

TARTU ÜLIKOOL  
Arvutiteaduse instituut  
Informaatika

Helena-Kris Rebane

**Programm paronüümide õige kasutamise õppimiseks**

Bakalauerusetöö (9 EAP)

Juhendaja Sven Aller

Tartu 2021

## **Programm paronüümide õige kasutamise õppimiseks**

### **Lühikokkuvõte:**

Selleks, et inimestevaheline suhtlus oleks sujuv ja edukas, on oluline korrektne keelekasutus. Käesoleva töö raames valminud keeleõppeprogramm aitab sellele kaasa, õpetades õige tähendusega paronüümi kasutamist õiges kohas. Töös kirjeldatakse, mis on paronüümid ning kuidas loodi programm nende õppimiseks.

**Võtmesõnad:** paronüüm, keeletehnoloogia, EstNLTK

**CERCS:** P175 Informaatika, süsteemiteooria

## **A tool for learning the correct usage of paronyms**

### **Abstract:**

Correct use of language is important for smooth and effective interpersonal communication. The program created for this thesis aims to help with this by teaching the correct usage of paronyms. This thesis explains the definition of paronyms and how the learning tool was created.

**Keywords:** paronym, language technology, EstNLTK

**CERCS:** P175 Informatics, systems theory

# Sisukord

1. Sissejuhatus	4
1.1 Paronüümid	5
2. Sarnased programmid	6
2.1 KHK Enesekontrollitest paronüümidest	6
2.2 Tartu Ülikooli sõnavaraharjutused kõrgtasemele	6
2.3 Miksikese paronüümide harjutus	6
2.4 Taskutarga paronüümiharjutused	7
3. Kasutatud keeleressursid ja programmi kirjeldus	8
3.1 Väike paronüümisõnastik	8
3.2 Korpus	8
3.3 Eeltöötlus	9
3.4 Veebirakendus	11
4. Kasutatud tehnoloogiad	12
4.1 EstNLTK	12
4.2 Flask	12
4.3 PyYAML	13
4.4 Pickle	13
4.5 React	13
4.6 Material-UI	14
4.7 CSS	14
5. Programmi kasutamine	15
6. Testimistulemused	17
7. Edasiarendusvõimalused	18
8. Kokkuvõte	19
Viidatud kirjandus	20
Lisad	22
I. Litsents	22

# 1. Sissejuhatus

Järjest rohkem on noori, kes loevad vähe eesti keelset toimetatud kirjandust, kust saada eeskujul korralikust keelekasutusest, mille abil arendada sõnavara ja keelevaistu ning tarbivad peamiselt võõrkeelset meediat. Sellega kaasnevad ohud: tekib võimalus, et inimene kas minetab oskuse õigesti kirjutada ja ennast kõnes selgelt väljendada või ei omandagi seda.

Koolinoorte seas aastal 2003 läbiviidud uuringu andmetel peavad kooliõpilased emakeeleoskust küll oluliseks, kuid kahtlevad selle perspektiivikuses ning oleks nõus emakeele kasutamisest loobuma, kui sellega kaasneks eeliseid tööturul või soodustusi ostude tegemisel. Levinud oli laialdane inglise keele kasutus igapäevaelus. [1]

Kaheksa aastat hiljem viidi läbi sarnane uuring, mille tulemused näitasid, et vahepeal on inglise keele kasutamine veelgi levinumaks muutunud ning endiselt ollakse nõus pragmaatilisel kaalutlustel eesti keele asemel inglise keelt kasutama. Uuringu andmetel kasutas 2011. aastal igapäevaselt inglise keelt 24% küsitletutest, aastal 2003 oli see arv 18%, seega tõusis kaheksa aasta jooksul inglise keelt igapäevaselt kasutavate koolinoorte arv 6%. [2]

Martin Ehala kinnitab, et keeleoskus on oluline, tuues välja, et seda on vaja, et omada täit juurdepääsu ühiskonna pakutavatele hüvedele, teha akadeemilist karjääri ja palju muudki. Samas kritiseerib ta liigset ranget ja segadusseajavat reeglipõhist keeleõpet ning peab paremaks intuiitivsemaid lahendusi. Korrektse keelekasutuse õpetamiseks peab ta vajalikuks leida kasutuselolevatele õppemeetoditele paremaid alternatiive. [3]

Eelnevalt refereeritu põhjal võib selgelt järeldada, et Eesti noored vajavad rohkem praktikat oma emakeele kasutamisel. Selle töö raames valmib õppeprogramm, mis seda võimaldab. Kuna eesti keele valdkond on lai ja kogu keelt pole võimalik selle töö raames käsitleda, siis keskendutakse käesolevas töös paronüümide õige kasutamise harjutustele.

Töös sisaldub lisaks sissejuhatusetele ja kokkuvõttele kuus peatükki. Kõigepealt kirjeldatakse sarnaseid programme ning iseloomustatakse neid. Kolmandas peatükis räägitakse töös vaja läinud keeleressurssidest ning kirjeldatakse lähemalt eeltöötusalgoritmi ja kasutajaliidest. Kasutatud tehnoloogiate peatükk sisaldab kasutatud tehnoloogilisi vahendeid, nende tutvustust ning lühikirjeldust nende funktsioonist töös. Peale tehnilise poole selgitamist näidatakse läbi ekraanitõmmiste ning sõnalise kirjelduse õppeprogrammi kasutamist. Kuues peatükk kirjeldab

testimistulemusi ning saadud tagasisidet. Seitsmendas peatükis on välja toodud võimalusi valminud töö edasiarendamiseks.

