследователями Института прикладных математических исследований (ИПМИ) в течение 15 лет удалось не только создать уникальный электронный ресурс – Открытый корпус вепсского и карельского языков (ВепКар), но и сделать его основной базой для исследования прибалтийско-финских языков Северо-Запада России [22, 30]. Корпус ВепКар¹ содержит большое количество инструментов, позволяющих успешно использовать этот электронный ресурс. В настоящее время

в корпусе размещено более 6 тыс. карельских и вепсских текстов с 2 млн. слов, словарь корпуса включает в себя более 68 тыс. словарных статей.

ВепКар представлен 30 собственно карельскими, 11 ливвиковскими, 4 людиковскими и 4 вепскими диалектами из традиционных классификаций. Кроме этого в корпус включены пять письменных норм: севернокарельская, тверская, ливвиковская, людиковская и вепсская. В 2019 г. в корпусе была разработана и добавлена функция генерации полной парадигмы по базовым словоформам вепсского языка, а чуть позже, в 2020 г. для новописьменных вариантов карельского языка – собственного

¹ см. <http://dictorpus.krc.karelia.ru>

карельского и ливвиковского [12, 681]. Разработанные программы помогли ускорить работу по разметке текстов и пополнению словарей вышеперечисленных карельских подкорпусов.

Процесс возрождения карельского языка, который начался в конце 1980-х гг., практически не затронул людиковского наречия. Если более тридцати пяти лет назад письменность на собственно карельском и ливвиковском наречиях начала активно развиваться, то на людиковском наречии издавалась только духовная и художественная литература. Таким образом, можно констатировать, что людиковское наречие практически оказалось на периферии ревитализационных процессов и, как следствие, в настоящее время исследовано наименее детально [9, 83]. Особый интерес представляют работы по фонологии [1; 6; 18; 23; 24], отдельным разделам морфологии [6; 7; 9; 10; 11; 18] людиковского наречия карельского языка. Попытки развития письменной формы людиковского наречия осуществлялись на основе михайловского [21] и святозерского (южнолюдиковского) диалектов [6, 61; 13, 48–50].

К сожалению, людиковское наречие находится под угрозой полного исчезновения. Согласно последним оценочным данным, число говорящих на родном людиковском наречии может не превышать 150 человек¹. Главными причинами быстрого угасания людиковского наречия карельского языка, кроме малочисленности его носителей, являются отсутствие его преподавания в учебных заведениях (школах, ВУЗах), а также нехватка специалистов, владеющих им, в образовании, культуре, СМИ и других официальных сферах Республики Карелия [8, 84; 20, 24]. Именно поэтому разработка правил генерации является исключительно важным шагом для создания грамматических норм людиковского наречия карельского языка.

Материалы и методы

Исследование проведено на базе материалов электронного словаря корпуса ВепКар на людиковском наречии карельского языка. В настоящий момент в южнолюдиковский

подкорпус внесено около 2000 лемм из «Сопоставительно-ономасиологического словаря диалектов карельского вепского, саамского языков» [17] и словаря М. Д. Георгиевского [16], что позволило исследователям провести эксперимент.

В качестве методов были использованы: метод словообразовательного анализа, сравнительный и статистический, что позволило определить алгоритм разработки системы правил автоматической генерации словоформ.

Результаты

Людиковское наречие в Открытом корпусе вепского и карельского языков

Людиковское наречие распространено в ряде деревень и посёлков юго-восточной части Республики Карелия в Олонецком, Пряжинском, Прионежском и Кондопожском районах. В российской науке исследователи традиционно классифицируют людиковские диалекты следующим образом: севернолюдиковский (кондопожский), среднелюдиковский (мунозерский), южнолюдиковский (святозерский) и михайловский, которые характеризуются своими языковыми особенностями [5, 27]. Людиковские диалекты представлены в ВепКаре согласно этой классификации.

В настоящий момент в корпусе размещено 326 людиковских текстов (рис. 1), подавляющее большинство из которых это – диалектные тексты (267). Людиковские диалектные тексты были включены в ВепКар из опубликованных образцов речи, и представляют собой расшифровки магнитофонных записей, передают при этом естественную речь информанта со всеми её индивидуальными и общеязыковыми особенностями [2; 19]. В последние годы людиковский подкорпус стал пополняться публицистическими текстами, опубликованными в изданиях издательства «Периодика», в корпусе их 58. Все тексты корпуса даны с подробной паспортизацией, где указаны диалект, данные информанта (имя, год рождения, место рождения), год и место записи, данные исследователя, записавшего текст, что крайне необходимо для работы диалектологов и других специалистов.

¹ Малые языки России (iling-ran.ru): <https://minlang.iling-ran.ru/lang/lyudikovskoe-narechie-karelskogo-yazyka>

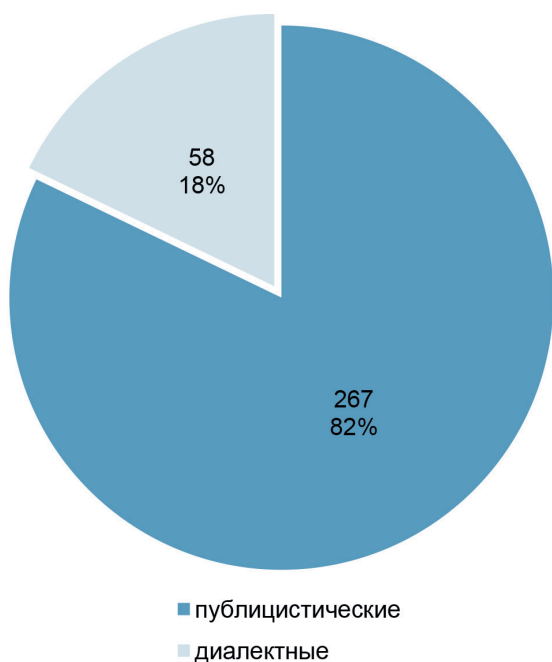


Рис. 1. Соотношение лудиковских текстов по источникам формирования в корпусе ВепКар

ВепКар также включает в себя электронные словари на всех основных наречиях карельского языка и на вепском языке. Лудиковский электронный лексикон насчитывает более 6300 лемм, и он в разы меньше, чем электронные словари на ливвиковском и собственно карельском наречиях. Это связано с тем, что, например, в электронный ливвиковский словарь включены материалы двух больших академических словарей [3; 4]. К большому сожалению, на лудиковском наречии подобных академических словарей в России до сих пор не издавалось. В связи с этим, было принято решение включить в лудиковский электронный словарь материалы «Сопоставительно-ономазиологического словаря диалектов карельского вепского, саамского языков» [16]. Уникальность данного издания заключается, прежде всего, в том, что оно отражает состояние диалектов того времени, так как сбор материала для словаря осуществлялся в полевых условиях в 1979–1981 гг. и количество носителей в то время было больше, нежели сейчас.

