



Corpus Literacy in the Latgalian Language Learning Process: Proper Noun Acquisition Opportunities

Antra Kļavinska

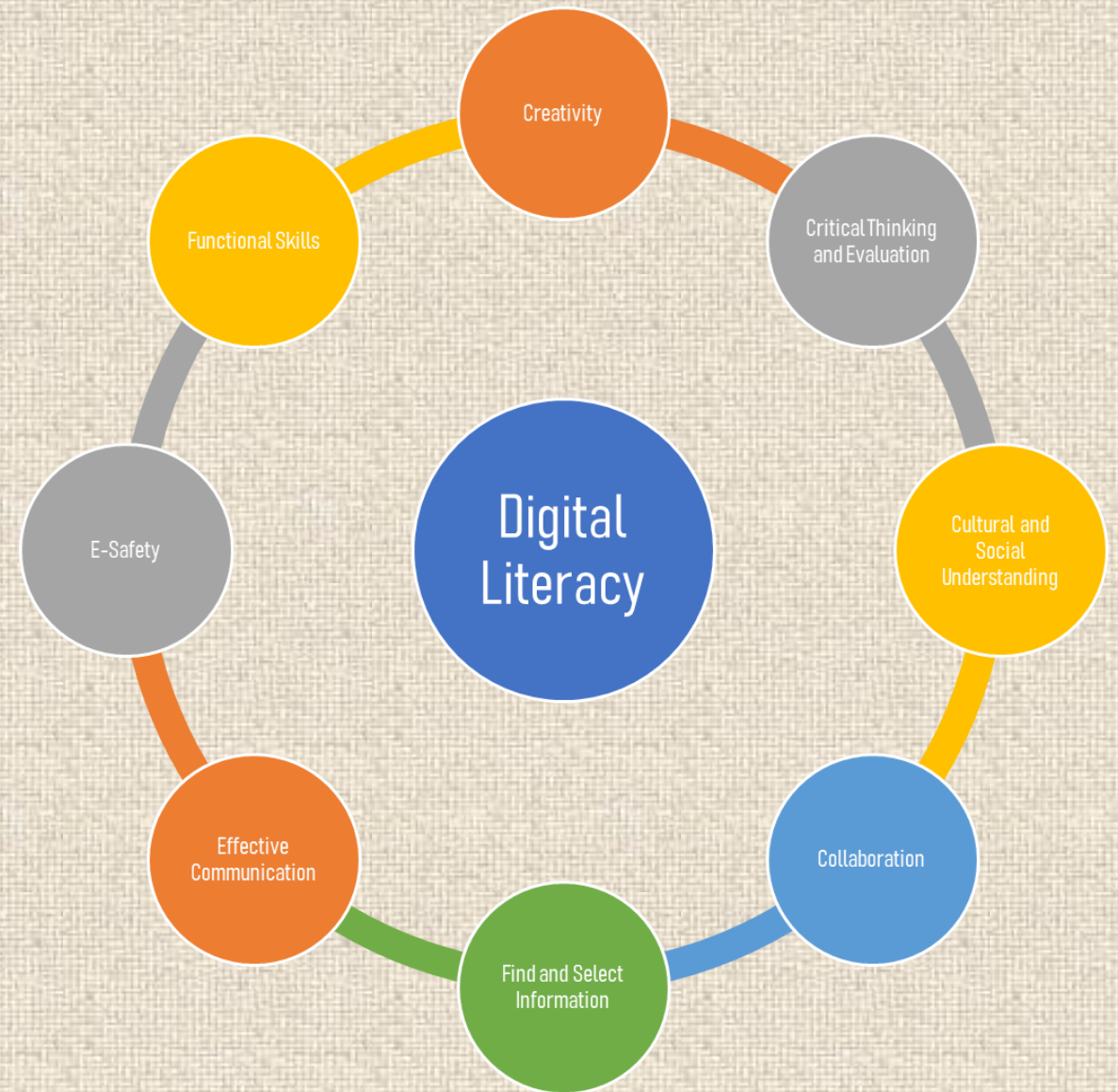
Storptautyskuo zynuotniskuo konference “Regionaluos volūdys izgleiteibā: nu teksta leidz tekstprateibai” 2022. gods 16.-17. junī Rēzeknē

Overview

- Corpus literacy & digital literacy.
- Modern Latgalian Language Corpus (MuLa) and other digital resources. Are they used in the Latgalian language learning process?
- Teacher education and corpus literacy.
- Some practical examples for learning proper names.

Corpus Literacy -

“the ability to use the technology of corpus linguistics to investigate language and enhance the language development of students” (Heather & Helt, 2012, 417).



