



Valsts pētījumu programmas «Humanitāro zinātņu digitālie resursi» projekts «Atvērtas un FAIR principiēm atbilstīgas digitālo humanitāro zinātņu ekosistēmas attīstība Latvijā»
(Nr. VPP-IZM-DH-2022/1-0002)

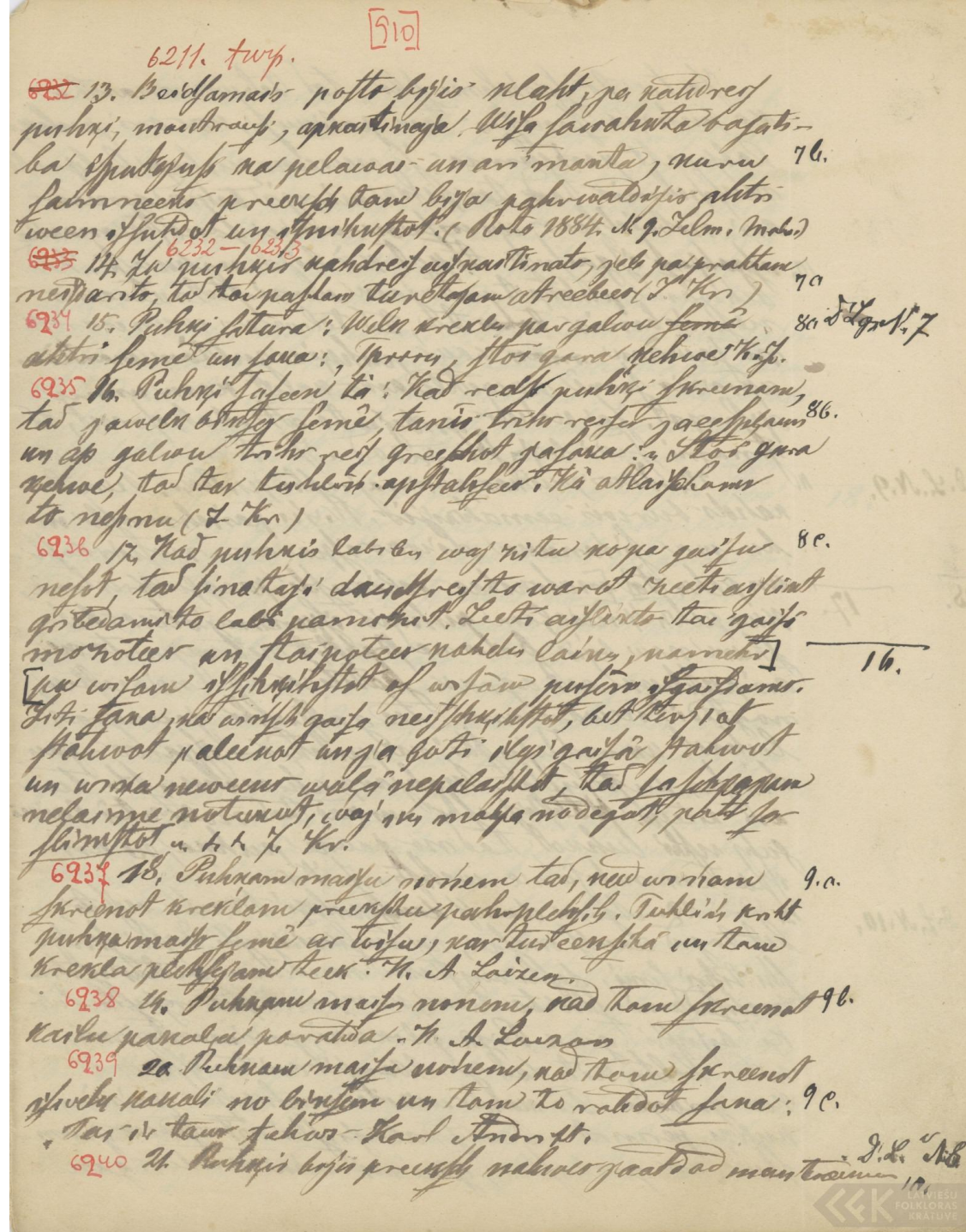
Metodiskas pieejas meklējumi digitalizēto latgaliešu folkloras kolekciju normalizēšanai

Angelika Juško-Štekele (RTU RA), Sandis Laime (LU LFMI)

XIV Starptautiskais baltistu congress
Viļņa, 24.10.2025.

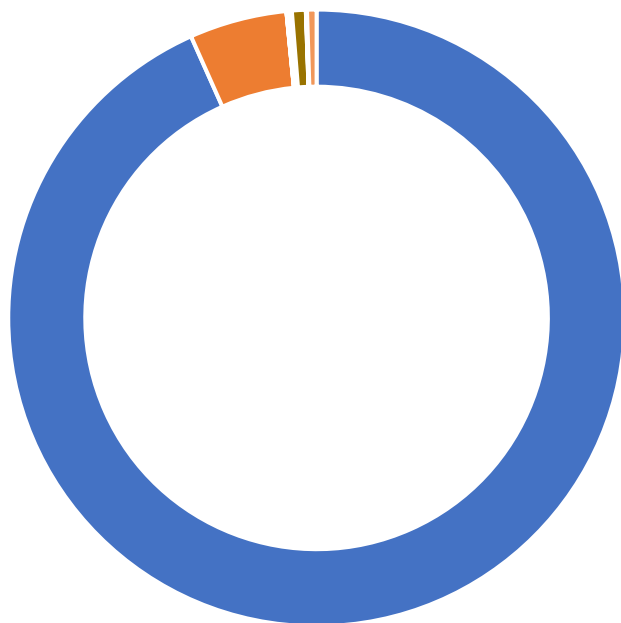
LFK arhīva saturs

- 3,000,000 folkloras tekstu
- 2570 kolekcijas
- Vismaz 360 kolekcijās sastopami teksti latgaliešu valodā (14% no visām kolekcijām)