## 1.1 Paronüümid

Paronüümid ehk samassõnad on sõnad, mis on väga sarnase kõlaga, kuid erineva tähendusega [4]. Paronüümidega puutub eesti keelt kõnelev inimene kokku igapäevaselt, kuna nende hulka kuuluvaid sõnu on nii kirja- kui argikeeles kasutusel palju. Vale paronüümi kasutamine jätab inimesest harimatu mulje ning tekitab arusaamatusi öeldu tähenduse osas.

Näiteks on paronüümid sõna *ennetama*, mis tähendab millegi ära hoidmist, ja *ennistama*, mis tähendab restaureerimist või ametisse vm õiguslikku olukorda seadmist. Nende sarnase kõla tõttu aetakse neid omavahel sageli sassi. Sellega kaasneb ilmsege probleem, kuna lause tähendus võib oluliselt muutuda, kui kasutada valet sõna:

- Ülekuulata **tunnistas** kuritegu
- Ülekuulata **tunnustas** kuritegu

Esimese lause ütleb, et ülekuulata võttis kuriteo omaks. Teine lause väidab, et ülekuulata hindas kuritegu kõrgelt.

## **2. Sarnased programmid**

Paronüümide õppimiseks on olemas varasemalt loodud programme ja internetist leitavaid harjutusülesandeid, kuid üldiselt ei ole nendes sisalduvate harjutuslausete arv väga suur. Järgnevalt on toodud välja mõned avalikult ligipääsetavatest variantidest.

### **2.1 KHK Enesekontrollitest paronüümidest**

Tartu Kutsehariduskeskuse valikvastustega enesekontrollitest on kõigile ligipääsetav KHK e-õppe veebilehel [5]. See koosneb kahekümnest küsimusest ning on vormistatud lihtsa valikvastustega lünktekstina, kus paronüümi asukohta tähistab kolm punkti. Kohe kui mõni vastustest on valitud, ütleb programm testi tegijale, kas ta valis õige variandi.

Programm on mõeldud enesekontrolliks ning see sisaldab ainult kakskümmend küsimust, seega pole see kordamiseks ja õppimiseks kõige parem. Erinevalt käesoleva töö raames valminud programmist näidatakse selles testis kohe, kas vastus on õige, mis ei soodusta vastuse pikemat läbimõtlemit.

### **2.2 Tartu Ülikooli sõnavaraharjutused kõrgtasemele**

Tartu Ülikooli keeleveebis on avalikkusele kättesaadavad „Eesti keele sõnavara“ kursuse alla kuuluvad harjutusülesanded [6], mille hulgas on ka paronüümidega seonduvaid ülesandeid. Paronüümide ja homonüümide osa sisaldab viite alapunkti, mis jagunevad kahekümne viieks küsimuseks. Alla poole neist kontrollivad paronüümide tundmist.

Sarnaselt eelmainitud KHK enesekontrollitestile pole nendes harjutustes piisavalt ülesandeid, et neid peamise õppe- ja kordamisvahendina kasutada.

### **2.3 Miksikese paronüümide harjutus**

Miksikese lehel on saadaval mitmeid keeleõppemänge [7], nende hulgas ka üks paronüümide kohta [8]. See test koosneb kolmekümne kuuest lünktekstina vormistatud küsimusest, mis käsitlevad kümmet paronüümipaari. Harjutuse teeb huvitavaks mänguline element – skoori lugemine. Lünkadesse tuleb lohistada õige valikuvariant, kusjuures nii kiirus kui täpsus loeb, kuna nii vale vastuse katsetamine kui ka möödud aeg võtavad punkte maha. Ülesannetes tuleb sobitada paronüümid õigete definitsioonidega ning kasutada õigetes lausetes. Testi lõpus on võimalus vaadata parimate tulemuste edetabelit.

Miksikese paronüümide harjutus on lahendatud huvitavalt, kuid nagu teisteski programmides, ei ole lausete ja erinevate paronüümide arv ülesannetes kuigi suur.

## **2.4 Taskutarga paronüümiharjutused**

Taskutarga veebilehel on saadaval kaks paronüümide õige kasutamise testi: esimeses testis [9] tuleb valida, kas väide on õige või vale, teises [10] tuleb lünkteksti valida õige paronüüm.

Esimesene test koosneb kahekümnest küsimusest. Iga küsimus sisaldab paronüümi definitsiooni, näiteks „VAHEL tähendab ‘ehk’, ‘võib-olla’“, mille kohta tuleb märkida, kas see on õige või vale. Küsimused on iga kord samad, kuid erinevas järjekorras. Iga paronüümi kohta on kaks küsimust. Kui kõik küsimused on vastatud, kuvatakse õigete vastuste osakaal nii protsentuaalselt kui murdarvu kujul ning valitud vastused, mille kohta näidatakse, kas need olid õiged.

Teises testis on kokku neliteist ülesannet. Küsimused on vormistatud valikvastustega lünktekstina, kus valikuteks on kaks paronüümi. Sarnaselt esimesele testile kuvatakse õigete vastuste osakaal protsentuaalselt ning murdarvuliselt ja lahendaja saab kontrollida oma vastuste õigsust. Teises testis esindatud paronüümid on samad, mis esimeseski.

### **3. Kasutatud keeleressursid ja programmi kirjeldus**

Programmi ülesanded on koostatud Eesti ilukirjanduse koondkorpuse lausetest, kasutades paronüümisõnastikku, et välja sorteerida sobivad laused. Eeltöötlus ning veebirakendus on teostatud eri projektides.

