



Vietvārdu *Reiga*, *Daugovpiļs* un *Rēzekne* diskursīvie modeļi: Mūsdienu latgaliešu runas korpusa (MuLaR) datu analīze

Antra Kļavinska

XIV Tarptautinis baltistų kongresas, Vilniaus universitetas, 2025 m. spalio 22–25 d.

Mērķis:

atklāt, kādus priekšstatus par *Mūsdienu latgaliešu runas korpusā* (MuLaR) biežāk pieminētajām pilsētām (*Reiga, Rēzekne, Daugovpiļs*) pauž latgaliski runājošā kopiena, kādus vietvārdu variantus lieto. Salīdzināt latgaliešu runas korpusa datus ar latgaliešu tekstu korpusa (MuLa2022) un līdzsvarotā mūsdienu latviešu valodas korpusa (LVK2022) datiem.

Pētījuma jautājumi

1. Kādi vietvārdu *Rīga*, *Daugavpils*, *Rēzekne* varianti lietoti MuLaR, MuLa2022 un LVK2022?
2. Kādi diskursīvie modeļi raksturo katru no šīm pilsētām?
3. Kā šie modeļi atšķiras viens no otra?
4. Kā korpusu dati palīdz atklāt toponīmu kontekstuālo informāciju?

Teorētiskais pamats

- ***Corpus-assisted onomastics*** (CAO) uses text corpora as a data source. The analysis is based on corpus linguistics methods, and the subject of the study is the distribution of proper names, their use in contexts, as well as regional and local variations (Jantunen et al. 2022).
- ***Corpus-assisted discourse studies*** (CADS) combines qualitative and corpus linguistics methods to study the frequency, concordance (contextual usage) and collocations (words that appear statistically frequently) of keywords with the aim of revealing discourse types (Partinton et al. 2013).

Mūsdienu latgaliešu runas korpus (MuLaR, 2024) <https://korpus.lv/id/MuLaR>

MuLaR     Korpus

Myusu dīnu latgalīšu runys korpus

Runuotuos volūdys audio īroksti i tūs atšifriejumi

Par korpusu

MuLaR ir Rēzeknis Tehnoloģeju akademejis (RTA) humanitarūs zynuotpu pietniku sataiseitais latgalīšu runys korpus.

MuLaR nūsaukums ir atvasynojuums nu MuLa – Mūsdienu latgaliešu tekstu korpus, kurā ir iklauti rokstu volūdys teksti. Tekst ite.

 Nacionālā korpusu kolekcija **Rādītājs** Meklēšana Jaunumi Materiāli Par NKK

MuLaR

🔍 Meklēt korpusā

☰ Vārdu biežuma saraksts

Mūsdienu latgaliešu runas korpus

Korpusu veido audioieraksti un to transkripcijas. Tas dokumentē dabisku, spontānu runu,

Audio ieraksti (27 h) un to transkripti (220 844 vārdlietojumi):

- lauka pētījumos Latvijā iegūtas intervijas (**2009-2022**);
- TV (Latgales Reģionālā televīzija) un radio (Latvijas Radio Latgales studija) raidījumu ieraksti (**2018 - 2024**);
- lauka pētījumos Sibīrijā iegūti ieraksti (2017-2018);
- audio ieraksti un transkripcijas no TriMCo dialektu korpusa projekta.
- Metadati: dzimums, vecums, ieraksta veikšanas vieta.
- Korpus **nav lemmatizēts, nav morfoloģiski marķēts.**

Salīdzinājumam izmantotie tekstu korpusi

Mūsdienu latgaliešu tekstu korpus (MuLa 2022)

<https://korpuss.lv/id/MuLa2022>

- Latgaliešu rakstu valodā publicēti (1988–2021) teksti (2 milj. vārdlietojumu).
- Metadati: teksta autors, nosaukums, publicēšanas vieta un gads, kā arī informācija par teksta veidu un žanru (publicistika (55%), daiļliteratūra (37%), populārzinātniski, zinātniski teksti(8%)).
- Korpus **nav lemmatizēts un nav morfoloģiski marķēts.**

Līdzsvarotais mūsdienu latviešu valodas korpus (LVK 2022)