Components of Digital Literacy (Hague & Payton 2010, 19)

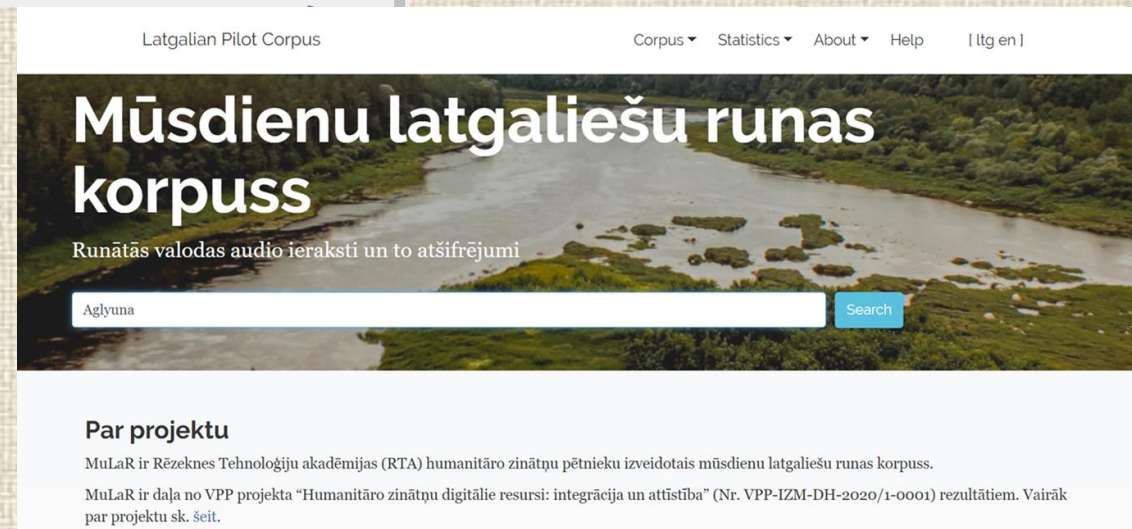
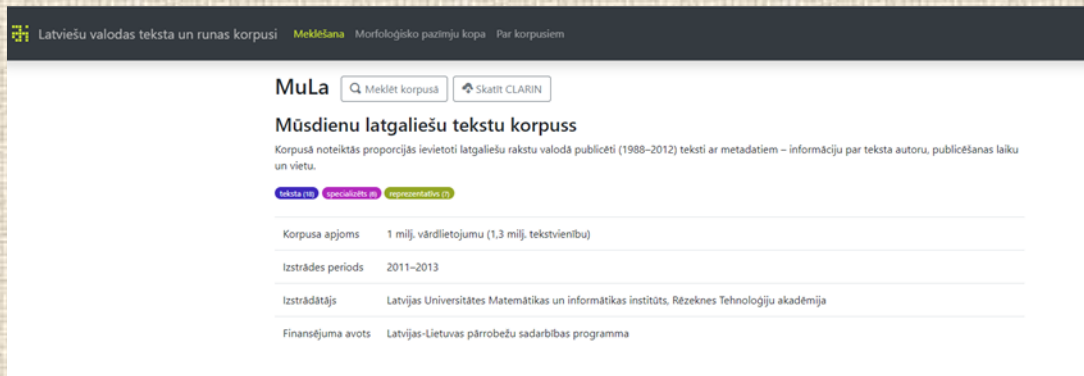
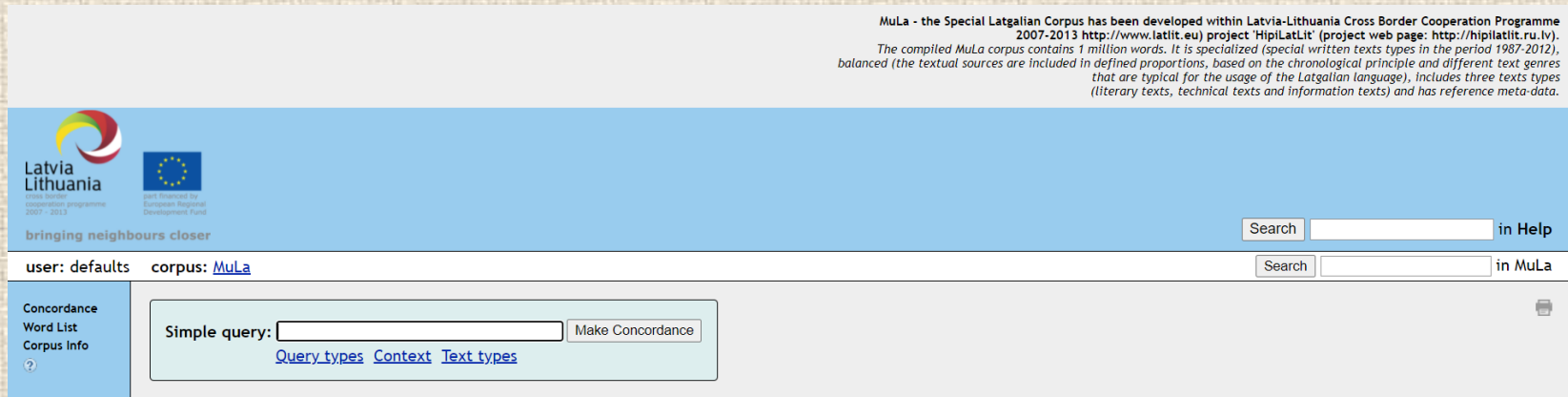
Storptautyskuo zynuotniskuo konference “Regionaluos volūdys
izgleiteibā: nu teksta leidz tekstprateibai”