#### **3.1 Väike paronüümisõnastik**

Raamat „Väike paronüümisõnastik“ [11] (edaspidi „paronüümisõnastik“) sisaldab nimekirja paronüümidest koos definitsioonide ja näidetega. Sõnad on oma samassõnadega grupeeritud ning grupe eraldavad nende vahel olevad tühjad read. Programm leiab paronüümisõnastikust sada kolmkümmend paronüümigruppi, kokku kolmsada kolmteist sõna.

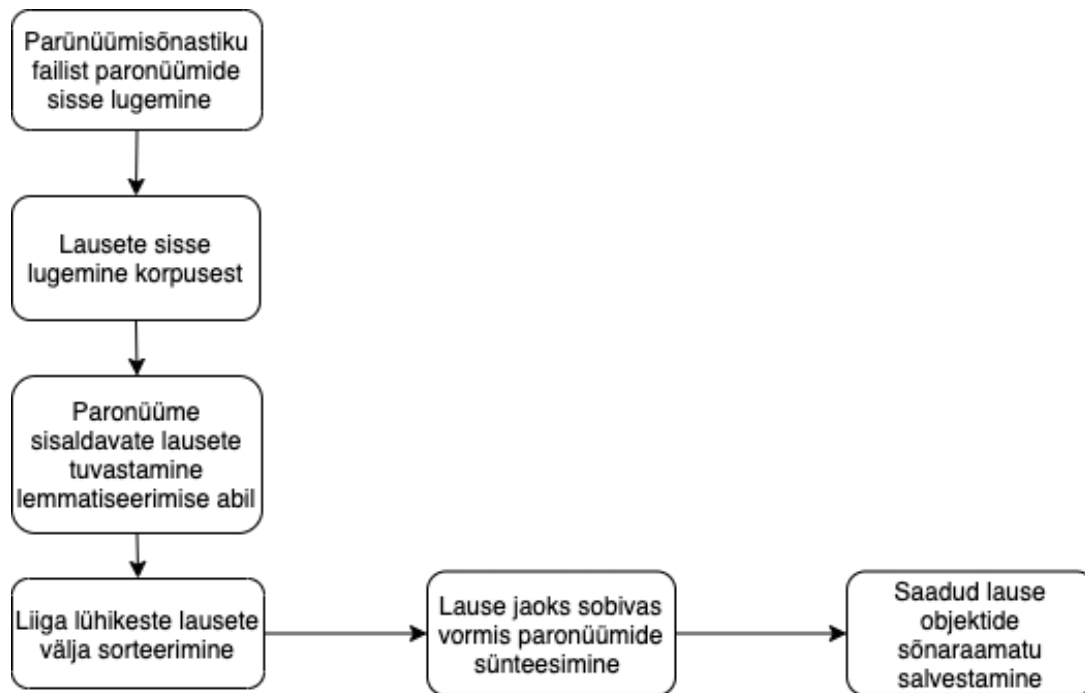
#### **3.2 Korpus**

Korpus on keeleressursside liik, mis sisaldab märgendatud teksti. Märgendamine võimaldab kergemat automaatset töötlemist kui tavateksti puhul. "Keeleressurss on masinloetaval kujul keeleline andmekogum, mida kasutatakse arvutites loomuliku keele (sh. eesti keele) uurimiseks või keeletehnoloogia arendamiseks." [12]

Selle programmi jaoks kasutati Eesti ilukirjanduse segakorpusest pärinevaid laused. See korpus sisaldab ilukirjanduslikke tekste, millest suurem osa on avaldatud peale 1990. aastat ning on vabalt kasutatav mitteärilistel eesmärkidel [13].



### 3.3 Eeltöötlus



Joonis 3.1 Veebirakendusele sobivate lauseobjektide koostamine

Programmi jaoks vajalike lausete genereerimine toimub veebirakendusest eraldi, teises projektis, kuna seda on vaja teha ainult ühe korra.

Paronüümisõnastiku põhjal on koostatud kahemõõtmeline järjend, kus omavahel on grupeeritud paronüümid. Ühe grupi moodustavad näiteks *edastama*, *edendama*, *edestama* ja *etendama* ning iga grupp sisaldab vähemalt kahte sõna. Programmisisene järjend genereeritakse tekstidokumendi põhjal, milles on paronüümisõnastikus sisalduvad paronüümid ning gruppide eraldamiseks kasutatakse tühja rida. Selle alusel genereeritakse kahemõõtmeline järjend, kus iga paronüümigrupp moodustab eraldi alamjärjendi. Kokku saadakse nii sada kolmkümmend kolm gruppi, mis kokku sisaldavad 313 paronüümi.

Laused pärinevad Eesti ilukirjanduse koondkorpusest ning nende hulgast on välja otsitud programmile tuttavaid paronüüme sisaldavad laused. Paronüüme sisaldavate lausete tuvastamiseks kasutati programmi sõnastikku ning EstNLTK abil lemmatiseerimist ehk sõna algvormile viimist.

Sobivate lausete korpusest välja sorteerimiseks vaatab programm ükshaaval läbi kõik korpuses sisalduvad laused, lemmatiseerib vaadeldava lause sõnad ning võrdleb saadud lemmasid paronüümisõnastikus olevate sõnadega. Kui lemmatiseeritud lausest leitakse paronüüm, siis märgitakse see lause paronüümi sisaldavaks, viiakse kõik vastava grupi paronüümid lause jaoks

morfoloogilist analüüsi ja sünteesi kasutades sobivale kujule ning arvutatakse välja paronüümi positsioon lauses, et võimaldada lünga kuvamine õiges kohas. Kuna väga lühikesest lausest on ilma kontekstita raske aru saada, siis alla neljasõnalised laused jäetakse välja. Sama tehakse ka liiga pikkade lausetega, mis selle töö raames tähendab üle viieteistkümnest sõnast koosnevaid lauseid. Kokku leiab programm ilukirjanduskorpusest 2473 sobiva pikkusega ja paronüümi sisaldavat lauset.