Лудиковское наречие в словаре представляют четыре говора населённых пунктов: д. Койкары и с. Галлезеро Кондопожского района,

с. Михайловское Олонецкого района и с. Святозеро Пряжинского района [16]. Однако, в настоящее время в ВепКар диалект д. Койкары не учитывается как лудиковский, а относится к мяндусельгскому диалекту собственно карельского наречия¹.

В ВепКар на диалектах с. Михайловское и д. Галлезеро представлено чуть более 1600 лемм, количество лемм на южнолудиковском (святозерском) диалекте составляет более 1800 лемм (рис. 2), поскольку в него были также включены материалы словаря М. Д. Георгиевского [17].

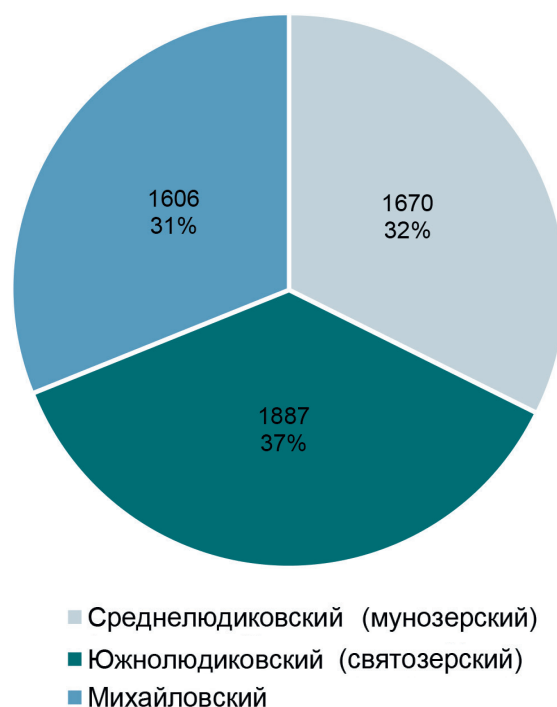


Рис. 2. Соотношение диалектных слов в лудиковском словаре ВепКар

Генерация именных словоформ южнолудиковского диалекта

Святозерский диалект в российской науке относят к южнолудиковскому [15, 8]. К особенностям этого диалекта следует отнести: 1) переход первичных конечных гласных *a, ä* в *e* в двусложных словах: *adre* ‘соха’, *päive* ‘день’, *akke* ‘женщина, баба’ (ср. с-к. *adra, päivä, akka*, мхл. *adr, päiv, akk*) [24, 151–154]; 2) отсутствие качественного вида чередования согласных: *lugeda / lugen* ‘читать / читаю’ [6, 121–125]; 3) явление трифтонгизации, т. е. перехода

¹ <http://dictorpus.krc.karelia.ru/ru/dict/dialect>

стяжённых дифтонгов на *i* в трифтонги (**ai* > *uai*, *äi* > *iäi*, *oi* > *uoi*, *ei* > *iei*, *öi* > *yöi*) в середине и на конце слова, в формах множественного числа именного словоизменения односложных имён, а также формах имперфекта индикатива и кондиционала односложных глаголов, напр. свт.: *rebuoi*, *saniui* ‘он сказал’) [6, 69].

Для разработки правил генерации людиковского наречия южнолюдиковский диалект был выбран не случайно. В Архиве КарНЦ РАН хранится рукопись монографии «Людиковский диалект карельского языка (фонетика и морфология Святозерского ареала), общим объёмом 9 а. л., подготовленная известным исследователем людиковского наречия карельского языка А. П. Баранцевым¹. Рукопись представляет

собой монографическое синхронное описание фонетики и морфологии людиковского наречия карельского языка на материале южнолюдиковского (святозерского) говора. В монографии уделено внимание, в том числе, именным категориям, их падежной парадигме, степеням сравнения, категории числа в карельско-людиковской диалектной речи и др.

Людиковскому наречию присуща вся падежная парадигма, свойственная карельскому языку. Основные различия между наречиями и диалектами проявляются, главным образом, в оформлении грамматических показателей и наличием в ней падежей. В людиковском наречии исследователями насчитывается до 19 продуктивных падежей [14, 285] (табл. 1).

Таблица 1

Парадигма именного словоизменения людиковского наречия карельского языка

Падеж	Одноосновное слово		Двуосновное слово	
	Ед. ч.	Мн. ч.	Ед. ч.	Мн. ч.
номинатив	<i>kirikkö</i> ‘церковь’	<i>kirikö-d</i> , <i>kirikö-t</i>	<i>alaine</i> ‘рукавица’	<i>alaiže-t</i> , <i>alaiže-d</i>
аккузатив	<i>kirikkö</i> , <i>kirikö-n</i>	<i>kirikö-d</i> , <i>kirikö-t</i>	<i>alaine</i> , <i>alaiže-n</i>	<i>alaiže-t</i> , <i>alaiže-d</i>
генитив	<i>kirikö-n</i>	<i>kirikyöi-den</i> , <i>kiriköi-den</i>	<i>alaiže-n</i>	<i>alaiži-den</i>
партитив	<i>kirikkö-d</i>	<i>kirikkyöi-d</i> , <i>kirikköi-d</i>	<i>alaš-te</i>	<i>alaiži-d</i>
эссив	<i>kirikö-n</i> , <i>kirikö-nny</i>	<i>kirikyöi-n</i> , <i>kirikyöi-nny</i> , <i>kiriköi-n</i> , <i>kiriköi-nny</i>	<i>alaiže-n</i> , <i>alaiže-nnu</i>	<i>alaiži-n</i> , <i>alaiži-nnu</i>
транслатив	<i>kirikö-ks</i> , <i>kirikö-kse</i>	<i>kirikyöi-ks</i> , <i>kirikyöi-kse</i> , <i>kirikö-ks</i> , <i>kiriköi-kse</i>	<i>alaiže-ks</i> , <i>alaiže-kse</i>	<i>alaiži-ks</i> , <i>alaiži-kse</i>
абессив	<i>kirikö-tä</i>	<i>kirikyöi-tä</i> , <i>kiriköi-tä</i>	<i>alaiže-ta</i>	<i>alaiži-ta</i>
инессив	<i>kirikö-s</i>	<i>kirikyöi-s</i> , <i>kiriköi-s</i>	<i>alaiže-s</i>	<i>alaiži-s</i>
элатив	<i>kirikö-s</i> , <i>kirikö-spiäi</i>	<i>kirikyöi-s</i> , <i>kirikyöi-spiäi</i> , <i>kiriköi-s</i> , <i>kiriköi-spiäi</i>	<i>alaiže-s</i> , <i>alaiže-spiäi</i>	<i>alaiži-s</i> , <i>alaiži-spiäi</i>
иллатив	<i>kirikkö-h</i>	<i>kirikkyöi-h</i> , <i>kirikköi-h</i>	<i>alaiže-h</i>	<i>alaiži-h</i>
адессив	<i>kirikö-l</i>	<i>kirikyöi-l</i> , <i>kiriköi-l</i>	<i>alaiže-l</i>	<i>alaiži-l</i>