Subcomponents of corpus literacy

(Mukherjee, 2006, 14):

- understanding basic concepts of corpus linguistics (What is a corpus and what types of corpora are available and how? What can you do—and cannot do — with a corpus?);
- searching corpora and analysing corpus data by means of corpus software tools, e.g. concordancers (What is corpus software and how can it be used to search a corpus? How can corpus output be analysed?);
- interpreting corpus data (How may general trends in language use/change be extrapolated from corpus data?);
- using corpus output to generate teaching material and activities.

Modern Latgalian Language Corpus (MuLa 2013) and Latgalian Speech Corpus (MuLaR 2022 – pilot corpus)

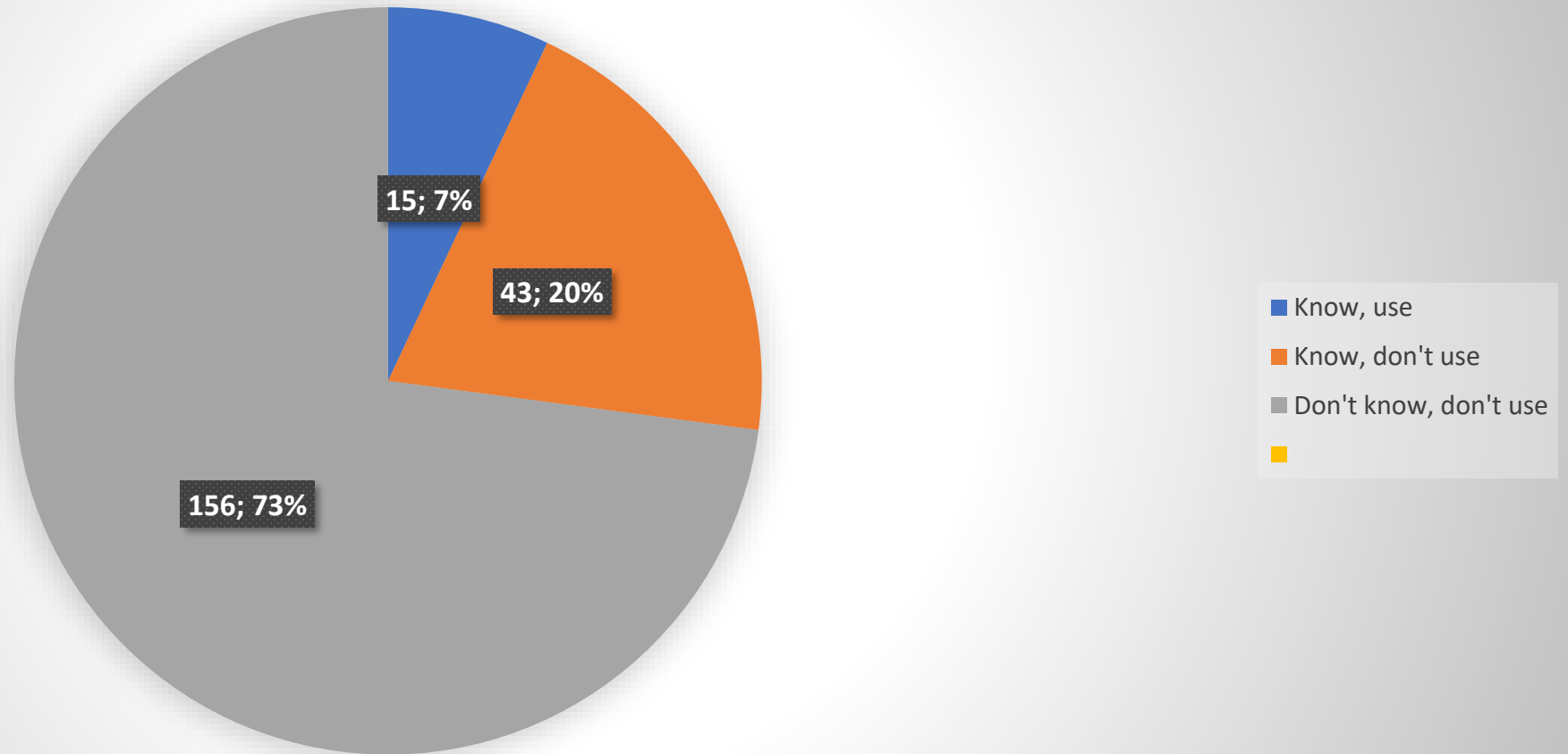


Storptautyskuo zynuotniskuo konference “Regionaluos volūdys
izgleiteibā: nu teksta leidz tekstprateibai”

