

Dzejas analīze ar statistiku

Anotācija

Šajā dokumentā salīdzinātas O.Vācieša un Raiņa dzejoļu statistiskās īpašības, kuras viegli izmērīt ar datoru. Analizēti divi dzejoļu krājumi — O.Vācieša *Dzegužlaiks* (ap 100 tūkstoši rakstzīmju) un Raiņa *Tālas noskaņas zilā vakarā* (ap 60 tūkstoši rakstzīmju). Turpmāk apzīmēsim tos attiecīgi ar *Dz-L* un *TNZV*.

Abos krājumos skaitīts burtu biežums, meklēti dzejoļi vai to fragmenti, kuros ir bieži sastopams kāds burts/skaņa, salīdzināts vārdformu biežums kā arī veikts salīdzinājums krāsas izsakošiem vārdiem - to relatīvajam biežumam un morfoloģiskajām īpašībām.

1 Ievads

Ludis Bērziņš grāmatā *Ievads latviešu tautas dzejā* [4] aprakstījis latviešu tautasdziesmu dzejas formas īpatnības un to stilistiku. Izmantojot lielu tautasdziesmu korpusu, viņš noteicis tur sastopamos pantmērus, saskaitījis dažādu krāsu nosaukumus, skaitļu vārdus utml., vienlaikus skaidrojot mākslinieciskos izteiksmes līdzekļus saistībā ar šiem vārdiem.

Mums nav zināmas L.Bērziņa datu vākšanas procedūras, bet 1930-to gadu statistikas pētījumos parasti skaitīja piemērus, velkot svītņiņas vai izrakstot paraugu koordinātes uz papīra. Bija pazīstami arī mehāniski risinājumi (sal. [3]): Kartotēkas kartītes salika precīzā kaudzītē un izdūra tām cauri metāla stienīti. Kartītes, kurām attiecīgajā pozīcijā izgriezts robiņš, izkrīt ārā.

Agrāk pētniecība prasīja ievērojamu pacietību un darba ētiku, savukārt datoru izplatība to padara pieejamu daudz plašākam ļaužu lokam. Var izmantot valodniekiem īpaši paredzētas informatīvas sistēmas, piemēram, konkordances un datorizētus literatūras korpusus. Vai arī universālus bezmaksas rīkus, piemēram, R un RStudio. RStudio ir ērta vide, kurā ģenerēt statistiskus pārskatus programmēšanas valodā R.

Dž.Ridbergs-Kokss [7] uzskaita vairākas literatūras teorijas grāmatas, kurās izmanto statistikas metodes. Interese par datoru izmantošanu literatūras pētniecībā bijusi jau 1987. gadā, kad Dž.F.Barouzs nāca klajā ar Dž.Ostinas romānu matemātisku analīzi [1]. Šajā grāmatā salīdzināta Dž.Ostinas varoņu lietotā valoda - prievārdu, saikļu, vietniekvārdu un artikulu lietošanas biežums, lai pētītu romānu (piemēram *Northanger Abbey*) varoņu raksturu attīstību romāna gaitā. Ar valodas kvantitatīvo analīzi parasti nodarbojas universitātēs, tomēr daudzi jautājumi pieejami arī skolēniem.

Šī raksta mērķis nav nodarboties ar mūsdienu datorlingvistiku, kas ir strauji augoša un tehniski sarežģīta nozare, ko Latvijā praktizē LU Matemātikas un informātikas institūta Mākslīgā intelekta laboratorijā. Arī literatūras teorijā atsaucamies tikai uz vienkāršām lietām. Esam centušies apskatīt vienkāršas parādības dzejā: burtu un vārdformu biežumu, aliterāciju meklēšanu, skaitot dzejoļu burtus, vienkāršas konkordances. Ievaddatus šim rakstam sagatavoja datorprogramma; vienīgi krāsu nosaukumu morfoloģiskā analīze prasīja roku darbu.

O.Vācieša un Raiņa izvēle šim pētījumam bija sagādīšanās. Neesam pamatojuši, kādēļ vajadzētu salīdzināt tieši šos autorus vai dzejoļu krājumus. Raiņa svarīgākie darbi ir digitalizēti - sk. <http://www.korpuss.lv/klasika/Senie/Rainis/rainissat.htm>. Tādēļ eksperimenti ar Raini neprasa ievadīt datorā pētāmo tekstu. Savukārt O.Vācieša dzejas statistiska analīze palīdz atrast dzejoļus, kuri ir formas ziņā neparasti, rada papildu iemeslu tos lasīt, mācīties un mēģināt saprast. O.Vācieša dzejas krājums *Dz-L* no kopotu rakstu 2.sējuma ievadīts datoram viegli lasāmajā *Markdown* [5] formātā. Pēc tam datorprogramma (R scenārijs jeb skripts) nolasa abu krājumu tekstus, saskaita tur burtus, vārdformas vai ko citu – un iegūtos rezultātus apkopo tabulās vai attēlo diagrammās.

Kādas lietas tika pētītas:

- **Burtu relatīvais biežums krājumos kopumā.** Tas ļauj mums iegūt pirmo priekšstatu par skaņu atšķirībām. Diemžēl, neesam atsevišķi saskaitījuši līdzskani "dz" (to ieskaitām kā divus burtus - "d" un "z"). Nešķirojam arī šauros un platos "e" un "ē", kā arī īso vai garo patskani "o" vai divskani "uo".
- **Pēdējo burtu relatīvais biežums.** Atsevišķi skaitām tos burtus, kas ir vārda vidū un burtus, kuri ir vārda pašas beigās. Vārdu pēdējie burti mēdz norādīt uz lietvārdu, īpašības vārdu u.c. dzimti, skaitli un locījumu.

Arī atskaņu tipi atšķiras atkarībā no tā, vai dzejas rinda beidzas ar atvērtu vai slēgtu zilbi (t.i. beidzas ar patskani/divskani vai arī ar līdzskani).

- **Burtu relatīvais biežums atsevišķos dzejoļos.** Burtam “S” izveidojām histogrammas, kas attēlo tā koncentrācijas O.Vācieša un Raiņa dzejoļos. Noskaidrojot dzejoļus ar neparasti augstu attiecīgās skaņas koncentrāciju, varam meklēt iespējamās *aliterācijas* vai *asonances*. Vispusīga aliterāciju meklēšana prasītu dziļāku analīzi; jo līdzīgu skaņu sablīvējums var neizpausties visa dzejoļa garumā.
- **Garākie kopīgie fragmenti.** Šajā sadaļā algoritmiski meklējam garākās kopīgās vārdu virknītes abu autoru darbos. Tas ir naivs mēģinājums atrast *reminiscences*.
- **Biežāk sastopamās vārdformas katram no autoriem.** Viens dzejoļu krājums ir nepietiekams, lai saprastu autora leksiku; bet rodas vismaz priekšstats par lietotajiem saikļiem, prievārdiem, vietniekvārdiem kā arī populārākajiem pārstāvjiem no citām vārdšķirām.
- **Krāsu nosaukumi.** Katram no autoriem apkopoti krāsu nosaukumi kopā ar dzejas rindām, kurās tie parādās; apkopota statistika par dažādu krāsu locījumu formām.

Ko vajadzētu darīt savādāk:

- Skaitīt skaņas nevis burtus. Skaņas "dz" nesaskaitīšana skaidrojama ar slinkumu, savukārt šauros un platos "e", "ē" rakstībā neatšķir. To skaitīšanai jālieto algoritms, kurš tos atšķir pēc kaut kādiem pozicionāliem likumiem (un vajadzīgs arī dzīvs cilvēks, kurš tiktu galā ar sarežģītākiem gadījumiem). Tāpat arī - atsevišķi no patskaņiem būtu jāskaita divskaņi, jo tie mēdz veidot paši savas asonances.
- Skaitot vārdformas būtu interesanti ņemt vērā lokāmo vārdu morfoloģiju, piemēram, skaitot kopā viena vārda vairākus locījumus.
- Aliterācijas un asonances būtu jāmeklē nevis dzejoļu līmenī, bet aplūkojot dzejas rindu grupas.
- Pagaidām nav priekšstats par to, cik "Dz-L" un "TNZV" ir tipiski O.Vācieša un Raiņa dzejoļu krājumi. Lai secinājumi būtu par autoriem (nevis tikai konkrētajiem krājumiem) vajadzētu plašākus ievaddatus. Piemēram, aplūkot kādus krājumus, ko attiecīgie autori sarakstījuši vēlāk.

2 Burtu biežums

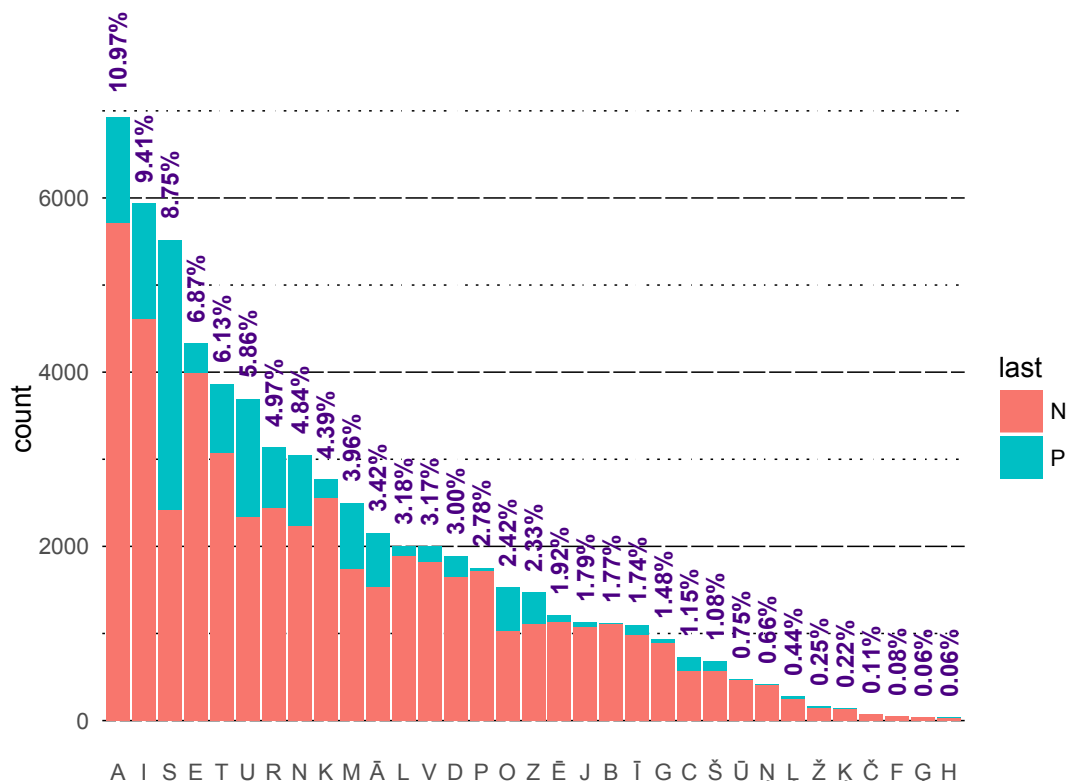
Katrā tekstā var saskaitīt, cik tajā ir dažāda veida burtu. Lielos un mazos burtus skaitām kopā. Ņemot vērā to, ka garākā tekstā visa veida burtu ir vairāk, mūs interesēs nevis burta *absolūtais biežums*, bet gan tā *relatīvais biežums* – t.i. cik reizes attiecīgais burts caurmērā parādās starp katriem 100 burtiem teksta. Absolūto biežumu izsaka ar veselu skaitli (cik reizes burts parādās), savukārt relatīvo biežumu mēra procentos.

Latviešu valodai (tāpat kā jebkurai citai) raksturīgi noteikti burtu relatīvie biežumi, kas ir visai stabili, neatkarīgi no teksta temata un autora. Teksta tematu un autora īpatnības parasti atklāj nevis burtu biežumi, bet gan leksika – t.i. vārdu un vārdu savienojumu biežumi. Tomēr dzejā pat skaņu biežums mēdz atšķirties no citiem valodas funkcionālajiem stiliem.