¹ Научный архив КарНЦ РАН ф.1 оп. 6. д. 164: Баранцев А. П. Людиковский диалект карельского языка. Фонетика и морфология Святозерского ареала (рукопись) / А. П. Баранцев. Петрозаводск, 1985. 207 с.

аблатив	kirikö-l, kirikö-lpiäi	kirikyöi-l, kirikyöi-lpiäi, kiriköi-l, kiriköi-lpiäi	alaiže-l, alaiže-lpiäi	alaiži-l, alaiži-lpiäi
аллатив	kirikö-l, kirikö-le	kirikyöi-l, kirikyöi-le, kiriköi-l, kiriköi-le	alaiže-l, alaiže-le	alaiži-l, alaiži-le
комитатив	kirikö-nke	kirikyöi-neh, kiriköi-neh	alaiže-nke	alaiži-neh
пролатив	kirikö-či	kirikyöi-či, kiriköi-či	alaiže-či	alaiži-či
инструктив	-	kirikyö-in, kirikö-in	-	alaiži-n
аппроксиматив	kirikö-llyö, kirikö-llöh, kirikö-llöö	kirikyöi-llyö, kirikyöi-llöh, kirikyöi-llöö, kiriköi-llyö, kiriköi-llöh, kiriköi-llöö	alaiže-lloh, alaiže-lloo, alaiže-lluo	alaiži-lloh, alaiži-lloo, alaiži-lluo
терминатив	kirikkö-ssuai	kirikkyöi-ssuai, kirikköi-ssuai	alaiže-ssuai	alaiži-ssuai
адитив	kirikkö-hpiäi	kiriikkyöi-hpiäi, kirikköi-hpiäi	alaiže-hpiäi	alaiži-hpiäi

При разработке правил генерации людиковских именных словоформ была использована аналогичная методика, что и для новописьменных форм ливвиковского и собственно карельского наречий [12, 683].

Как и для первых двух наречий были выбраны характерные для выявления именного словоизменения формы: словарная форма, гласная слабая основа в генитиве ед. ч. (основа 1), глас-

ная сильная основа в иллативе ед. ч. (основа 2), форма партитива ед. ч. (основа 3), слабая основа в генитиве мн. ч. (основа 4) и сильная основа в иллативе мн. ч. (основа 5). В качестве примеров для заполнения представлены как одноосновные, так и двуосновные имена, выявленные из примеров частотных лемм на южнолюдиковском диалекте карельского языка (табл. 2).

Таблица 2

Словоизменительные типы имён южнолюдиковского диалекта карельского языка

Гл. осн.	Ном. ед.	Ген. ед.	Илл. ед.	Парт. ед.	Ген. мн.	Парт. мн.
Одноосновные имена						
O ¹ -	halgo ‘полено’	halgo-n	halgo-h	halgo-d	halguoi-den	halguoi-d
U ² -	čuppu ‘угол’	čupu-n	čuppu-h	čuppu-d	čupui-den	čuppu-d
Oi-	dänöi ‘заяц’	dänöi-n	dänöi-h	dänöi-d	dänölyöi-den, dänölöi-den	dänölyöi-d, dänölöi-d
A ³ i	astii ‘посуда’	ast’ai-n	ast’ai-h	ast’ai-d	ast’oi-den, ast’uoi-den	ast’oi-d, ast’uoi-d
i-	počči ‘свинья’	poči-n	počči-h	počči-d	počii-den	poččii-d
e-	d’ogi ‘река’	d’oge-n	d’oge-h	d’oge-d	d’ogi-den	d’ogi-d
A-	akke ‘женщина’	aka-n	akka-h	akka-d	akuoi-den, akoi-den	akkuoi-d, akko-i-d
V ¹ V-	mua ‘земля’	mua-n	mua-h	mua-d	mualuoi-den	mualuoi-d

¹ O - по гармонии гласных o или ö.

² U - по гармонии гласных u или y.

³ A - по гармонии гласных a или ä

VVV-	tiäi ‘вошь’	tiäi-n	tiäi-h	tiäi-d	tiäilyöi-den	tiäilyöi-d
Двуосновные имена						
he-	pereh ‘семья’	perehe-n	perehe-h	pereh-te	perehi-den	perehi-d
le-	hiili ‘уголь’	hiile-n	hiile-h	hiil-te	hiili-den	hiili-d
me-	härkin ‘мутовка’	härkime-n	härkime-h	härkin-te	härkimi-den	härkimi-d
ne-	pieni ‘маленький’	piene-n	piene-h	pien-te	pieni-den	pieni-d
re-	tytär ‘дочь’	tyttäre-n	tyttäre-h	tytär-te	tyttäri-den	tyttäri-d
te-	yksi ‘один’	yhte-n	yhte-h	yh-te	yksi-den	yksi-d
de-	vezi ‘вода’	vede-n	vede-h	vet-te	vezi-den	vezi-d
že-	naine ‘женщина’	naiže-n	naiže-h	naiš-te	naiži-den	naiži-d
se-	lapsi ‘ребёнок’	lapse-n	lapse-h	las-te	lapsi-den	lapsi-d
	verez ‘свежий’	verekse-n	verekse-h	veres-te	vereksi-den	vereksi-d
dA-	lyhyd ‘короткий’	lyhydä-n	lyhydä-h	lyhyt-te	lyhydi-den	lyhydi-d
hA-	barbaz ‘палец ноги’	barbaha-n	barbaha-h	barbas-te	barbahi-den	barbahi-d
ttOmA-	iänetyöi ‘безголосый’	iänettömä-n	iänettömä-h	iänetön-te	iänettämi-den	iänettämi-d