Survey on the use of the MuLa corpus (2021)

(Martena, Naua, Briška, 2022)

214 respondents in total



Expert opinion on digital resources of Latgalian language

MuLa naasu lītuojuse, bet tagad ir interese pasaskatēt! (E1, 2022)

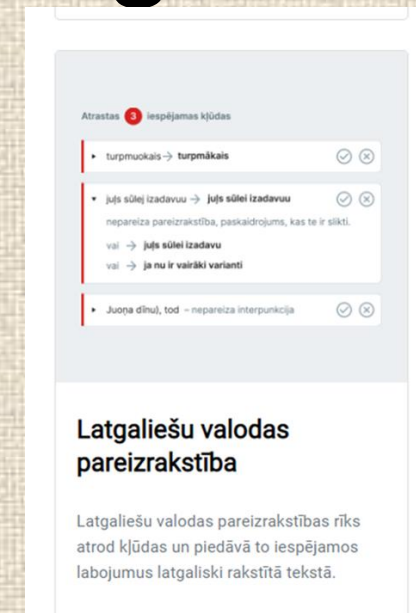
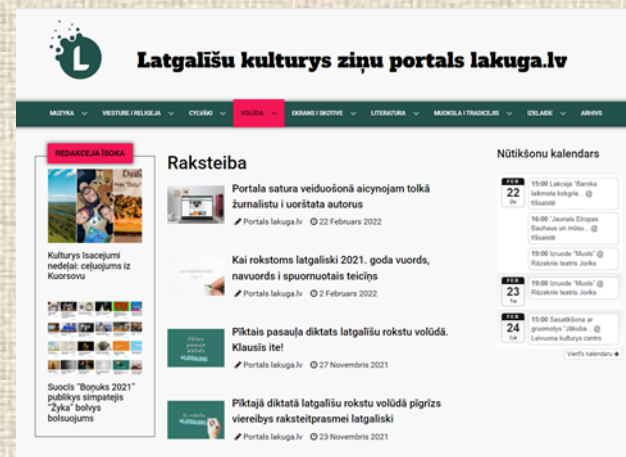
Ja arī izmontoju korpusa datus i kod ir atšķirīgys formys, tod maņ pošai svareigs kriterejs ir teksta autors - kurs tū ir rakstejs voi izdeviejs. (E4)

“Pyрмаis nūteikti ir HipLatLit projekta latgalīšu-latvīšu-lītuvišu vuordineica. Tod ira vuordineica.lv, kas ira taidš breivpruoteigū projekts (..)” (Husare, “Kolnasāta” 04.04.2020.)



Expert opinion on digital resources of Latgalian language

- (**..**) **hugo** pyрма loba laika **nazkas interesants** beja izliecs, par k  pasasmieju, **i vaira nasavieru** (E2, 2022).
- Kod juopasaver kaidi **v tvuordi** latgaliski, ir **Lakug  vol dys sada a "Raksteiba"** (**..**) Dzerau u n saukumus pasaveru **Vietv rdu datub z ** (E3, 2022).



Storptautyskuo zynuotniskuo konference “Regionaluos volūdys
izgleiteibā: nu teksta leidz tekstprateibai”

Educating Latgalian Language Teachers

- **The aim** is to prepare video lectures (conduct seminars) promoting the opportunities offered by the MuLa and MuLaR corpora in the Latgalian language teaching / learning process, thus also raising the overall corpus literacy level.
- The description of the linguistic competence in the example of the curriculum for the subject **“Latgalian written language for grades 1 – 9”** states that in grades 7 – 9 students **“are able to use (..) a Latgalian language corpus (..) for learning to obtain required information”** (Andrejeva et al., 2020, 94).

Some examples for the acquisition of proper nouns (video lecture developments)

- **Topic “Our Names” (grades 4 – 6)** (Andrejeva et al., 2020, 42)
- **Communicative competence:** Is learning to obtain information from visually and auditory presentation of text and to use it in spoken communication. Discusses the use of personal name variants in everyday life in different communicative situations.
- **Linguistic competence:** Is learning to differentiate between older and modern first names, their meaning, and the full and shortened versions of first names. Acquires the spelling of forenames, their diminutive forms and hypocorisms in accordance with the norms of Latgalian written language grammar and orthography.
- **Sociocultural competence:** Understands that a personal name is a mark of a person's identity and can provide information about the person's belonging to a specific culture.

How do you understand this joke?

- LTG *klāvs* 'cattle-shed'
- *īpasazeisim* / *īsapazeisim*
- What other personal names known in Latvia could cause misunderstandings in the Latgalian language? (Olga, Nadīna, Monta...)
- Do you know people who are called *Pīters*, *Juons*, or *Muora*?
- What names were popular in Latgale before and what are popular now?
- Is the pronunciation of your name different in Latvian and Latgalian?
- How do you address each other in the family, among friends?