<https://korpuss.lv/id/LVK2022>

- Latviešu literārajā valodā publicēti (2000 – 2021) teksti (101 milj. vārdlietojumu).
- Žanri:periodika (60%), daiļliteratūra (10%), zinātniski teksti (10%), teksti no „Vikipēdijas” (7%), normatīvie akti (7%), Saeimas stenogrammas (3%) un subtitri (3%)
- Automātiski **gramatiski marķēts.**

Pētījuma metodoloģija

1. **Frekvence** – atslēgvārdu biežums MuLaR, MuLa2022 un LVK2022.
2. **Kolokācijas** – MuLaR un MuLa2022– vietvārdu piecas labās un kreisās puses kolokācijas (MuLaR ; LVK2022 – trīs labās un kreisās puses kolokācijas (100 biežākās)).
3. **Diskursīvo modeļu noteikšana** - diskursa prosodijas (*discourse prosodies*) jeb izteikumu īpašības parādīties noteiktos diskursīvos kontekstos (Baker, 2006). Dažāda korpusu apjoma dēļ - %.
4. **Konkordanču analīze** (lietojums kontekstā).

Frekvence

MuLaR (2024): *Rēzekne* (95), *Riezekne* (77); **Reiga** (166 vārdlietojumi), *Rīga* (2), *Ryga* (2); **Daugovpiļs** (80), *Daugavpiļs* (11), *Dauguopiļs* (3) un *Dinaburga* (1).

MuLa2022: *Rēzekne* (2,502), *Riezekne* (26), *Režica* (12), *Rositten* (5); **Reiga** (1,991), *Rīga* (52); **Daugovpiļs** (684), *Daugavpiļs* (615) *Daugpiļs* (80), *Dinaburga* (60), *Dvinska* (63), *Borisoglebska* (2).

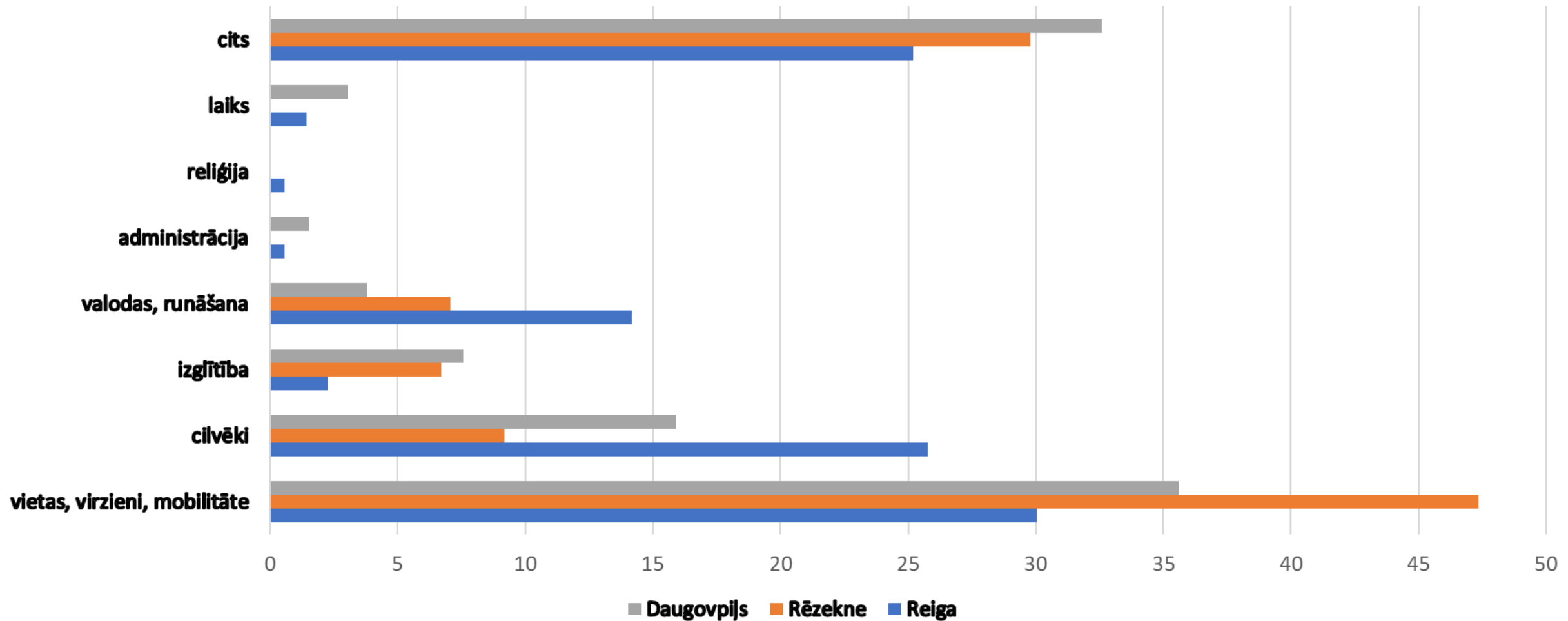
LVK2022: *Rīga* (203,946); *Reiga* (2); **Daugavpils** (14,817), *Dinaburga* (214), *Dvinska* (67), *Borisoglebska* (5), *D-pils* (2), *Dpils* (1); **Rēzekne** (8,459) *Režica* (10), *Rositten* (4).