Peale sobivate lausete tuvastamist luuakse nendest sõnastik, kui võtmeks on lause ning väärtusteks kahemõõtmeline järjend, milles elementideks on järjend õigesse käändesse või pöördesse seatud ning vajadusel suure algustähega paronüümidega, lauses esinenud paronüüm ning paronüümi algus- ja lõpuindeks lauses ehk:

```
{„Leitud lause.“: „paronüüm lauses“, [„paronüüm1“, „paronüüm2“], <indeks1>, <indeks2>}
```

Saadud sõnastik näeb välja selline:

```
{"Ja mina tunnen sind juba ligi kolmkümmend aastat," õhkas itaallane.': ([ 'õhkis', 'õhkas'], 'õhkas', 52, 58), 'Lugeja leiab neilt lehekülgedelt ka Karl Asti ja Friedebert Tuglase nimed.': ([ 'nimetused', 'nimed'], 'nimed', 68, 75), 'Enamiku vööasteroidide kaugus üksteisest on 2 — 4 astronoomilist ühikut.': ([ 'Enamiku', 'Enamuse'], 'Enamiku', 0, 8)}
```

Valmis sõnastik salvestatakse YAML-faili.

Automaatset töötlust kasutades jäid välja valitud lausete hulka mitmed sobimatud laused. Selliste lausete hulgas olid laused,

- mis ei olnud üheselt mõistetavad, näiteks: „Kuni "Tallinna" kohviku taganurga seltskondlik vanaraualadu pakkus mulle sobivamana paistva juhuse.“ (juhu, juhtumi, juhuse)
- mis sisaldasid hoopis mõne programmile tuttava paronüümi homonüümi ehk sõna, mis on sama kirjapildiga, kuid teise tähendusega, näiteks: „Kõhus jänesekondid, -maks, -süda ja mõned soolikad.“ („makse“, „maks“)
- mis olid õppeprogrammi jaoks sobimatu sisuga, näiteks: „Rahvas väänles kogu vaatajaskonda vallanud võimsa orgasmilaine käes.“
- mis korpuse eripärade tõttu sisaldasid valesti paiknevaid sümboleid, näiteks üksikuid jutumärke lause alguses või lõpus.

Ebasobivatest lausetest vabanemiseks või võimalusel parandamiseks on programmi laused enne veebirakendusele edasi andmist käsitsi üle vaadatud, eemaldades vales tähenduses homonüümi sisaldavad laused, arusaamatud laused ning laused, mis sisaldavad liiga vähe konteksti või muul põhjusel lubaks mitut paronüümi grupist õigeks lugeda. Veel kustutati laused, mille sisu oli mingil viisil kohatu ning vähendati lausete arvu, mis sisaldasid kõige levinumaid paronüüme, et erinevate paronüümide suhe ülesannetes rohkem tasakaalus oleks.

Korpuse piirangute tõttu jäi siiski üle poole paronüümisõnastikus toodud paronüümidest välja. Kõige populaarsemad paronüümid olid: „enne“, „andma“, „kogu“, „jälgima“, „õieti“, „teatama“, „eelmine“, „tunnistama“, „täna“ ja „kinnitama“. Igaühte neist esines ülesannetes üle kolmekümne korra.

Lõpuks salvestatakse Pickle mooduli abil käsitsi valitud 1295 lausest sõnastiku objekt, mida veebirakendus kasutada saab.

### **3.4 Veebirakendus**

Valminud programm aitab õppida paronüümide õiges tähenduses ja kontekstis kasutamist. See on teostatud veebirakendusena, mis võimaldab õppijal sellele mugavalt ning lihtsalt ligi pääseda, kuna selleks ei ole vaja midagi alla laadida, piisab veebiaadressi teadmisest.

Eeltöötlemise tulemusel saadud lausetest kuvab veebirakendus õppijale kümne ülesande kaupa lünkteksti, kus õppija peab lüngad täitma õigete paronüümidega. Selleks pakitakse *back-endis* lahti Pickle mooduli abil salvestatud sõnaraamatu objekt, milles sisalduva üle tuhande lause hulgast valitakse iga harjutuskorra jaoks juhuslikult kümme. *Front-endis* kuvatakse Sentence objekti põhjal laused, kus paronüümi kohal on lauses hüpikmenüü võimalike paronüümidega. Programmil on selliseid lauseid kokku 1295. Kui lüngad on täidetud, saab kasutaja oma vastuseid kontrollida ning näha õigeid, kui ta on valesti vastanud.

## 4. Kasutatud tehnoloogiad

Alusprogrammi keeleks on Python, kuna ainus töö kirjutamise ajal olemasolev eestikeelsete vabatekstide töötlemise jaoks mõeldud tarkvarateek EstNLTK nõuab seda. Lausete genereerimisel on kasutusel veel moodulid Pickle ning PyYAML.

Back end ehk tagarakendus on rakenduse osa, mis tegeleb resursside ja andmete haldusega [14]. Valminud programmi *back end* on kirjutatud Pythonis ning kasutab Flaski.

*Front end* ehk eesrakendus on kasutajale nähtav osa. Veebirakenduse *front end* on kirjutatud keeles React, kuna see võimaldab lihtsalt ja kiiresti luua kasutajasõbraliku ning professionaalse välimusega veebilehe. Suurem osa *front endis* kasutusel olevatest elementidest pärinevad raamistikust Material-UI. Veel on kasutatud CSS-i.