Searching examples by means of corpus software tools, e.g. concordancers, generating teaching materials (LVK 2018 – MuLa 2013)

CONCORDANCE LVK2018

simple **Klāvs** 80 (6.51 per million)

☐ Details ☐ Left context ☒ KWIC ☐ Right context

61 ☐ doc#15482 n, paņēmuši bises, gāja medībās. </s><s> Mātes nomierināja **Klāvu** un visādi darīja, lai aizdzītu lielo pārdzīvojumu, citādi šeit arī

62 ☐ doc#15482 adrūmuši, tukšā. </s><s> Ilgi runājās ar mātēm. </s><s> Tikai **Klāvs** nosēdās un rotaļājās ar Lauras lellēm. </s><s> Tagad tās pied

63 ☐ d

64 ☐ d

65 ☐ d

66 ☐ d Klinti, Doku Ati, aktierus un gleznotājus, kas guldīti Mākslinieku kalniņā, un vēl un vēl. </s><s> Īpaši skaists skats pavērās, krēslā

67 ☐ d uzkāpjot kalnā pie Jāņa Poruka. </s><s> Savukārt Raiņa kapos svecītes iededza, protams, pie Raiņa pieminēkļa, godināja

68 ☐ d Aleksandru Čaku, Leo Svempu, **Klāvu** Elsbergu un desmitiem citu rakstnieku, aktieru un režisoru. </s><s> Jā, turpat blakus

69 ☐ d apglabāti arī padomju politiskie darboņi, taču tiem bija jāapmierinās ar pāris " kroņa" sarūpētām svecēm. </s><s> Jo tautas

mīlestību un cieņu nevar nedz nopirkt, nedz izkalpoties. </s><s> Arī tagad Mirušo piemiņas dienā cilvēku

LVK 2018

MuLa 2013

7 ☐ doc#101 un zemeiti daleit. naboqūs vin Īzakulsi. </s><s> Pi tam akminu **klāvs** kai aqrōk. tai taqad daudzūs mudvnoīa skaudeibu. ari Zuzanny

8 ☐ d

9 ☐ d tys gribēja atsasaceit. — Tevi nūdzirdeit nivīns nasataisa, un "Lapaču" saimē nav īroduma dzert dzeršonas pēc. </s><s> Bet pēc

10 ☐ d kōrteiga dorba pa kauseņam loba monta varim vysī. </s><s> Pēc pīcom stuņdem ražeiga un labi organizeita dorba **klāvs** nu

11 ☐ d māsīm beja tukšs, kai izslauceits. </s><s> Puiši vēl izkaiseja vysu ar sausom cysom, tod gōja mozgōt zōbokus iz tyvejū upeiti, un

12 ☐ d tod jau Katre vysus aicynōja pi golda. </s><s> Cik kārdynūši smaržōja sutynōti kōpusti ar cyukas

13 ☐ d

i kur tod

gūvis ar te

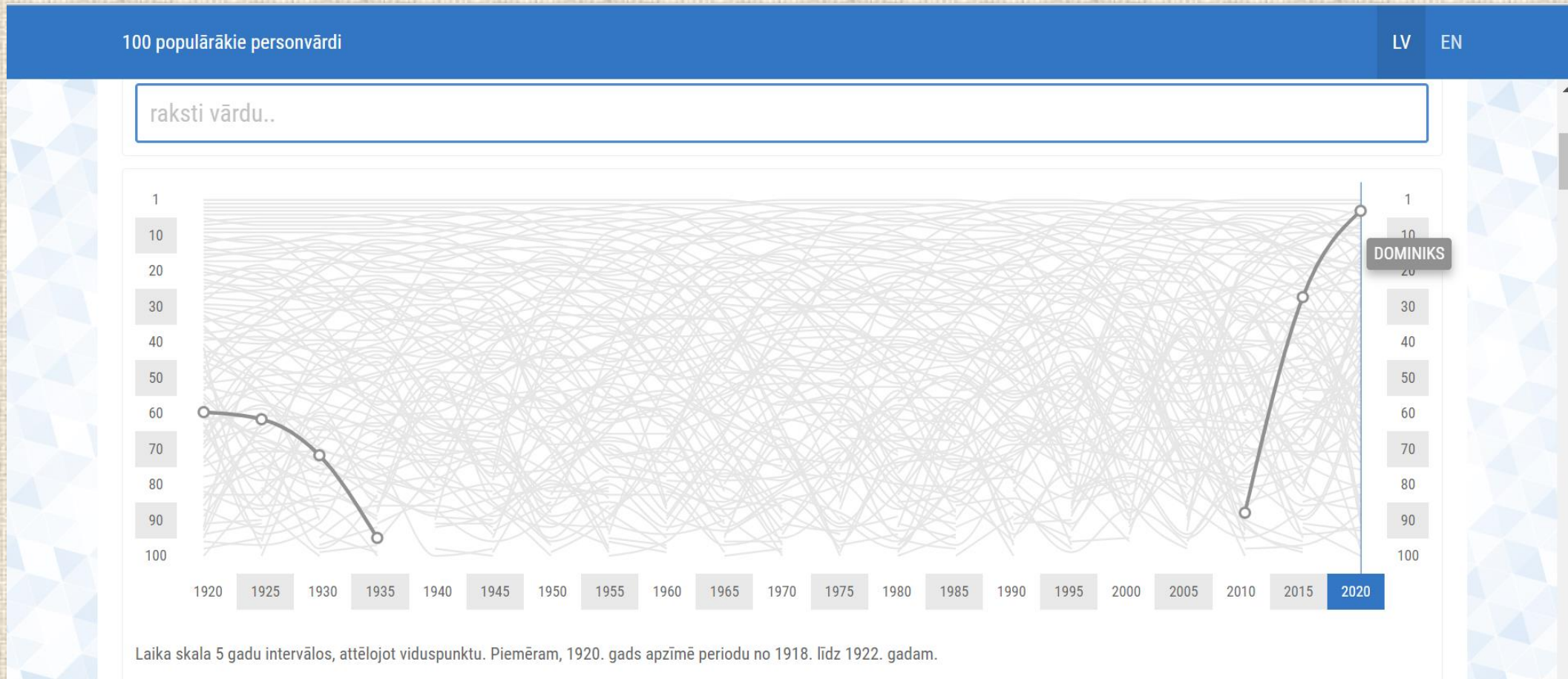
> Pi klāv:

ībuozta i

vītas. </s>

tā ir maza

Other digital tools for generating activities, e.g., <https://tools.csb.gov.lv/names/lv> **The 100 most popular personal names in Latgale**



Topic “Latgalian Cities and Towns and Municipalities” (grades 4 – 6). (Andrejeva et al., 2020, 65)

- **Communicative competence:** Working with a partner or in a group, gets acquainted with one settlement in Latgale and using different information sources prepares a virtual tour for classmates.
- **Linguistic competence:** Knows how to write the names of settlements and their residents in Latgalian.
- **Sociocultural competence:** Is aware that toponyms are an important part of the linguistic landscape and carriers of historical and cultural information, a cultural treasure.

Create the city sign / city slogan in Latgalian!

- How are place names pronounced in Latgalian (Latgalian dialects)?
- How to write a settlement name (place name form) in Latgalian?
- How to name the residents of the settlement?



Karteņa: jauno Bolvu nūvoda rubeža zeime, foto nu Bolvu nūvoda pošvaļdeibys publicitatis materialu



Publicitātes foto

How to pronounce / to write the name of a settlement in Latgalian?

- [MuLa 2013](#)

CONCORDANCE MuLa

character **Rēzeknī** 1,089 (821.61 per million)

☐ Details ☐ Left context ☐ KWIC ☐ Right

321	☐	☐	doc#372 "lyudžānī" i olimpiadi vydsškolānim latgaļu volūdā i kulturviesturē. </s><s> Olimpiade nūtkis š.g. 28. aprēlī Rēzeknēs 1. v. š.
322	☐	☐	doc#375 "Kai dzejas komuļš pa dzeiporam teits," — tai par izstodi saceja pr. </s><s> V. Naglis tās atklāšonas breidī Rēzeknē , kod s
323	☐	☐	doc#380 ias </s><s> NABEIDZAMŌ DISKUSIJA </s><s> Nupat izšķirstēju jaunūkū literari mōksliniškū almanahu " Rēzekne 2006"
324	☐	☐	doc#381 kas cyts. </s><s> Vinim - pasaule olimpiade voi pasaule šovs ar Lady Gaga, cytim - pasaule latgaļu saits Rēzeknē </s>
325	☐	☐	doc#381 daudz ministru i Saeims deputatu Latgolys reģionā organizātā konferencē nabeja saguojs redzēt. </s><s> Rēzekne metafi
326	☐	☐	doc#381 saguojs redzēt. </s><s> Rēzekne metaforiski teik saukta par Latgolys sirdi. </s><s> Reiga ir golvyspiļšāta, Rēzekne - sirds
327	☐	☐	doc#381 dspļšāta </s><s> I tai laikam ir - na vysu, kas nūteik ap latgaļu volūdū, var izskaidruot ar pruoitu. </s><s> Rēzeknē ir radz
328	☐	☐	doc#383 :iksnis, žurnala "Katōlu Dzeive" atbildeguo redaktore Maruta Latkovska, literate i žurnaliste Līga Seiksta nu Rēzeknis , Juo f
329	☐	☐	doc#385 ss kaida apzineiga dāma, kura uzklōs gultu. </s><s> "Vuolyudzānī" i olimpiade </s><s> ltuo goda 24. aprēlī Rēzeknis 1. vyd
330	☐	☐	doc#385 ure, </s><s> trešo vīta – Dagdys vydsškola, sagataviejs Juons Platacs. </s><s> Aizdavumus sagatavuoja Rēzeknis Augst
331	☐	☐	doc#385 reofesore Lideja Leikuma. </s><s> Atbalstieja Kulturkapitāla fonds, Latgolys Kultūrys centra izdevnīceiba i Rēzeknis Augst

