# Meklēšana korpusos un datu kārtošana (noSketchEngine platformā)

Ilze Auziņa, Baiba Saulīte

2022. gada 12. janvārī

# Kopsavilkums

- Korpusi noSketchEngine platformā
- Biežuma saraksti
- Konkordances
  - Vienkāršā meklēšana
  - Meklēšana morfoloģiski marķētā korpusā
- Darbības ar konkordanču rindiņām

#### Kā nokļūt noSketchEngine platformā?

1. var. http://nosketch.korpuss.lv/

#### 2. var. <u>http://korpuss.lv</u>

Latviešu valodas teksta un runas korpusi Meklēšana Morfoloģisko pazīmju kopa Par korpusiem

Filtrēt (teksta (16) (morfoloģija (13) reprezentatīvs (8) vispārīgs (8) specializēts (6) (manuāli pārbaudīts (4) (sintakse (3) runas (3) (apguvēju (3) (kjūdas (2) (semantika (1) (tīmekja (1) (autora (1) (paralēis (1) (paraleis (1) (paraleis

#### Populārākie korpusi

LVK2018 Līdzsvarotais mūsdienu latviešu valodas tekstu korpuss 2016–2018, 10 milj. vārdlietojumu (12 milj. tekstvienību)

MuLa Mūsdienu latgaliešu tekstu korpuss 2011–2013, 1 milj. vārdlietojumu (1,3 milj. tekstvienību)

🗮 Vairāk informācijas 🛛 🔾 Meklēt korpusā

#### Korpusa izvēle

- http://nosketch.korpuss.lv/
- Šobrīd pieejami 14 korpusi (divi korpusa *Senie* varianti un divas korpusa *LiLa* daļas – latviešulietuviešu un lietuviešu-latviešu)

FREE CO	RPORA		
Latvian	Saeima	19,174,103 words	OPEN
Latvian	LVK2013	4,300,348 words	OPEN
Latvian	LVK2018	9,813,011 words	OPEN
Latvian	Rainis	1,600,249 words	OPEN
Latgalian	MuLa	1,018,962 words	OPEN
Latvian	Tīmeklis2007	99,419,089 words	OPEN
Latvian	Emuāri	6,689,145 words	OPEN
Latvian	Pārspriedumi	185,461 words	OPEN
Latvian	LaVA	192,714 words	OPEN
Latvian	Subtitri	5,505,743 words	OPEN
Latvian	LiLa-LV	4,306,984 words	OPEN
Lithuanian	LiLa-LT	4,142,977 words	OPEN
Latvian	SENIE sākotnējais	1,920,285 words	OPEN
Latvian	SENIE Unicode	2,058,974 words	OPEN
Latvian	UD_Latvian-LVTB	215,547 words	OPEN
Latvian	LRK100 (alpha)	889,343 words	OPEN



# Biežuma saraksti



<b>5</b> 5	WORD	LIST	LVK2018	Q i		
	BASIC	ADVANO	ED ABO	UT		
						Ð
	find ?	empty		all		
				starting with		
$\odot$				ending with		
00				containing		
•=				GO		
•=						
<b>=•</b> =						

Ì≡

<b>5</b> 9	WORDLIST	-	LVK2018	৫ (	i					(	?		° O
	Vārdforma	(109,	121 items I 9,378,8	892 total frequency)						হ	Ŧ	0	<b>(</b> )
			Vārdforma	↓ Frequency ?			Vārdforma	↓ Frequency ?					
		1	un	326,361	•••	11	vai	54,246	•••				
: :		2	ir	153,416	•••	12	bet	51,723	•••				
$\odot$		3	ar	122,850	•••	13	bija	45,915	•••				
0 0		4	par	104,286	•••	14	to	43,784	•••				
•=		5	ka	94,850	•••	15	tā	38,026	•••				
•		6	no	87,430	•••	16	nav	36,157	•••				
•=		7	arī	76,675	•••	17	es	35,779	•••				
		8	kā	73,130	•••	18	lai	35,139	•••				
Ì≣		9	kas	64,988	•••	19	pēc	32,548	•••				
NE		10	uz	62,607	•••	20	tas	32,123	•••				
£Ξ	First 20,000 rows are o	displ	ayed. You can	download whole w	wordlist	<b>±</b>							
المر				Rows per pag	e: 2	0 .	✓ 1–20 of 2	20,000 K	<	1		>	>I