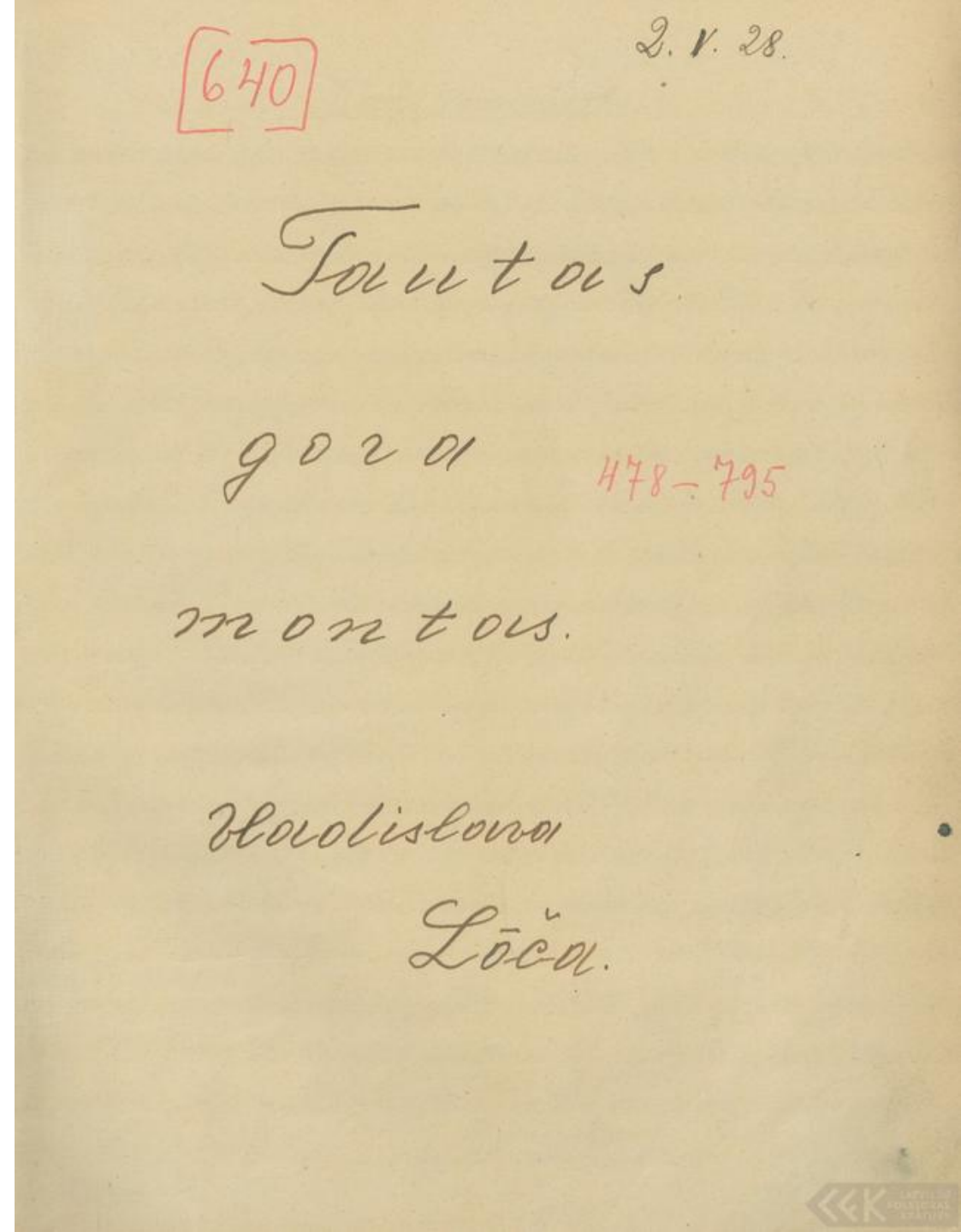


Latgaliešu manuskripti digitālajā arhīvā garamantas.lv

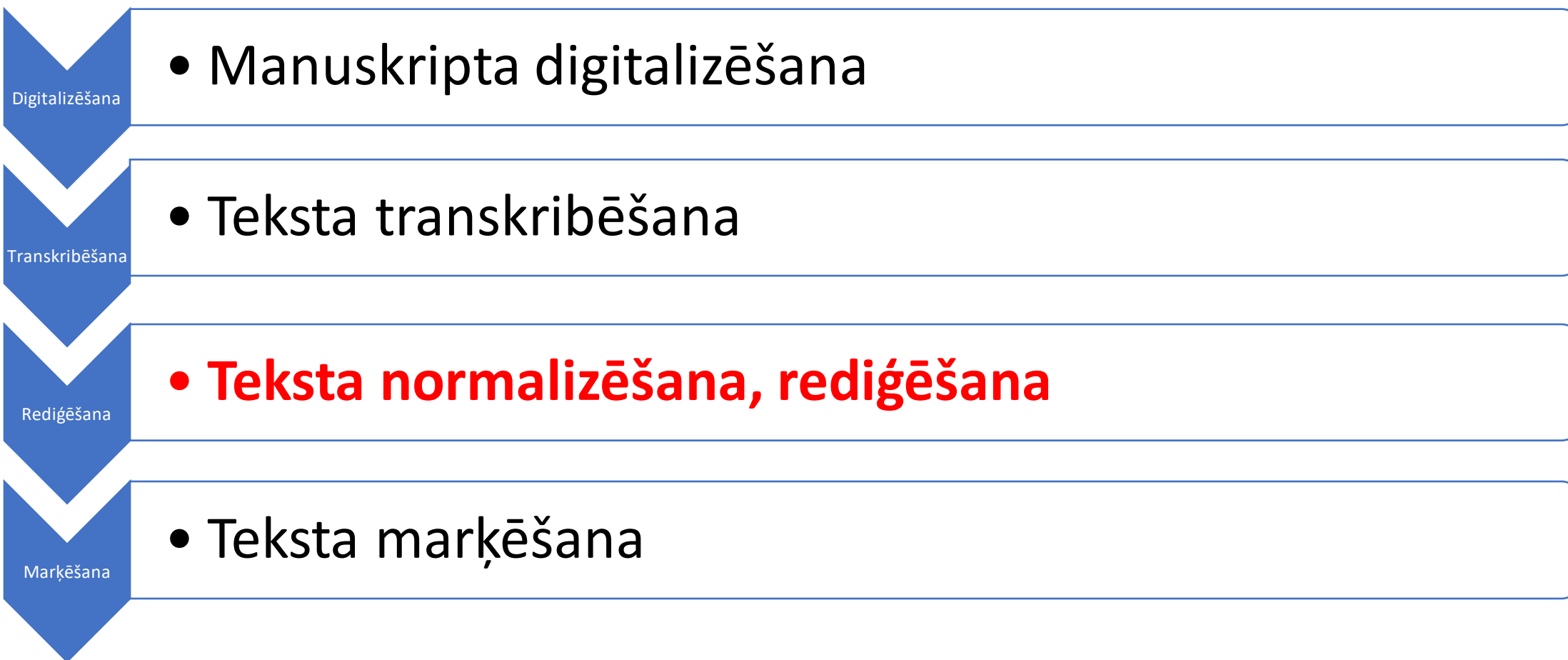
>1,015,500 digitalizētu manuskriptu failu
>55,000 no tiem latgaliešu valodā (5.5%)