CONCORDANCE MuLa

character **Riezeknī** 17 (12.83 per million)

☐ Details ☐ Left context ☐ KWIC

1	☐	☐	doc#29 ruoda jauna tradiceja – egleitis dadzynuošona. </s><s> Storp cytym nūsaukta ari dryanīte, tūlaik viel Riezeknis
2	☐	☐	doc#94 du, a Zīta tikū kai beja atsagrīzuse nu „Avikstis svātuku” Lubuonā. </s><s> Duovonā nu Germanišku – Riezeknis
3	☐	☐	doc#172 pasaruoda, paspeist, pasamaniuos - a vysam puoni Atbreivuošonys alejis mašynu ryukšona. </s><s> Riezeknis
4	☐	☐	doc#172 iošonys alejis mašynu ryukšona. </s><s> Riezeknis pilsātā i ļaužu pruoūtōs -pavasars. </s><s> LgCS Riezeknis
5	☐	☐	doc#172 kumā sastuodeitais dorbu plans ir izpildēits. </s><s> Prica pošim i cytim! </s><s> Asam dasadalejušs Riezeknis
6	☐	☐	doc#172 mdinēs, ūtardinēs, trešdinēs nu 16.00 da 19.00. </s><s> Aktivu i radūšu pavasari įjusim vīelej LgSC Riezeknis
7	☐	☐	doc#722 ī". </s><s> Piec lekcejis - čajs, dzīsmis i vysī soviejī. </s><s> Catūrtīn, nojabra 28 d., 18.00 stuņdēs, Riezeknes
8	☐	☐	doc#722 is A.Rancāne, O.Kōjojs i cyti. </s><s> Sastīn, dekabra 28 d., vokorā vysī Čyguoni byus meili gaidieti Riezeknis
9	☐	☐	doc#756 ar vokorā ļaužu ass bejs gona. </s><s> Redz, kū Anna pošā roksta par LaLaDra atskanem. </s><s> * Riezeknē

- [MuLaR pilot corpus](#)

MuLaR

Korpuss Statistika Par Palīdzība [ltg en]

Query: [word="Rēzeknī.*%c"]
Number of results: 7

☐ Hide metadata

Recording	Transcription
☐ 0:00 / 0:00 ☐	janvari febrali sa - sasatikt Rēzeknī pi Edeitis ha Person ID: Year of birth: Sex: Place: ☐
☐ 0:00 / 0:00 ☐	vedem uz Rēzeknī puordūt maizi graudus i Person ID: M104-1928 Year of birth: 1928 Sex: m Place: Andrupene ☐
☐ 0:00 / 0:00 ☐	es pats braukovu uz Rēzeknī vodovu ar zyrgu Person ID: M104-1928 Year of birth: 1928 Sex: m Place: Andrupene ☐

MuLaR

Korpuss Statistika Par Palīdzība [ltg en]

Query: [word="Riezeknī.*%c"]
Number of results: 7

☐ Hide metadata

Recording	Transcription
☐ 0:00 / 0:00 ☐	pūdņika skūla kas ar įusu įusu pūlīm įusu pūlī Riezeknī galā ar pūlīm struođoj kūpā i pats pateikamuokais ir tys ka Person ID: M107-1972 Year of birth: 1972 Sex: m Place: Audeja ☐
☐ 0:00 / 0:00 ☐	Riezeknī jū ryna laikam pīļņeigi vysī tai ka Person ID: M315-xxxx Year of birth: 1993 Sex: m Place: unknown ☐
☐ 0:00 / 0:00 ☐	tolka rozaleiguoja i suoka Riezeknī bombardēt Person ID: F124-1929 Year of birth: 1929 Sex: f Place: Čibla ☐

Storptautyskuo zynuotniskuo konference “Regionaluos volūdys izgleiteibā: nu teksta leidz tekstprateibai”

Check which forms of place name are found in the corpus? Do they comply with contemporary spelling rules?

- MuLa 2013

✓ Show frequency per million

Word	↓ Frequency	Frequency per million
1 Rēzeknis	314	236.90
2 Rēzeknes	303	228.60
3 Rēzeknē	296	223.32
4 Rēzekne	78	58.85
5 Rēzekni	56	42.25
6 Rēzeknei	13	9.81
7 Rēzeknī	9	6.79
8 Rēzeknes-Agļunas	6	4.53
9 Rēzekneit	3	2.26
10 Rēzeknis-Agļunys	2	1.51
11 izsokaRēzeknei	1	0.75
12 Sita-Rēzekne	1	0.75
13 Rēzeknīšim	1	0.75
14 Rēzeknīši	1	0.75
15 Rēzeknītei	1	0.75
16 Rēzeknīte	1	0.75
17 Rēzekne-Kuorsova	1	0.75
18 Rēzekneitē	1	0.75
19 Rēzekne-Ludza	1	0.75