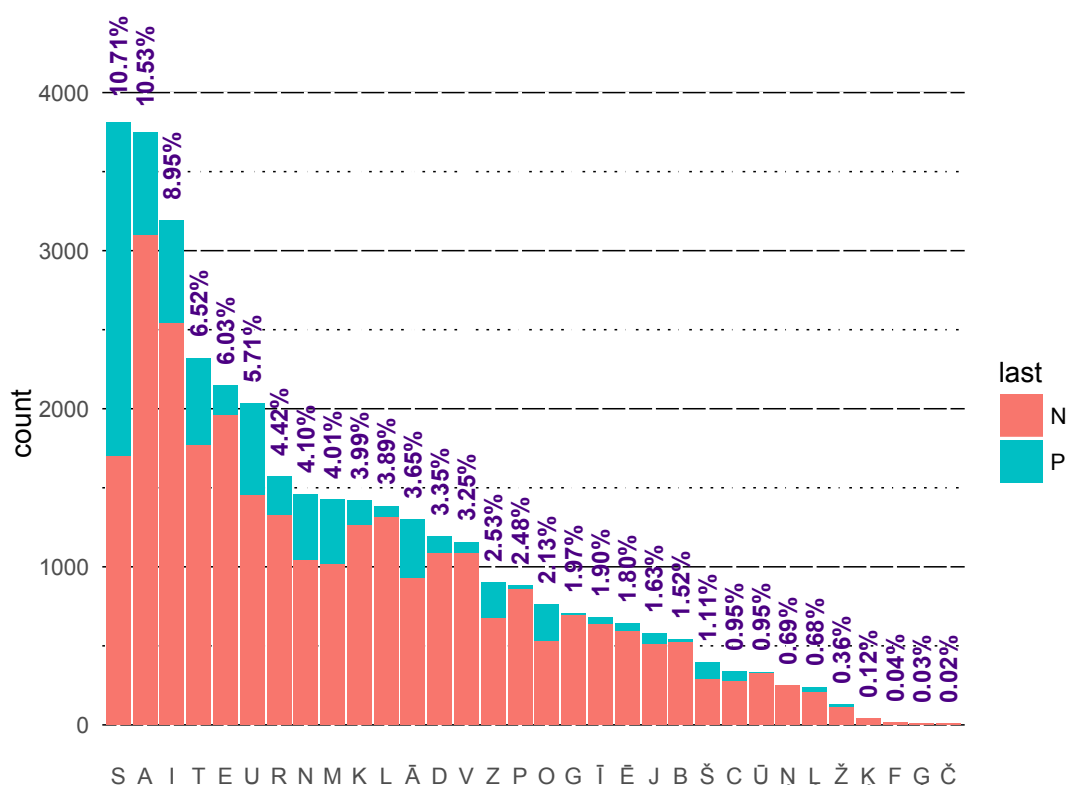
Dz-L satur izķēmotas citvalodu frāzes “*nach Westen drang*”, “*nach Süden drang*”, tādēļ O.Vācieša dzejoļos ir pa pa vienam burtam “ü” un “w”. Tos ignorējam. Tāpat arī neskaitām tukšumus un pieturzīmes. Paturam tikai 33 latviešu alfabēta burtus. Pēc tam līdzīgi saskaitām burtus no *TNZV*. Tajā ir teksta gabaliņš latīniski “*Fiat justitia ruat mundus*” (atsauce uz pazīstamu izteicienu “*Fiat justitia ruat caelum*” – “dāriet taisnīgumu [kaut] sabrūk debesis”), bet tas pierakstīts ar latviešu alfabēta burtiem. Krājuma nodaļu nosaukumus, dzejoļu virsrakstus, apakšvirsrakstus, veltījumus un epigrāfus burtu skaitīšanā (un arī visur citur šajā rakstā) neņemam vērā.

Burtus vārdu galotnēs ir pamats aplūkot atsevišķi no citiem burtiem, jo to parādīšanās biežumu iespaido ne tik daudz autora leksika, cik viņa dzejā izplatītākās gramatiskās formas. Tā kā mums nav programmatūras, kas ļautu nodalīt vārdu galotnes, tad skaitīsim atsevišķi katra vārda pēdējo burtu. Diemžēl, šādi nebūsim uzskaitījuši priekšpēdējos u.c. burtus, kas arī varbūt ietilpst galotnēs, būs arī nevajadzīgi pieskaitījuši pēdējos burtus nelokāmajām vārdšķirām. Bet šādi vismaz var uzzināt pirmo tuvinājumu par dažiem populāriem burtiem (teiksim, “a”, “i” un “s”): cik bieži tie ir vārda celmā un cik bieži – galotnē.

Secinājumi:



Attēls 1: Burtu biežums Dz-L (burti vārda beigās apzīmēti zili)



Attēls 2: Burtu biežums TNZV

- O.Vācietim un Rainim atšķiras pat pirmais biežāk lietotais burts. O.Vācietim visizplatītākais burts ir “A”, savukārt Rainim – burts “S”. Sk. Attēlus 1 un 2. Ņemot vērā analizēto teksta paraugu ievērojamo garumu (desmittūkstoši burtu), ir pamats uzskatīt, ka šī atšķirība raksturo abu autoru dzeju kopumā. Atšķirība “S” burtā biežumos ir nepilni 2% (*Dz-L* to ir 8.75%, bet *TNZV* to ir 10.71%) var nešķīst liela, bet tā nozīmē, ka vidēji katrās 4 dzejas rindās no *TNZV* ir par vienu “S” burtu vairāk nekā līdzīga garuma 4 dzejas rindās no *Dz-L*.
- O.Vācieša dzejā sastopamo burtu biežums daudz precīzāk atspoguļo burtu biežumu mūsdienu zinātniski tehniskajā literatūrā. E.Riekstiņš un A.Andžāns [6] publiskoja šādu Tabulu 1, kas iegūta, analizējot ap 70000 rakstzīmju no zinātniskas literatūras latviski.

A	11.1%	K	4.1%	J	2.1%	Ļ	0.2%
I	9.3%	M	4.1%	Z	2.0%	F	0.2%
S	8.3%	O	4.0%	B	1.6%	Ž	0.2%
T	6.9%	P	2.9%	C	1.4%	Ķ	0.1%
E	6.4%	D	2.7%	G	1.3%	H	0.1%
R	5.7%	Ī	2.3%	Š	1.1%	Č	0.02%
U	5.0%	L	2.3%	Ģ	0.4%		
N	4.9%	V	2.2%	Ū	0.4%		
Ā	4.3%	Ē	2.1%	Ņ	0.3%		

Tabula 1: Burtu biežumi zinātniski tehniskajā literatūrā (no 1)

Grāmatā 1 burtu biežumi iegūti, analizējot samērā īsu (un, iespējams, nereprezentatīvu) teksta gabalu. Būtu jānoskaidro burtu biežumi kādā mūsdienu latviešu tekstu korpusā un arī jānosaka tipiskas burtu biežumu novirzes atkarībā no autora un žanra. Tikai tad varēsim uzzināt, vai *TNZV* sastopamie burtu biežumi ir kaut kas neparasts.

- *TNZV* vispār neparādās burts “H”, savukārt nedaudz garākajā *Dz-L* burts “H” ietilpst veselas 35 reizes (vārdformas: “Mínhenes”, “čehu”, “Heil!”, “hospitāļa”, “himnu”, “cehu”, “Bēthovens”, “haubices”, “horizontiem”, “saharīnu”, “Donkihota”, “cehos”, “he”, “haosus”, “horizontiem”, “arhitektiem”, “horizonts”, “Hitlers”, “tuhačevskiem”, “remontmehāniķi”, “mehāniķi”, “hipnozi”, “haizivs”, “Herkules”, “orhideja”, “arheologiem”). Tas lieku reizi apliecina, ka O.Vācietis dzejā ienesis daudz jaunas “nepoētiskas” leksikas.
- Kā jau minējām, Raina *TNZV* ir paaugstināts burtā “S” īpatsvars (ne vien, salīdzinot ar O.Vācieša dzeju, bet arī, ņemot vērā burtu “S” izplatību latviešu valodas tekstos kopumā). Tam iespējami divi izskaidrojumi: vai nu Rainis bieži lieto vārdus, kuru celmā ir burts “S”, vai arī viņam patīk gramatiskas formas, kur lietvārdu, īpašības vārdu, vietniekvārdu galotnēs parādās līdzskanis “S”. Tādēļ saskaitījām atsevišķi burtus “S” vārdu “iekšienē” un arī tos “S”, kas ir vārdu pēdējie burti. (Tie attēloti arī Attēlos 1 un 2 – pēdējo burtu īpatsvars atbilst stabiņa augšējai, zili iekrāsotajai daļai.) Iegūtie rezultāti apkopoti Tabulā 2.

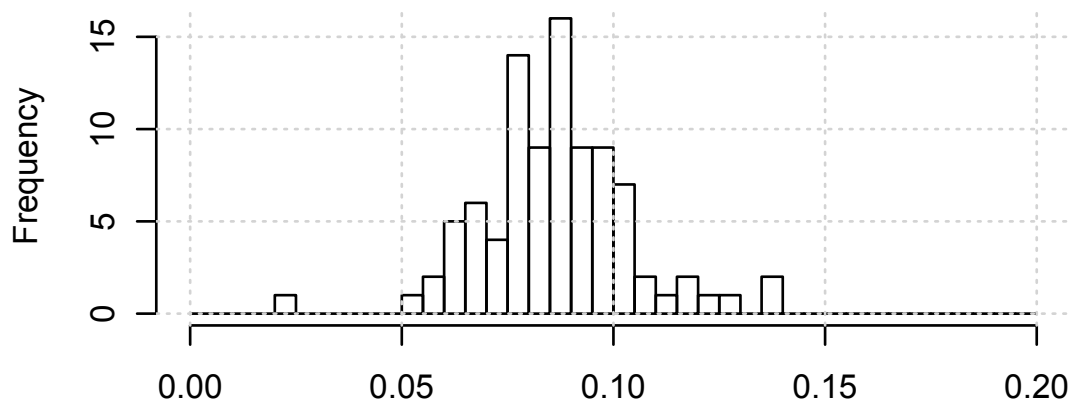
Krājums	“S” nav vārda beigās	“S” ir vārda beigās
Dz-L	2424 (43.9%)	3098 (56.1%)
TNZV	1704 (44.7%)	2111 (55.3%)

Tabula 2: Burtu “S” biežums *Dz-L* un *TNZV*

Krājumā *Dz-L* ir pat nedaudz raksturīgāk lietot “S” vārda beigās nekā *TNZV* (attiecīgi 56.1% un 55.3%). Varam secināt, ka vārdu morfoloģija “S” biežuma atšķirību neizskaidro. Burtu “S” paaugstināto biežumu Raina dzejā pilnībā nosaka Raina leksika – viņš vienkārši biežāk raksta vārdus, kuru celmi satur “S”.

3 Vienādu skaņu sablīvējumi

Apzināti nesaucām tos par aliterācijām vai asonansēm, jo iespējami gadījumi, kad vienādas skaņas kādā dzejas pantā sastopamas biežāk nekā tekstā kopumā, bet tās nerada stila efektu. Jau minējām, ka datorprogramma mērīja burtu (nevis skaņu) biežumu. Arī tas neļauj ar pārliecību kaut ko apgalvot par aliterācijām un asonansēm. Šajā sadaļā apkopoti dzejoļu fragmenti, kuru burtu biežumi ir neparasti – neatkarīgi no mākslinieciskās izteiksmes līdzekļiem.



Attēls 3: Burta S biežumu histogramma (Dz-L)

Burtu “S” salīdzināšanai izvēlējamies tādēļ, ka *TNZV* šī burta biežums ir lielāks – sk. secinājumus 2. nodaļas beigās. Skaņai “S” ir arī īpaša fizioloģiska loma – no visām izrunājamām skaņām tai mēdz būt visaugstākā frekvence, tā ir vistuvākā ultraskaņai. Skaņas “S” un “Š” (to palatalizēti varianti) plaši sastopami arī mēles mežģos. Pat nedaudz paaugstinātu “S” vai “Š” īpatsvaru var sadzirdēt ar neapbruņotu ausi, par ko drīz pārliecināsimies. Sal. [9].

3.1 Fragmenti Dz-L dzejoļiem ar $\rho(S) \geq 12\%$

Turpmāk ar $\rho(A), \rho(B), \dots$ apzīmēsim burta koncentrāciju dzejolī, t.i. konkrētā burta “A”, “B” u.c. skaita attiecību pret visiem burtiem dotajā dzejolī. Koncentrācijas ir daļskaitļi starp 0 un 1. *Dz-L* pavisam ir 92 dzejoļi vai dzejoļu daļas ar atsevišķiem apakšvirsrakstiem. Attēls 3 ir histogramma, kas parāda, “S” koncentrāciju katrā no šiem 92 dzejoļiem.

Histogrammā redzams vientuļš kvadrātiņš kreisajā pusē – viens dzejolis ar neparasti zemu “S” burtu koncentrāciju (“Ar labo roku\ Apņemt var sievu,\ Ar labo roku\ Nolauzt var ievu...”), kuram $\rho(S) = 2.36\%$. Turpmāk pievērsīsimies dzejoļiem, kuros “S” burtu koncentrācija ir liela – stipri pārsniedz burta biežumu krājumā kopumā (šos biežumus sk. 1). Piemēram O.Vācieša krājumā “S” burta biežums ir vidēji 8.75%. Bet atsevišķos dzejoļos vai dzejas pantos tas var būt 12% vai vēl lielāks. Seko piemēri.

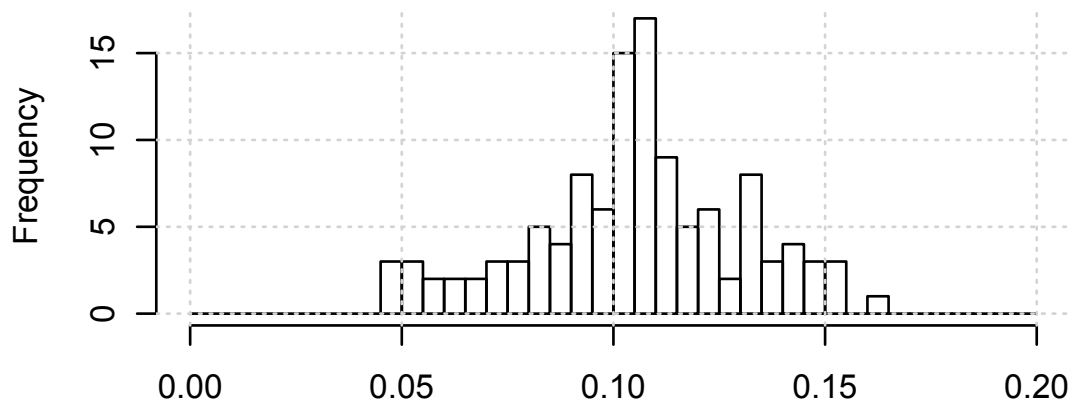
Pie nātres skāries,
 Neraudi, ka skāries, —
 Kas tevi dzina
 To ziedu plūkt!
 Reiz sāpes pāries,
 Visas sāpes pāries,
 Tik nevajaga
 Nevienam lūgt. ...

(Dz-L, "Dziesma", $\rho(S) = 13.9\%$)

Šajos ceriņos rīt ziedi būs —
 Mēļas liesmas no tumšzaļas pekles.
 Un pa debesīm eņģelus — mūs
 Un pa zemi mūs — grēcīgos meklēs. ...