- latviešu valoda
- latgaliešu valoda
- angļu valoda
- baltkrievu valoda
- romu valoda
- franču valoda
- igauņu valoda
- ivrits
- jidišs
- krievu valoda
- lietuviešu valoda
- lībiešu valoda
- poļu valoda



LFK arhīva tekstu apstrāde



Iepriekšējā vienība

Nākamā vienība



Atšifrēt

Iepriekšējā vienība

Nākamā vienība

Vienība

Faili

Teksts

Vietas

Nosaukums

Nakts raganas velējas

Vienības Nr.

19963

Krājums

Latviešu folkloras krātuve

Organizācija

Rencēnu pamatskola

Klasifikācijas

Mitoloģiskā teika, Tautas gudrība un leksika: vietu vārdi, Ragana, Ragana - pārdabiska būtne, Raganu kalns, D4.1.3. Nakts raganas velējas, D3.4.8. Nakts raganu uzturēšanās vieta – kalns

Valodas

Latviešu valoda

Faili

0116-Rencenu-pamatskola-25-0009

Iesūtītājs

Artūrs Tenis Lapsiņš

Teicējs

Kārlis Barīss

Pierakstītājs

Mirdza Pētersone

Pierakstīšanas laiks un vieta

1939, Stukūži
"Stukūži", Rencēnu pagasts, Valmieras novads, LV-4232

Atslēgvārdi

Pievienot atslēgvārdus

Vienības profils:

Metadati

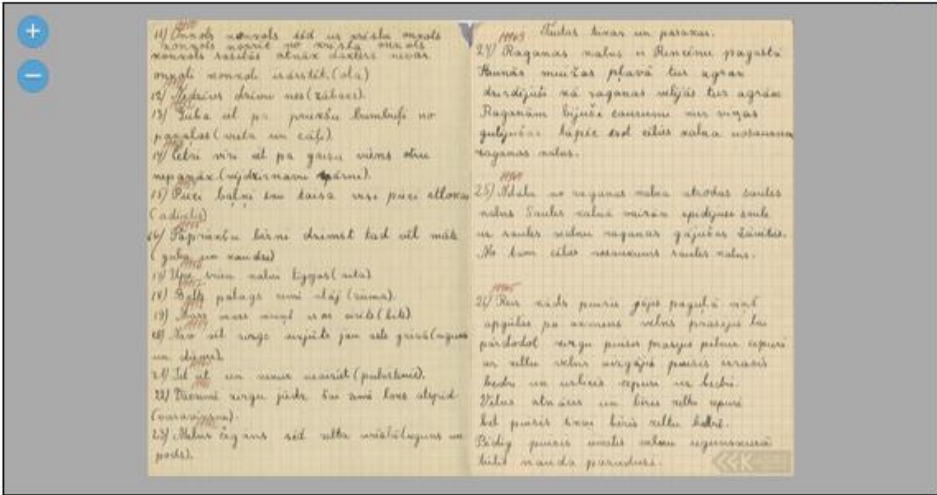
Iepriekšējā vienība Nākamā vienība



Atšifrēt

Vienība Faili Teksts Vietas

Fails



Vienības profils:

Origināla fails un tā transkripcija

Atšifrējums	
Atšifrēja	Sandis Laime
Pabeigts	Jā
Pārbaudīts	Jā
Izveidots	2017-10-10 16:11:12
Labots	2024-11-15 08:17:23
Atšifrējums	19950 11) Onkols konkols sēd uz krēslu onkols konkols nokrīt no krēsla onkols konkols sasitās atnāk dakters nevar onkoli konkoli izārstēt. (ola) 19951 12) Nedzīvs dzīvu nes (zābaks). 19952 13) Guba iet pa priekšu bumbuli no pakaļas (vīsta un cāļi). 19953 14) Četri vīri iet pa gaisu viens otru nepanāk. (vējdzimavu spēmi). 19954

Vienības profils:

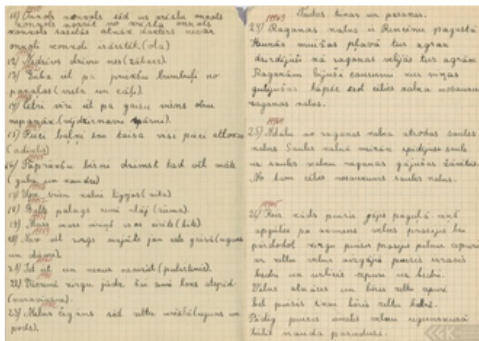
Rediģētais un markētais teksts

#LFK 116, 19963



Iepriekšējā vienība

Nākamā vienība



Atšifrēt

Iepriekšējā vienība

Nākamā vienība

Vienība

Faili

Teksts

Vietas

Atvērt

Raganas kalns ir Rencēnu pagastā Jaunās muižas pļavā. Agrāk dzirdējuši, ka tur raganas velējas. Tur agrāk raganām bijuši caurumi, kur viņas gulējušas, tāpēc esot cēlies kalna nosaukums – Raganas kalns.

Atvērt

Dažādas ortogrāfijas + fonētiskā transkripcija

Dzīdūt dzymu, Dzīdūt augu,
Dzīdūt dzējvi nūdzējivoju,
Dzīdūt nōvi īraudzēju,
Lowus godus pōrdzējivojsīd.
Pr. Smelters. un J. J. J. N. Kers.