При исследовании было установлено, что для формирования полной парадигмы слова достаточно знать изменяемую и неизменяемую часть словарного слова, слабую/гласную и сильную/согласную основы. Для дальнейшей работы информация о слове записывается в виде шаблонов:

неизм.|ок. [п.о.1] для одноосновных слов и

неизм.|ок. [п.о.1, п.о.2] для двуосновных слов, где

неизм. — неизменяемая часть слова, т. е. часть словарного слова до | или слово целиком, если нет |;

ок. — часть словарного слова после |²;

п.о.1 — первое (или единственное) буквосочетание в скобках, после «[» и перед «,» (ген. ед. ч. без неизменяемой части слова и конечного -n);

п.о.2 — второе буквосочетание в скобках, после «,» и перед «[» (парт. ед. ч. без неизменяемой части слова и конечного -te).³

В таблице 3 приведены примеры шаблонов и правила выделения основ из описанных шаблонов. Таким образом, из представленного шаблона последовательно формируются словарная форма (с.ф.), четыре основы (о.1, о.2, о.3, о.6), далее, с учётом полученных основ, формируются две вспомогательные основы (о.4, о.5).

Таблица 3

Правила формирования основ для генерации словоформ⁴

с.ф.	с.с. ⁵ – и	
		aba i [ja]: abai ‘тоня’ ak ke [a]: akke ‘женщина’ ala ine [iže, š]: alaine ‘рукавица’ ast ii [’ai]: astii ‘посуда’ d’alga oi [toma, on]: d’algatoi ‘безногий’ dänö i [i]: dänöi ‘заяц’ eläv [ä]: eläv ‘живой’ emänd e [ä]: emände ‘хозяйка’ halgo []: halgo ‘полено’

¹ V – любая гласная.

² Если ок. пусто, то | не пишется.

³ Даже если п.о.2 пусто, обязательно ставится запятая: tuoh|i [e,].

⁴ В таблицах 3 и 4 приняты следующие обозначения: | – отделение изменяемой части слова от неизменяемой, || – разделение составных слов в сложном слове, – – убрать, + – добавить, = – равно, > – заменить, / – две основы; V – любая гласная, C – любая согласная, Cч – двойная согласная; с.с. – словарное слово, может содержать | и ||, с.ф. – словарная форма (без | и ||).

⁵ Предварительно, если слово сложное, определяется гармония по последнему составному слову и удаляется || в шаблоне.

		<p>kirik kō [ō]: kirikkō ‘церковь’ leib e [ä]: leibe ‘хлеб’ lindu []: lindu ‘птица’ mua []: mua ‘земля’ nuot te [a]: nuotte ‘невод’ par ži [de,]: parži ‘бревно’ pedä i [jä]: pedäi ‘сосна’ poč či [i]: počči ‘свинья’ tiä i [i]: tiäi ‘вошь’ tuoh i [e,]: tuohi ‘береста’ ve zi [de, t]: vezi ‘вода’ y ksi [hte, h]: yksi ‘один’</p>
о.1 (сл. / гл. осн.)	А. Если скобки пустые, то = с.ф. Б. Если скобки не пустые, то = неизм.+п.о.1	<p>mua []: mua tuoh i [e,]: tuoh+e=tuohe ak ke [a]: ak+a=aka</p>
о.2 (осн. иллат)	А. Если неизм. оканчивается и ок. начинается на одну и ту же согласную C _ч (k, p, t, č, s, š, s или h), то удваиваем эту согласную, т. е. = неизм.+C_ч+п.о.1 Б. В остальных случаях = о.1	<p>ak ke [a]: ak+k+a=akka nuot te [a] : nuot+t+a = nuotta poč či [i] : poč+č+i = počči</p> <p>tuoh i [e,]: tuohe</p>
о.3 (парт. ед.)	А. Если о.6 оканч. на V, то = о.6+d Б. Если о.6 оканч. на C, то = о.6+te	<p>mua+d=muad kirikkö+d=kirikköd alaš+te=alašte</p>
о.4 (вспом. осн. мн.ч., сл.)	А. Если о.1 состоит из одного слога и оканчивается на трифтонг, дифтонг или долгую гласную (CVVV или CVV), то о.1+IUOi	<p>tiäi+lyöi=tiäilyöi mua+luoi=mualuoi</p>
	Б. Если о.6 оканч. на C, то = о.1 с конеч. заменами 1) hte>ksi 2) rde>rži 3) de>zi 4) A>i	<p>yh, yhte>yksi (1) par, parde>parži (2) vet, vede>vezi (3) d’algaton, d’algattoma>d’algattomi (4)</p>
	В. Если о.6 оканч. на O, A, Oi, или Ai, то = о.1 с заменами, если о.1 оканч. на 1) CO, то O>Oi / O>UOi 2) Ca, то a>i, если в о.6 а) два слога, в первом слоге есть u, o ¹ ИЛИ б) более 2 слогов и о.6 оканч. на mpa/mba, ma ИЛИ является прилагательным с о.6 на va 3) Ca, то a>oi / a>uoi, если в о.6 а) два слога и в первом слоге есть a, e, i ² ИЛИ б) более двух слогов и о.6 НЕ оканч. на mpa/mba, ma и НЕ является прилагательным с о.6 на va	<p>halgo, halgo>halgoi/halguoi (1) kirikkö, kirikö>kiriköi/kirikyöi (1) nuotta, nuota>nuoti (2a) elävä, elävä>elävi (2б) akka, aka>akoi, akuoi (3a) abaja, abaja>abajoi/abajuai (3б) leibä, leibä>leibi (4a) pedäjä, pedäjä>pedäji (4б) emändä, emändä>emändöi/emändyöi (5) ast’ai, ast’ai>ast’oi/ast’uoi (6) dänöi>dänölöi/dänölyöi (7)</p>

¹ В том числе в составе дифтонгов ui, uo, ui, ou, oi.

