

VarEs: eesti keele morfosüntaktiline varieerumine



Üles ja ülesse varieerumine kõnes

Eesti Rakenduslingvistika Ühingu aastakonverents
24.04.2025

Margetta Pajuste, Miia Malm, Brigitte Pärn, Taavi Terasmaa, Annika
Kängsepp, Maarja-Liisa Pilvik, Liina Lindström



Mida ja miks me uurime?

- *Üles ja ülesse* varieerumine kõnes
 - Missugune on varieerumise suhe?
 - Millest sõltub vormi valik?



Etümoloogia

Uurali algkeel – *wülä

üles:

- muistne **s**-lõpuline latiiv
 - soome – **ylös**
 - põhjasaami – **bâjas**

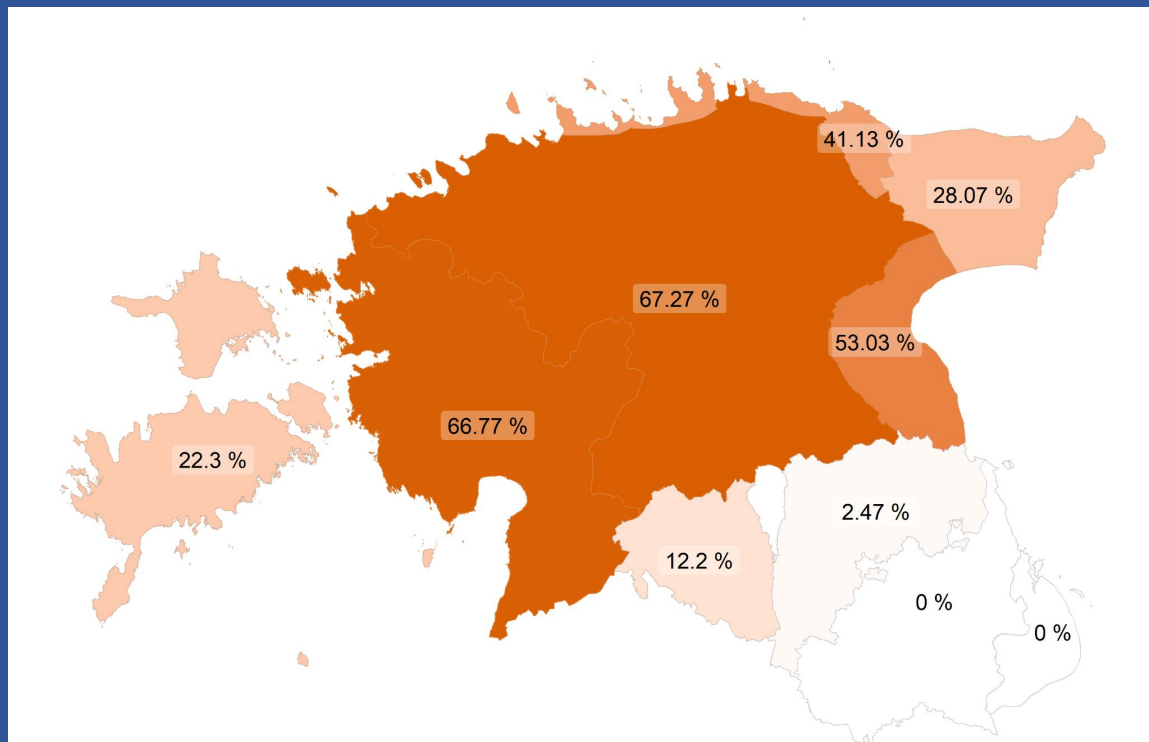
ülesse:

- uuem, murdeline
- tekkinud sisseütleva käände eeskujul



Varieerumine murretes

Ülesse osakaal
murretes



Andmed: eesti murrete korpus (Lindström jt 2022)