MuLa2022 un LVK2022 diskursa prosodijas

(A.Kļavinska, S.Martena, H.Martens, A. Āboliņa, 2025)

	MuLa2022			LVK2022		
	Daugovpils	Rēzekne	Reiga	Daugavpils	Rēzekne	Rīga
Places, Directions, Mobility	40,01%	46,99%	49,01%	48,48%	58,13%	35,56%
Other	14,06%	10,48%	15,55%	14,54%	9,02%	8,62%
Education	12,11%	13,32%	4,39%	4,08%	4,80%	3,93%
Time	9,08%	10,28%	11,65%	4,38%	5,49%	17,86%
Culture and Arts	8,94%	5,37%	4,54%	3,43%	1,70%	2,92%
People	5,26%	4,48%	4,54%	1,54%	1,58%	3,81%
Religion	2,74%	2,70%	1,88%	0,00%	0,00%	0,31%
Economics and Business	2,67%	0,69%	2,81%	2,41%	3,41%	3,35%
Administration	2,45%	1,66%	0,00%	15,46%	12,05%	6,60%
Institutions and Organizations	1,73%	2,67%	4,10%	4,17%	2,91%	11,68%
Languages and Speaking	0,65%	0,00%	0,74%	0,34%	0,00%	0,00%
Sport	0,29%	0,00%	0,00%	0,77%	0,31%	4,13%
Politics	0,00%	1,38%	0,79%	0,41%	0,59%	1,23%

MuLaR diskursa prosodijas (%)



Vietas, virzieni, mobilitāte (kolokācijas)

	MuLaR	MuLa2022	LVK2022
Rēzekne	47,36 % <i>nūvods, pogosts, pilsāta, aprņkis; Latveja, Varakļuoni, Daugovpiļs, Ludza, Sylajuoni, Rēzekne; braukt, staiguot, atvest; traktors; iz/uz (10); nu (18);</i>	46,99% <i>nūvods, piļsāta, pogosts, centrs, rajons, (Atbreivuošonys) aleja; upe, Latgola, Daugovpiļs, Ludza, Reiga, Latveja; iz (102) nu (116)</i>	58,13% <i>novads, pilsēta, pagasts, rajons, valsts, iela, centrs, aleja, teritorija; Latgale, Latvija Daugavpils, Ludza, Liepāja, Valmiera, Ventspils; tikt, atrasties, uz, no</i>
Rīga	30,03 % <i>uorzemis, regions; Latveja, Latgola, Daugovpiļs; braukt, atbraukt, aizbraukt, nabraukt, īt, aizīt; iz / uz (26), nu (19)</i>	49,01% <i>valsts, centrs; golvyspiļsāta, piļsāta, īla, ceļš, ūsta; Latveja, Latgola, Rēzekne, Daugovpiļs, Jelgova; dzeivuot, aiz-/nū-braukt, tikt, atsarast, puorsaceļt; viļcīņš; iz (176), nu (146)</i>	35,56% <i>lela, pilsēta, satiksme, centrs, rajons, reģions, priekšpilsēta, teritorija, valsts; Latvija, Jelgava, Liepāja, Vidzeme; atrasties, dzīvot, tikt, doties, uz, no</i>
Daugavpils	35,61 % <i>vide, centrs, aprņkis, saleņa; Grīva, Reiga, Rēzekne; aizbraukt, dzeivuot; iz (1) – nu (6)</i>	40,01% <i>nūvods, aprņčs, piļsāta, pogosts, rajons; centrs, īla; Latgola, Latveja, Reiga, Ludza, Kruoslova, Daugova, Preiļi, Vabale,dzeivuot, apsamest; iz (54), nu (57)</i>	48,48% <i>pilsēta, novads, iela, rajons, pagasts, centrs, šoseja, dzelzceļš, Rīga, Rēzekne, Liepāja, Jelgava, Daugava, Latvija, Latgale, Jēkabpils, dzimt, dzīvot, uz, no</i>