(Dz-L, " * * * ", $\rho(S) = 13.6\%$)

...Gaujas balss ilgi meklēts un saukts,



Attēls 4: Burta S biežumu histogramma (TNZV)

Baltu mākoņu kopenās jaukts,
 Ziedu sildīts un viesuļu nests, —
 Ja tu būsi. Bet saredzu es —
 Katrā solī, ko tālāk tu liec,
 Pelēks, silts, samīts zaķītis klie dz.
 Tā jau nav kāda varena balss —
 Tā klie dz viss,
 Kur iet sals,
 Kur iet sals.
 (Dz-L, " * * * ", $\rho(S) = 12.9\%$)

Kad graudu samaj miltos
 Un miltus mājās ved,
 Lai briedīgs zemes siltums
 Tev kaulos spēku met.
 (Dz-L, "Visvaldim Lāmam", $\rho(S) = 12.4\%$)

3.2 Fragmenti TNZV dzejoļiem ar $\rho(S) \geq 15\%$

Iet dienas tukšas, vilkdamās,
 Drīz austrā, drīz dziesna tās krāso...
 Bez spēka tās, bez jēgas tās,
 Kā asins iz brūces tās lāso...
 (TNZV, "Mūsu dienas", $\rho(S) = 16.5\%$)

Man mīļāks godīgs pesimists,
 Kas dzīves tukšumu nes, zobus griezdamš,
 Ne vieglis filistris, saukts optimists,
 Kurš šķiet, ka, dzīves nastas projām sviezdams,
 Viņš visus mezglus ir jau atšķetījis
 Tik gudris palikdams, cik agrāk bijis.
 (TNZV, "Pesimists", $\rho(S) = 15.3\%$)

Viss vīlis, viss kritis un zemē rakts,

Apkārt tev noslēpums, šausmas un nakts —

Sakost tik zobus un dzīvi nest,

Kamēr vēl krūtīs spēj dvašu dvest, —

Sakost tik zobus un ciest, un ciest,

Lai arī pēdējā cerība dziest,

Lai visi locekļi lūst un līkst:

Tumsas dziļumos gaisma dzimst!

(TNZV, "Ziedu sapņi", $\rho(S) = 15.1\%$)

Uz debess dziļi gulies mākonis

Un smagā tumsā sauli žņaugdams sedzis;

Aiz segas uguns, kurš tik verdošs dedzis,

Kā matu zelts līst šķiesnām iziris. (TNZV, "Bula laiks", $\rho(S) = 15.0\%$)

4 Vārdformu biežums

Katram rakstītājam ir viņa stilu raksturojošs relatīvais vārdu un vārdformu biežums. Šo īpašību nereti izmanto, lai noskaidrotu, vai kādu tekstu ir rakstījis viens un tas pats cilvēks. Precīzākai diagnostikai ir jāizmanto nedaudz sarežģītāki algoritmi - piemēram, atsevišķo vārdu vietā jāaplūko N -vārdu virknītes jeb t.s. N -grammas. Ir arī SVM (*Support Vector Machine*), Beijesa (*Bayes*) klasifikatori un citi paņēmieni, lai ar datora palīdzību atrastu plaģiātu vai spoku-rakstniecības (*ghostwriting*) pazīmes.

Protams, Rainim un O.Vācietim lietotie vārdi un to biežumi būtiski atšķiras. Tabulā 3 apkoposim 120 biežākās vārdformas no $Dz-L$, norādot katrai no tām "biežuma rang", sastapto reižu skaitu (iekavās), kā arī promiles — cik reizes šī vārdkopa sastopama starp katriem 1000 vārdiem šajā tekstā. Blakus stabīnā norādām arī šīs pašas vārdformas biežumu krājumā $TNZV$ (arī norādot sastapto reižu skaitu un promiles). Savukārt Tabulā 4 apkopojam 120 biežākās vārdformas no $TNZV$ — un salīdzinām to biežumu ar $Dz-L$. Protams, vārdu saraksti abās tabulās daļēji pārklājas.

Lai būtu vieglāk noteikt vārdus, kas vairāk raksturīgi vienam vai otram dzejniekam, ieviešam apzīmējumus:

1. ○— Raiņa darbā vārdformas relatīvais biežums ir vismaz divreiz lielāks;
2. ●— Raiņa darbā relatīvais biežums lielāks nekā O.Vācietim, bet mazāk nekā divas reizes;
3. ●— O.Vācieša darbā biežums pārsniedz biežumu Raiņa darbā, bet mazāk nekā divas reizes;
4. ●— O.Vācieša darbā biežums vismaz divreiz pārsniedz vārda biežumu Raiņa darbā.

Secinājumi:

- Krājumā $Dz-L$ ir 13216 vārdformas, starp tām 4869 ir unikālas (t.i. 4869 dažādi vārdi vai to locījumi). Krājumā $TNZV$ ir 7422 vārdformas, starp tām 3270 ir unikālas. No šejienes galīgi neizriet, ka vienam autoram būtu bagātāka valoda nekā otram — vienkārši garākā krājumā ($Dz-L$) ir vairāk unikālu vārdu. Ja skaita unikālās vārdformas līdzīga garuma O.Vācieša un Raiņa dzejas tekstos (piemēram, "nogriežot" $Dz-L$ īsāku, lai tas būtu vienādā garumā ar $TNZV$), sanāk, ka unikālo vārdformu skaits ir aptuveni vienāds: unikālo vārdformu ir aptuveni 40% no visām (citādi sakot, katra vārdforma aptuveni 50,000 rakstzīmju garā dzejas gabalā atkārtojas vidēji 2.5 reizes — gan O.Vācietim, gan Rainim).
- Vārdu biežuma sadalījums $Dz-L$ ir krietni nevienmērīgāks nekā biežuma sadalījums $TNZV$: Ja skatāmies uz abiem vārdu sarakstiem (Tabulas 3 un 4), tad ieraugām, ka biežākie vārdi O.Vācieša dzejā tiek lietoti biežāk, bet retākie vārdi — retāk. Teiksim, O.Vācieša vārds nr.1 ("un") ir ar biežumu 47.22‰: vairāk nekā Raiņa vārds nr.1 ar biežumu 40.96‰ (arī "un"). Līdzīgi O.Vācieša vārds nr.2 ("es") ir ar biežumu 16.57‰: vairāk nekā Raiņa vārds nr.2 ar biežumu 10.51‰ ("kā"). Šī tendence turpinās pirmajiem 52 vārdiem sarakstā. O.Vācieša vārds nr.52 ("tevi") ir ar biežumu 2.04‰: joprojām nedaudz vairāk nekā Raiņa vārds nr.52 ("būs") ar biežumu 2.02‰. Pēc tam tendence apvēršas otrādi — un Raiņa vārds nr.53 ir biežāks nekā O.Vācieša vārds nr.53, utt.
- $Dz-L$ vārdformu sarakstā salīdzinoši augstās vietās parādās netipiski vārdi ("atbalss", "trako", "Līz"), kas tur nonākuši pateicoties bagātīgai atkārtošanai vienā dzejolī.

- *Dz-L* satur daudz vairāk 1. personas vietniekvārdu formu (“es”, “man”, “mēs”, “mums”) nekā *TNZV* – atšķirība relatīvajā biežumā (sk. promilles iekavās) ir vairākkārtīga. Šis ir gadījums, kad vajadzētu analizēt dziļāk: varbūt tā ir O.Vācieša liriskā varoņa īpatnība – runāt pirmajā personā. Tā būtu līdzība ar S.Jeseņinu – sal. “Vēstule vectēvam” – “*Pokinul ja rodimoye žilišće. Golubczik! Deduszka! Ja vnov’ k tebe piszu.*”. Varbūt proporciju iespaido tikai atsevišķi neparastā formā uzrakstīti dzejoļi. Piemēram, “Balsij bez pavadījuma”, kas ir visai garš un var izmainīt līdz nepazīšanai statistiku par krājumu kopumā.
- *TNZV* krietni biežāk nekā *Dz-L* parādās 2. personas daudzskaitļa vietniekvārdi (“jūs”, “jums”).
- Raksturīgs to vārdu saraksts, kas visai izplatīti *TNZV*, bet nav vai ļoti maz sastopami *Dz-L* (ietilpst tur 0 vai 1 reizi): “nau”, “iz”, “ik”, “dvēsele”, “plūst”, “lēni”, “aizvien”, “domas”, “dvēsele”, “gala”, “nedz”, “saule”, “ļaudis”, “mūžam” “prāts”, “spīd”, “tāles”. Ja atmetam dažus vārdus vai gramatiskas konstrukcijas, ko mūsdienās parasti izsaka citādi (“nau”, “iz”, “ik”), tad paliek tieši tie vārdi, kuri apraksta cilvēka gara gaismu, racionālu optimismu, apceri par galīgo un bezgalīgo. Tā varētu būt 1960-to gadu rakstītāju kopīga iezīme, PSRS dzejas redaktoru gaume, O.Vācieša tālaika ekspresionisms, ko varētu salīdzināt ar agrāku ekspresionistu, piemēram, L.Paegles leksiku. Bet tas var būt arī O.Vācieša ticības zudums modernisma un progresā ideālam, “jaunajam laikam”.

5 Līdzīgas vārdu virknītes

Dažreiz dzejas pētnieki aplūko *alūzijas* (norādes uz ārējiem faktiem) vai *reminiscences* (tās uzskatām par alūziju atsevišķu gadījumu – tiešu atsauci vienā tekstā no kāda agrāk radīta teksta). Naivs mēģinājums atrast reminiscences būtu – meklēt divos tekstos līdzīgas frāzes. Protams, tādā veidā nevar atrast lielāko daļu to kopsakarību, kuras pamana vērīgs lasītājs – toties var gūt zināmu priekšstatu par dzejas valodu, ko Ruta Veidemane [8] (un daudzi dzejas pētnieki pirms viņas) raksturojuši kā informatīvi ietilpīgu – ja birokrātiskās un zinātniski tehniskās publikācijās ir parasta lieta sastapt līdzīgus garus formulējumus (ko neatkarīgi viens no otra radījuši dažādi autori), tad dzejas tekstā katrs nākamais vārds ir lielā mērā negaidīts. R.Veidemane to ilustrē ar nespēju atjaunot dzejas tekstu, ja visi vārdi tajā aizstāti ar saīsinājumiem vai no tiem izdzēsti visi patskaņi. Savukārt zinātniski tehniskās literatūras izteiksmes stils ir tik prognozējams, ka var atjaunot bez zudumiem arī tekstu, no kura daļa informācijas ir pieminētajos veidos “saīsināta”.

Nav pārsteidzoši, ka Raiņa un O.Vācieša darbos (*TNZV* un *Dz-L*) ir tikai ļoti īsi kopīgi teksta fragmenti (2–3 vārdu virknītes, kuras eksistē abu autoru dzejā). 3-vārdu virknītes ir tikai četras; tādēļ varam aplūkot tās visas Tabulā 5.

(1) guļ mana dzimtene	
...Daudz simtu jūdžu tālumā, Aiz tīreļiem, purviem un siliem, Guļ mana dzimtene diendusā Tā aizsegta debešiem ziliem, Zil-saulainiem debešu palagiem Pret dvesmām un strāvām, un negaisiem...	Te guļ mana dzimtene. Šeit dzīvoja mana māte Tajā naktī, Kad rados.
(2) un kas tad bija	
Un kas tad bij šis viņa tēvs? Kas māte? Teic: gars un mūžība! — Tie tikai vārdi: Tēvs — vecais pilsoniskais gadsimts, Un māte — staltā Vaļība? — par kuru Viņš karstām jaunasinīm cīnījās.	Nu un kas tad būtu? Mēs te gaudojam kā divas vientuļas kuces, Un es tomēr piederu pie vīriešu kārtas. Pie kā mēs palikām?
(3) es sen jau	
Lai sļauj tev acīs, tu noslauki vien, Jo pilsonim pretī sist nepiedien! — Es sen jau šo mācību dzirdēju, To sauc par “prātīgu rīcību”.	Bezspalvaini Bezdvēseles Zirgi Dūc ap manu stalli, Dūc to pašu, ko es sen jau zinu: “Var-r-ram, var-r-ram, var-r-ram...”
(4) līdz sirmiemi matiem	

Visa viņa dzīve Līdz sirmiem matiem, Pūlēs un sviedros, Bijusi velta — Nebija spējis Sievai un bērniem Sagādāt vairāk Kā nabagu māju.	Visu laiku viņš runā Kā plate: — He, viņš skatās pa logu! Visus kokus bez lapām Noskatīsi. Visas meitenes līdz sirmiem matiem Izskatīsi.
---	---

Tabula 5: Sakrītošas vārdu virknītes

Secinājumi:

- Iespējami garas kopīgas teksta virknītes abu autoru krājumos sastāv no 3 vārdiem (ir ļoti daudz virknīšu no 2 vārdiem, bet nav virknīšu no 4 vārdiem).
- Kopīgās virknītes parasti satur bieži izmantotus vārdus (“un kas tad”; arī praktiski visas no 2 vārdiem sastāvošās kopīgās virknītes), tajās esošais teksts nav pietiekams, lai varētu paredzēt atsauci uz citu tekstu.
- “Guļ mana dzimtene” ir vienīgais fragments, kas varētu būt reminiscence – t.i. Raiņa dzejoļa zinātajam, lasot O.Vācieša dzejoli (*Nodedzinātā Rūjiņa > Piektais kapakmens*), ir iespēja to uztvert “stereoskopiski”. Jau Raiņa dzejoli dzimtene aprakstīta kā vienīga un sargājama (*pret dvesmām un vētrām un negaisiem*) un arī kā attālināta vieta (*Daudz simtu jūdžu tālumā...*). O.Vācieša dzejolis pievieno dzimtenes izjūtas noturību arī pēc dzimtas vietas fiziskas iznīcināšanas (*Un, ja arī sabruks šis pusapdegušais mūrītis ... Te būs mana dzimtene. Jauna mirušās vietā, bet — dzimtene. Jo nevar izkustināt no vietas pirmo blāvienu...*).
- Garāko kopīgo apakšvirkņu meklēšana parasti nav efektīva, lai atklātu reminiscences – t.i. mūsu primitīvā pieeja neder. Cilvēki tekstu asociācijas uztver elastīgāk – nav obligāti, lai tiktu lietoti precīzi tie paši vārdi tajā pašā secībā. No otras puses, reminiscenču (un arī plašā nozīmē – mākslinieciskās izteiksmes līdzekļu) datorizētai meklēšanai ir ievērojama praktiska nozīme, piemēram, lai atvieglotu filozofisko izmeklētāju darbu (sal. G.Čestertons (*Gilbert Chesterton*) *The Man Who Was Thursday > The Tale of a Detective* [2]): “*The ordinary detective goes to pot-houses to arrest thieves; we go to artistic tea-parties to detect pessimists. The ordinary detective discovers from a ledger or a diary that a crime has been committed. We discover from a book of sonnets that a crime will be committed. We have to trace the origin of those dreadful thoughts that drive men on at last to intellectual fanaticism and intellectual crime.*”
- Jau tagad ar teksta klasificēšanas algoritmiem iespējams atpazīt un bloķēt *spamu* jeb *mēstules* – nevēlamu korespondenci, kas piedāvā kaut ko nopirkt. Ja mašīnmācīšanās (*machine learning*) algoritmu apmācītu atpazīt jau zināmus reminiscenču piemērus, tad tas, iespējams, spētu atrast tos arī līdz tam neanalizētos tekstos. Interneta forumos teksta klasifikāciju var izmantot, lai atpazītu un moderētu komentārus, kas ir pesimistiski, norāda uz rakstītāja pašdestruktīvām tieksmēm (*suicidal thoughts*) vai līdzcilvēku vajāšanu Internetā (*cyberbullying*). Vienkāršākos piemērus var izķert ar “vārdnīcas analīzi”, atrodot dažu raksturīgu vārdu paaugstinātu skaitu, bet kvalificētam uzbrucējam šos ierobežojumus viegli apiet.