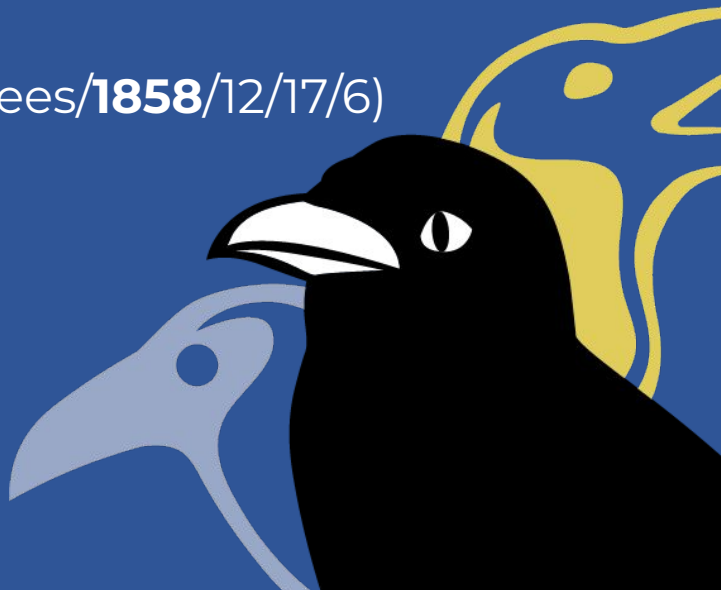
Varieerumine 19. saj kirjakeeles (DIGAR)

- (1) **Laudad olid nii ilusaste ja mönusaste kiwidest üles ehitud, nagu arwa mõisades olen näinud.**

(<https://dea.digar.ee/page/eestipollomees/1870/09/01/5>)

- (2) **Nüüd töusis üks preester ülesse.**

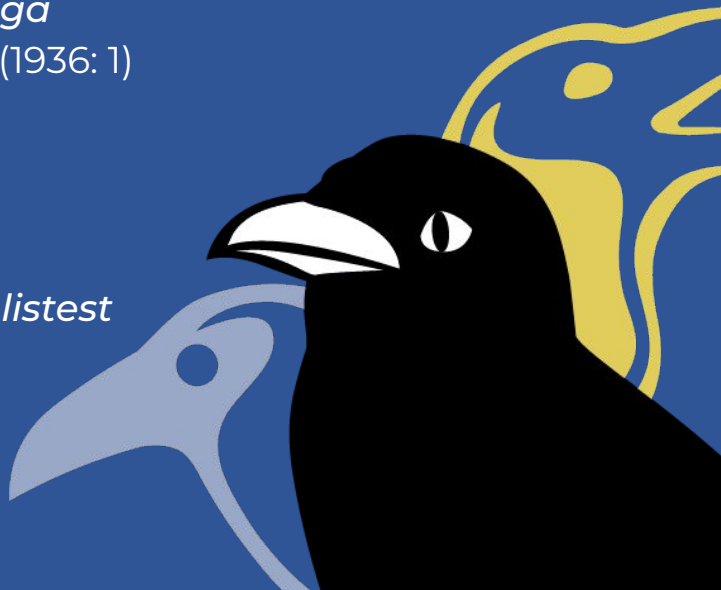
(<https://dea.digar.ee/article/pernopostimees/1858/12/17/6>)



Normeerimine

Kaks leeri:

- Tugevalt ülesse vormi vastu
 - “‘ülesse’ on üldse viga ja on murdekeeles tekkinud sse-lõpulise sisseütleva mõjul; ‘üles’ ei ole sees- ega sisseütlev, vaid muistne s-lõpuline latiiv.” – Aavik (1936: 1)
- Neutraalsem hoiak
 - “Keelelist olukorda ärgem muutkem keeleajaloolistest vaatekohtadest.” – Saareste (1937: 8)



Normeerimine

ÕS 1918: ainult *üles* variant

üle; -anne|tu²; -anne|tus; -
(-homme); -kute; ül|em; -mal
(olema); -male (minema);
-|mane; -|mine; -|mus; -n|da-
ma; -n|dus; -|nema; -ni;
-s; -stikku; -tsi; ül|ew (=

waimliselt kõrge); -wal (ole-
ma); -walt (tulema); -|wane;
-üldi|ne³; -üldis|tama; -ül-
dis|tus; -üld|ne²; -üldse;
-üld|sus

ÕS 2018: *üles*

Sõnaveeb: *ülesse* kõnekeelne ning *üles* normikeelne

et **ülesse** 

EKI ÜHENDSÕNASTIK 2024

et KÕNEKEELNE üles

Näited

Regi jooksis kui nool mäest alla ja ülesse. 

Vahel ärkan keset ööd ülesse. 