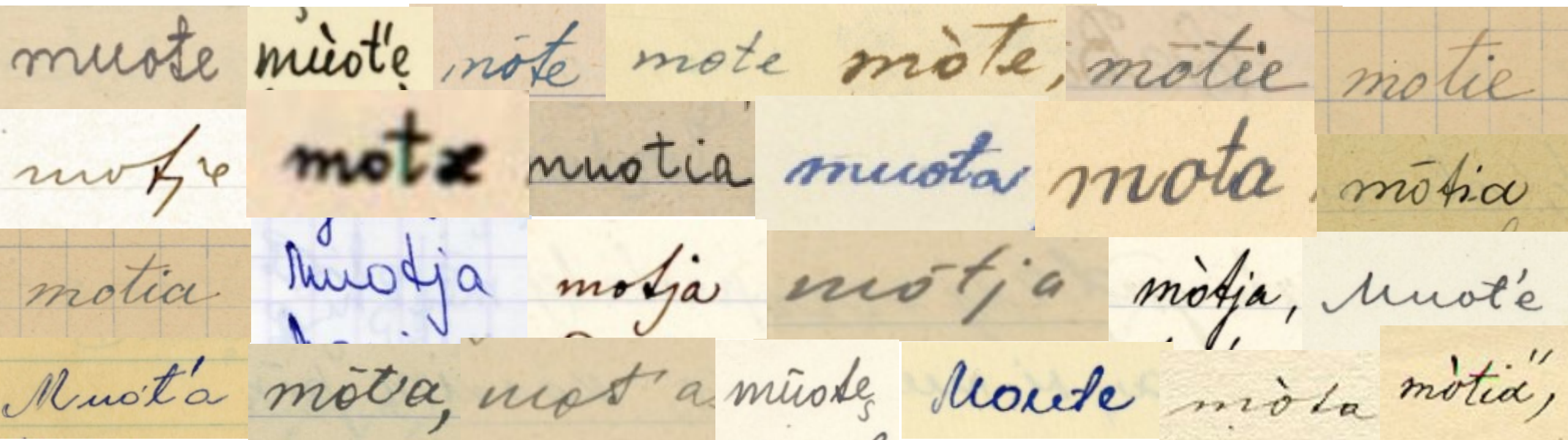
5352
Hydins jura uz arminā,
Prizās putu gabalins.
Hydins lautu bolelinis,
Stow mōsiga raudāma.

Lyma, lyma t'ei agl'eit'e,
Kur s'aul'eit'e nark'i guļ.
Suōfu, suōfu lys c'imeņis,
Kur aūk muna k'auda'ena.

Nu tolins es redzeju,
Saksām sagtu istabīnu,
Atved manim bolelini,
Nu to cīva leigavinu.

Pieraksta daudzveidība: *muote* (26 varianti)

muote, mūt'ē, mūotē, mōte, mōte, mote, mōtie, motie, motje, motæ, muota, mota, mōta, muotia, mōtia, motia, mōtiā; muotja, mōtja, mōtja, motja; muot'e, muot'a, mōt'a, mot'a; moute



Terminu izpratne

- **NORMALIZĒŠANA** – dažādos periodos, atšķirīgās ortogrāfijas tradīcijās un ar atšķirīgu lingvistisko kompetenci fiksētu rokraksta tekstu vienādošana pēc vienotiem kritērijiem.
- Normalizēšanas **MĒRĶIS** – folkloras teksta pieraksta nestandartizētu versiju pielāgošana oficiālās ortogrāfijas prasībām (tai skaitā, saglabājot izloksnes īpatnības), lai to varētu izmantot digitālajā pētniecībā, izglītojošiem nolūkiem, lai teksts lietotājam būtu lasāms vai saprotams, tai skaitā novēršot kļūdainas, nesalasāmas vai nekonsekventas un līdz ar to nesaprotamas valodas formas, saglabātu vai atjaunotu oriģinālu informāciju, kas tehnisku iemeslu dēļ pazaudēta vai nesaprotama.

Problēmjaautājumi

- **Vai un ciktāl pieļaujama** konkrētā dialektālā areālā pierakstītu pirmšķietami dialektāli autentisku **tekstu pielāgošana** attiecīgās izloksnes īpatnībām vai oficiālās ortogrāfijas normām?
- Kā panākt, lai normalizēšana **neklūst par teksta konstruēšanu?**

MĒRĶIS: uz divu LFK Latgales folkloras kolekciju piemēra vērtēt latgaliešu valodā pierakstīto tekstu normalizēšanas izaicinājumus:

- no vienas puses, respektējot jebkura teksta kā avota vērtību;
- no otras puses, tos publicējot tā, lai teksts būtu pilnvērtīgi izmantojams digitālajā pētniecībā, izglītības vajadzībām un citējams publikācijās.

Latgaliešu folkloras tekstu kvantitāti un atveidi ietekmējošie faktori

- Sava laika sabiedrības informēšanas akcijas – aicinājumi un norādes par folkloras pierakstīšanas nozīmi un prasībām.
- Folkloras pierakstītāju latgaliešu rakstu valodas kompetence – latgaliešu ortogrāfijas veidošanās process un latgaliešu valoda skolā.
- Folkloras pierakstītāju un teicēju piederība izloksnes areālam.

Kultūrvēsturiskais konteksts

- Sava laika sabiedrības informēšanas akcijas – aicinājumi un norādes par folkloras pierakstīšanas nozīmi un prasībām.

stòstu un dzismiu. — Nu Rēzeknes Tērdzības
školnikim šū pīmāru wajadzātu jemt ari cytōs
školōs, cytōs myusu Dzimtines pilsātōs. Myusu
tautas gorigajs monts: dzismes, pōsokas, parūnas
gaist ar kotru godu, jo waci cylwāki izmērst un
jauni jau atsoluši uz senejōm dzismem. Rēzeknes
mōcekli ar pulku war izglōbt nu pazusšonas daudz
dzišmu, daudz pōsoku. Un nōkūtne par tū jīm byus
pateiciga. R ē z e k n i t s.