Rows per page: 20 1-19 of 19

- Latgaliešu valodas pareizrakstības rīks

Valodu rīki ▼ Par Hugo Palīdzība HUGO.LV

Sākums / Valodu rīki / Latgaliešu rakstu valodas pareizrakstības rīks

Rēzekne

Vārda analīze

Rēzekne

Rēzekne

iešu dzimte; vienskaitlis; vokatīvs
iešu dzimte; vienskaitlis; nominatīvs

nas tabulu

Locīšanas tabula

Lietvārds Rēzekne

	Vienskaitlis	Daudzskaitlis
Nominatīvs	Rēzekne	-
Ģenitīvs	Rēzeknis	-
Datīvs	Rēzeknei	-
Akuzatīvs	Rēzekni	-
Instrumentālis	ar Rēzekni	-
Lokatīvs	Rēzeknē	-
Vokatīvs	Rēzekne!	-

Storptautyskuo zynuotnyskuo konference “Regionaluos volūdys
izgleiteibā: nu teksta leidz tekstprateibai”

In what texts and contexts is the word used?(Collocation, metadata.) [MuLa 2013](#)

a luoču meitine. </s><s> Oskars Seiksts (1973), dzimts **rēzeknīts** . </s><s> Itei ir catūrtuo gruomota, kurā jys pīsadola. </

Document number 🔍 (1)

☐ Token number	293175
☐ Document number	319
☒ Autors	Seiksts Oskars, Gagaine Līga
☐ Avots	Reitišku mads
☐ Nosaukums	Reitišku mads
☐ Publicēts	2000.gads

1	☐	☐	doc#39	ītas dzīsmes vārdus drūši var saceit par HELENU LĀCI, kuru rēzeknīši pazeist kai energisku, dzeivesprīceigu un mērtīceigu cylvāku	
2	☐	☐	doc#39	ī pi posta styura aug. </s><s> Tī beja citi laiki. </s><s> Dēļ tō rēzeknīši varēja byut lapni par sovu piļsātu, kas beja nūvērteita par vyst	
3	☐	☐	dc		
4	☐	☐	dc		
5	☐	☐	dc	, kod beja īplānōts kū byuvēt, bet gon projektu pōrtaiseja un objektu par pōrs metrim pastyume sōņš, lai, pīmāram, tam ūzulam natraucej, kas pi posta styura aug. </s><s> Tī beja citi laiki. </s><s> Dēļ tō rēzeknīši varēja byut lapni par sovu piļsātu, kas beja nūvērteita par vysteirōkū un vyszaļōkū vysā Latvijā, par vīnu nu zaļōkōm Padūmjū Savīneibā, prūtams, pasateicūt Helenai. </s>	
6	☐	☐	dc	<s> Un vīns ūtrs čīnavnīks Reigā atsaļōve Helenu pavīļkt iz zūba:	
7	☐	☐	dc		

Summary

- It is necessary to popularize the digital resources of the Latgalian language and to promote digital literacy.
- Spelling and pronunciation of Latgalian proper nouns may be activated by comparing the data of one's native dialect, Latgalian written language and Latvian standard language, thus also activating the comparative aspect of using different corpus materials.

Future plans – corpus literacy training

- Survey on teachers' understanding basic concepts in corpus linguistics, using corpus output to generate teaching material and activities.
- Video lectures / tutorials.
- Hands-on exercises / home assignments.
- Survey on teacher's attitudes.

References

- Andrejeva, A., Dundure, V., Vulāne, A., & Žulina, V. (2020). *Latgaliešu rakstu valoda 1. – 9. klasei. Mācību priekšmeta programmas paraugs*. Retrieved from : <https://www.visc.gov.lv/lv/media/11407/download>
- Hague, C., & Payton, S. (2010). Digital literacy across the curriculum. FutureLab. Retrieved from <http://www.nfer.ac.uk/publications/FUTL06/FUTL06.pdf>
- Heather, Julian & Helt, Marie. 2012. Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education* 20(4): 415–440.
- Li, S. & Xu, M. (2022). Corpus literacy empowerment: taking stock of research to look forward for practice. *Journal of China Computer-Assisted Language Learning*, 2(1), 126-155. <https://doi.org/10.1515/jccall-2022-0002>
- Martena, S., Naua, N., Briška, A. (2022). Latgaliešu valodas korpuss citu Eiropas mazāk lietoto valodu kontekstā: korpusa raksturojums, lietojums un potenciālā iespējošana. (iesniegts publicēšanai žurnālā “Letonica”)
- Mukherjee, J. 2006. Corpus linguistics and language pedagogy: The state of the art – and beyond. In Braun, S., Kohn, K. & Mukherjee J. (eds). *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, Frankfurt: Peter Lang, 5–24.