² В том числе в составе дифтонгов ai, au, ua, ei, ie, iu, ie, ee, eu.

	4) Cä, то ä>i, если в о.6 а) два слога ИЛИ б) более двух слогов и о.6 оканч. на mä, vä, žä, sä, jā, pä, bā 5) Cä, то ä>yöi / ä>öi, если в о.6 более двух слогов и о.6 оканч. на dä, tä, gä, kä, čä, hä, lä, nä, rä 6) CAi, то Ai>Oi / Ai>UOi 7) COi, то i>IOi / i>IUOi	
	Г. Если о.1 оканч. на Ce, то = о.1 с заменой e>i	alaiže > alaiži
	Д. В остальных случаях, если о.6 оканч. на V, то о.1+i	lindu, lindu+i=lindui
о.5 (вспом. осн. мн., сильн.)	Аналогично правилам для о.4 , только вместо о.1 берётся о.2	tiäi+lyöi=tiäilyöi (A) mua+luoi=mualuoi (A)
		yh, yhte>yksi (Б.1) par, parde>parži (Б.2) vet, vede>vezi (Б.3) d'algattoma>d'algattomi (Б.4)
		halgo>halgoi/halguoi (B.1) kirikkö>kirikköi, kirikkyöi (B.1) nuota>nuoti (B.2a) elävä>elävi (B.2б) akka>akko/akkuoi (B.3a) abaja>abajoi/abajuai (B.3б) leibä>leibi (B.4a) pedäjä>pedäji (B.4б)
		emändä>emändöi/emändyöi (B.5) ast'ai>ast'oi/ast'uoi (B.6) dänöi>dänölöi/dänölyöi (B.7)
		alaiže > alaiži
		lindu+i=lindui
о.6 (сильн. / согл. осн.)	А. Если в скобках нет п.о.2 , то = о.2	mua []: mua=mua kirik kö [ö]: kirikkö=kirikkö
	Б. Если в скобках два п.о., то неизм. + п.о.2	ala ine [iže, š]: ala+š=alaš hiir i [e,]: hiir

Выделенные основы (табл. 3) используются для создания словоформ по правилам, описан-

ным в таблице 4, путём присоединения падежного окончания к подходящей основе.

Таблица 4

Правила генерации именных словоформ святозерского диалекта карельского языка

Падеж	Единственное число		Множественное число	
	Правило	Примеры	Правило	Примеры
ном.	с.ф.	akke tuohi	о.1 + d, о.1 + t	akad, akat tuohed, tuohet
акк.	с.ф., о.1 + n	akke, akan tuohi, tuohen		

ген.	o.1 + n	akan tuohen	o.4 + den	akoiden, akuoiden tuohiden
парт.	o.3	akkad tuohete	o.5 + d	akkoid, akkuoid tuohid
эсс.	o.1 + n, o.1 + nnU	akan, akannu tuohen, tuohennu	o.4 + n, o.4 + nnU	akoин, akoинну, akuоин, akuоинну tuohin, tuohинну
транс.	o.1 + ks, o.1 + kse	akaks, akakse tuoheks, tuohekse	o.4 + ks, o.4 + kse	akoiks, akoikse, akuoiks, akuoikse tuohiks, tuohikse
абес.	o.1 + tA	akata tuoheta	o.4 + tA	akoita, akuoita tuohita
инес.	o.1 + s	akas tuohes	o.4 + s	akois, akuois tuohis
элат.	o.1 + s, o.1 + spiäi	akas, akaspiäi tuohes, tuohespiäi	o.4 + s, o.4 + spiäi	akois, akoispiäi, akuois, akuoispiäi tuohis, tuohispiäi
илл.	o.2 + h	akkah tuoheh	o.5 + h	akkoih, akkuoih tuohih
адес.	o.1 + l	akal tuohel	o.4 + l	akoil, akuoil tuohil
абл.	o.1 + l, o.1 + lpiäi	akal, akalpiäi tuohel, tuohelpiäi	o.4 + l, o.4 + lpiäi	akoil, akoilpiäi, akuoil, akuoilpiäi tuohil, tuohilpiäi
алл.	o.1 + l, o.1 + le	akal, akale tuohel, tuohеле	o.4 + l, o.4 + le	akoil, akoile, akuoil, akuoile tuohil, tuohile
ком.	o.1 + nke	akanke tuohenke	o.4 + neh	akoineh, akuoineh tuohineh
прол.	o.1 + či	akači tuoheči	o.4 + či	akois, akuois tuohiči
инст.	-	-	o.4 + n	akoин, akuоин tuohin
аппр.	o.1 + lloh, o.1 + lloo, o.1 + lluo,	akalloh, akalloo, akalluo tuohelloh, tuohelloo, tuohelluo	o.4 + lloh, o.4 + lloo, o.4 + lluo	akoilloh, akoilloo, akoilluo, akuoilloh, akuoilloo, akuoilluo tuohilloh, tuohilloo, tuohilluo
терм.	o.2 + ssuai	akkassuai tuohessuai	o.5 + ssuai	akkoissuai, akkuoissuai tuohissuai
адит.	o.2 + hpiäi	akkahpiäi tuohehpiäi	o.5 + hpiäi	akkoihpiäi, akkuoihpiäi tuohihpiäi

Таким образом, были разработаны правила генерации словоформ для людиковского наречия карельского языка. Эти правила запрограммированы в корпусном менеджере Dictorpus корпуса ВепКар. При работе со словарной статьёй о людиковском слове *kirikkö* 'церковь' (рис. 3) редактор вводит шаблон *kirik|kō* [ö] и в словаре корпуса ВепКар автоматически создаётся 78 словоформ для людиковского наречия (рис. 3).

Создание таких правил позволило добавить в словарь ВепКар 75 тыс. людиковских словоформ, количество размеченных токенов в текстах

людиковского подкорпуса на 4 %. Такое незначительное увеличение обусловлено тем, что далеко не все именные словоформы встречаются в текстах корпуса. Кроме этого, некоторые падежи в текстах малоупотребительны, напр., аппроксиматив, терминатив, пролатив, адитив.