### 4.1 EstNLTK

EstNLTK on Tartu Ülikooli loodud *Pythoni* teekide kogumik, mille eesmärgiks on hõlbustada eestikeelsete vabatekstide keeletehnoogilist töötlemist. EstNLTK liidestab erinevaid varem olemasolevaid abivahendeid. [15] Teegis on olemas sõnestaja, lausestaja, osalausestaja ning selle abil saab teostada morfoloogilist ja süntaktilist analüüsi ning sünteesi. [16] Lisaks mitmekülgsel keeletehnoogilisele võimekusele on EstNLTK ka põhjalikult dokumenteeritud [17].

Programmis on kasutusel EstNLTK versioon 1.4. Selle abil toimub lausete sõnade lemmatiseerimine ehk algvormile viimine, et neid saaks sõnastiku paronüümidega võrrelda ning lause lünga jaoks sobivate vormide sünteesimine paronüümide algvormidest.

### 4.2 Flask

Flask on mikroraamistik Python *back-endi* loomiseks. Kuigi Flask on minimalistlik, võimaldab ta vastavalt vajadusele mitmesuguste laienduste kasutamist [18]. Flaski kood on loetav ning puhas ja lisaks põhjalikule dokumentatsioonile on selle raamistiku kasutamise kohta React *frontendiga* internetis palju õppematerjale.

Õppeprogrammis tegeleb Python Flask *backend* React *frontendi* päringutega täitmisega. *Backend* edastab *frontendi* päringu vastuseks PKL-failist saadud sõnaraamatust kümme juhuslikult valitud lause objekti, mida *front-end* kasutajale ülesannetena kuvab.

### 4.3 PyYAML

YAML on keel, mis võimaldab programmi objekte salvestada nii, et inimene saab loodud faili lugeda ning selle sisu modifitseerida. YAML jõudis avalikkuse ette esimest korda aastal 2001 [19]. PyYAML on Pythoni moodul, mille abil saab erinevaid Pythoni objekte inimesele loetaval kujul salvestada [20].

PyYAML-i kasutamine programmis näeb välja nii: peale seda, kui programm on lausetega tehtavad automaatsed sammud – paronüüme sisaldavate lausete korpusest välja otsimine, õigete vormide sünteesimine paronüümide grupile ning valmis lause objektidest sõnaraamatu loomine – kirjutatakse see YAML-faili, et võimaldada inimesel fail üle kontrollida, vajadusel vead parandada ja sobimatud laused eemaldada.

### 4.4 Pickle

Pickle on Pythoni moodul, mille abil saab erinevaid Pythoni objekte faili salvestada [21]. Pickle'i eeliseks on kasutajamugavus ning laialdaselt saadavalolevad koodinäited ning õpetused. Kahjuks ei ole Pickle'iga loodud failid aga inimloetavad.

Töös kasutatakse seda lause objektide sõnaraamatu salvestamiseks 'handpicked\_sentences\_clean.pkl' faili, kust veebirakendus hiljem lauseid valib.

### 4.5 React

React on väga laialdaselt kasutuselolev Facebooki loodud JavaScripti põhine *frontend*-raamistik, mis võimaldab efektiivselt ning paindlikult veebirakendusi luua [22]. Alusrakenduse loomine toimub vaid ühe käsuga ning saadaval on lai valik lihtsasti installeeritavaid teeke. Tänu React'i populaarsusele on internetis ohtralt õpetusi ja koodinäiteid ning ka teegi loojad on seda põhjalikult dokumenteerinud.

Reacti kasutamise eelis ise HTML-CSS-JavaScript'is lehe elementide kirjutamise ees seisneb selles, et tänu erinevatele teekidele on suur osa tööd juba varem ära tehtud ning programmeerijal on vaja vaid vaja alla laadida valmis kasutajaliidese komponentide teek ning neile parameetrid ette anda.

## 4.6 Material-UI

Material-UI [23] on Reacti rakenduste komponentide teek, mis pakub mitmeid erinevaid *frontend* elemente. Educative'i Material-UI artiklis [24] on Material-UI eelistena välja toodud põhjaliku dokumentatsiooni olemasolu, sagedased uuendused ning komponentide välimuse ühtsus, mis tagab nende kokkusobimise ning aitab veebilehel meeldiva välimuse säilitada.

Käesolevas programmis on Material-UI komponente kasutatud nii veebilehe joondamisel kui ka hüpikmenüüde, eraldajate ning nuppude loomisel.

## 4.7 CSS

CSS ehk *Cascading Style Sheets* on väga laialdaselt kasutusel olev keel, mis võimaldab kujundada HTML ning XML keeltes kirjutatud dokumente [25]. CSS muudab terve lehe välimust – samale HTML-ile uue CSS-faili lisamine muudab kogu lehte drastiliselt [26]. Kuna CSS-i hoitakse eraldi failis, siis võimaldab see hoida lahus ning taaskasutada stiilielementide koodi.

Suurem osa selles programmis sisalduvatest komponentidest ei vajanud tänu Reactile ja Material-UI-le lisa CSS-i, kuid see on siiski mõnes kohas kasutusel.

## 5. Programmi kasutamine

Lihtsuse huvides on kasutajaliides loodud veebirakendusena, mis võimaldab mugava ligipääsu mitmetest arvutitest ja kaotab vajaduse programmi kasutamise eeldusena midagi alla laadida.

Kasutajaliides kuvab õpilasele ühe harjutuskorra jooksul kümme lünklauset, kuhu tuleb õige paronüüm valida. Kui valikud on tehtud, saab vastust kontrollida ja laadida uued ülesanded.