Ģ	WORD	LIST [	LVK2018	৫ ি					0
	BASIC	ADVANCE	D ABOU	т					
88	find <sup>?</sup>	empty		all		(	valsts	Ð	
				starting with					
$\odot$				ending with		Vārdforma	↓ Frequency <sup>?</sup>	Vārdforma	↓ Frequency ?
õ				containing		valsts	22,963	12 valstspiederīgie	9
$\odot$				_		2 valstsvienības	190 …	13 valstsvienībām	7 ***
•=				GO		3 valstsvienību	77 …	14 valstsnācijas	7 …
					_	4 valstsvienība	63 …	15 valstspiederīgā	6 …
<b>-</b>						5 valstsvienībā	51 …	16 valstsvīriem	6 …
=•=						6 valstsvīru	28 …	17 valstsvīra	5 …
<b>=•</b> =						7 valstsvienībai	26 …	18 valstsvīrs	5 …
						8 valstspiederīge	0 24 …	19 valstspiederīgajar	n 5
+=						9 valstspiederīga	ajiem 17 …	20 valstspiederības	5 …
						10 valstsvīri	15 …	21 valststiesību	5 ***
						11 valstspiederīga	ais 11 •••	22 valstsnācija	5

# Vārdu un vārdformu biežuma saraksts (1/9)

Vārdu biežuma sarakstā ir vārdi (vai vārdformas), kuri kaut reizi ir minēti korpusā, dots to lietojuma biežums.

Vārdus sarakstā var grupēt:

- pēc alfabēta var viegli atrast konkrētus vārdus;
- pēc biežuma parāda korpusa vispārējās iezīmes.

Biežuma saraksts ļauj atlasīt biežāk lietotos vārdus turpmākiem pētījumiem.

## Vārdu un vārdformu biežuma saraksti (2/9)

**Absolūtais biežums** – parāda kāda valodas elementa skaitu izlasē (konkrēti – korpusā).

**Relatīvais biežums** – absolūtā biežuma attiecība pret visu datu skaitu; relatīvā biežuma rādītājus iespējams salīdzināt ar citiem analoģiskiem datiem.

# Vārdu un vārdformu biežuma saraksti (3/9)

#### Skatījuma iespējas

- Automātiski tiek parādīts absolūtais biežums
- Relatīvo biežumu var iegūt, izvēloties *Show frequency per million*
- Datus var izkārtot vienā kolonnā (*Single column*)
- Iespējams aplūkot, cik dokumentos parādās biežāk lietotie vārdi vai vārdformas (*Document frequency*)

• VIEW OPTIONS				
<ul> <li>Single column ?</li> <li>Show line numbers ?</li> </ul>	$\checkmark$	Show absolute frequency ? Show frequency per million ?	• 0 0	Frequency ? Document frequency ? ARF ?
	$\checkmark$	Show bars ?		

#### Vārdu un vārdformu biežuma saraksti (4/9)

**Uzdevums.** Noskaidrot, kuras ir biežāk lietotās vārdformas korpusā "Saeima", "Rainis", "LaVA", "Subtitri".

Saeima: *ir, un, par, ka, es, to, ar, arī, mēs, priekšlikums* Rainis: *un, es, ir, no, par, bet, ar, kā, arī, ka* LaVA: *es, ir, un, man, mana, patīk, ar, mans sauc, esmu* Subtitri: *ir, un, ka, par, ar, es, tas, to, no, tā* 