Tänapäev

Eesti keele ühendkorpus (2023)

	<i>üles</i>		<i>ülesse</i>		Kokku
academic	3892	97,89%	84	2,11%	3976
blogs	260235	93,60%	17798	6,40%	278033
e-commerce	23478	94,58%	1346	5,42%	24824
encyclopedia	12663	98,56%	185	1,44%	12848
fiction	203516	98,84%	2391	1,16%	205907
forums	194964	85,43%	33246	14,57%	228210
non-fiction	35557	99,80%	72	0,20%	35629
periodicals	599474	98,46%	9362	1,54%	608836
	1333779	95,39%	64484	4,61%	1398263

Andmed

- Kaks suulise kõne korpust:
 - Eesti keele spontaanse kõne foneetiline korpus (EKSKFK, Lippus jt 2023): täiskasvanud vanuses 20–80;
 - Eesti teismeliste keele korpus (TeKE, Vihman jt 2023): teismelised vanuses 9–18.
- Kõik *üles/ülesse* kasutusjuhud koos kontekstiga (TeKE 447; EKSKFK 340).



Vaadeldud tunnused

- Korpuste metaandmetest kõneleja vanus ja sugu.
- Automaatselt arvutatud lokaalne kõnetempo (silpi/sek).
- Käsitsi on kodeeritud
 - sõna asukoht osalauses (alguses, keskel, lõpus),
 - pausi järgnemine sõnale (jah, ei),
 - põhitegusõna algvorm.



Metoodika

1. Varieerumise ühetunnuseline analüüs:

- hii-ruut-testid kategooriliste tunnuste (asukoht, pausi järgnemine) mõju analüüsiks;
- t-test arvuliste tunnuste (vanus, kõnetempo) mõju analüüsiks.

2. Varieerumise mitmetunnuseline analüüs:

- logistilise regressiooni segamudelid.

Eesti teismeliste keele korpus (TeKE)

- Sisaldab 9–18-aastastelt kogutud suulist ja kirjalikku keelematerjali.
- Maht 97 tundi ja 741 175 sõna.
- Materjali 131 noorelt.
- Koostatud aastatel 2020–2022.

okei (...) ma tahtsin lihtsalt teada sest ma mõtlesin eile et mis film on kirjutan	ülesse	teen sihukest segusööki siin et eile oli ka mul olid nagu (...) mitte lagrits	ülesse	lõpp	jah	kirjutama
mida ta teeb need kaardid aa ma pean need teised dokumendid ka	üles	otsima ju või ma võin emale saata lihtsalt ja olla et prindi veel	üles	keskel	ei	otsima
ees oli nagu mis mõttes pall tead m vau ma ärkasin	üles	ja olin et kes see on @(--)@ @nagu palun@ jah	üles	lõpp	ei	ärkama
miks mitte kassi peal @ma räägin see on raske see ei tõuse	ülesse@	(vaata jah) panen (...) kahepoolse teibiga ennast sinna külge (nii aga) kuidas see peaks	ülesse	lõpp	jah	tõusma

Eesti keele spontaanse kõne foneetiline korpus (foneetikakorpus)

- Sisaldab kokku umbes 135 tundi kõnet.
- Ligikaudu miljon sõnatasandi segmenti (ka pausid ja üneemid).
- Andmestikus on esindatud 207 kõnelejat üle Eesti.
- Kogutud aastatel 2006–2023.

oli see asi mida ta # .e kõige lähem mälestus või seal kontserdil .öö/.kärin oli see kui ta kuskil ärkas	ülesse	ja prügi+auto lihtsalt # ku- .mm # .m/.kärin # visat- .e/.kärin kukkus talle mingi prügi peale/.naer või/.naer nagu/.naer .sisse et	ylesse	ülesse	keskel	ei	ärkama
tean küll mihuke see on alguses see pole üldse lauge see on ikka päris # mõnus järsk seal # jooksed	ülesse	.h # .sisse # aga siis vist on edasi on jah sihuke hästi nagu tasakesi aga tõuseb kogu aeg .h	yl{s:se_k	ülesse	lõpp	jah	jooksma
jah .m # nad jaa natuke võiks # .sisse # .h jah ma ei tea kuidas see rada on seal	üles	ehitatud Barcelonas/.kohanimi muidugi .välja # mõnna jah .m # hmm # mhmh # äge .välja # .sisse # .naer #	yles	üles	keskel	ei	ehitama

Andmed

Esialgsest foneetikakorpusest 340 vaatlust ning TeKE-st vastavalt 447.