6 Krāsu nosaukumi

Doma par krāsu nosaukumu analīzi aizgūta no L.Bērziņa kvantitatīvajiem pētījumiem tautasdziesmu poētikā [4]. Viņš rakstīja arī par skaitļu vārdiem, deminutīviem, ritma īpatnībām – daudz ko tādu, kam šajā rakstā nepietika laika.

Rainis dzejolī *Senatne* (luga *Pusideālists*, arī *TNZV*) raksta šādi - “*Dies saulīte atkal trīskrāsaina: Drīz zila, drīz zaļa, drīz sarkana.*” Mūsdienās šis 3 pamatkrāsu uzskaitījums šķiet pašsaprotams, bet Rainis to rakstīja ilgi pirms RGB (red-green-blue) ekrāniem. Pirms krāsaino lampu, gaismas reklāmu un krāsainās televīzijas krāsu saskaitīšana (priekšstats par to, ka aplūkojot cieši tuvu novietotus vai pārmaiņus mirgojošus zilus, zaļus un sarkanus gaismas punktiņus, gaismas krāsa šķitīs balta) nebija ikdienas parādība. Gleznotāji, jaucot krāsas, saskārās ar krāsu atņemšanu (t.i. pievienojot jaunu pigmentu, krāsa kļūst tumšāka nevis gaišāka). Raiņa laikabiedri redzēja, ka, sajaucot zilo, zaļo un sarkano krāsu, var iegūt netīri brūnu. Bet dzejnieka intuīcija ļāva saskatīt balto.

Citā dzejolī no *TNZV* – “*Kur iespiesties nespēj sarkan-zeltainā, karstā, septiņkrāsainā pasauls saule*” – ir norāde uz saules spektra sadalīšanu 7 varavīksnes krāsās, kas labāk atbilst populārzinātniskiem priekšstatiem 1903. gadā –

Ņūtona eksperimentam ar stikla prizmu. Spēja ieraudzīt varavīksnē tieši 7 krāsas nav saistīta ar latviešu valodu, kurā nav atsevišķu vārdu “gaiši zilam” un “tumši zilam” – atšķirībā no angļu valodas, kur ir “blue” un “indigo”. Faktiski varavīksnē ir neierobežoti daudz grūti nosaucamu krāsu (neskaitāmas toņu pārejas no sarkana uz oranžu utt.). Varavīksnes septiņkrāsainība drīzāk raksturo Īzaka Ņūtona priekšstatus par numeroloģiju nevis fizisku realitāti.

Apkopojam tabulīnā kontekstus, kur minētas krāsas. Parasti krāsas apzīmē īpašības vārdi, bet arī daži apstākļa vārdi (“**balti** klāja”), lietvārdi (“**manu pilno zelta** matu”), darbības vārdi (“desmitstaraina austriņa **sārtojas**”) un darbības vārdu divdabju formas (“uz ar sūnu **apzaļojuša** akmens”) norāda uz krāsu. Aiz katras krāsu vārdformas norādām locījumu (F/M - sieviešu/vīriešu dzimte; I/D - nenoteiktā/noteiktā galotne; S/P - vienskaitlis/daudzskaits; N, G, D, A, I, L, V - locījums). Piemēram FISN apzīmē sieviešu dzimtes nenoteiktās galotnes vienskaitļa nominatīvu (“balta”, “melna” u.c.). Ja no krāsas nosaukuma darināts apstākļa vārds, lietvārds, darbības vārds vai divdabjs, rakstām aiz vārda attiecīgi **adv, noun, verb, part**.

No krāsām atvasinātus vārdus, kas attiecīgā dzejoļa kontekstā neveido vizuālu tēlu (“zilskābe”, “Sarkandaugava”) pārskatā neiekļaujam; savukārt “zilums” (t.i. sasitums, kura zilā krāsa ir būtiska Dz-L dzejolī) tiek pieskaitīts. Dzejoļiem ar apla kompozīciju nepieskaitām divas reizes tos fragmentus ar krāsu nosaukumiem, kuri pilnībā vai gandrīz pilnībā atkārtojas (*Ap manu bērniņu apkārt ir apkritis pelēku lapu pelēks lapkritis* un *Kad ap bērniņu bija apkritis pelēku lapu pelēks lapkritis* – dzejolis “Dzērves” no Dz-L). Ja krāsas nosaukums izmantots atkārtoti pastiprinājuma nozīmē, to pieskaita kā vienu piemēru (*Drīz būs zili, zili ūdeņi, drīz būs balti, balti pūpoli* – te vienreiz pieskaitīts “zils” un vienreiz “balts”). Savukārt, ja piemērā tā pati krāsa parādās divās dažādās lomās, to pieskaita divreiz (*Baltā māt, veļu māt, mazgā baltu paladziņu* un *Pāri man ar visbaltāko baltumu debess jūra savas bangotnes putotos mākoņus nes*).

Krāsas ņemtas no šāda saraksta – “melns”, “balts”, “sarkans”, “zaļš”, “zils”, “oranžs” (arī “ruds”, “rūsgs”, “rūsgans”, “rūsas”), “dzeltens” (arī “zelta”, “zeltains”, “bēšs”, “krēmkrāsas”), “violets” (arī “mēļš”, “purpura”, “lillā”), “sārts” (arī “rozā”), “brūns”, “pelēks” kā arī “raibs”. Kopā ar krāsu pamatnosaukumiem aplūkoti arī atvasināti vārdi (“zaļgans” kopā ar “zaļš”, “sniegbalts” kopā ar “balts”, “purpurkvēls” kopā ar “violets”). Konteksts katrai no krāsām izraudzīts tā, lai dzejas citāta garums nepārsniegtu 100 rakstzīmes – parasti ņemtas dažas blakus rindas, retāk pilns pants.

Krāsa		Konteksts
melnā	FDSN	Bet melnā čūska maļ jūrā miltus, \ Tos jāēd būs tiem, kam vara un viltus
melns	MISN	Pār zemi pārlaidās pāri \ Melns mākonis valguma
melni	MIPN	Un gaisā celsies melni putekļi \ No izžuvašā asins plūduma
melns	MIPA	Pa klājiem laukiem aukas kauc, \ Pie debesīm melns mākoņus jauc
melnos	MIPL	Gan melnos dumbros \ Jūs skriesat līst \ Pret jums tie balti, \ Jūs neatzīst
melna	MDSN	Sēd tā zelta krēslā, \ Acis drūmi kvēlo, \ Sēd un vērpj; tai melna suns pie kājām
melna	MDSN	Melna suns pa laikam \ Norūc, paceļ galvu, \ Tālu augšā juzdams asins smaku
balta	FISN	Un mīļā dzimtene guļ manā priekšā \ Kā plaši izplatīta balta sega
balta	FISN	Tur balta laime man liekas, \ Tur tu, tur nogrimst daurs
balta	FISN	Gars brīvus spārnus dzirdē cilā liegi, \ Tik balta dvēse mirdz kā kalnu sniegi
baltā	FISL	Tā staro klusi un pastāvīgi \ Baltā vienkrāsas gaismā
baltas	FIPN	Lai baltas viņām rokas \ Un tīnes no pilnības tūkst
baltas	FIPN	Zaļoši tīrumi, \ Puķotas pļavas, \ Zib baltas kājiņas, \ Sīk-žigli tekot
baltās	FDSG	Zem baltās segas svaigi kapi - - \ No svaigiem kapiem svaigas asins kūp
sniegabaltās	FDSG	Tad beigās izaugs pastars saules dzimums \ Uz skaidrās sniegabaltās svētku zemes
baltā	MISL	Rasas klātu zemes vaigu, \ Baltā asru autā tītu
balti	MIPN	Zib balti zirnīši, \ Bērti pa plānu:
balti	MIPN	Un balti kauli \ Vēl laukus klās, \ Daudz asins plūdīs, \ Pirms vaidi stās
balti	MIPN	Gan melnos dumbros \ Jūs skriesat līst \ Pret jums tie balti , \ Jūs neatzīst
balti	adv	Visu nakti rasa rita, \ Balti klāja zemes vaigu, \ Zaļā sagšā mēmi krita
balti	adv	Drīz kaiva krīt, drīz balti spīd, \ Kā dvēsele bez miera šaudās
balti	adv	Debesīs putnu ceļš \ Balti zib zvaigznēm
sarkana	FISN	Dies saulīte atkal trīskrāsaina: \ Drīz zila, drīz zaļa, drīz sarkana
sarkana	FISN	Aiz tālām zemēm šie pirkstiņi mīt, \ Bet sarkana saule caur tālēm spīd
sarkanas	FIPN	Div' sarkanas saulītes aust šobrīd, \ Kad gaisma caur pirkstiem cauri spīd
sarkan-zeltainā	FDSN	Kur iespiesties nespēj \ Sarkan-zeltainā , karstā, \ Septiņkrāsainā pasaule saule
sarkans	MISN	Un acu priekšā sarkans izplešas \ Kā ziemeļblāzma - jaunais gadsimtenis
sarkans	MISN	Arī viņam kauss \ Bij rokā - glums un sarkans receklis, - \ Bet otrā rokā viņam bija vējš
sarkani	adv	Tā tērpta sarkani ar zelta audiemi, \ Smags zīda mētels tai pār pleciem čaukst
sarkanšvītras	noun	Ak, jāsteidz izteikt, — gaust nau laiks, nau telpas, \ Drīz sarkanšvītras pantos iegriežas
zaļa	FISN	Dies saulīte atkal trīskrāsaina: \ Drīz zila, drīz zaļa , drīz sarkana
zaļa	FISN	Atplešat tad logu slēģus: \ Visur apkārt skaidrā tālē \ Jauna, zaļa gaisma bālē, \ Rīta rasa mirgo zālē
zaļa	FISN	Pili dzintarsienas, \ Zaļa vara grīda, \ Gliemežgriesti, sijas tēraudzilas
zaļai	FISD	Ka elpa aizrāvās no ļauniem tvaikiem \ Un zaļai dzīvībai bij jānosmok
zaļā	FISL	Visu nakti rasa rita, \ Balti klāja zemes vaigu, \ Zaļā sagšā mēmi krita
zaļā	FDSN	Pirms vēl sākās pavasara \ Saules siltā, zaļā vara
zaļ-zaļā	FDSN	Un galviņu lieca \ Pie zemes krūts \ Zaļ-zaļā zāle
zaļgans	MISN	Viz bāli zaļgans zaigums debess tālē, \ No saules rīta zelta migla plūst
zaļi	MIPN	Bet trūdiem cauri zaļi asni dīgst, \ Un jauna pasaule virs vecās ceļas