#### Vārdu un vārdformu biežuma saraksti (5/9)

**Uzdevums.** Noskaidrot, kuri LVK2018 ir biežāk lietotie vārdi, kas beidzas ar *šanās*.

≡	WORI	OLIST [	LVK2018	Q ()	
	BAS	SIC		ADVANCED	
find <sup>?</sup>	•	empty		all	šanās
				starting with	
				ending with	
				containing	
			GC		

	Vārdforma	$\checkmark$	Frequency ?	
1	tikšanās		1,024	•••
2	vēlēšanās		916	•••
3	vienošanās		747	•••
4	atrašanās		501	•••
5	stāšanās		419	•••
6	šķiršanās		337	•••
7	uzstāšanās		280	•••
8	atgriešanās		275	•••
9	veidošanās		267	•••
10	uzturēšanās		228	

#### Vārdu un vārdformu biežuma saraksti (6/9)

Kā iegūt korpusa vārdu pamatformu biežuma sarakstu?

- Padziļināta meklēšana (Advanced)
- Var norādīt minimālo pieminējumu skaitu

WORDLIS	WORDLIST LVK2018 (beta)						
BASIC ADVANCED							
			-11				
find ?	Värdform	a	all				
	Morfoloģi	skais tag	starting with				
	Lemma		ending with				
			containing				
			matching regex				
			from this list:				
Exclude the	se words:						
Include non	words ?						
√ A = a ?							
Frequency min ? 5	Frequence 0	cy max ?	0				
result format <ul> <li>Simple list <sup>2</sup></li> <li>Display as <sup>2</sup></li> </ul>							

WC Lei	VORDLIST LVK2018 (beta) (i Lemma (48,804 items 1 9,635,656 total frequency)									
	Lemma	↓ Frequency ?				Lemma	↓ Frequency ?			
1	un	326,361			14	es	61,410	•••		
2	büt	270,827			15	tā	59,294			
з	ar	122,850			16	vai	54,246			
4	tas	109,387			17	nebūt	52,083	•••		
5	par	104,286			18	bet	51,723			
6	ka	94,850			19	latvija	43,174	•••		
7	kas	92,146			20	tikt	43,001			
8	no	87,430			21	varēt	42,384	•••		
9	arī	76,675			22	šis	35,625	•••		
10	kā	72,922			23	lai	35,139			
11	viņš	65,322			24	laiks	33,813	•••		
12	uz	62,607			25	valsts	33,647	•••		
13	gads	62,390			26	kurš	33,632			

# Vārdu un vārdformu biežuma saraksti (7/9)

Kā iegūt kādas **vārdšķiras vārdu sarakstu**? [vārdformas identifikators] + [.\*], piemēram, **n.**\*

- Skatījuma iespējas:
  - Simple list morfoloģiskie tagi
  - Display as –
     vārdformas, morfoloģiskie tagi, lemma



### Vārdu un vārdformu biežuma saraksti (8/9)

WORDLIST LVK2018 (beta) 🤇 🖟			Vaicāju	Vaicājums: <b>n.*</b>				0 🖪 😤
Morfoloģi	iskais tags (29,682 items )							୧ 🛓 🗿 🛈
Lemma	Frequency	Lemma	Frequency	Lemma	Frequency		Lemma	Frequency
1 gads	62,390	14 pasaule	11,043	27 pilsēta	8,922		40 attīstība	7,750
2 Latvija	43,174	15 māja	10,875	28 darbība	8,771		41 vārds	7,730
3 laiks	33,813	16 informācija	10,825	29 sabiedrība	8,556		42 sistēma	7,713
4 valsts	33,646	17 persona	10,713	30 pants	8,433		43 rezultāts	7,663
5 darbs	26,450	18 Eiropa	10,671	31 pasākums	8,308		44 tiesa	7,530
6 cilvēks	22,886	19 dzīve	10,011	32 reize	8,251		45 lēmums	7,443
7 vieta	18,274	20 pašvaldība	9,948	33 ministrs	8,214		46 veids	7,413
8 bērns	18,225	21 uzņēmums	9,915	34 piemērs	8,107		47 punkts	7,401
9 Rīga	17,719	22 grupa	9,880	35 noteikums	8,041		48 attiecība	7,375
10 diena	15,536	23 sieviete	9,667	36 izglītība	8,027		49 iela	7,308
11 likums	12,272	24 jautājums	9,027	37 projekts	7,963		50 iestāde	7,300
12 daļa	12,083	25 spēks	9,021	38 ģimene	7,831			
13 iespēja	11,550	26 gadījums	8,972	39 eiro	7,816			