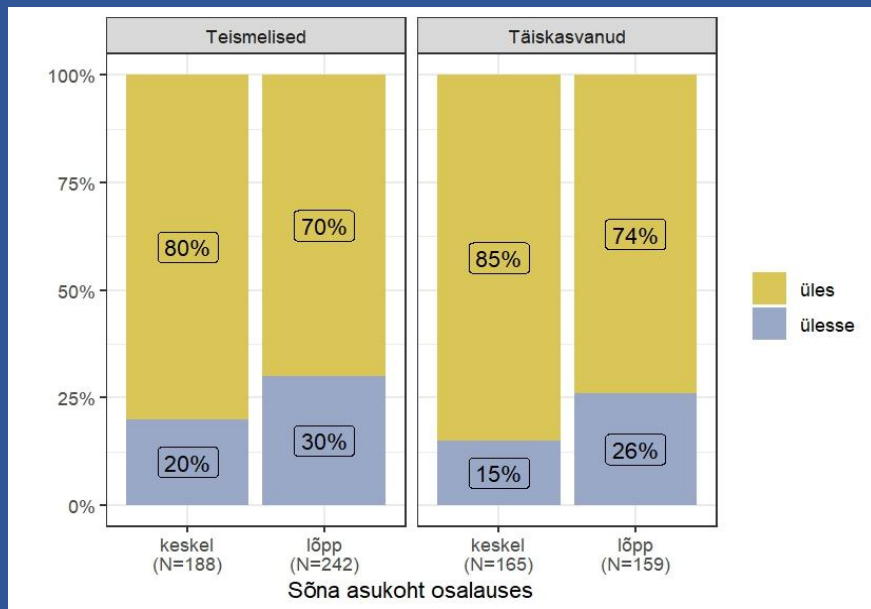
Analüüsi käigus jäeti välja juhud, kus:

- kõneleja sugu või vanus polnud teada;
- polnud võimalik määrata sõna vormi (*üles/ülesse*);
- polnud võimalik määrata asukohta osalauses;
- juhud, kus positsiooni väärtus oli „algus” (foneetikakorpuses kaks ja TeKE-s üks juhtum).

Kahe korpuse võrdlus

Teismeliste ja täiskasvanute
üles/ülesse **proportsioonid** on
sarnased (statistiliselt oluline
erinevus puudub).

Sugu ei mõjuta vormi kasutust.

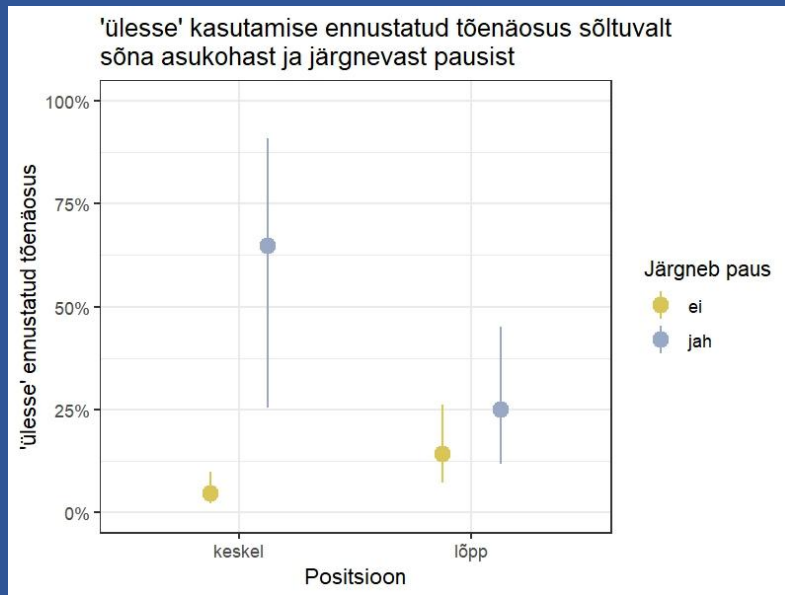


Foneetikakorpus (täiskasvanud)

Parima mudeli puhul sõltus vormi valik positsiooni ja järgneva pausi koosmõjust ning kõneleja vanusest.

Positsioon ja järgneva pausi olemasolu kui peamised vormivalikut määravad tegurid.

Kuigi vormi esinemine osalause lõpus ja järgnev paus iseseisvalt suurendavad *ülesse* tõenäosust, põhjustab nende samaaegne esinemine ootuspärasest väiksema tõenäosuse valida see vorm.