Vietas, virzieni, mobilitāte (konkordanču piemēri)

- *pirmeit Riezeknī beju </s><s> nu Riezeknis kai brauču tai </s><s> te motors maņ sātā beja benzinovais </s><s> a tagad kai benzins duorgs (MuLaR, M102-1938, LTG-E-ANDR-2011)*
- *Vysim stuostu, ka na jau Latgola ir tuoli nu Reigys, bet Reiga ir tuoli nu Eiropys Savīneibys suokuma. (MuLa 2022, publicistika, 2020)*
- *Patlaban Rīgā, Olainē, Jelgavā, Salaspilī, Rēzeknē un Daugavpilī tiek apsaimniekoti kopumā astoņi industriālie parki. (LVK2022, nacionālā prese, 2015)*

Cilvēki

	MuLaR	MuLa2022	LVK2022
Rēzekne	9,19% <i>es, mes, riezeknīši, latgalīši</i>	4,48% <i>latgalīši, tauta, grupa, Juons, Ivanovs, Rancāns, dorbuotīs</i>	1,58% <i>iedzīvotājs, Jānis, dzimt, dzīvot</i>
Rīga	25,78% <i>latgalīši, latvīši; mozbārni, bārni, tāvs, bruoļs, muosa, es, mes, juos.</i>	4,54% <i>latvīši, (Latgolys) latvīši, latgalīši, jaunīši, reidzinīki, cylvāks, Rancāne, dzimt</i>	3,81% <i>latvieši, iedzīvotājs, Ušakovs, Stradiņš, dzimt</i>
Daugavpils	15,91% <i>es, mes, jī</i>	5,26% <i>latvīši, latgalīši, krīvi, ebreji, franči, Luocs/Lōcis, Soms, Rotko, Marks</i>	1,54% <i>iedzīvotājs, Rotko, Marks, dzimt</i>

Cilvēki

- *Riezekne beja pilsāta kura ir **latgalīšu pilsāta** </s><s> es dīmžāl juos **nauzskotu par Latgalis golvyspylsātu** partū ka tī ir ee divvalūdeiba ir jau ļūti krasa (MuLaR, M402-1963, LTG-E-AGLO-2021)*
- ***Reigys Latvīšu** bīdreibā ar sovīm referatim **Trasuns** uzastuoja leidz pat Pyrmajam pasaulā karam. [MuLa2022, publicistika, 2013]*
- *Daudzi **latgalīši Reigā** pruvej puorvarēt gravitacejis spāku i lidoj tik augši, augši, ka aizmierst par savom saknem. (MuLa2022,)*
- *Tā **Rīgā** 42,2 % ir **latvieši**, 41,7 % **krievi**, otrajā lielākajā pilsētā **Daugavpilī** 17,4 % ir **latvieši**, 53,2 % **krievi**, **Rēzeknē** - 44,4 % **latvieši**, 48,1 % **krievi**. (LVK2022, disertācija, 2010)*

Valodas un runāšana

	MuLaR	MuLa2022	LVK2022
Rēzekne	7,07% <i>saceit, volūda, izlūksne runuot</i>		
Rīga	14,16% <i>latgaliski, latviski, runuot, volūda, saceit</i>	0,74% <i>runuot, latgtaliski, vuords</i>	
Daugavpils	3,79 % <i>runuot, saceit</i>	0,65% volūda, vuords	0,34% vārds

Valodas un runāšana

- *tak vacuoki centēs obi atbalstē tū **latvīšu volūdu** </s><s> lai īsamuocās **pareizi runuot** lai naapsasmej **Reigā** i tai tuoļuok </s><s> a tagad (..) **Reigys** vydā tu vari **latgaliski runuot** </s><s> nivīns tev napīvērss uzmanību (MuLaR, F407-1943, E-AGLO-2021)*
- *dīzkuo es nagrybu latgalīšu vairs portālus skateitīs </s><s> jū naz parkū **Riezeknis izlūksne** teik uzbuožta viersā (MuLaR, F407-1943, LTG-E-AGLO-2021)*
- *Līpuojis em **krīvvalūdeigī latviski runoj** </s><s> **Daugovpiļs narunoj** (MuLaR, M402-1963, LTG-E-AGLO-2021)*