Корпусный менеджер Dictorpus корпуса ВепКар представляет собой набор PHP-скриптов, работающих на платформе Laravel. Код программы распространяется по открытой лицензии, размещён на сайте GitHub¹.

¹ См. <https://github.com/componavt/dictorpus>

kirikkö

язык: карельский: людиковское наречие

часть речи: существительное

1 значение

понятие:

церковь

- русский: церковь



перевод

вепсский: *cerkof; cerkov; pagast*

карельский: ливвиковское наречие:

kirikkö

карельский: собственно карельское

наречие: *kirikkö; kirikk; čerkkova; čerkku*

диалекты употребления:

Южнолюдиковский (святозерский)

Примеры (20)

★ лучший ★ отличный ★ хороший

1. ★ Mänöu kirikköh, časounah, zvoniu, siid malittut panou sigä miitutte hänele, armiis kous sluužii.

Пойдет в часовню, звонит, молитвы читает, какие он [выучил], когда в армии служил. (Külän časoune)

2. ★ Muštan ned ol'd'ih vie kirikös storožannu ol'd'ih.

Помню, они в церкви сторожами были. (Lähikülden eläjiden nimet)

3. ★ Čerkov' (kirikkö).

Церковь. (MATKUSTAMIINE)

4. ★ Müö pogostale kilometrin vie astumme, sigä d'ö kai kellod zvonitah kirikös, oi d'ö vuotetah bohatištod dai häntte valmehekse selgäz otettih.

Мы до погоста еще километр прошли, там в церкви уже во все колокола звонят, уже ждут богатые. (Voine tuli Pelduoizih)

5. ★ Nu siide nouzimme, emme üöd olnu, mänimme kirikköh kävüimme da senker

словоформы (78)

kirik|kö [õ]

No	грамматические признаки	Южнолюдиковский (святозерский) (78)
Единственное число		
1.	номинатив	<i>kirikkö</i>
2.	аккузатив	<i>kirikkö, kirikön</i>
3.	генитив	<i>kirikön</i>
4.	партитив	<i>kirikköd</i>
5.	эссив	<i>kirikön, kirikönny</i>
6.	транслатив	<i>kiriköks, kirikökse</i>
7.	абессив	<i>kirikötä</i>
8.	инессив	<i>kirikös</i>
9.	элатив	<i>kirikös, kiriköspiäi</i>
10.	иллатив	<i>kirikköh</i>
11.	адессив	<i>kiriköl</i>
12.	аблатив	<i>kiriköl, kirikölpäi</i>
13.	аллатив	<i>kiriköl, kiriköle</i>
14.	комитатив	<i>kirikönke</i>
15.	пролатив	<i>kiriköči</i>

Рис. 3. Пример людиковского слова *kirikkö* ‘церковь’ в словаре ВепКар

Обсуждение и заключение

В рамках нашего исследования именной словоизменительной системы людиковского наречия карельского языка (южнолюдиковского диалекта) с целью создания программы автоматической генерации словоформ удалось выделить и унифицировать правила образования всех грамматических форм для имён основных словоизменительных типов. Нами была минимизирована грамматическая информация об имени, которую необходимо вносить вручную для дальнейшего автоматического заполнения полной парадигмы слова: мы сократили её до нуля или одного псевдоокончания для одноосновных имён (*diä* [], *lindu* [], *ak/ke* [a]) и до двух или трёх

– для двуосновных (*barbalz* [ha, s], *velzi* [de, t], *kynäbrylz* [kse, s]). Для людиковского наречия полностью автоматизировать процесс генерации именной словоизменительной парадигмы стало вполне возможно благодаря отсутствию в нём качественного чередования согласных. Использование программы генерации именных словоформ позволило в несколько раз сократить время заполнения одной именной парадигмы: до 10–20 секунд. Правила образования форм единственного и множественного числа станут существенным шагом вперёд для разработки письменной нормы людиковского наречия, правил орфографии и подготовки пособий и учебников по людиковской грамматике.

Список источников и литературы

1. Баранцев А. П. Фонологические средства людиковской речи (дескриптивное описание). Л.: Наука, 1975. 280 с.
2. Баранцев А. П. Образцы людиковской речи (образцы корпуса людиковского идиолекта). Петрозаводск: Карелия, 1978. 287 с.
3. Бойко Т. П. Большой карельско-русский словарь (ливвиковское наречие). Петрозаводск: Периодика, 2016. 352 с.
4. Бойко Т. П., Маркианова Л. Ф. Большой русско-карельский словарь: (ливвиковское наречие). Петрозаводск: Периодика, 2018. 399 с.
5. Зайков П. М. Глагол в карельском языке. Петрозаводск: Изд-во ПетрГУ, 2000. 293 с.
6. Карельский язык в грамматиках / И. Новак, М. Пенттонен, А. Руусканен, Л. Сиилин. Петрозаводск: КарНЦ РАН, 2019. 479 с.
7. Ковалёва С. В., Родионова А. П. Традиционное и новое в лексике и грамматике карельского языка (по данным социолингвистического исследования). Петрозаводск: ИЯЛИ КарНЦ РАН, 2011. 138 с.
8. Патроева Н. В., Пашкова Т. В. Система междометий в людиковском наречии карельского языка: опыт классификации // Вестник Московского университета. Серия 9. Филология. 2023. № 6. С. 81–92.