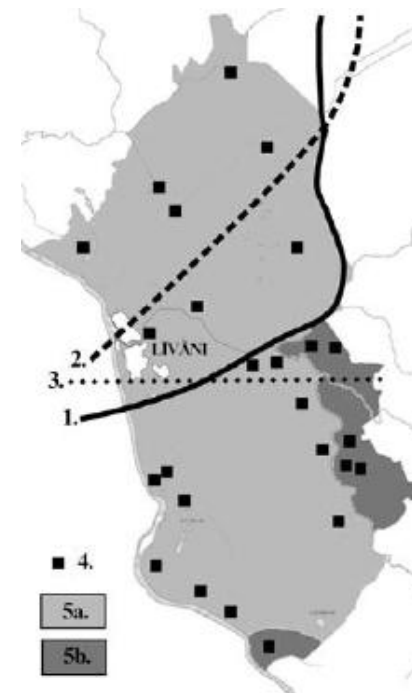
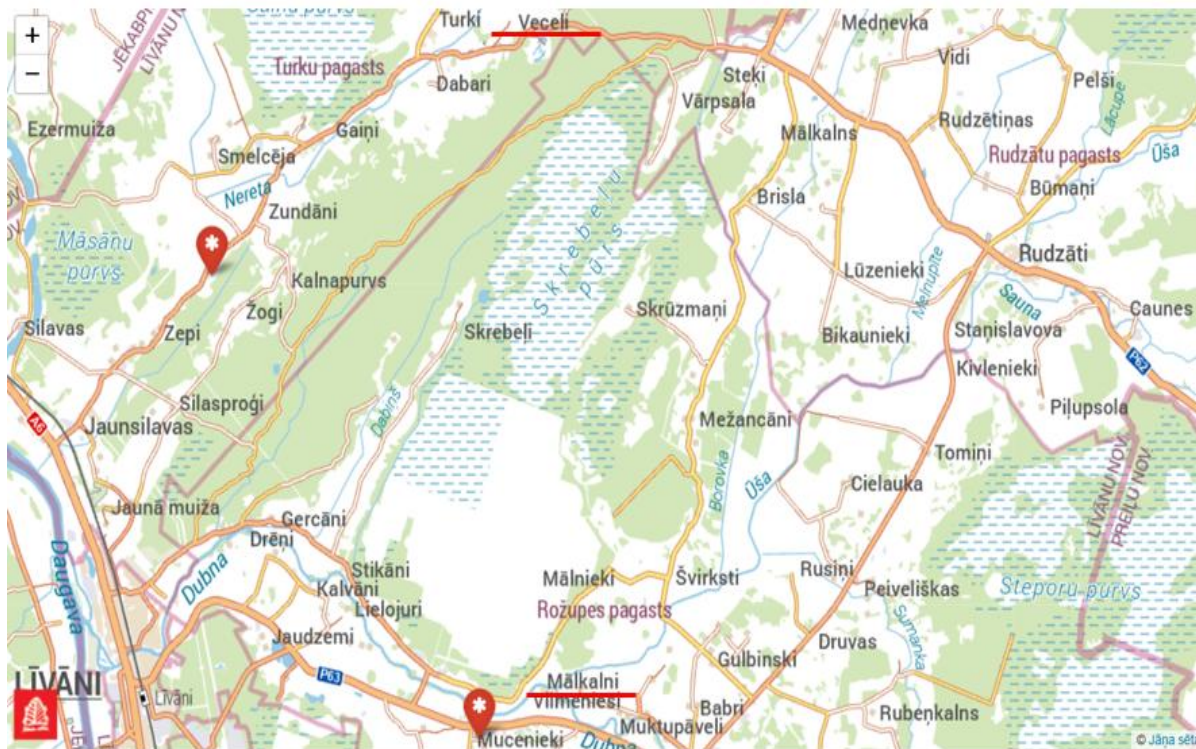
Jaunas Zinias. Nr.15. 22.03.1913.

- Berzkalne, Anna. Turpynojit tautas dzīsmu uzrakstīšonu. *Latgolas Words*. Nr.28. 08.07.1925.
 - Pierakstīt tikai dzirdētas, nevis jau nodrukātas tautasdziesmas;
 - Sniegt ziņas par folkloras teicēju, pierakstītāju, teksta teikšanas laiku un vietu,
 - Pierakstīt tekstu tieši tā, kā teicēji saka, neko ne pieliekot, ne izlaižot.
- Šmits, Pēteris. Lyugums lasītōjim. *Zīdūnis*. Nr.1-2. 01.01.1925. 80 (880).lpp.
 - Uzrunā «latgaliešu patriotus» iesūtīt latgaliešu pasakas un tradīcijas, pierakstot tās vietējā izloksnē, kādā tās paredzēts arī publicēt.

Latgaliešu valodas apguve skolās

- Likums par Latvijas izglītības iestādēm. (1919; 40.pants) Obligātā izglītība Latvijā iegūstama ģimenes valodā. Par skolēna ģimenes valodu atzīstama valoda, kuru norāda bērnu vecāki, piesakot bērnu skolā un kurā viņš var brīvi izteikt savas domas.
- Likums par valsts valodas lietošanu Latgales skolās (1922). Obligātais mācību saturs Latgales skolās paredzēja, ka divus pirmos gadus skolēni mācījās latgaliešu valodā, sākot ar trešo mācību gadu, sāka apgūt latviešu valodu.
- Latgaliešu valodas mācību stundās iekļautas mācību programmā arī vidējās izglītības posmā. (Aglonas reālgimnāzijas gatavības apliecība 1933.g.)
- Likums par valsts valodu (1935). Valsts valoda ir latviešu valoda.
- Latgaliešu valodas apmācība vēl turpinās zemākajās klasēs un ģimnāzijās, arī ticības mācība varēja notikt latgaliešu valodā (Izglītības ministra biedris J.Čamanis par Latgales skolām. *Valdības Vēstnesis*. Nr.36. 13.02.1937.)
- 1937.g. latgaliešu izloksnes skolotāji apstiprināti Jaunaglonas sieviešu ģimnāzijā, Aglonas ģimnāzijā. *Latgales Vēstnesis*. Nr.108. 24.09.1937.