Rows per page: <u>50</u> ▼ 1–50 of 50

17

## Vārdu un vārdformu biežuma saraksti (9/9)

**Uzdevums.** Noskaidrot, korpusā LVK2018 biežāk lietotos darbības vārdus.

I	Lemma	Frequency	Lemma Fi	requency
	1 būt	247,908	11 teikt	7,983 ••••
Vaicājums: <b>v.*</b>	2 varēt	38,843 •••	12 norādīt	7,767 ••••
, ,	3 tikt	37,163	13 kļūt	7,388 ••••
	4 notikt	11,767 ••••	14 saņemt	7,222
	5 zināt	10,745 ••••	15 iespēt	7,115 ••••
	6 noteikt	9,478 ••••	16 sacīt	6,949 ••••
	7 veikt	9,410 ••••	17 paredzēt	6,785 ••••
	8 izmantot	8,358 ••••	18 gribēt	6,218 ••••
	9 sākt	8,321 ••••	19 radīt	6,089 ••••
	10 redzēt	8,054 ••••	20 domāt	6,061 ••••

# Konkordances

Examples of use in context

#### Konkordance ir

- korpusā atrasto vārdu vai vārdu savienojumu lietojumu saraksts kontekstā;
- korpusa analīzes rezultāts, kas parāda vaicājumā norādītā vārda vai vārdu savienojuma lietojumu korpusa tekstos.

*Keyword in Context* (KWIC) – vaicājumā ierakstītais vārds vai vārdu savienojums un tā apkaime.

, uz kurieni tie cilvēki , tie strīdnieki nākotnē ies . Uz tām pašām tieslietu ministra lobētajām šķīrējtiesām termināļu ... Labi , viņi paši arī ražo to gāzi , bet tas iet uz leju - sāk izsīkt tie avoti . Bet Anglija ir savienota to , ka sadursmju laikā Kijevā ir cietuši un bojā gājuši cilvēki , kā arī pausta līdzjūtība cietušo un bojāgājušo stāstīt , ka viņš tur ir aizņemts , ne tiesnesis varēs iet pēkšņi atvaļinājumā . Lūk , būs trīs mēnešu grafiks beidzot tie jautājumi ir jāatrisina , ja mēs gribam iet uz priekšu . Paldies . ( Aplausi . ) Cienījamais ldīgu lēmumu pieņemšanu ?! Es domāju , ka mums ir jāiet uz saprātīgiem lēmumiem , lai mēs varētu piesaistīt lītājam vajadzētu kaut vai pusminūti apsvērt iespēju iet strādāt amatā , kurā katru dienu ir jāuzņemas milzīga

#### Vienkāršā meklēšana

Pēc vārdformas atrod konkrēto vārdformu:  $ziema \rightarrow ziema; eju \rightarrow eju$ 

Pēc pamatformas atrod visas vārdformas:  $iet \rightarrow eju, iet, ejam, gāja$ 

Pēc vairāku vārdu pamatformām atrod piemērus ar šiem vārdiem dažādos locījumos: *lasīt grāmata → lasa/lasīdama grāmatu, lasāma grāmata*..



Pamatforma (*lemma*):  $garšot \rightarrow garšo, garšoja, garšos ..$ .\* $garšot \rightarrow garšo, negaršo..$  $(at | iz)braukt \rightarrow izbraucu, atbraucu..$ 



Frāze (phrase): man garšo; man .\*garšo  $\rightarrow$  .. velns viņu zina  $\rightarrow$  .. stāstīt pasakas  $\rightarrow$  ..