9. Пашкова Т. В., Родионова А. П. Образование глагольных форм императива в ливвиковском и людиковском наречиях карельского языка // Научный диалог. 2020. № 4. С. 100–112.
10. Пашкова Т. В., Родионова А. П. Словообразовательные суффиксы -NDU/-NDY/- ND(E) И -HINE в ливвиковском и людиковском наречиях карельского языка (на примере наименований болезней) // Учёные записки Петрозаводского государственного университета. Петрозаводск: Изд-во ПетрГУ, 2021. Т. 43. № 3. С. 17–22.
11. Пашкова Т. В., Родионова А. П. «Быть» или «Не быть» потенциалу в карельском языке (на материале ливвиковского и людиковского наречий карельского языка) // Ежегодник финно-угорских исследований. 2022. Т. 16. № 3. С. 397–404.
12. Разработка правил генерации именных словоформ для новописьменных вариантов карельского языка / И. П. Новак, Н. Б. Крижановская, Т. П. Бойко, Н. А. Пеллинен // Вестник угроведения. 2020. Т. 10. № 4. С. 679–691.
13. Родионова А. П., Нагурная С. В. Людиковское наречие карельского языка: проблемы становления литературной нормы // Финно-угорский мир. 2018. № 2. С. 48–57.
14. Родионова А. П. О некоторых особенностях именного словоизменения в карельско-людиковской диалектной речи // Вестник угроведения. 2018. Т. 8. № 2. С. 284–293.
15. Родионова А. П. О первом печатном словаре на людиковском наречии карельского языка // Ежегодник финно-угорских исследований. 2021. Т. 15. № 1. С. 6–15.
16. Русско-карельский словарь / сост. учитель Святозерского одноклассного уч-ща Михаил Дмитриевич Георгиевский; под ред. и с предисл. Ф. Фортунатова. СПб.: Русский начальный учитель, 1908. 53 с.
17. Сопоставительно-ономасиологический словарь диалектов карельского, вепсского, саамского языков / под общей ред. Ю. С. Елисева и Н. Г. Зайцевой. Петрозаводск: Изд-во КарНЦ РАН, 2007. 346 с.
18. Karelian in Grammars: a study of phonetic and morphological variation: scientific electronic edition / I. Novak, M. Penttonen, A. Ruuskanen, L. Siilin. Petrozavodsk: KarRC RAS, 2022. 452 p.
19. Näytteitä karjalan kielestä. 1: Karjalan tasavallan ja Sisä-Venäjän karjalaismurteet = Образцы карельской речи. 1: говоры Республики Карелии, тихвинских и тверских карел / ред. В. Д. Рягоев, М. Есканен. Joensuu; Petroskoi: Joensuun yliopiston monistuskeskus, 1994. 459 p.
20. Pahomov M. Непризнанный народ: древний этнос Карелии: [о карелах-людиках] // Lüüdilaine: lyydiläisen seuran vuosijulkaisu 2013. Helsinki: Helsingin yliopisto, 2013. С. 22–26.
21. Pahomov M. Lüüdi-venän, venä-lüüdin sanakird': lähes 19 000 sana-artikkelid / Miikul Pahomov. 1. painuz. Helsingi : Lüüdilaine Siebr, 2022. 385 с.
22. The Open corpus of the Veps and Karelian languages: overview and applications / T. Boyko, N. Zaitseva, N. Krizhanovskaya, A. Krizhanovsky, I. Novak, N. Pellinen, A. Rodionova // KnE Social Sciences. 2022. № 7 (3). Pp. 29–40.
23. Turunen A. Lyydiläismurteiden äännehistoria. I. Konsonantit. Helsinki: SUST, 1946. 338 с. (Suomalais-Ugrilaisen Seuran Toimituksia, № 89).
24. Turunen A. Lyydiläismurteiden äännehistoria. II. Vokaalit. Helsinki: SUST, 1950. 266 с. (Suomalais-Ugrilaisen Seuran Toimituksia, № 99).

References

1. Barantsev A. P. *Fonologicheskie sredstva lyudikovskoi rechi (deskriptivnoe opisanie)* [Phonological means of Ludic speech (descriptive analysis)]. Leningrad: Nauka Publ., 1975. 280 p. (In Russian)
2. Barantsev A. P. *Obrazcy ljudikovskoj rechi (obrazcy korpusa ljudikovskogo idiolekta)* [Samples of Ludic speech (samples of the corpus of the Ludic idiolect)]. Petrozavodsk: Karelia Publ., 1978. 287 p. (In Russian, Karelian)
3. Boyko T. P. *Bol'shoj karel'sko-russkij slovar' (livvikovskoe narechie)* [Great Karelian-Russian dictionary (Livvi dialect)]. Petrozavodsk: Periodika Publ., 2016. 351 p. (In Russian, Karelian)
4. Boyko T. P., Markianova L. F. *Bo'shoj russko-karel'skij slovar' (livvikovskoe narechie)* [Great Russian-Karelian dictionary (Livvi dialect)]. Petrozavodsk: Periodika Publ., 2018. 399 p. (In Russian, Karelian)
5. Zaykov P. M. *Glagol v karel'skom jazyke* [Verb in the Karelian language]. Petrozavodsk: PetrGU Publ., 2000. 293 p. (In Russian)
6. Novak I., Penttonen M., Ruuskanen A., Siilin L. *Karel'skiy yazyk v grammatikakh* [Karelian language in grammars]. Petrozavodsk: KarRC RAS Publ., 2019. 479 p. (In Russian)
7. Kovaleva S. V., Rodionova A. P. *Traditsionnoe i novoe v leksike i grammatike karel'skogo yazyka (po dannym sotsiolingvisticheskogo issledovaniya)* [Traditional and new in the vocabulary and grammar of the Karelian language (according to sociolinguistic research)]. Petrozavodsk: IYaLI KarNTs RAN Publ., 2011. 137 p. (In Russian)
8. Patroeva N. V., Pashkova T. V. *Sistema mezhdometij v lyudikovskom narechii karel'skogo yazyka: opyt klassifikacii* [System of interjections in the Ludik dialect of the Karelian language: classification experience]. *Vestnik Moskovskogo universiteta. Seriya 9. Filologiya* [Lomonosov Philology Journal. Series 9. Philology], 2023, no. 6, pp. 81–92. (In Russian)
9. Pashkova T. V., Rodionova A. P. *Obrazovanie glagol'nykh form imperativa v livvikovskom i lyudikovskom narechiyakh karel'skogo yazyka* [Formation of verb forms of the imperative in the Livvi and Ludic dialects of the Karelian language]. *Nauchnyi dialog* [Scientific Dialogue], 2020, no. 4, pp. 100–112. (In Russian)
10. Pashkova T. V., Rodionova A. P. *Slovoobrazovatel'nye suffiksy -NDU/-NDY/-ND(E) I -HINE v livvikovskom i lyudikovskom narechiyakh karel'skogo yazyka (na primere naimenovanii boleznei)* [Derivative suffixes -NDU/-NDY/-ND(E) AND -HINE in the Livvik and Ludic dialects of the Karelian language (on the example of disease names)]. *Uchenye*

zapiski Petrozavodskogo gosudarstvennogo universiteta [Proceedings of the Petrozavodsk State University]. Petrozavodsk: PetrSU Publ., 2021. Vol. 43 (3). Pp. 17–22. (In Russian)