zaļi	MIPN	Daudz karstas sirdis\ Pukstēt rims,\ Daudz zaļi stāvi\ Vēl lauzti jims
jaunzaļi	MIPN	Ilgi ilgotas veldzošas vēsmas,\ Lai jaunzaļi asni dīgst
zaļoši	MIPN	Zaļoši tīrumi,\ Puķotas pļavas:\ Zib baltas kājiņas,\ Sīk-žigli tekot
zaļais	MDSN	Man sirds ir pilna mīlas\ Kā majja zaļais ārs,\ Kā smaršas puteklīšu\ Pilns ziedošs ābeļdārzs
zaļi	adv	Ko jūs vēl skumstat,\ Jūs jaunās dvēseles,\ Tik zaļi sārtas?
zaļos	verb	No dusmas mostas viss pasauls ārs,\ Drīz zaļos zeme viens vienīgs dārzs
zila	FISN	Dies saulīte atkal trīskrāsaina:\ Drīz zila , drīz zaļa, drīz sarkana
zilu	FISA	Un mīlā dzimtene guļ manā priekšā\ Kā plaši izplatīta balta sega,\ Ar zilu jūras lentu apšūta
tēraudzilas	FIPN	Pili dzintarsienas,\ Zaļa vara grīda,\ Gliemežgriesti, sijas tēraudzilas
zilganu	FIPG	Met zilganu dzirksteļu mirgas\ Iz spēka pilnās sirds
zilās	FDPN	Pret mīksto miesu šļākstot sper kā lode,\ Visapkārt sprakstot zilās pupas plīst
zils	MISN	Pus-pelēks un pus-tumši zils \ Spīd ūdens pretī debess jumam
zil-augstu	MIPG	Zil-augstu kalnāju\ Sniegotā gultā\ Guldīju galvu,\ Lūkoju tālē
ziliem	MIPD	Bez gala upe sniecas\ Uz ziliem kalnājiem.\ Tur debess pret zemi liecas\ Aiz mūžu tālumiem
ziliem	MIPI	Aiz tīreļiem, purviem un siliem,\ Guļ mana dzimtene diendusā\ Tā aizsegta debešiem ziliem
zil-saulainiem	MIPI	Zil-saulainiem debešu palagiem\ Pret dvesmām un strāvām, un negaisiem
zilos	MIPL	Cieti ieaudzis stāv viņš savā zemē,\ Lidot nejaudzams augstos, zilos gaisos
zili	adv	Tu, tāles plīvurs,\ Zili blāvs,\ Jel paveries\ Šo brīd', šo brīd'
zeltītas	FIPN	Lūk, zeltītas lapas\ Trīs kalna bērzu\ Galotņu zaros
sarkan-zeltainā	FDSN	Kur iespiesties nespēj\ Sarkan-zeltainā , karstā,\ Septiņkrāsainā pasauls saule
zelta	noun	Sēd tā zelta krēslā,\ Acis drūmi kvēlo,\ Sēd un vērpj; tai melnais suns pie kājām
zelta	noun	Pāri upei mēness meta\ Tiltu zelta stariem:\ Pārņākt sapņu gariem\ Miglai līdz iz tumšiem mežiem
zelta	noun	Viz bāli zaļgans zaigums debess tālē,\ No saules rietā zelta migla plūst
zelta	noun	Manu pilno zelta matu\ Un ziedošo vaigu tām trūkst
zelta	noun	Tie vijas ap tevi un viņo\ Kā kūpošs zelta tvaiks
zelta	noun	Tie klust un klist, un saplok,\ Viss zelta mirdzums tumst
zelta	noun	Tā tērpta sarkani ar zelta audiem,\ Smags zīda mētels tai pār pleciem čaukst
zelta	noun	Bez žēluma, bez pagātnes, bez vainas\ Tu skati priekšā zelta rīta ainas
zelts	noun	Tev birzē ritēs rasas zelts ,\ Man vecais slogs būs plecos velts
zelts	noun	Aiz segas uguns, kurš tik verdošs dedzis,\ Kā matu zelts līst šķiesnām iziris
purpurkvēls	noun	Ir lepni nesalauzts, ap viņa vājām\ Un kurlām miesām skrandains purpurkvēls ?
sārta	FISN	Tālos mežos saule slīka:\ Rasa klāja zemes vaigu,\ Sārta blāzma vaigā tvīka
sārta	FISN	Timeklā sārta \ Ir mīlas dvēsele,\ Irst, tikko skārta
sārtas	FIPN	Ko jūs vēl skumstat,\ Jūs jaunās dvēseles,\ Tik zaļi sārtas ?
sārtos	MIPL	Viņš savu dvēseli jums pretim nesa\ Kā rīta sauli sārtos mākoņos
sārtos	MIPL	Bet savu dvēseli viņš tālāk nesa\ Kā rīta sauli sārtos mākoņos
sārtais	MDSN	Mans sārtais ziedīņš, mans vientulīts,\ Kas mīlā saulē plaucis
sārtais	MDSN	Mans sārtais ziedīņš, starodams\ Tu saulīti sirdī man nesi
sārtojas	verb	Jūs esat skaistās lauka lilijas. -\ Kā viņas zied un sārtojas !
sārtojas	verb	Tu izplēt viņus smiedamās\ Desmitstaraina austriņa sārtojas
sārtojas	verb	Viss, ko tā teic, tiek tūdaļ izpildīts,\ Līdz zeme pārplūzdama sārtojas \ Un asins atspīd debesīs kā blāzma.
pelēka	FISN	Tā patiesa dzīve bez apsega:\ Pelēka , netīra, pūstoša
pelēks	MISN	Sen saule aug, sen gaisi silst,\ Sen marta ledus pelēks dilst
pus-pelēks	MISN	Pus-pelēks un pus-tumši zils\ Spīd ūdens pretī debess jumam
pelēkzīda	noun	Vēsmas drēbe\ Pelēkzīda ,\ Miglas šūta;\ Pērļu smaga\ Vīle velkas,\ Liecas rūta

Tabula 6: Krāsu nosaukumi TNZV

Tā pati tabula krājumam Dz-L:

Krāsa	Konteksts
melnai	FISD Rīti.\ Melnai pusnaktij dzimušie bērni\ Iet no loga pie loga,\ Un Rīgai logu ir daudz
melnu	FISA Un kāpēc man visu\ Pa melnu nakti,\ Pa šautriem sapņiem\ Vajag izdomāt, izdomāt
melnā	FDSN Un neatsedzas\ Ne trauks, no kā lej,\ Ne melnā kapuce,\ Zem kuras slēpjas\ Tā draudīgā seja
melno	FDSA Kā sev no atmiņas un acīm izsviest\ To balto, sarmas mežģīņoto rītu\ Un melno kara hospitāļa bedri
melnās	FDPN Jo bedrē bija sasviests juku jukām\ Gan asiņainas marles melnās skrandas,\ Gan kaut kas sarkans, notraipīts ar brūnu
melns	MISN Tad dienu baržas —\ Smagi krautās baržas —\ Melns , miris ūdens\ Var tālāk vest
melnajiem	MDPD Stiepijas vilcienu stigas\ Ar gulšņu melnajiem taustiņiem,\ Stiepijas tālruņa vadu\ Tūkstoš vijoļu stigas
balta	FISN Tu teici,\ Ka paliksi balta , —\ Un dzīvoši\ Mirkli
baltu	FISA Vai tu atceries vēl, kalēj,\ Večuku ar baltu bārdu
baltu	FISA Es tevi iemīlēju\ Vasarā, redzēdams baltu
baltā	FDSN Stāmerienas muižas liesmās\ Baltā bārda rūsga svila\ Vēl no Piekta gada dūmiem\ Gadu pamale ir zila
baltā	FDSN Un tāpēc nāvei — nē! Un salam — nē!\ Lai baltā nāves dvaša kūp no mutes
balto	FDSA Bet bariņš zilišu jau metās bedrē\ Un sāka knābāt tieši balto kāju
balto	FDSA Nenomest ziemas glābjošo\ Balto spalvu\ Un nekļūt rudai
baltā	FDSV Baltā māt, veļu māt,\ Mazgā baltu paladziņu\ Jau ganās mani zirgi\ Viņas saules pļaviņā
baltā	FDSV Baltā māt, veļu māt,\ Ņem mani savā daļā,\ Tur jau īd manas govīs\ Neganītas, neizslauktas
baltā	FDSV Baltā māt, veļu māt,\ Aizver man acu vārtus,\ Kad atnācu, tad atvēru,\ Nu ir reize aizdarīt
baltā	FDSV Mana baltā un mierīgā māt,\ Es tev zilganu asteri atnesu.\ Ir pie Kalnciema lielceļa dārzs
baltā	FDSV Mana baltā un mierīgā māt,\ Drīz jau leduspuķes vien ziedēs te,\ Es tās ņemšu un nesīšu tev
baltu	MISA Baltā māt, veļu māt,\ Mazgā baltu paladziņu\ Jau ganās mani zirgi\ Viņas saules pļaviņā
balti	MIPN Drīz būs zili, zili ūdeņi,\ Drīz būs baltī, baltī pūpoli,\ Drīz es beigšu sevi mierināt