### Vali lünka sobiv paronüüm

1. Mööda dändidest, kes  saksa turistidest.
2. Trevor kasutas  ja lasi jalga, otsustades edaspidi ettevaatlikum olla.
3. Algasid talvised eksamid ja .
4. Eminents oli teed mekkides silmad sulgenud, nüüd avas ta need aeglaselt ja  tassi lauale.
5. Armastuses tahetakse ajada äri, kus  hinge, vastu aga pakutakse keha.
6.  kui tema surnulõustast soolasambaks tardunud naine piuksatadagi jõudis, haarasid koolnu luised sõrmed tal kõrist.
7. "Oh, tikime ja heegeldame nagu muistegi,"  Tom vagusalt.
8. "Ja te soovite  tõlget ka hollandikeelsest tekstist?" palusin üle korrata.
9. Seetõttu  korterivaliku suures osas vana logisev lift.
10. Rahva Meele Teadusliku  Instituudi direktor Madis Lambert, luges ta ühelt ukselt.

KONTROLLI

UUED KÜSIMUSED

Joonis 5.1. Programmi kasutajaliides

### Vali lünka sobiv paronüüm

1. Mööda dändidest, kes  elatavad saksa turistidest.
2. Trevor kasutas  ja lasi jalga, otsustades edaspidi ettevaatlikum olla.
3. Algasid talvised eksamid ja .
4. Eminents oli teed mekkides silmad sulgenud, nüüd avas ta need aeglaselt ja  tassi lauale.
5. Armastuses tahetakse ajada äri, kus  hinge, vastu aga pakutakse keha.
6.  kui tema surnulõustast soolasambaks tardunud naine piuksatadagi jõudis, haarasid koolnu luised sõrmed tal kõrist.
7. "Oh, tikime ja heegeldame nagu muistegi,"  Tom vagusalt.
8. "Ja te soovite  tõlget ka hollandikeelsest tekstist?" palusin üle korrata.
9. Seetõttu  korterivaliku suures osas vana logisev lift.
10. Rahva Meele Teadusliku  Instituudi direktor Madis Lambert, luges ta ühelt ukselt.

KONTROLLI

UUED KÜSIMUSED

Joonis 5.2 Lünka täitmine

Joonisel 5.1 on kujutatud programm algseisus, kus on programmi lausete seast valitud juhuslikult kümme. Igas lauses asub paronüümi kohal lünk, mille menüüst tuleb valida sinna sobiv paronüüm (vt joonis 5.2). Veebilehe alumises osas asub kaks nuppu: „KONTROLLI“, mis kontrollib, kas lüngad on korrektselt täidetud „UUED KÜSIMUSED“, mis värskendab lehte ning kuvab uued lünkülesanded.

## Vali lünka sobiv paronüüm

1. Mööda dändidest, kes elatuvad  saksa turistidest. ✓ Õige!
2. Trevor kasutas juhust  ja lasi jalga, otsustades edaspidi ettevaatlikum olla. ✓ Õige!
3. Algasid talvised eksamid ja arveldused  ✗ Õige vastus: **arvestused**
4. Eminents oli teed mekkides silmad sulgenud, nüüd avas ta need aeglaselt ja  tassi lauale. ✗ Õige vastus: **asetas**
5. Armastuses tahetakse ajada äri, kus nõutakse  hinge, vastu aga pakutakse keha. ✓ Õige!
6. Enne  kui tema surnulõustast soolasambaks tardunud naine piuksatadagi jõudis, haarasid koolnu luised sõrmed tal kõrist. ✓ Õige!
7. "Oh, tikime ja heegeldame nagu muistegi," kinnistas  Tom vagusalt. ✗ Õige vastus: **kinnitas**
8. "Ja te soovite täna  tõlget ka hollandikeelsest tekstist?" palusin üle korrata. ✓ Õige!
9. Seetõttu määras  korterivaliku suures osas vana logisev lift. ✓ Õige!
10. Rahva Meele Teadusliku Uurimise  Instituudi direktor Madis Lambert, luges ta ühelt ükselt. ✓ Õige!

KONTROLLI

UUED KÜSIMUSED

### Joonis 5.3 Vastuste kontrollimine

Joonis 5.3 näitab programmi peale lünkade täitmist ja „KONTROLLI“ vajutamist. Korrektselt täidetud lünga puhul ilmub lause järgi roheline linnuke kirjaga „Õige!“, tühja või valesti täidetud lünga puhul punane ristike ning kuvatakse õige vastus. Seejärel on võimalik saada nupule „UUED KÜSIMUSED“ vajutades järgmised ülesanded või muuta vastuseid ja uuesti kontrollida.



## 6. Testimistulemused

Rakendust testisid viis inimest vanuses 26-47, kellel paluti lahendada paronüümirakenduses viiskümmend ülesannet ning seejärel vastata küsimustikule oma kogemuse kohta. Küsimustik koosnes järgmistest küsimustest:

- Vastaja vanus.
- Kas veebirakendus oli mugavalt kasutatav?
- Kas veebirakenduse kasutamisel esines vigu?
- Kui esines vigu, palun kirjuta siia.
- Kas laused olid üheselt mõistetavad ja arusaadavad?
- Kas lausete sisu oli õppeprogrammi jaoks kohane?
- Kas vastuste kontrollimine toimus korrektselt?
- Kui esines vigu, siis palun täpsusta.
- Kas veebirakenduses oli midagi puudu?
- Kui oli puudu, siis palun täpsusta.
- Palun hinda üldist raskustaset.
- Kas on lisamärkusi?