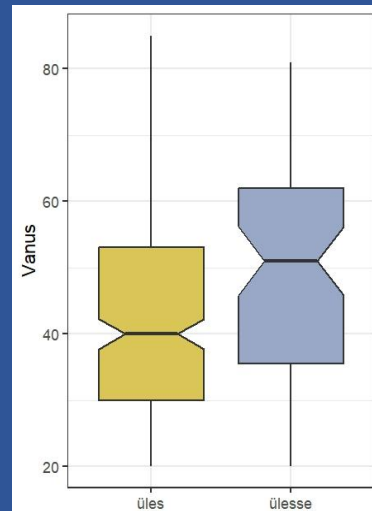
Foneetikakorpus (täiskasvanud)

Vanusel väiksem, kuid siiski statistiliselt oluline mõju.

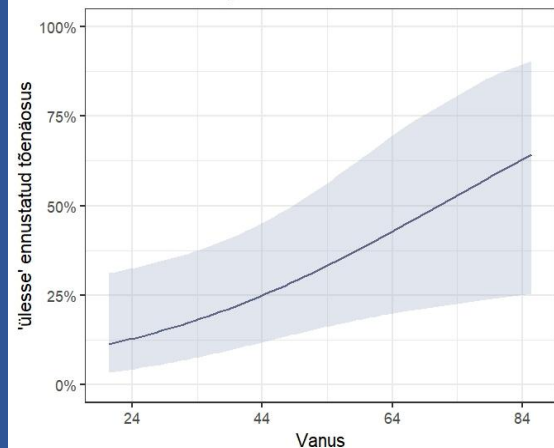
Vanuse kasvades suureneb *ülesse* kasutamise tõenäosus.

Vanemaid kõnelejaid on andmestikus vähem, seega on hinnangus rohkem ebakindlust.

Vormi valik sõltub palju individuaalsest kõnelejast – 56%.



'ülesse' kasutamise ennustatud tõenäosus sõltuvalt kõneleja vanusest

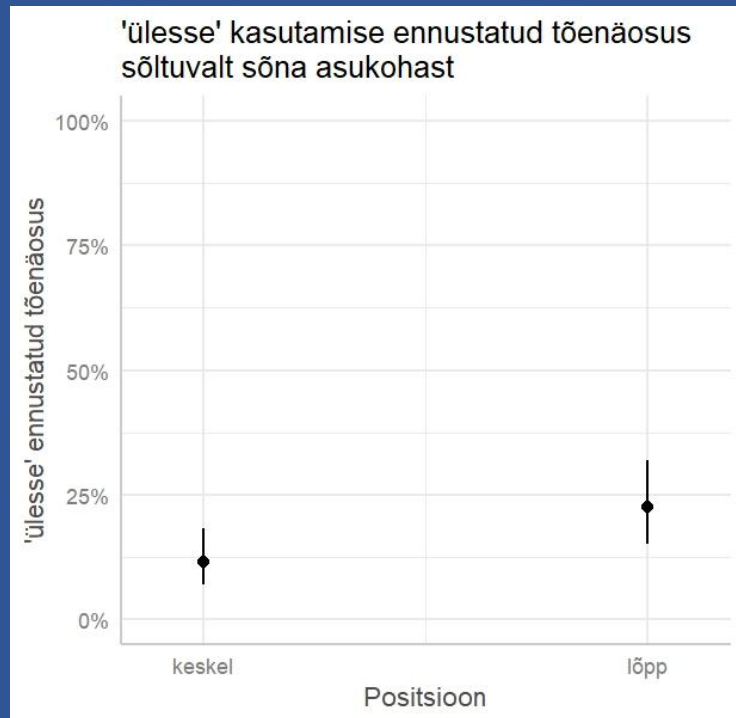


TeKE (teismelised)

Vormi valik sõltub vaid **positsioonist osalauses**.

Sõna paiknemine osalause lõpus suurendab *ülesse* kasutamise tõenäosust.