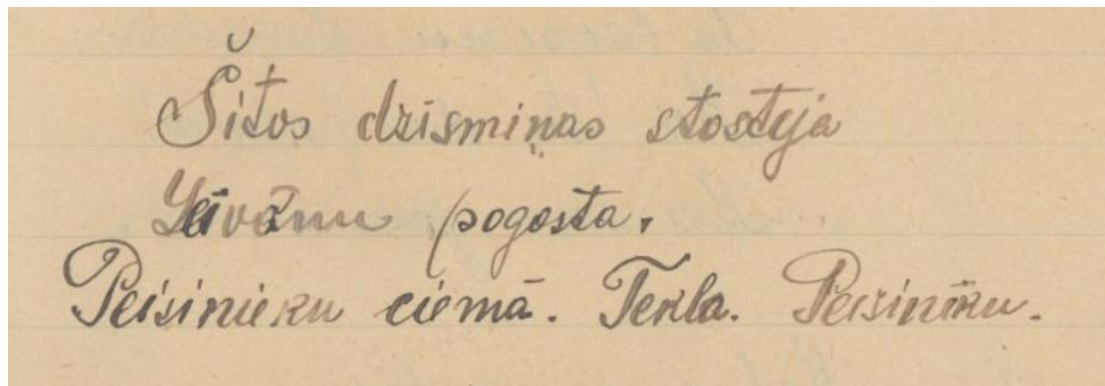
Folkloras pierakstītāju un teicēju piederība Līvānu izloksnes areālam: kolekcijas LFK [251] un LFK [935]



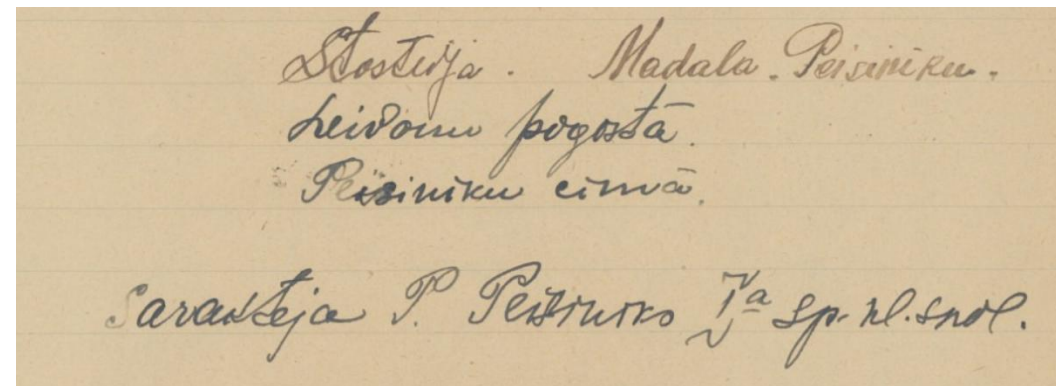
Kāpjošās (5a) un lautzās intonācijas (5b) areāli Līvānu izloksnē (Trumpa, 2012, 68). 1.līnija – V. Rūķes fiksētā kāpjošās un lautzās intonācijas izoglosa. 2.līnija – M. Rudzītes izoglosa. 3. līnija – Latviešu valodas dialektu atlanta izoglosa.

Kolekcija LFK [251]

- Pierakstīta un iesūtīta LFK 1926.g.
- Iesūtītājs – Aglonas reālgimnāzijas I.a speciālās klases skolnieks Pīters (Peters) Peisinīks.
- Kolekciju veido 85 tautasdziesmas. 78 dziesmas teikusi Līvānu pagasta Peisinīku ciema iedzīvotāja Tekla Peisinīka, 7 dziesmas – šī paša ciema iedzīvotāja Madaļa (Magdalena) Peisinīka.

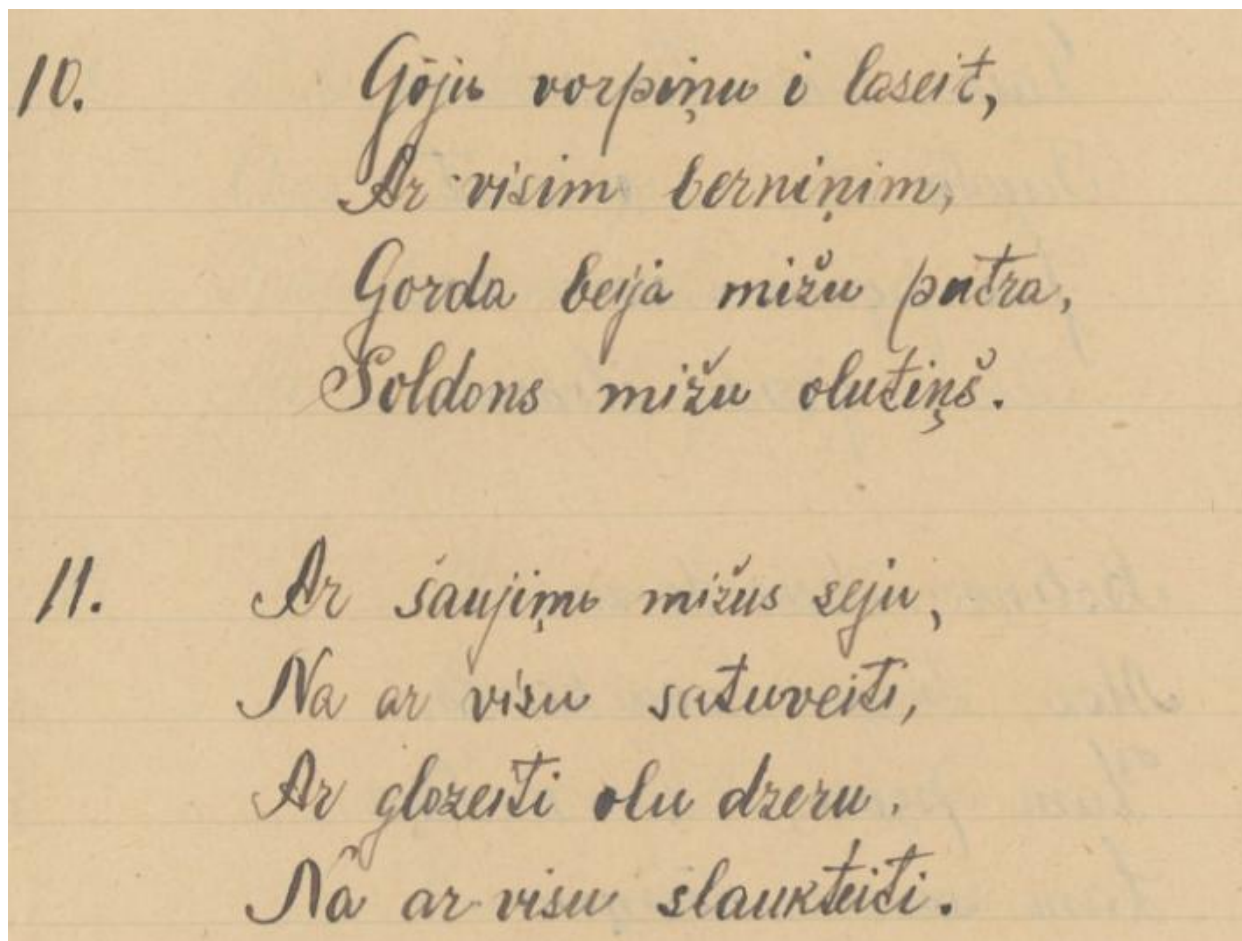


Šīs dziesmas stāstīja
Līvānu pagasta,
Peisinīku ciemā. Tekla. Peisinīka.