ADVANCED	ABOUT
phrase	
abc	
V A=a	
_	
	$\frac{\text{ADVANCED}}{\text{phrase}}$

Vārdforma (*word*):

 $soda \rightarrow soda$  (pamatf. sods), soda (pamaf. sodīt), soda (pamatf. soda)  $m\bar{a}su \rightarrow m\bar{a}su$  (vsk. akuz.),  $m\bar{a}su$  (dsk. ģen.)  $iet \rightarrow iet$  (nenoteiksme), iet (tag. 3.pers.)

Tiek meklēta tikai ierakstītā vārdforma.



Rakstzīme (character): skol → augstskolā, internātskolā, pamatskola, skola, skolā, skolēni, skolotāja ..

BASIC	ADVANCED
Query type ⑦	character
simple	skol
lemma	
phrase	
word	
character	
CQL	

#### Meklēšana pēc morfoloģiskajām pazīmēm (1/4)

**CQL** (*Corpus Query Language*)

- Izmanto sarežģītiem vaicājumiem
- Iespējams meklēt, izmantojot pamatformu (*lemma*) un/vai morfoloģisko pazīmi (*tag*), vārdformu (*word*)
- Konstrukciju meklēšana, vaicājumā nenorādot konkrētu vārdu

CQL: [tag="a.\*"][lemma="valoda"] → *jauka, jauna, grūta, dzimtā, nākamā, lieliska .. valoda* 



#### Meklēšana pēc morfoloģiskajām pazīmēm (2/4)

- Var meklēt vārdformas pēc konkrētām morfoloģiskajām pazīmēm (tag) CQL: [tag= "\_ "]
- Vaicājuma veidošanai nepieciešama latviešu valodas morfoloģisko pazīmju kopa, skat. <u>http://www.korpuss.lv/LV\_TagSet\_v.2.2.1\_22092021.pdf</u>
- Piemēram:

Var meklēt darbības vārdus vajadzības izteiksmē CQL: [tag="v..d.\*"]

Var meklēt 4. deklinācijas lietvārdus CQL: [tag="n....4"]

#### Meklēšana pēc morfoloģiskajām pazīmēm (3/4)

- Var meklēt konstrukcijas, norādot konkrētas **morfoloģiskās pazīmes** (*tag*) un konkrēta vārda **pamatformu** (*lemma*)
- Piemēram, var meklēt jebkuru darbības vārdu savienojumā ar lietvārdu *valoda*

CQL: [tag="v.\*"][lemma="valoda"]

#### Meklēšana pēc morfoloģiskajām pazīmēm (4/4)

- Var meklēt konkrēta vārda vārdformas
- Vaicājumā jāizmanto gan morfoloģiskās pazīmes (*tag*), gan pamatforma (*lemma*), ko apvieno ar & CQL: [tag=" " & lemma=" "]
- Piemēram, var atrast lietvārdu *valsts* visās vienskaitļa formās CQL: [tag="n..s.\*" & lemma="valsts"]

# Datu kārtošana un skatījums



- 1. Meklēšanas kritēriji
- 2. Mainīt kritērijus
- 3. Lejuplādēt rezultātu
- 4. Atgriezties
- 5. Mainīt skata iespējas
- 6. Nejauši atlasīt rindas
- 7. Mainīt rindu secību
- 8. Kārtot rindas
- 9. Pievienot kritēriju 10. GDEX

11. Aprēķināt biežumu 12. Aprēķināt kolokācijas 13. Meklētās vienības izplatība korpusā 14. Mainīt skatu 15. Kārtot (konteksts pa kreisi) 16. Kārtot (meklētais vārds) 17. Kārtot (konteksts pa labi) 18. Metadatu skats 19. Paplašināt kontekstu