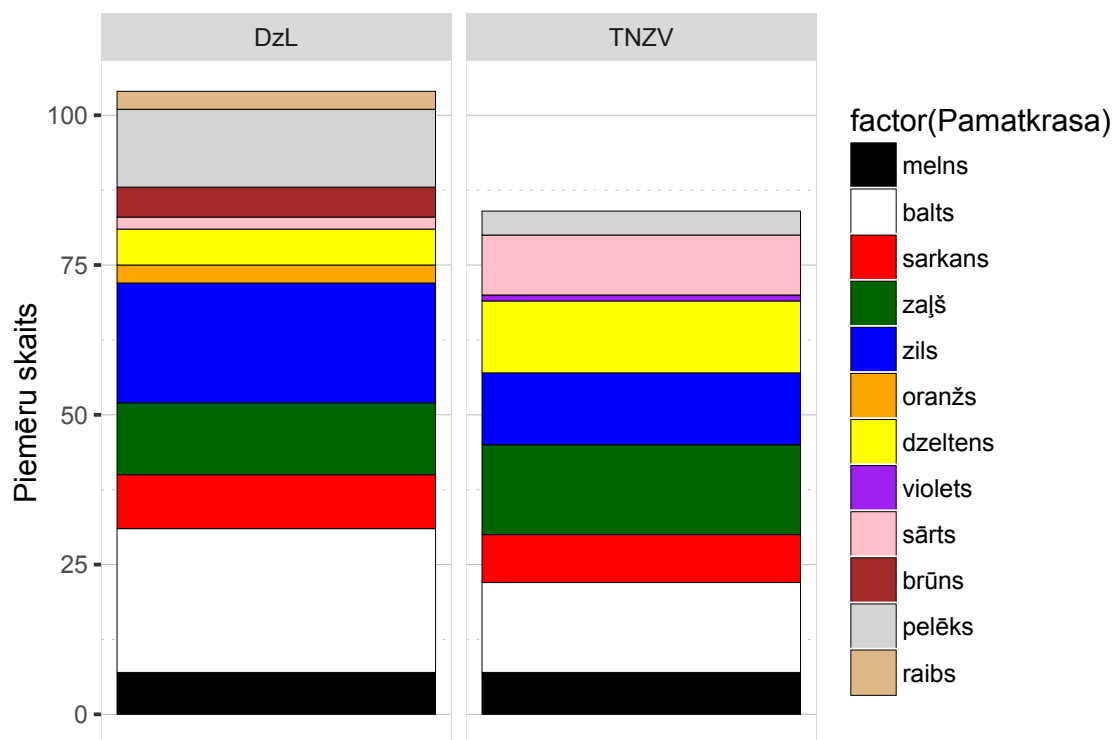
balti	MIPN	Rūciet, kamēr vēl nav zili ūdeņi,\ Rūciet, kamēr vēl nav baltī pūpoli
baltu	MIPG	Gaujas balss ilgi meklēts un saukts,\ Baltu mākoņu kupenās jaukts,\ Ziedu sildīts un viesuļu nests
baltiem	MIPD	Ceļi kūp.\ Ar baltiem , baltiem dūmiem ceļi kūp.\ Uz galda ceļamaizes rieciens drūp
balto	MDSA	Kā sev no atmiņas un acīm izsviest\ To balto , sarma mežģīņoto rītu\ Un melno kara hospitāļa bedri
visbaltāko	MDSA	Pāri man\ Ar visbaltāko baltumu\ Debess jūra\ Savas bangotnes putotos mākoņus nes
baltais	MDSV	Es tev piegāju klāt,\ Baltais brīnum,\ Un redzēju,\ Ka bises tu esi saskaitījusi
baltie	MDPN	Tikai pēc tam lietus varēja līt vai nelīt —\ Mani interesēja tikai zibeņu baltie raksti
baltie	MDPN	Pland matu baltie lini,\ Iet meitene iekš mini\ Un modri skatās apkārt
baltums	noun	Bet vectēvam pašam\ Gar deniņiem baltums \ Un rokas uz spieķa\ Kā puķes pēc salnas
baltumu	noun	Pāri man\ Ar visbaltāko baltumu \ Debess jūra\ Savas bangotnes putotos mākoņus nes
sarkanai	FISD	Un nekas, ja jūs saskrāmbās zari,\ Nolis pa sarkanai lāsei...\ Toties jūs zināsit
sarkanu	FISA	Ja manam vectēvam būtu pietrūcis\ Tavu\ Ar sarkanu uguni degošu skalu
sarkanām	FIPD	Kad uzspīdēs saule,\ Asaras nožūs\ No bērzu zeltaini smalkām\ Un apšu sarkanām notīm
sarkanām	FIPD	Pāri sirds malām kūšās —\ Un uz tepiķa gulēs\ Runcis ar sarkanām ūsām
sarkans	MISN	Gan kaut kas sarkans , notraipīts ar brūnu,\ Gan brūns ar sarkanu. Un viena kāja
sarkanu	MISA	Gan kaut kas sarkans, notraipīts ar brūnu,\ Gan brūns ar sarkanu . Un viena kāja
sarkanie	MDPN	Kuras stacijām jāpieņem\ Straujie un sarkanie asiņu vilcieni
sarkani	adv	Liku tā kā pie nepazīstama suņa purna,\ Kā pie sarkani sakurinātas plīts
sarkstiet	verb	Sarkstiet ,\ Lai zin, kādā krāsā\ Pēc tam\ Degt jūsu mājām
zaļas	FISG	Te beidzās viena no ieliņām,\ Un te es atradu šķiltavas\ No zaļas čaulītes
tumšzaļas	FISG	Šajos cerīgos rīt ziedi būs —\ Mēļas liesmas no tumšzaļas pekles
zaļas	FIPN	Paskaties palmā pār viņiem —\ Palmas lapas kā zaļas rokas ar izplestiem pirkstiem
zaļās	FIPD	Un es šūpojos šinīs bizēs,\ Gan tumšās un zaļās , gan ziedošās un sirmās
zaļajā	FDSL	Pēc tam bija krusa.\ Tā krita patvertnes uzbērums zaļajā zālē,\ Ripoja lejup pa Doma zaļgano jumtu
zaļgans	MISN	Un, ka vasarā birzī no četrām pusēm\ Mūs sargāja zaļgans un caurspīdīgs aizkars,\ Atceries
zaļi	MIPN	Un pirmā latviešu elektrība,\ Un zaļi matī\ Uz kāpu skausta
zaļgani	MIPN	Tad iešalcas meža skaņa,\ Pa gravām kāpelē\ Zaļgani paegļu vīriņi
zaļgana	MDSN	Baznīcu torņu zaļgana is pelējums\ Un jauno kvartālu\ Augšanas untums
zaļgano	MDSA	Pēc tam bija krusa.\ Tā krita patvertnes uzbērums zaļajā zālē,\ Ripoja lejup pa Doma zaļgano jumtu
zaļgano	MDSA	Mašīnu jūklis.\ Reklāmas uguns.\ Taksometru pietura.\ Tas ar pelēko, tas ar zaļgano mēteli
apzaļojuša	part	Nenosēdēt\ Pat uz paša mīkstākā\ Ar sūnu apzaļojuša akmens
zila	FISN	Stāmerienas muižas liesmās\ Baltā bārda rūsga svila,\ Vēl no Piekta gada dūmiem\ Gadu pamale ir zila
zila	FISN	Jo lāse ir viršu zila .\ Aiz prieka par sauli\ Tos vēji ir sajaukuši
zila	FISN	Kur tikko redzami pulsē\ Zila bērniņa tava,\ Sārta jaunība tava\ Un zils vakars tavš
zila	FISN	Tikai āboli krīt\ Rudens rasainā zālē.\ Rasa zila kā tāle,\ Kurp tu aizbrauksi rīt
zilu	FISA	Zilu velvi pār galvu es esmu saņēmis,\ Steigā dunošu zemi\ Tēvs un vectēvs zem kājām man nolīcis
zilganu	FISA	Mana baltā un mierīgā māt,\ Es tev zilganu asteri atnesu.\ Ir pie Kalnciema lielceļa dārzs
zilā	FISL	Kvēlos liesmā un plaiksnīšies sārts\ Zilā debesu sejā mums pāri
zilu	FIPG	Nakti, kurā it kā miljons mieru,\ Miljons zilu zvaigžņu pāri spīd
zilganu	FIPG	Es neprotot laist ziepju burbuļu krāsainās domas pie kafijas tases un zilganu actiņu bezdibens tukšumā
zilacaino	FDSA	Un skaisto, zilacaino netikli rudzupuķi,\ Uz kuru rudeņos visas vārpas sāk skatīties,\ Atceries
zilās	FDPN	Paegļu zilās ,\ Apsūbējušās ogas,\ Kaut arī tās\ Ne reizi nav runājušās ar zemenēm
zils	MISN	Krūtīs sāpēs kāds aromāts zils ,\ Un viens atlidos, un otrs atbrauks
zils	MISN	Kur tikko redzami pulsē\ Zila bērniņa tava,\ Sārta jaunība tava\ Un zils vakars tavš
zili	MIPN	Drīz būs zili , zili ūdeņi,\ Drīz būs balti, balti pūpoli\ Drīz es beigšu sevi mierināt
zili	MIPN	Rūciet, kamēr vēl nav zili ūdeņi,\ Rūciet, kamēr vēl nav balti pūpoli
zili	MIPN	Kas par miežiem!\ Mieži — tie ir tur —\ Aiz tiem zilgiem torņiem\ Zili meži
ziliem	MIPD	Aiz tiem ziliem miežiem\ Tāli lauki,\ Tur, tur — akotu pie akota\ Ilgūciema alus mieži aug
zilgiem	MIPD	Kas par miežiem!\ Mieži — tie ir tur —\ Aiz tiem zilgiem torņiem\ Zili meži
zila	MDSN	Jauno dienu zila is asmens nodila.\ Večuks tirgojās ar cīrvja kātiem
ziluma	noun	Kritiet, mani bērni,\ Es papūtīšu uz ziluma —\ Tā nav nemaz tik briesmīga krāsa
rūsga	FISN	Stāmerienas muižas liesmās\ Baltā bārda rūsga svila,\ Vēl no Piekta gada dūmiem\ Gadu pamale ir zila
rudai	FISD	Nenomest ziemas glābjošo\ Balto spalvu\ Un nekļūt rudai
rudā	FDSN	Viņš nomirs, jo liesmā rudā \ Jauns sprādziens ar zemi to sakausēs
dzeltena	FISN	Vai tev ir viena dzeltena lapu kaudze\ Priekš diviem blandoņām
dzeltenas	FISG	Bet ne velti pats gadsimts\ Klieudz pēc māla un koka\ Un pēc vienas dzeltenas \ Bērzu ataudzes
dzeltenas	FIPA	Šķir vēsture vecas dzeltenas lapas\ Un ierauga\ Kardinālu
dzeltenīgs	MISN	Bet tagad līst lietus.\ Līst dzeltenīgs lieplapu lietus.\ Par ko tu domā\ Šai pelēkā, pelēkā lietū
dzeltenu	MIPG	Dzeltenu kļavziedu pilni matī,\ Un bites pinas kā trakas,\ Un pirmie negaisi danco apkārt\ Kā līgavai
zeltaini	adv	Kad uzspīdēs saule,\ Asaras nožūs\ No bērzu zeltaini smalkām\ Un apšu sarkanām notīm
sārta	FISN	Kur tikko redzami pulsē\ Zila bērniņa tava,\ Sārta jaunība tava\ Un zils vakars tavš
rozā	FISN	Tā nav nemaz tik briesmīga krāsa.\ Briesmīga ir tā rozā \ Uz taukiem purniem
brūns	MISN	Gan kaut kas sarkans, notraipīts ar brūnu,\ Gan brūns ar sarkanu. Un viena kāja
brūnu	MISA	Tos ķeizarkroņus,\ Kurus es ostīju kādreiz\ Un nosmērēju sev degunu brūnu \ Ar putekšņiem
brūnu	MISA	Gan kaut kas sarkans, notraipīts ar brūnu ,\ Gan brūns ar sarkanu. Un viena kāja
brūnu	MISA	Mirguļo kupenas,\ Un upēs ir daudz ūdens.\ Lieli kuģi var peldēt\ Pa tādu dziļu un brūnu ūdeni
brūni	adv	Džimlaurīdi, puto brūni \ Sveša alus stops,\ Spārdās svešas zemes šķūnī\ Svešas zemes lops
pelēkas	FIPN	Es redzēju, kā krita dzērves,\ Kā sniga līdz pašam tālākam mežam\ Pelēkas pārslas
pelēku	FIPG	Ap manu bērniņu apkārt ir apkritis\ Pelēku lapu pelēks laprkītis
pelēkās	FDSG	Atbalss, bet paskaties apkārt,\ Vai tu redzi pelēkās ikdienas pīkus,\ Ikdienas rūpju trulos asmeņus
pelēkās	FDPN	Pār visiem, kas krita,\ Dzervju pelēkās spalvas krita
pelēks	MISN	Ap manu bērniņu apkārt ir apkritis\ Pelēku lapu pelēks laprkītis
pelēks	MISN	Katrā solī, ko tālāk tu liec,\ Pelēks , silts, samīts zaķītis klieudz.\ Tā jau nav kāda varena balss

pelēks	MISN	Bet tikai, uz darbu ejot, \ Vēss, pelēks vējš, \ Spuldžu zvaigznes \ Un kvartālu dzirkstošie mākoņi \ Man palīdzēja
pelēkā	MISL	Ko tu domā šai pelēkā, pelēkā lietū, \ Mana dzimtā un bezgala tālā puse
pelēkā	MISL	Manai dzimtajai pusei \ Pelēkā lietū ir auksti
pelēkā	MISL	Bet tagad līst lietus. \ Līst dzeltenīgs lieplapu lietus. \ Par ko tu domā \ Šai pelēkā, pelēkā lietū
pelēki	MIPN	Pār krūmu, kurš mani slēpa, \ Krita pelēki sniegi, \ Pār vāgūzi krita, \ Pār kaķa raibajiem sāniem
pelēko	MDSA	Mašīnu jūklis. \ Reklāmas uguns. \ Taksometru pietura. \ Tas ar pelēko , tas ar zaļgano mēteli
pelēko	MDSA	Es pats redzēju, \ Kā tu baroji \ Tur to pelēko kaķi, \ Kas pieder visiem
raibā	FISL	Es ķeros koku zaros \ Un raibā šlipsē karos, \ Un izveļos no gultas, \ Un tūlīt mostos ar
raibajā	FDSL	Lai blakus karājošies \ Ar mīlīgu smaidu \ Neietriec dunci tā raibajā mugurā
raibajiem	MDDP	Pār vāgūzi krita, \ Pār kaķa raibajiem sāniem, \ Pār visiem, kas krita

Tabula 7: Krāsu nosaukumi Dz-L

Secinājumi:

- Melno krāsu Raiņa *TNZV* daudz biežāk lieto vīriešu nekā sieviešu dzimtē (6 pret 1), savukārt balto – biežāk sieviešu dzimtē (8 pret 4). Tas nozīmē, ka šajā krājumā melnās un baltās krāsas simbolisms izpaužas citādi nekā lugā *Zelta zirgs* ar Melno Māti un Balto Tēvu. Ķīniešu valodā ietvertajam simbolismam tuvāks būtu *Zelta zirgs*, jo tur baltums, gaisma parasti saistās ar vīriešu dzimti (*Tàiyáng* jeb “Lielais *Yang*” nozīmē “Saule”; savukārt *Tàiyīn* jeb “Lielā Yin” nozīmē “Mēness”). Rainis šādu dzimtes/krāsu simbolismu konsekvēnti neievēro pat *Zelta zirgā*: Melnās Mātes pavadoņi ir melnie kraukļi, bet savdabīgs kraukļu pretstats ir Taurenīte-Dvēselīte.
- O.Vācieša dzejoļos “melnaiss” un “baltais” nav tik stipri asociēti ar vienu dzimti. Toties apzīmējums “pelēks” daudz biežāk saistās ar vīriešu nekā sieviešu dzimti (9 pret 4).
- Raiņa dzejolī sastopama arhaiska forma – instrumentāla locījums: “*Tā aizsegta debešiem ziliem, zil-saulainiem debešu palagiem*”. Mūsdienās biežāk raksta prievārdu kopā ar akuzatīvu vai datīvu: “aizsegt **ar** kaut ko”.
- TNZV* neizmanto īpašības vārdu “dzeltens”, toties nereti veido vārdkopas ar “zelta” (krāsas nevis materiāla nozīmē) – “*sēd tā zelta krēslā*”, “*tiltu zelta stariem*”.
- Dz-L* atšķirībā no *TNZV* bieži atkārtoti krāsu vārdus pastiprinājuma nozīmē. (“Drīz būs zili, zili ūdeņi, drīz būs balti, balti pūpoli”, “Ko tu domā šai pelēkā, pelēkā lietū, Mana dzimtā un bezgala tālā puse”, “Pāri man ar visbaltāko baltumu”). Citādi latviešu valodā pastiprinājumu parasti neizsaka ar vārda atkārtotāšanu, teiksim, Vecās Derības grāmatas nosaukumu, kas burtiski būtu “Dziesmu Dziesma”, parasti tulko “Zālamana Augstā Dziesma”.
- Abi autori tikpat kā neizmanto salīdzināmās pakāpes (pārāko un vispārāko pakāpi) krāsu adjektīviem. Vienīgais izņēmums ir no *Dz-L*: “Pāri man ar **visbaltāko baltumu** debess jūra savas bangotnes putotos mākoņus nes”. Salīdzināmo pakāpju trūkums liecina par loģisku valodas lietošanu, jo krāsa ir kvalitatīva nevis kvantitatīva pazīme: vai nu kaut kas ir sarkans vai nav. Salīdzināt vienas krāsas intensitātes ir jēga tikai retos gadījumos.
- Kā redzams Attēlā 6, krāsu adjektīviem ar noteikto galotni ir gluži cits krāsu sadalījums nekā adjektīviem ar nenoteikto galotni. Starp krāsu vārdiem ar noteikto galotni ir lielākas izredzes sastapt krāsas “melns” un “balts” (t.i. “melnaiss” un “baltais” dažādos locījumos) – attiecīgi 15.2% un 34.8%. Starp vārdiem ar nenoteikto galotni (“melns” un “balts” dažādos locījumos) veido attiecīgi 6.3% un 16.2% no visiem krāsu nosaukumiem. (Melnie un baltie laukumiņi Attēlā 6 izskatās gandrīz vienādi – bet piemēru ar noteiktajām galotnēm ir krietni mazāk: 46 pret 111).
- Kā redzams Attēlā 8, krāsas adjektīvs “zils” ir ar īpaši augstu īpatsvaru starp daudzskaitļa formām (t.i. “zili” vai “zilas” dažādos locījumos).
- Kā redzams Attēlā 5, O.Vācieša dzejas valodā ir dažas krāsas, kuru Raiņa *TNZV* nav. Piemēram, “rūsgrs/ruds” (diagrammā zīmēts oranžs), “rozā” (Rainim tikai “sārts”), “brūns” un “raibs”. Toties Rainim ir “purpurkvēls”.
- Dažreiz atšķirīgi krāsu nosaukumi (vai norādes uz tiem) lietoti tās pašas lietas vai parādības apzīmēšanai. Ir aizdomas, ka ne katras divas krāsas varētu kombinēt vienā dzejas tēlā.
 - Piemēri no *TNZV*: “Ko jūs vēl skumstat, Jūs jaunās dvēseles, Tik **zaļi sārtas**?”, “Kur iespieties nespēj **Sarkan-zeltainā**, karstā, Septiņkrāsainā pasaulē saule”, “**Zil-saulainiem** debešu palagiem”.