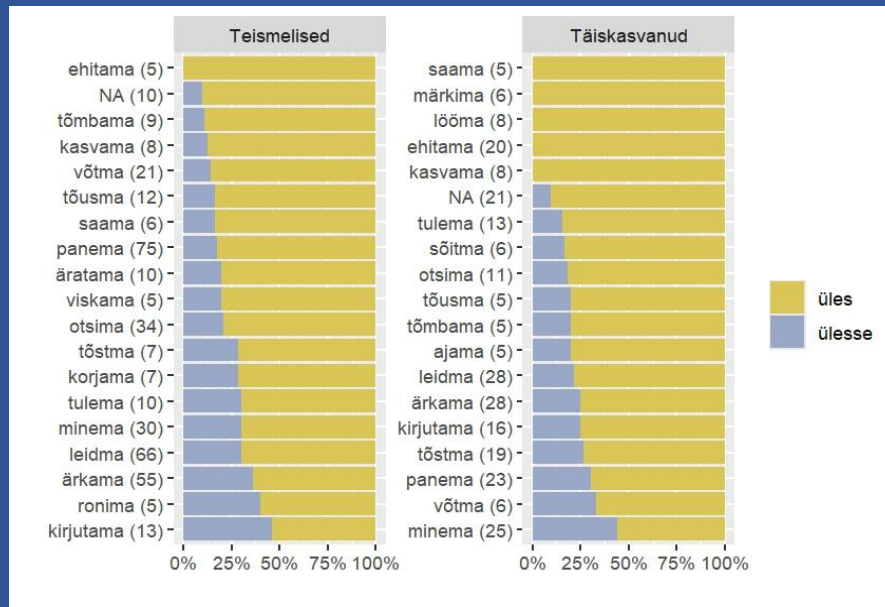
Vormi valik sõltub palju individuaalsest kõnelejast – 37%.



Andmeanalüüsi tulemused

Selgeid mustreid vormivaliku ja tegusõna lemmade vahel ei ilmne.

Mõlemad vormid esinevad nii konkreetsete (*tõstma üles, ronima üles*) kui ka abstraktsete tegevuste (*kirjutama üles, leidma üles*) väljendamisel.



Kokkuvõte

- Kirjakeelne vorm *üles* on vanem, kõnekeelne *ülesse* uuem.
- Täiskasvanute ja teismeliste *üles* ja *ülesse* vormide kasutusproportsioonid ei erine oluliselt (*ülesse* 21–25%).
- Vormi kasutamine on täiskasvanutel seotud järgneva pausiga, vormi asukohaga osalauses ning vanusega, teismelistel seos ainult asukohaga osalauses. Mõjud võivad peegeldada ka korpuse transkribeerimispõhimõtete erinevusi.
- Väga suur individuaalne varieerumine.

Aitäh!



Viited

- Aavik, Johannes** 1936. Eesti õigekeelsuse õpik ja grammatika. Tartu: Noor-Eesti Kirjastus.
- DIGAR Eesti artiklid.** Eesti Rahvusraamatukogu digitaalarhiiv. <https://dea.digar.ee>
- Eesti keele ühendkorpus 2023.** Kristina Koppel, Jelena Kallas, Madis Jürviste, Helen Kaljumäe. Lexical Computing Ltd. / Eesti Keele Instituut.
- Lindström, Liina, Triin Todesk, Maarja-Liisa Pilvik.** 2022. Eesti murrete korpus. Tartu Ülikooli eesti ja üldkeeleteaduse instituut. <https://datadoi.ee/handle/33/492>.
- Lippus, Pärtel, Kätlin Aare, Anton Malmi, Tuuli Tuisk & Pire Teras.** 2023. Phonetic Corpus of Estonian Spontaneous Speech v1.3. Institute of Estonian and General Linguistics, University of Tartu.
- Saareste, Andrus** 1937. Eesti õigekeelsuse päevaküsimustest. Eesti Keele Arhiivi toimetised 9. Tartu: Eesti Keele Arhiiv, lk 8.
- Vihman, Virve-Anneli, Maarja-Liisa Pilvik, Aive Mandel, Annika Kängsepp, Mari Aigro, Kadri Koreinik, Kristiina Praakli & Liina Lindström.** 2023. Estonian Teen Language Corpus.
- ÕS 1918 = Eesti keele õigekirjutuse-sõnaraamat 1918.** Koost. J. Tammemägi. EKS Keeletoimkonna toimetused nr 5. Tallinn: KÜ Rahvaülikooli Kirjastus.
- ÕS 2018 = Eesti õigekeelsussõnaraamat ÕS 2018.** Toim. Maire Raadik. Koost. Tiia Ereli, Tiina Leemets, Sirje Mäearu, Maire Raadik. Tallinn: Eesti Keele Sihtasutus.