Suurem osa vastajatest nõustus, et veebirakendust on mugav kasutada. Üks inimene tõi välja, et telefonis pole võimalik viimasele küsimusele vastata, kuna nupud katavad selle ära. Esinenud vigade all toodi välja sama probleem. Üks vastaja ütles, et vastuste kontrollimine ei toimunud tema hinnangul iga kord õigesti, kuid seda väidet ei saanud kinnitada ega ümber lükata, kuna ta ei kirjutanud potentsiaalselt vigast lauset tagasisidesse ning ei suutnud seda ka hiljem meelde tuletada.

Kõik vastajad leidsid, et ülesanded on üldiselt üheselt mõistetavad ning et nende nähtud laused olid õppeprogrammi jaoks kohase sisuga. Veebirakenduste puudustena toodi taas välja telefonis kasutamiseks sobimatus, ühele vastajale ei meeldinud hüpukmenüüde laius ning üks vastaja leidis, et võiks „äkki mingi õnnetlusakna ka ette visata kui kõik õige või mõnitada, kui kõik vale“.

Raskustaset hinnati viie palli süsteemis kolmeks. Veel tõi üks vastaja välja, et tal hakkasid mõned laused kohe korduma.

## 7. Edasiarendusvõimalused

Testimistulemusi ja töö käigus ilmunud probleeme arvesse võttes annaks valminud programmi täiendada mitmel viisil.

Alustuseks annaks palju juurde rohkemate paronüümide ning eriti lausete lisamine. Programmis on üle tuhande lause, kuid enamik lausetest sisaldavad samu kõige levinumaid paronüüme, mida õnnestus ilukirjanduskorpusest leida. Suurem osa paronüüme on kas esindamata või väheesindatud. Edasiarendamisel tasuks lisada programmi selliseid lauseid, mis sisaldavad vähemlevinud paronüüme. Selleks oleks vaja korpust, kus esinevate sõnade valik on laiem.

Edasiarenduse teostamine kujul, kus on vähemesindatud paronüüme rohkem, oleks küllaltki lihtne. Vaja läheks suuremat korpust, nagu näiteks Eesti keele koondkorpus [27]. Selleks, et saavutada ka parem tasakaal sõnade vahel ning lisada lauseid, kus on esindatud just sellised paronüümid, mida ei praegu ei ole või esineb liiga vähe, tuleks kõigepealt koostada nimekiri paronüümidest, mille kohta on juba piisavalt ülesandeid ning vastavaid lauseid mitte juurde lisada.

Lisaks ei sisalda „Väike paronüümisõnastik“ kõiki eesti keeles leiduvaid paronüüme, seega saaks ka programmile tuttavat paronüümide sõnastikku laiendada. Selleks oleks vaja rohkemaid paronüüme sisaldavat sõnaraamatut.

Veel saaks lisada paronüümide definitsioonid, et õppija saaks peale kontrollimist rakendusest valede vastuste puhul näha, mida tema valitud paronüüm tegelikult tähendab ja miks see lausesse ei sobi. Tegemist on küllaltki lihtsa laiendusega, mis nõuaks paronüümisõnastiku sisselugemist kujul, kus definitsioonid säilivad ja lause objektile definitsioonide lisamist. Siis saab neid kasutajale näiteks *tooltipi* kaudu kuvada.

Veebirakenduse edasi arendamiseks võiks panna seda ühe sessiooni vältel salvestama ja väljastama saadud punktisummat või lisada kasutajakontode registreerimine, et võimaldada lisafunktsioone, näiteks tulemuste salvestamine ja oma arengu jälgimine. Selline arendus on veidi mahukam, kuna nõuab minimaalselt ühe vaate lisamist ning kasutajasüsteemi loomist.

## 8. Kokkuvõte

Käesoleva bakalaureusetöö käigus valmis paronüümide õppeprogramm, mis sisaldab üle 1295 lünkülesande. Ülesanded on koostatud automaatselt paronüümisõnastiku põhjal Eesti ilukirjanduse koondkorpuse lausetest, kasutades morfoloogilist analüüsi ja sünteesi. Enne veebirakendusele andmist on ülesanded käsitsi üle kontrollitud ning vajadusel parandatud või eemaldatud. Saadud laused on üheselt mõistetavad ja selged.

Töö eesmärk sai täidetud: õppeprogramm töötab nagu soovitud, kuid tekstikorpuse sisust tulenevalt on programmis ülekaalus mõned üksikud levinumad paronüümid ning mitmeid „Väikses paronüümisõnastikus“ esindatud paronüüme programmi lausetes ei ole või esineb harva. Ka ei ole kasutatud sõnastikus kaugeltki kõiki eesti keeles olemas olevaid paronüüme, mille tõttu ei tuvastatud ilmselt kõiki korpuses sisalduvaid paronüüme.

Testimistulemuste põhjal võib väita, et üldiselt on kasutajad programmiga rahul. Laused olid vastajate hinnangul arusaadavad ja üheselt mõistetavad ning enamik vastajatest nõustusid, et vastuste kontroll toimub korrektselt. Kasutamiskogemus oli enamiku testijate jaoks meeldiv ning ülesannete raskustase keskmine. Üldiselt hinnati rakendust positiivselt.

Rakenduse edasi arendamiseks on mitmeid võimalusi ning suundasid. Programmi sisu saaks täiendada laiemal sõnastiku põhjal ning suurema korpus võimaldaks lisada vähemlevinud paronüüme sisaldavaid lauseid. Veebirakendusele saaks lisada võimekuse kuvada ka paronüümide definitsioonid ning võimaluse kasutajaid luua ja oma arengut jälgida.