### Meklēšanas kritēriji un biežums

CONCORDANCE	LaVA	Q (j
simple iet 1,048 (4,347.32 per million)		

- Vaicājuma tips
- KWIC (*Key Word in Context*) jeb meklētais vārds vai vārdu savienojums
- Atrasto piemēru (rindu skaits), resp., absolūtais biežums
- Relatīvais biežums
- Korpuss, kurā tiek meklēts

#### Rezultātu lejupielāde un saglabāšana



Saglabājot .txt, rezultātā parādīts: 1) no kura korpusa ņemts, 2) cik rezultātu, 3) kāds bijis vaicājums, 4) arī pamatforma un tags.

Lejuplādēt visus piemērus (vai pirmās rindas) kādā no piedāvātajiem formātiem.

#### corpus: lava subcorpus: size: 1048 query: Query:[lc="iet" | lemma\_lc="iet"] doc#3| , bet | , kopš es atbraucu uz Rīgā | Rīgu | , man | es nav speleju | nespēlēju futbolu . Es eju/vmnipi11san uz sporta zāli , | " peoples | Peoples fitness " . Es parasti eju uz sporta zāli 2 vai 3 reizes nedēla | nedēlā . doc#3| es nav speleju | nespēlēju futbolu . Es eju uz sporta zāli , | " peoples | Peoples fitness ". Es parasti eju/vmnipillsan uz sporta zāli 2 vai 3 reizes nedēla | nedēlā . Veselīgs dzīvesveids Mans vards | vārds ir Donata un uzvards | uzvārds ir Hoolz doc#4| es ēdu ar draugiem kafejnīcā | , un vakarinas es gatavoju mājā | mājās . Man patik | patīk gulēt ilgi , jo es eju/vmnipi11san gulēt diezgan vēlu . Es zinu , ka iet gulēt agri ir veselīgāks | veselīgāk , bet es vienmēr eju gulēt vēlu . Lai gan es daudz doc#4| | , un vakariņas es gatavoju mājā | mājās . Man patik | patīk gulēt ilgi , jo es eju gulēt diezgan vēlu . Es zinu , ka iet/vmnn0i1000n gulēt agri ir veselīgāks veselīgāk , bet es vienmēr eju gulēt vēlu . Lai gan es daudz mācos , seviški | seviški | , doc#4| patik | patīk gulēt ilgi , jo es eju gulēt diezgan vēlu . Es zinu , ka iet gulēt

#### Skata iespējas (1/2)

#### Var izvēlēties, lai rāda papildu pazīmes (katrai tekstvienībai vai meklētajai): pamatforma, tags, rādīt visu laiku vai kā uznirstošo logu.

5

#### **•** VIEW OPTIONS

Show attributes ? (3)	Show structures ? $=$ (1)	Show full context ?		
type to search Q	Teikums Q	✓ Show line selection		
Vārdforma (1)	Dokuments	✓ Show line numbers <sup>?</sup>		
✓ Morfoloģiskais tags (2)	Rindkopa	✓ Use glue ?		
Jemma (3)	Teikums			
words (lowercase)		LINE DETAILS		
lemmas (lowercase)				
<ul> <li>For each token</li> <li>For KWIC only</li> </ul>				
Show as tooltips				

### Skata iespējas (2/2)



1	i	doc#93 > Valtrauta tad cēla galvu no savām smilšu	kūkām ncfpd4/kūka	un jautāja: — Ko tu gribi zināt vāciski?  📄
2	i	doc#93 ēšos. <s> Nekad! </s> <s> Cep smilšu</s>	<b>kūkas</b> ncfpn4/kūka	ar Valtrautu! <s> Kaujies ar puikām! </s>
3	i	doc#93   pārdot, tos paši nevarētu ēst. >Bet	<b>kūku</b> ncfpg4/kūka	atgriezumi, kas bija tikpat gardi un vēl gardā 下
4	i	doc#93 em savu māju, pusdienās viņiem allaž būtu	<b>kūku</b> ncfpg4/kūka	atgriezumi. <s> Valtrautai dotu — Valtra</s>
5	i	doc#93 Vanda griež. <s> Stasis un Vanda ēstu</s>	<b>kūku</b> ncfpg4/kūka	atgriezumus, cik lien. <s> Lecīgā Brigit</s>
6	i	doc#93 >> <s> Un Billei nebūtu draugu, ko cienāt ar</s>	<b>kūku</b> ncfpg4/kūka	atgriezumiem. <s> Vai viņa maz drīkstē</s>