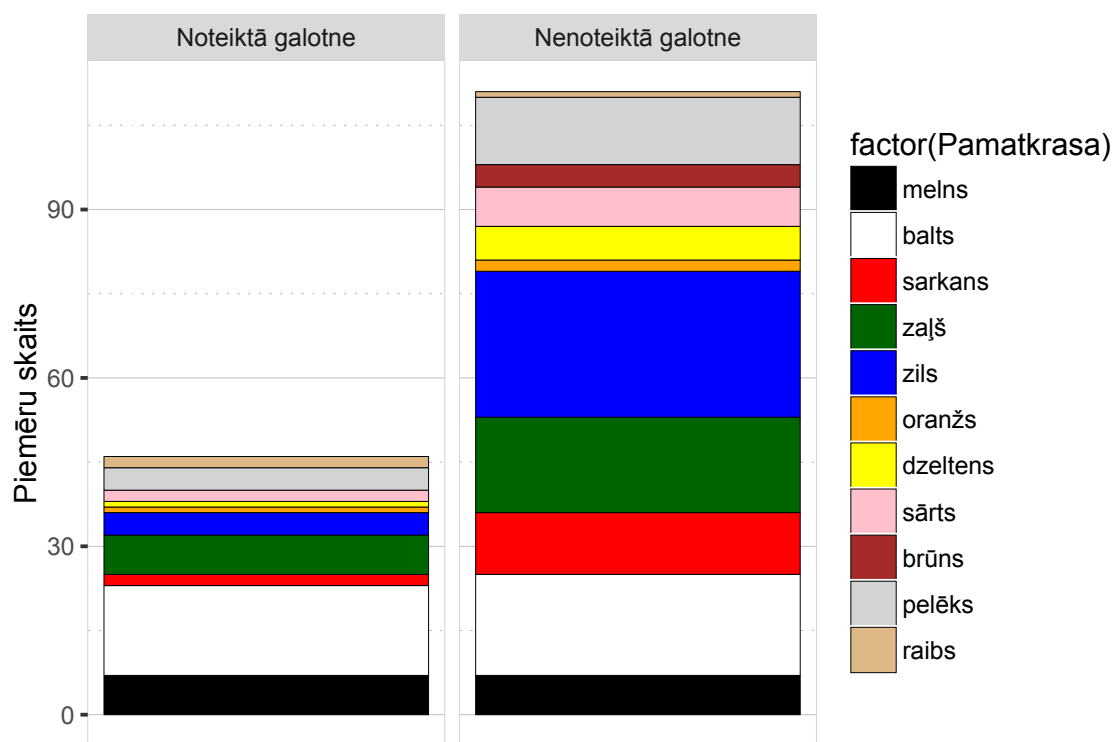


Attēls 5: Krāsu (arī neadjektīvu) sadalījums pēc autora

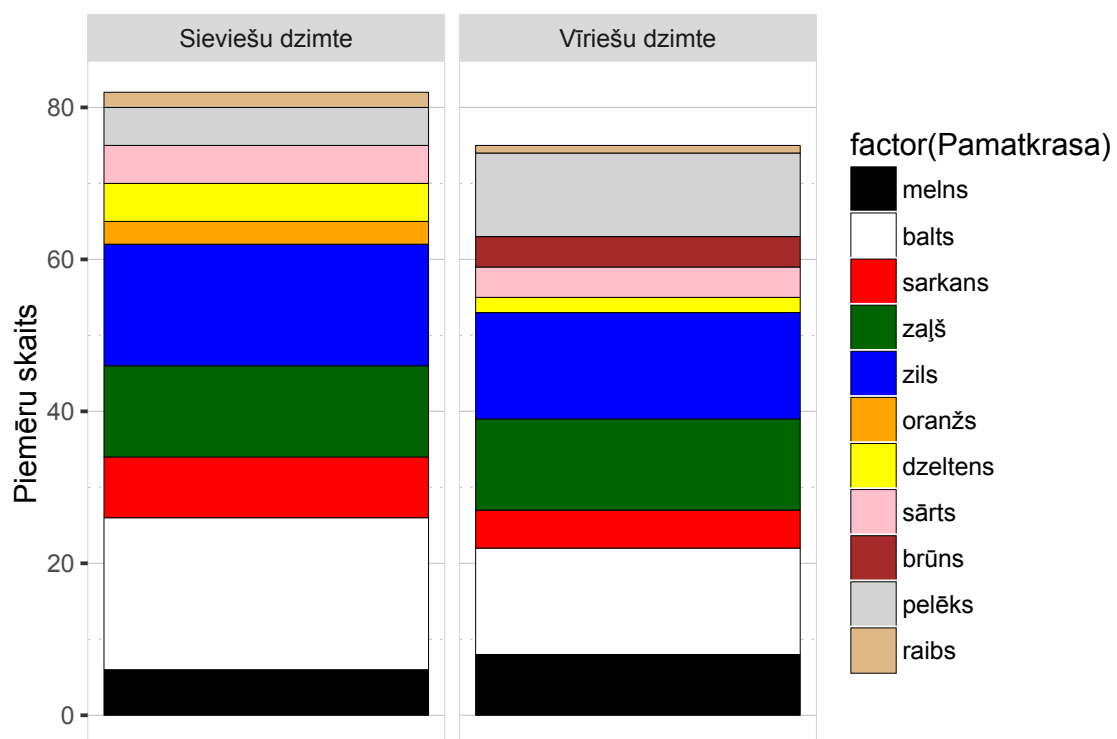
- Piemēri no *Dz-L* “**Baltā** bārda **rūsga** svila”, “Gan kaut kas **sarkans**, notraipīts ar **brūnu**, Gan **brūns** ar **sarkanu**”, “Ka rudenos tev ir tāds milzums **vara** un **zelta** naudas”.

Atsauces

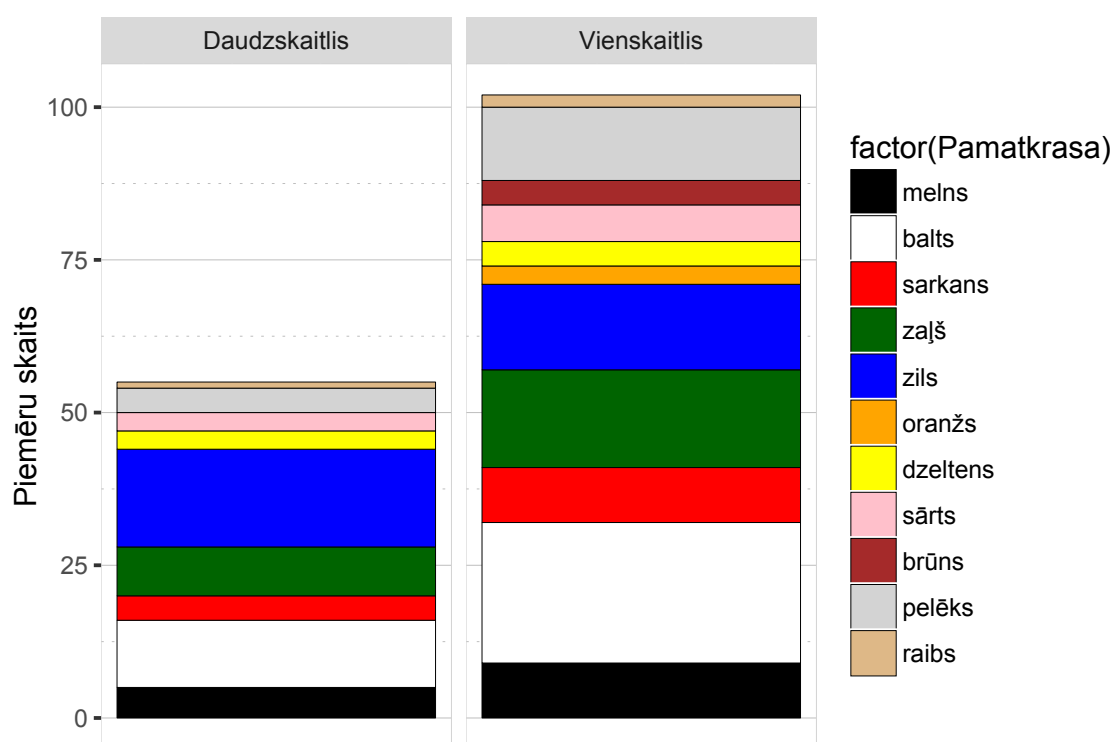
- [1] J.F.Burrows. *Computation into Criticism: A Study of Jane Austen's Novels*. Oxford University Press. 1987.
- [2] G.Chesterton. *The Man Who Was Thursday*. 1908. Lejuplādēts 2016-10-28 no <http://www.bartleby.com/158/4.html>.
- [3] The Technium. *Edge-notched cards*. Lejuplādēts 2016-10-27 no <http://kk.org/thetechnium/one-dead-media/>.
- [4] Ludis Bērziņš *Ievads latviešu tautas dzejā: Metrika un stilistika, Volume 1*. Latvijas Universitāte. 1940.
- [5] Lasar Liepins. *Markdown: Syntax*. Lejuplādēts 2016-10-27 no <http://markdown.de/>.
- [6] E.Riekstiņš, A.Andžāns. *Atrisini pats! Zvaigzne* 1984, 271 lpp.
- [7] Jeff Rydberg-Cox. *Statistical Methods for Studying Literature using R*. Lejuplādēts 2016-10-27 no <http://www.chlt.org/StatisticalMethods/>.
- [8] Ruta Veidemane. *Izteikt neizsakāmo*. Liesma. 1977.
- [9] Yuen Ren Chao. *Lauvas ēdošais dzejnieks*. Lejuplādēts 2016-10-27 no <https://www.youtube.com/watch?v=9jtiw721RAg>.



Attēls 6: Noteiktās un nenoteiktās galotnes krāsu adjektīviem (abi autori)



Attēls 7: Dzimte krāsu adjektīviem (abi autori)



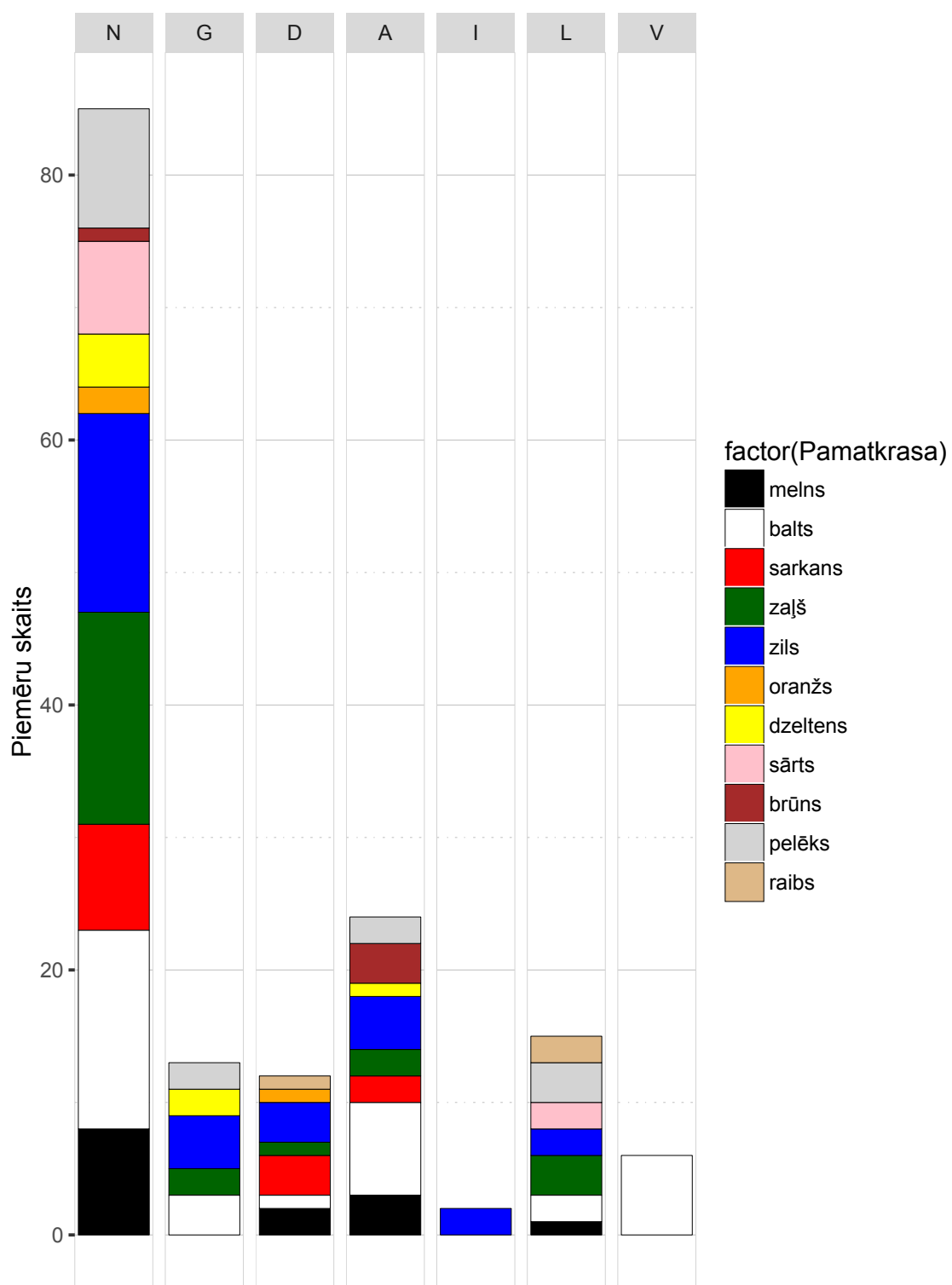
Attēls 8: Skaitlis krāsu adjektīviem (abi autori)