Stāstīja. Madala. Peisinīka,
Līvānu pagasta,
Peisinīku ciemā.
Sarakstīja P. Peisnīks 1926.g. 12. sept.

P. Peisinīka kolekcijas piemērs

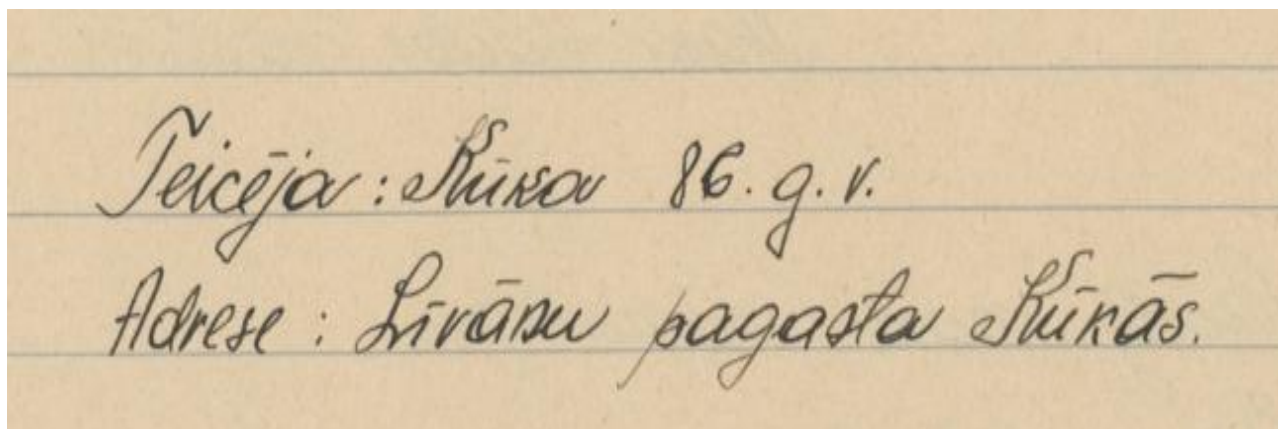


LFK 251, 10-11

<https://garamantas.lv/lv/file/1444429>

Kolekcija LFK [935]

- LFK saņemta 1937.g.
- Iesūtītājs – Daugavpils skolotāju institūta Vladislavs? Ancāns
- Kolekciju veido 40 tautasdziesmas burtnīcā ar nosaukumu «V.Ancāna vasaras darbs latviešu valodā»



Teicēja un pierakstīšanas vietas norāde LFK kolekcijā [935]

<https://garamantas.lv/lv/file/1126624>

Līvānu izloksnes apraksti

- Detalizētāko zināmo Līvānu izloksnes aprakstu 1940.gadā publicējis latgaliešu etnogrāfs, folklorists, valodnieks un rakstnieks Leonards Latkovskis (1905-1991). Izloksnes aprakstā Latkovskis balstās uz tai raksturīgākajām sēliskajai izloksnei atbilstošām fonētiskajām īpatnībām, kas aptver: intonāciju; -e- skaņas izrunu; latviešu rakstu valodas skaņai -ū- atbilstošos diftongus -yu-, -eu-, -ou-; latviešu literārās valodas skaņai -ā- atbilstošos refleksus -uo-, -ō- un formas *kô, tô, kôda, tôda*. Tā kā kolekciju pierakstā nav iespējams identificēt intonācijas un -e- skaņas izrunas savpatību, turpmāk rakstā salīdzinoši analizētas trīs no L.Latkovska nosauktajām Līvānu izloksnes fonētiskajām īpatnībām.
- Latkovskis Leonards. Līvānu izloksne. Filologu biedrības raksti. 1940.g. 20.sēj. Rīga: Latviešu filologu biedrības izdevums. 134.-158.lpp.

Diftongizēšanās:

-ū- pārvēršas par -yu-, -eu-, -ou-

Leksēma latviešu literārajā valodā	LFK [251]		LFK [935]	
	vienības Nr. kolekcijā	leksēmas atveide vārdformā	vienības Nr. kolekcijā	leksēmas atveide vārdformā
‘būt’	84	<i>byut</i>	12604	<i>beušu</i>
‘pūst’	55	<i>pyuta</i>	12636	<i>nūsapeute</i>
‘ūdens’	38	<i>yudinī</i>	12628	<i>eudins</i>
‘cūka’	76	<i>ciukom</i>	12614	<i>ceuka</i>
‘mūsu’	70	<i>myusu</i>	12606	<i>meusu</i>

- Papildus kolekcijā LFK [251] sastopami arī citi -*ū*- diftongizēšanās piemēri, kur *ū* > *yu*: 'mūžs' > *myužu* (28), 'lūgt' > *lyudza* (56), 'mūžs' > *myuža* (28), 'šūpot' > *šyupoj* (38), 'pūrs' > *pyurs* (47) u.c., kas ļauj šo pazīmi atzīt par tipisku Peisinīka latgaliešu rakstu valodas atveidē;
- Arī kolekcijā LFK [935] sastopami citi -*ū*- diftongizēšanās piemēri, kur *ū* > *eu*: 'grūti' > *greut* (12605), 'jūgt' > *jeudze* (12611), 'dūmi' > *deumi* (12615), 'brūklene' > *breukleneiti* (12633), 'grūst' > *greud* (12609) u.c.