11. Pashkova T. V., Rodionova A. P. “Byt” ili “Ne byt” potentsialu v karel’skom yazyke (na materiale livvikovskogo i lyudikovskogo narechii karel’skogo yazyka) [“To be” or “Not to be” potential in the Karelian language (on the material of the Livvik and Ludic dialects of the Karelian language)]. *Ezhegodnik finno-ugorskikh issledovanii* [Yearbook of Finno-Ugric Studies], 2022, no. 16 (3), pp. 397–404. (In Russian)

12. Novak I. P., Krizhanovskaya N. B., Boyko T. P., Pellinen N. A. *Razrabotka pravil generacii imennykh slovoform dlya novopis’ennykh variantov karel’skogo yazyka* [Development of rules of generation of nominal word forms for new-written variants of the Karelian language]. *Vestnik ugrovedenia* [Bulletin of Ugric Studies], 2020, no. 10 (4), pp. 679–691. (In Russian)

13. Rodionova A. P., Nagurnaya S. V. *Lyudikovskoe narechie karel’skogo yazyka: problemy stanovleniya literaturnoi normy* [Ludic dialect of the Karelian language: problems of the formation of the literary norm]. *Finno-ugorskii mir* [Finno-Ugric World], 2018, no. 2, pp. 48–57. (In Russian)

14. Rodionova A. P. *O nekotorykh osobennostyakh imennogo slovoizmeneniya v karel’sko-lyudikovskoj dialektnoj rechi* [About some features of nominal inflection in the Karelian-Ludic dialect]. *Vestnik ugrovedenia* [Bulletin of Ugric Studies], 2018, no. 8 (2), pp. 284–293. (In Russian)

15. Rodionova A. P. *O pervom pechatnom slovare na lyudikovskom narechii karel’skogo yazyka* [About the first printed vocabulary on the Ludic dialect of the Karelian language]. *Ezhegodnik finno-ugorskikh issledovanij* [Yearbook of Finno-Ugric Studies], 2021, no. 15 (1), pp. 6–15. (In Russian)

16. *Russko-karel’skij slovar’* [Russian-Karelian vocabulary]. Comp. Mikhail Dmitrievich Georgievskiy; ed. F. Fortunatov. Saint-Petersburg: Russkij nachal’nyj uchitel Publ., 1908. 53 p. (In Russian, Karelian)

17. *Sopostavitel’no-onomasiologicheskij slovar’ dialektov karel’skogo, vepsskogo, saamskogo jazykov* [Comparative-onomasiological dictionary of dialects of the Karelian, Vepsian, and Sami languages]. Ed. Yu. S. Eliseev, N. G. Zaytseva. Petrozavodsk: KRC Publ., 2007. 346 p. (In Russian, Karelian, Vepsian, Sami)

18. Novak I., Penttonen M., Ruuskanen A., Siilin L. *Karelian in Grammars: a study of phonetic and morphological variation: scientific electronic edition*. Petrozavodsk: KarRC RAS Publ., 2022. 452 p. (In English)

19. *Näytteitä karjalan kielestä I*. Joensuu-Petrozavodsk: Joensuun yliopiston monistuskeskus, 1994. 455 p. (In Finnish)

20. Pahomov M. *Nepriznannyy narod: drevnii etnos Karelii: o karelahk-lyudikakh*. Lüüdilaine: lyydiläisen seuran vuosijulkaisu 2013. Helsinki: Helsingin yliopisto, 2013. Pp. 22–26. (In Karelian)

21. Pahomov M. *Lüüdi-venän, venä-lüüdin sanakird’*: lähes 19 000 sana-artikkelid. Miikul Pahomov. 1. painuz. Helsinki: Lüüdilaine Siebr, 2022. 385 p. (In Karelian, Russian)

22. Boyko T., Zaytseva N., Krizhanovskaya N., Krizhanovsky A., Novak I., Pellinen N., Rodionova A. The Open corpus of the Veps and Karelian languages: overview and applications. *KnE Social Sciences*, 2022, no. 7 (3), pp. 29–40. (In English)

23. Turunen A. *Lyydiläismurteiden äännehistoria. I. Konsonantit*. Helsinki: SUST, 1946. 338 p. (Suomalais-Ugrilaisen Seuran Toimituksia, no. 89). (In Finnish)

24. Turunen A. *Lyydiläismurteiden äännehistoria. II. Vokaalit*. Helsinki: SUST, 1950. 266 p. (Suomalais-Ugrilaisen Seuran Toimituksia, no. 99). (In Finnish)

ИНФОРМАЦИЯ ОБ АВТОРАХ

Родионова Александра Павловна, научный сотрудник Института языка, литературы и истории ФИЦ Карельского научного центра Российской академии наук (185910, Российская Федерация, Республика Карелия, г. Петрозаводск, ул. Пушкинская, 11), кандидат филологических наук.

santrar@krc.karelia.ru

ORCID ID: 0000-0001-5645-9441

Крижановская Наталья Борисовна, ведущий инженер-исследователь лаборатории информационных компьютерных технологий Института прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр» (185910, Российская Федерация, Республика Карелия, г. Петрозаводск, ул. Пушкинская, д. 11).

nataly@krc.karelia.ru

ORCID ID: 0000-0002-9948-1910

ABOUT THE AUTHORS

Rodionova Aleksandra Pavlovna, Researcher of the Institute of Language, Literature and History of the Karelian Research Centre of the Russian Academy of Sciences (185910, Russian Federation, Republic of Karelia, Petrozavodsk, Pushkinskaya Str., 11), Candidate of Philological Sciences.

santrar@krc.karelia.ru

ORCID ID: 0000-0001-5645-9441

Krizhanovskaya Natalya Borisovna, Leading Engineer Researcher of the Laboratory for Information Computer Technologies, Institute of Applied Mathematical Research, Karelian Research Center of the Russian Academy of Sciences (185910, Russian Federation, Republic of Karelia, Petrozavodsk, Pushkinskaya Str., 11).

nataly@krc.karelia.ru

ORCID ID: 0000-0002-9948-1910