#### Datu kārtošana: vienkāršā kārtošana







#### Kārtot rezultātu pēc labā vai kreisā konteksta (resp., blakus vārdiem) vai pēc meklētā vārda.

		De	etails	L	.eft context	KWIC ↓	Right context
1		G	) doc#	15825 R/	AINS MEŽĢĪŅU SPĻĀVIENSPALDIES JANVĀ	RI <mark>Ej</mark>	NU CITIEM SAVUS ROKDARBUS RĀDĪTMAN JAL 👕
2		G	) doc#	17673 as	s federāciju rīkoja pirmās sacensības cīņas vei	ā <mark>EJ</mark>	DAR. A. Kostirevs sacīja, ka viņš vēlējies, lai šāds e 下
3		[ (i	) doc#	17673,k	kuri arī senioru vecumā vēlas aktīvi darboties.F	ar <mark>EJ</mark>	DAR sākumu var uzskatīt sacensības pludmalē, sk; 下
4		i	) doc#	17673 <b>u</b>	ir jaunā sporta veida celmlauži. Gvido precizē,	a <mark>EJ</mark>	DAR ir balstīts uz trim pamatakmeņiem. Viens no tik 下
5		] (i	) doc#	17673 tie	ešu saliedēšanu un integrēšanu sabiedrībā. Vā	di <mark>EJ</mark>	DAR ir bijis seno latviešu cilšu karotāju uzmundrinoš 👕
			Details	6	Left context	KWIC	Right context ↓
	7341		i do	c#5885	vēl aizvien silda sirdi un liek rosīties, lai vērtīg	ais <b>neiet</b>	zudumā.Foto: Maksims Šaldajevs. Visu rakstu las 👕
	7341 7342		i do	oc#5885 oc#15010	vēl aizvien silda sirdi un liek rosīties, lai vērtīg 3I apmainīs pret eiro, bet pārējā nauda varētu	ais <mark>neiet</mark> pūt <b>gājus</b>	zudumā.Foto: Maksims Šaldajevs. Visu rakstu las 💿 i zudumā. " Līdz 15% varētu vispār neatgriezties ce 🕞
	7341 7342 7343		<ul> <li>i do</li> <li>i do</li> <li>i do</li> </ul>	0c#5885 0c#15010 0c#15373	vēl aizvien silda sirdi un liek rosīties, lai vērtīg šl apmainīs pret eiro, bet pārējā nauda varētu sna trešdaļa cūku kļūst par blakusproduktiem u	ais <b>neiet</b> pūt <b>gājus</b> n " <b>iet</b>	<ul> <li>zudumā.Foto: Maksims Šaldajevs. Visu rakstu las</li> <li>zudumā. " Līdz 15% varētu vispār neatgriezties ce</li> <li>zudumā". ASV fermās ik gadu 920 miljoni dzīvniel</li> </ul>
	7341 7342 7343 7344		<ul> <li>i do</li> <li>i do</li> <li>i do</li> <li>i do</li> <li>i do</li> <li>i do</li> </ul>	0c#5885 0c#15010 0c#15373 0c#15436	vēl aizvien silda sirdi un liek rosīties, lai vērtīg šl apmainīs pret eiro, bet pārējā nauda varētu ena trešdaļa cūku kļūst par blakusproduktiem u airāki vietējie veikalu tīkli tieši tāpēc ātri vien e	ais <b>neiet</b> pūt <b>gājus</b> n " <b>iet</b> sot <b>gājuš</b>	<ul> <li>zudumā.Foto: Maksims Šaldajevs. Visu rakstu las</li> <li>zudumā. " Līdz 15% varētu vispār neatgriezties ce</li> <li>zudumā". ASV fermās ik gadu 920 miljoni dzīvniel</li> <li>zudumā.Ja tu nemūrē — neko nevarēsi uzcelt. Tie</li> </ul>