Vārds			Dz-L ‰	TNZV ‰	Vārds			Dz-L ‰	TNZV ‰
1	un	●	624 (47.22)	304 (40.96)	61	vairs	●	24 (1.82)	7 (0.94)
2	es	●	219 (16.57)	25 (3.37)	62	viņa	●	24 (1.82)	18 (2.43)
3	ir	●	218 (16.50)	41 (5.52)	63	esmu	●	23 (1.74)	1 (0.13)
4	kā	●	154 (11.65)	78 (10.51)	64	pats	●	23 (1.74)	14 (1.89)
5	no	●	151 (11.43)	59 (7.95)	65	reiz	●	23 (1.74)	2 (0.27)
6	uz	●	138 (10.44)	42 (5.66)	66	sirds	●	23 (1.74)	22 (2.96)
7	bet	●	134 (10.14)	55 (7.41)	67	labi	●	22 (1.66)	2 (0.27)
8	tu	●	124 (9.38)	45 (6.06)	68	tāpēc	●	22 (1.66)	0 (0.00)
9	ar	●	121 (9.16)	49 (6.60)	69	viņu	●	21 (1.59)	10 (1.35)
10	ka	●	103 (7.79)	15 (2.02)	70	būtu	●	20 (1.51)	5 (0.67)
11	man	●	99 (7.49)	36 (4.85)	71	tiem	●	20 (1.51)	9 (1.21)
12	par	●	98 (7.42)	27 (3.64)	72	vajag	●	20 (1.51)	0 (0.00)
13	kas	●	97 (7.34)	60 (8.08)	73	viņi	●	20 (1.51)	4 (0.54)
14	nav	●	90 (6.81)	0 (0.00)	74	atkal	●	19 (1.44)	13 (1.75)
15	ko	●	85 (6.43)	42 (5.66)	75	gan	●	19 (1.44)	6 (0.81)
16	ja	●	83 (6.28)	7 (0.94)	76	iet	●	19 (1.44)	12 (1.62)
17	tad	●	81 (6.13)	29 (3.91)	77	Līz	●	19 (1.44)	0 (0.00)
18	kad	●	80 (6.05)	22 (2.96)	78	nevar	●	19 (1.44)	4 (0.54)
19	lai	●	80 (6.05)	38 (5.12)	79	trako	●	19 (1.44)	0 (0.00)
20	pie	●	75 (5.67)	18 (2.43)	80	būt	●	18 (1.36)	5 (0.67)
21	to	●	72 (5.45)	37 (4.99)	81	it	●	18 (1.36)	4 (0.54)
22	mēs	●	70 (5.30)	17 (2.29)	82	kāds	●	18 (1.36)	2 (0.27)
23	ne	●	68 (5.15)	35 (4.72)	83	pār	●	18 (1.36)	14 (1.89)
24	tā	●	68 (5.15)	53 (7.14)	84	savu	●	18 (1.36)	11 (1.48)
25	vēl	●	67 (5.07)	50 (6.74)	85	zemes	●	18 (1.36)	19 (2.56)
26	jau	●	65 (4.92)	40 (5.39)	86	jūs	○	17 (1.29)	55 (7.41)
27	vai	●	65 (4.92)	25 (3.37)	87	kamēr	●	17 (1.29)	1 (0.13)
28	pa	●	62 (4.69)	21 (2.83)	88	mūs	●	17 (1.29)	6 (0.81)
29	tev	●	62 (4.69)	46 (6.20)	89	pat	●	17 (1.29)	5 (0.67)
30	viņš	●	60 (4.54)	48 (6.47)	90	viena	●	17 (1.29)	3 (0.40)
31	pēc	●	59 (4.46)	15 (2.02)	91	atceries	●	16 (1.21)	0 (0.00)
32	var	●	57 (4.31)	5 (0.67)	92	tie	○	16 (1.21)	22 (2.96)
33	tas	●	54 (4.09)	28 (3.77)	93	mana	●	15 (1.13)	5 (0.67)
34	tik	○	42 (3.18)	58 (7.81)	94	mūsu	●	15 (1.13)	1 (0.13)
35	mums	●	41 (3.10)	15 (2.02)	95	pārāk	●	15 (1.13)	3 (0.40)
36	bija	●	40 (3.03)	8 (1.08)	96	pāri	●	15 (1.13)	13 (1.75)
37	kur	●	40 (3.03)	28 (3.77)	97	pret	●	15 (1.13)	14 (1.89)
38	tikai	●	40 (3.03)	10 (1.35)	98	visiem	●	15 (1.13)	4 (0.54)
39	būs	●	39 (2.95)	15 (2.02)	99	krīt	●	14 (1.06)	3 (0.40)
40	daudz	●	39 (2.95)	14 (1.89)	100	šo	●	14 (1.06)	10 (1.35)
41	mani	●	39 (2.95)	3 (0.40)	101	tajā	●	14 (1.06)	0 (0.00)
42	arī	●	37 (2.80)	10 (1.35)	102	vienīgi	●	14 (1.06)	0 (0.00)
43	tur	●	37 (2.80)	33 (4.45)	103	viens	●	14 (1.06)	9 (1.21)
44	jo	●	34 (2.57)	18 (2.43)	104	vienu	●	14 (1.06)	1 (0.13)
45	tās	●	33 (2.50)	29 (3.91)	105	visas	●	14 (1.06)	9 (1.21)
46	atbalss	●	32 (2.42)	0 (0.00)	106	aiz	●	13 (0.98)	11 (1.48)
47	cik	●	31 (2.35)	15 (2.02)	107	kuru	●	13 (0.98)	1 (0.13)
48	nu	●	31 (2.35)	15 (2.02)	108	nezinu	●	13 (0.98)	0 (0.00)
49	līdz	●	30 (2.27)	26 (3.50)	109	rokas	●	13 (0.98)	8 (1.08)
50	tam	●	30 (2.27)	8 (1.08)	110	sev	●	13 (0.98)	14 (1.89)
51	kaut	●	28 (2.12)	11 (1.48)	111	vien	●	13 (0.98)	7 (0.94)
52	tevi	●	27 (2.04)	18 (2.43)	112	viņiem	●	13 (0.98)	3 (0.40)
53	kāpēc	●	25 (1.89)	0 (0.00)	113	visi	●	13 (0.98)	7 (0.94)
54	viss	●	25 (1.89)	24 (3.23)	114	zeme	●	13 (0.98)	8 (1.08)
55	visu	●	25 (1.89)	18 (2.43)	115	zini	●	13 (0.98)	5 (0.67)
56	bez	○	24 (1.82)	31 (4.18)	116	nāk	●	12 (0.91)	3 (0.40)
57	esi	●	24 (1.82)	5 (0.67)	117	paliek	●	12 (0.91)	1 (0.13)
58	manu	●	24 (1.82)	1 (0.13)	118	varbūt	●	12 (0.91)	1 (0.13)
59	tagad	●	24 (1.82)	2 (0.27)	119	zinu	●	12 (0.91)	0 (0.00)
60	te	●	24 (1.82)	0 (0.00)	120	esam	●	11 (0.83)	1 (0.13)

Tabula 3: Dz-L 120 biežākās vārdformas

	Vārds		TNZV ‰	Dz-L ‰		Vārds		TNZV ‰	Dz-L ‰
1	un	●	304 (40.96)	624 (47.22)	61	pats	●	14 (1.89)	23 (1.74)
2	kā	●	78 (10.51)	154 (11.65)	62	pret	●	14 (1.89)	15 (1.13)
3	kas	●	60 (8.08)	97 (7.34)	63	saules	○	14 (1.89)	2 (0.15)
4	no	●	59 (7.95)	151 (11.43)	64	sev	●	14 (1.89)	13 (0.98)
5	tik	○	58 (7.81)	42 (3.18)	65	atkal	●	13 (1.75)	19 (1.44)
6	bet	●	55 (7.41)	134 (10.14)	66	dvēsele	○	13 (1.75)	0 (0.00)
7	jūs	○	55 (7.41)	17 (1.29)	67	pāri	●	13 (1.75)	15 (1.13)
8	tā	●	53 (7.14)	68 (5.15)	68	dienas	○	12 (1.62)	2 (0.15)
9	vēl	●	50 (6.74)	67 (5.07)	69	dzīve	○	12 (1.62)	10 (0.76)
10	ar	●	49 (6.60)	121 (9.16)	70	gala	○	12 (1.62)	1 (0.08)
11	viņš	●	48 (6.47)	60 (4.54)	71	iet	●	12 (1.62)	19 (1.44)
12	tev	●	46 (6.20)	62 (4.69)	72	plūst	○	12 (1.62)	0 (0.00)
13	tu	●	45 (6.06)	124 (9.38)	73	rīta	○	12 (1.62)	4 (0.30)
14	ko	●	42 (5.66)	85 (6.43)	74	acīm	○	11 (1.48)	8 (0.61)
15	uz	●	42 (5.66)	138 (10.44)	75	aiz	●	11 (1.48)	13 (0.98)
16	ir	●	41 (5.52)	218 (16.50)	76	jā	○	11 (1.48)	7 (0.53)
17	jau	●	40 (5.39)	65 (4.92)	77	kaut	●	11 (1.48)	28 (2.12)
18	lai	●	38 (5.12)	80 (6.05)	78	lēni	○	11 (1.48)	0 (0.00)
19	to	●	37 (4.99)	72 (5.45)	79	saule	○	11 (1.48)	4 (0.30)
20	man	●	36 (4.85)	99 (7.49)	80	savu	●	11 (1.48)	18 (1.36)
21	nau	○	36 (4.85)	0 (0.00)	81	tai	●	11 (1.48)	10 (0.76)
22	ne	●	35 (4.72)	68 (5.15)	82	tomēr	○	11 (1.48)	6 (0.45)
23	tur	●	33 (4.45)	37 (2.80)	83	aizvien	○	10 (1.35)	0 (0.00)
24	bez	○	31 (4.18)	24 (1.82)	84	ak	○	10 (1.35)	8 (0.61)
25	tad	●	29 (3.91)	81 (6.13)	85	arī	●	10 (1.35)	37 (2.80)
26	tās	●	29 (3.91)	33 (2.50)	86	debess	○	10 (1.35)	2 (0.15)
27	kur	●	28 (3.77)	40 (3.03)	87	domas	○	10 (1.35)	1 (0.08)
28	tas	●	28 (3.77)	54 (4.09)	88	dzīves	●	10 (1.35)	9 (0.68)
29	par	●	27 (3.64)	98 (7.42)	89	līdzī	○	10 (1.35)	3 (0.23)
30	līdz	●	26 (3.50)	30 (2.27)	90	šo	●	10 (1.35)	14 (1.06)
31	es	●	25 (3.37)	219 (16.57)	91	tikai	●	10 (1.35)	40 (3.03)
32	iz	○	25 (3.37)	0 (0.00)	92	viņu	●	10 (1.35)	21 (1.59)
33	vai	●	25 (3.37)	65 (4.92)	93	ap	●	9 (1.21)	9 (0.68)
34	viss	●	24 (3.23)	25 (1.89)	94	asins	○	9 (1.21)	4 (0.30)
35	kad	●	22 (2.96)	80 (6.05)	95	caur	○	9 (1.21)	6 (0.45)
36	sirds	●	22 (2.96)	23 (1.74)	96	dvēsele	○	9 (1.21)	1 (0.08)
37	tie	○	22 (2.96)	16 (1.21)	97	kurš	○	9 (1.21)	8 (0.61)
38	pa	●	21 (2.83)	62 (4.69)	98	miesas	○	9 (1.21)	3 (0.23)
39	zemes	●	19 (2.56)	18 (1.36)	99	nakts	○	9 (1.21)	3 (0.23)
40	drīz	○	18 (2.43)	7 (0.53)	100	nedz	○	9 (1.21)	0 (0.00)
41	jo	●	18 (2.43)	34 (2.57)	101	priekšā	○	9 (1.21)	7 (0.53)
42	jums	○	18 (2.43)	9 (0.68)	102	sauli	○	9 (1.21)	1 (0.08)
43	pie	●	18 (2.43)	75 (5.67)	103	tiem	●	9 (1.21)	20 (1.51)
44	tevi	●	18 (2.43)	27 (2.04)	104	viens	●	9 (1.21)	14 (1.06)
45	viņa	●	18 (2.43)	24 (1.82)	105	visas	●	9 (1.21)	14 (1.06)
46	visu	○	18 (2.43)	25 (1.89)	106	apkārt	○	8 (1.08)	7 (0.53)
47	bij	○	17 (2.29)	8 (0.61)	107	bija	●	8 (1.08)	40 (3.03)
48	mēs	●	17 (2.29)	70 (5.30)	108	galvu	○	8 (1.08)	5 (0.38)
49	zemi	○	17 (2.29)	11 (0.83)	109	guļ	○	8 (1.08)	7 (0.53)
50	acis	○	16 (2.16)	8 (0.61)	110	ļaudis	○	8 (1.08)	0 (0.00)
51	viņas	○	16 (2.16)	9 (0.68)	111	mūžam	○	8 (1.08)	2 (0.15)
52	būs	●	15 (2.02)	39 (2.95)	112	nekas	○	8 (1.08)	4 (0.30)
53	cik	●	15 (2.02)	31 (2.35)	113	prāts	○	8 (1.08)	0 (0.00)
54	ik	○	15 (2.02)	1 (0.08)	114	rokas	●	8 (1.08)	13 (0.98)
55	ka	●	15 (2.02)	103 (7.79)	115	sirdī	○	8 (1.08)	6 (0.45)
56	mums	●	15 (2.02)	41 (3.10)	116	spīd	○	8 (1.08)	1 (0.08)
57	nu	●	15 (2.02)	31 (2.35)	117	tāles	○	8 (1.08)	1 (0.08)
58	pēc	●	15 (2.02)	59 (4.46)	118	tām	○	8 (1.08)	7 (0.53)
59	daudz	●	14 (1.89)	39 (2.95)	119	tam	●	8 (1.08)	30 (2.27)
60	pār	●	14 (1.89)	18 (1.36)	120	visa	●	8 (1.08)	9 (0.68)

Tabula 4: TNZV 120 biežākās vārdformas



Attēls 9: Locījums krāsu adjektīviem (abi autori)