Rakendus on kättesaadav aadressil [http://prog.keeleressursid.ee/veebid/paronym\\_game/](http://prog.keeleressursid.ee/veebid/paronym_game/)

Eeltöötuse lähtekood on nähtav GitHub repositooriumis <https://github.com/helenarebane/generate-sentences> ning veebirakenduse kood asub aadressil <https://github.com/helenarebane/paronym-game>.

## Viidatud kirjandus

- [1] Ehala, Martin; Niglas, Katrin (2004). "Eesti koolinoorte keelehoiakud". Akadeemia, 10, 2115–2142. [http://www.tlu.ee/~kairio/7071/2%20klasteranalyyys/keelehoiakud\\_artikkel.pdf](http://www.tlu.ee/~kairio/7071/2%20klasteranalyyys/keelehoiakud_artikkel.pdf)
- [2] Martin Ehala, Anni Tammemägi. "Eesti õpilaste keelehoiakud 2011. aastal". Keel ja Kirjandus 04:241-260 (2012). <https://www.ceeol.com/search/viewpdf?id=250586>
- [3] Martin Ehala "Kui vajalik on grammatika?". Keel ja Kirjandus 08-09:614-628 (2012). <https://www.ceeol.com/search/viewpdf?id=127287>
- [4] Eesti Keele Instituut (18.04.2021) <https://www.eki.ee/books/ekk09/index.php?p=6&p1=4>
- [5] KHK e-õppe lehekülg (22.03.2021) [https://e-ope.khk.ee/ek/2013/eesti\\_sonavara/sona\\_tahendus/enesekontrollitest\\_paronmidest.html](https://e-ope.khk.ee/ek/2013/eesti_sonavara/sona_tahendus/enesekontrollitest_paronmidest.html)
- [6] Eesti keele ja kultuuri kursused (22.03.2021) <https://keeleweb2.ut.ee/kursused/sonavaraharjutused-korgtasemele>
- [7] Miksike veebileht (18.04.2021) <https://www.miksike.ee/en/gtests.html>
- [8] Miksike veebileht (18.04.2021) <https://www.miksike.ee/en/gtests.html?start=1&test=8074>
- [9] Taskutarga veebileht (18.04.2021) <https://www.taskutark.ee/test/paronuumid-ehk-sarnassonad-jargi-voi-jarele-vahest-voi-vahel-jne-vali-kas-vaide-on-oige-voi-vale/>
- [10] Taskutarga veebileht (18.04.2021) <https://www.taskutark.ee/test/paronuumid-ehk-sarnassonad-jargi-voi-jarele-vahest-voi-vahel-jne-vali-lauses-oige-puudev-sona/>
- [11] Plado H., Mandra K. "Väike paronüümisõnastik". Tartu: Keelehooldakeskus (2008). [http://engine.koduleht.net/templates/keelehooldakeskus/files/mdl\\_files.php/Mandra\\_Plado\\_Vaike\\_paronuumisonastik.indd.pdf](http://engine.koduleht.net/templates/keelehooldakeskus/files/mdl_files.php/Mandra_Plado_Vaike_paronuumisonastik.indd.pdf)
- [12] Eesti Keeleressursside Keskus (18.04.2021) <https://keeleressursid.ee/et/keeleressursid>
- [13] Lehekülg [www.cl.ut.ee](http://www.cl.ut.ee) (10.12.2020). [https://www.cl.ut.ee/korpused/segakorpus/eesti\\_ilukirjandus\\_1990/](https://www.cl.ut.ee/korpused/segakorpus/eesti_ilukirjandus_1990/)
- [14] Vallaste e-teatmik (07.05.2021) <http://www.vallaste.ee/>

- [15] Eesti Keeletehnoloogia veebileht (10.12.2020). <https://www.keeletehnoloogia.ee/et/ekt-projektid/estnltk-pythoni-teegid-eestikeelsete-vabatektside-lihtsamaks-tootlemiseks>
- [16] Eesti Teadusinfosüsteem (06.05.2021) <https://www.etis.ee/Portal/Projects/Display/0b4c5efa-6b2c-40cd-aff6-4a6b6668560a>
- [17] EstNLTK (06.05.2021) <https://estnltk.github.io/estnltk/1.4.1/tutorials/index.html>
- [18] Flaski kodulehekülg (16.03.2021) <https://flask.palletsprojects.com/en/1.1.x/foreword/>
- [19] YAML kodulehekülg (07.05.2021) <https://yaml.org/about.html>
- [20] PyYAML dokumentatsioon (17.04.2021) <https://pyyaml.org/wiki/PyYAMLDocumentation>
- [21] Pythoni dokumentatsioon (16.03.2021) <https://docs.python.org/3/library/pickle.html>
- [22] Reacti GitHub (16.02.2021) <https://github.com/facebook/react>
- [23] Material-UI (18.04.2021) <https://github.com/mui-org/material-ui>
- [24] Educative (06.05.2021) <https://www.educative.io/edpresso/what-is-material-ui-in-react>
- [25] CSS-i dokumentatsioon (18.04.2021) <https://developer.mozilla.org/en-US/docs/Web/CSS>
- [26] W3 Schools (07.05.2021)
- [27] Lehekülg [www.cl.ut.ee](http://www.cl.ut.ee) (06.05.2021) <https://www.cl.ut.ee/korpused/segakorpus/index.php?lang=et>

# Lisad

## I. Litsents

### **Lihlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks**

Mina, Helena-Kris Rebane,

1. annan Tartu Ülikoolile tasuta loa (lihlitsentsi) minu loodud teose Mäng paronüümide õige kasutamise õppimiseks, mille juhendaja on Sven Aller, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

*Helena-Kris Rebane*

**10.12.2020**