	7341 7342 7343 7344 7345		<ul> <li>i do</li> </ul>	oc#5885 oc#15010 oc#15373 oc#15436 oc#15721	vēl aizvien silda sirdi un liek rosīties, lai vērtīg šl apmainīs pret eiro, bet pārējā nauda varētu na trešdaļa cūku kļūst par blakusproduktiem u airāki vietējie veikalu tīkli tieši tāpēc ātri vien e spēja, ka daļa ar šo problēmu saistīto dokume	ais <b>neiet</b> pūt <b>gājus</b> n " <b>iet</b> sot <b>gājuš</b>	<ul> <li>zudumā.Foto: Maksims Šaldajevs. Visu rakstu las</li> <li>zudumā. " Līdz 15% varētu vispār neatgriezties ce</li> <li>zudumā". ASV fermās ik gadu 920 miljoni dzīvniei</li> <li>zudumā.Ja tu nemūrē — neko nevarēsi uzcelt. Tie</li> <li>zudumā vai pat apzināti iznīcināti. Tomēr jāpiekrīt</li> </ul>
	7341 7342 7343 7344 7345 7346		<ul> <li>i do</li> </ul>	0c#5885 0c#15010 0c#15373 0c#15436 0c#15721 0c#15799	vēl aizvien silda sirdi un liek rosīties, lai vērtīg 51 apmainīs pret eiro, bet pārējā nauda varētu na trešdaļa cūku kļūst par blakusproduktiem u airāki vietējie veikalu tīkli tieši tāpēc ātri vien e spēja, ka daļa ar šo problēmu saistīto dokume s.Laiku pa laikam cilvēki bailīgi sprieda, ka arm	ais neiet pūt gājus n" iet sot gājuš utu gājuš	<ul> <li>zudumā.Foto: Maksims Šaldajevs. Visu rakstu las</li> <li>zudumā. " Līdz 15% varētu vispār neatgriezties ce</li> <li>zudumā". ASV fermās ik gadu 920 miljoni dzīvniel</li> <li>zudumā.Ja tu nemūrē — neko nevarēsi uzcelt. Tie</li> <li>zudumā vai pat apzināti iznīcināti. Tomēr jāpiekrīt</li> <li>zudumā vairāk vai mazāk pazīstamu ļaužu dēls, lī</li> </ul>

#### Biežums

२. ± ∽ ⊙ < ४ = = ॥ ∎ … ₪

Piedāvā apskatīt meklētā vārda (vai pirmā vārda pa kreisi / pa labi) biežumu no dažādiem skatpunktiem: vārdforma, pamatforma vai vārdšķira. Var apskatīt arī biežumu noteiktos dokumentos.

				Vārdforma	↓ Frequency		
BASIC ADVANCED AE		1	iet	2,174	•••		
First word to the left	KWIC	More presets	L	2	gāja	889	•••
WORD FORMS	WORD FORMS	TEXT TYPES	L	3	jāiet	412 💻	•••
PART-OF-SPEECH TAGS	PART-OF-SPEECH TAGS		L	4	ejot	294 -	•••
				5	gājis	244 -	•••
LEMMAS			L	6	gājuši	232 -	•••
				7	gāju	230 -	•••
			L	8	eju	212 -	•••
PART-OF-SPEECH TAGS				9	Ej	205 -	•••
LEMMAS 🛛				10	ej	182 -	•••

# Korpusu vietne http://korpuss.lv



Ieskaties: <u>http://ailab.lv</u> Seko mums: @AiLab.lv