Tīra gara -o- skaņa

Leksēma latviešu literārajā valodā	LFK [251]		LFK [935]	
	vienības Nr. kolekcijā	leksēmas atveide vārdformā	vienības Nr. kolekcijā	leksēmas atveide vārdformā
‘gāju’	5	<i>gōju</i>	12626	<i>nūgōju</i>
	82	<i>goju</i>	12609	<i>goja</i>
‘māte’	18	<i>mote</i>	12630	<i>mōte</i>
			12600	<i>motes</i>
‘zāle’	64	<i>zolu</i>	12630	<i>zōle</i>
			12627	<i>zole</i>
‘kājas’	22	<i>kojas</i>	12637	<i>kōju</i>
	15	<i>kōjiņom</i>	12611	<i>kojys/kojas</i>
‘arājs’	2	<i>orojiņš</i>	12623	<i>orōjeņi</i>

- Lai gan **kolekcijā LFK [251]** parādās arī latviešu valodas patskaņa **-ā- atbilsme** burtam **-ō-** (*kōjas* ‘kājas’, *izgōju* ‘izgāju’, *tautiņōs* ‘tautiņās’, *ōrā* ‘ārā’, *mōsiņ* ‘māsiņ’), tomēr **dominē pieraksts ar īso -o-** (*broliņos* ‘brāliņos’, *orojins*, *Mortiņim* ‘Mārtiņiem’, *tautiņos* ‘tautiņās’, *toļu* ‘tālu’, *vōrpiņu* ‘vārpiņu’ u.c.). Tāpat kolekcijā LFK [251] burts **-o-** nav konsekventi pierakstīts viena un tā paša vārda saknē, piemēram, vārdos ‘gāju’ (*gōju* (3) / *goju* (17)); ‘kājas’ (*kojas* (22) / *kōjiņom* (15)); ‘māsiņa’ (*mosiņa* (31) / *mōsiņ* (30) u.c.
- Līdzīgi nekonsekvences vērojamas arī **kolekcijā LFK [935]**, piemēram, ‘māte’ (*motes* (12600) / *mōte* (12630)); ‘bāliņš’ (*boleņa* (12629) / *bōleņam* (12625), ‘zāle’ (*zole* (12627) / *zōli* (12627) u.c., tomēr **pierakstā dominē burts -ō-, kas ļauj pielaut vismaz daļēju kolekcijas atbilstību Līvānu izloksnes īpatnībām**. Šādi vārdi ir, piemēram, *tōs* ‘tās’, *tōli* ‘tālu’, *prōteņa* ‘prātiņa’, *mōseņa* ‘māsiņa’, *vōrnys* ‘vārnas’, *nōve* ‘nāve’, *nōc* ‘nāc’ u.c.
- Nevienā no abām kolekcijām nav sastopama L. Latkovska fiksētā īsā patskaņa **-a- atbilsme** garajam **-ō-**. Ņemot vērā, ka kolekcijā LFK [251] dominē **-ā-** reflekss ar īso **-o-**, un kolekcijā LFK [935] **-ā-** refleksi nav konsekventi lietoti, tikai nosacīti atzīstama kolekcijas LFK [935] atbilstība Līvānu izloksnes īpatnībām, skatot to citu izloksnes īpatnību kontekstā.

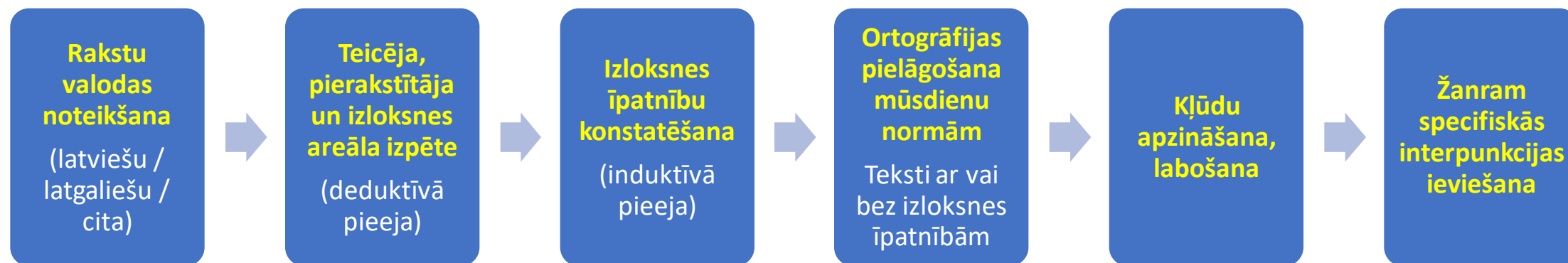
kô, tô, kôds, tôds lietojums

Leksēma latviešu literārajā valodā	Kolekcija LFK [251]		Kolekcija LFK [935]	
	vienības Nr. kolekcijā	leksēmas atveide vārdformā	vienības Nr. kolekcijā	leksēmas atveide vārdformā
‘kā’	81	<i>kai</i>	12623	<i>kai</i>
			12623	<i>ko (lōceiši)</i>
			12623	<i>kō (šeuneņi)</i>
‘kāda’	80	<i>kaida</i>		
‘tā’			12632	<i>to (saceja)</i>
‘tāda’			12614	<i>toda (kūka)</i>

Salīdzinājums rāda, ka **kolekcijā LFK [251]** konsekventi lietotas formas *kai* ‘kā’ (14 reizes) un *kaida* ‘kāda’ (divas reizes), kas atbilst latgaliešu rakstu valodas normatīvajam regulējumam un **neatspoguļo izloksnes īpatnības**.

Savukārt **kolekcijā LFK [935]** sastopamās formas *ko* (vienu reizi) / *kō* ‘kā’ (piecas reizes), *toda* ‘tāda’ (vienu reizi) **ir izloksnes īpatnība, par kuras atveidi rakstu formā jāņemj folkloras tekstu normalizēšanas procesā**.

Tekstu normalizēšanas process



Secinājumi

- Deduktīvā pieeja palīdz saprast kolekcijas kontekstu, bet tai salīdzinoši neliela nozīme tekstu normalizēšanas procesā
- Būtiska nozīme, pieņemot lēmumu par tekstu normalizēšanas principiem, ir induktīvajai pieejai
- Normalizējot tekstu, tiek respektēti arī izloksnes / oficiālās ortogrāfijas paralēlvarianti, kas nav pretrunā ar latgaliešu valodas normām
- Būtiskākie izaicinājumi: tekstu redaktoru kompetence (zināšanas dialektoloģijā), izlokšņu īpatnību atpazīšana, izlokšņu aprakstu pieejamība