Secinājumi (I)

- Paužot priekšstatus par Rēzekni, Daugavpili un Rīgu, latgaliski runājošā / rakstošā kopiena atspoguļo savu lokālo identitāti, vēsturisko atmiņu un kultūras kontinuitāti. Savukārt latviešu literārās valodas tekstos pamatā ir citas diskursīvās prioritātes — valstiskums un centralizācija.
- Latgaliešu runas korpusā (MuLaR) vairāk atklājas individuālās prakses un pieredzes, tekstu korpus (MuLa2022) liecina par latgaliešu kopienas pieredzi, savukārt latviešu tekstu korpus (LVK2022) konstruē kopīgās nacionālās prakses.

Secinājumi (II)

- Rīga gan latviešu, gan latgaliešu tekstos pozicionēta kā valsts administratīvais, ekonomiskais, kultūras un izglītības centrs, latgaliešu tekstos – vieta, kur Latgales latvieši skaidrāk apzinās un uzsver savu identitāti (*latgalīši – latvīši, rizeknīši - reidzinīki*), īpaši valodas diskursā.
- Rēzekne un Daugavpils latgaliešu runas un tekstu korpusa datos galvenokārt atklājas kā divi latgaliešu kultūrai nozīmīgi centri (gan vēsturiskā, gan mūsdienu kontekstā), latviešu tekstu korpusā šīs pilsētas reprezentētas salīdzinājumā ar citām valsts pilsētām dažādos rādītājos; nozīmīgu notikumu, pasākumu, norišu, ievērojamu personību kontekstā.
- Toponīmu kontekstuālo informāciju precīzāk un dziļāk var atklāt liela apjoma korpusi, MuLaR korpusa izmērs ļauj izteikt pirmos priekšstatus, ieskicēt tendences.

Avoti un literatūra

- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Jantunen, J.H., Ainiala, T., Jokela, S., & Tarvainen, J. (2022). Mapping Digital Discourses of the Capital Region of Finland: Combining Onomastics, CADS, and GIS. *Names: A Journal of Onomastics*. Vol. 70, Issue 1, Winter 2022, pp. 20-39. DOI 10.5195/names.2022.2289
- LVK2022 - Levāne-Petrova, K.; Dargis, R.; Pokratniece, K., Lasmanis, V. J., 2023, *Līdzsvarotais mūsdienu latviešu valodas tekstu korpuss*, CLARIN-LV digital library at IMCS, University of Latvia, <http://hdl.handle.net/20.500.12574/84>.
- MuLa2022 – Briška, A., Ziņģe, I., Dargis, R., Pokratniece, K., Martena, S., Kļavinska, A., Juško-Štekele, A. *Mūsdienu latgaliešu tekstu korpuss 2022*. CLARIN-LV digital library at IMCS, University of Latvia, <http://hdl.handle.net/20.500.12574/72>.
- MuLaR (2025). - Martena, S., Nau, N., Kļavinska, A., Juško-Štekele, A., Kociņš-Kūcenš, A., Sprukte, A., Briška, A., Gusāns, I., Mazure, L. *Mūsdienu latgaliešu runas korpuss (MuLaR)*. CLARIN-LV digital library at IMCS, University of Latvia, <http://hdl.handle.net/20.500.12574/118>.
- Partinton, A., Duguid, A., & Taylor, Ch. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-assisted Discourse Studies (CADS)*. Amsterdam: John Benjamin Publishing Company.



Paļdis par viereibu!
Paldies par uzmanību!
Ačiū už dēmesj!

Antra.Klavinska@rtu.lv

Izstrādāts ar ES ANM atbalstu projekta Nr. 5.2.1.1.i.0/2/24/I/CFLA/003 P&A grantā “Spatiotemporālā modeļa izstrāde un pielietošana Latgaliešu runas korpusa datu daudzveidīgas analīzes nolūkos”, Nr. RTU-PA-2024/1-0063