

III. BENDRINĖS LIETUVIŲ KALBOS FUNKCIONAVIMO VISUOMENĖJE POLINKIAI IR MODERNIŲ SKAITMENINIŲ LEKSIKOGRAFINIŲ SPRENDINIŲ PLĖTRA

1. **Programos vykdytojas** – Lietuvių kalbos institutas.

Etatų skaičius programai – 19,5.

2. **Programos trukmė** – 2022–2026 m.

3. **Programos vadovas** – dr. Jurgita Jaroslaviėnė, el. p. jurgita.jaroslaviene@lki.lt, tel. (8 5) 234 6054.

4. **Preliminarios programos lėšos darbo užmokesėiui** 1 858 568,40 Eur

5. **Programos tikslas** – kompleksiskai tirti bendrinės lietuvių kalbos ir kitų lietuvių kalbos atmainų funkcionavimą visuomenėje, kurti, taikyti ir plėtoti inovatyvias įvairių kalbos lygmenų tyrimo metodologijas ir modernius skaitmeninius leksikografinius sprendinius, įvertinant globalizaciją, lietuvių kalbos plėtros skaitmeninėje terpėje ir kalbos technologijų pažangą bei Europos Sąjungos integracijos procesų nulemtą kitų kalbų poveikį lietuvių kalbai.

6. **Programos uždaviniai:**

6.1. Tirti ir aprašyti dabartinę lietuvių kalbos leksiką (taip pat ir lietuvių terminiją), nustatyti svarbiausius jos funkcionavimo bei raidos polinkius, taikant išmaniąsias kalbos technologijas, ir kurti inovatyvius leksikografinius sprendinius.

Šis uždavinys apims:

6.1.1. Teorinius ir taikomuosius neologijos, bendrinės lietuvių kalbos žodyno ir kitų lietuvių kalbos atmainų leksikografinės medžiagos tyrimus ir modernių skaitmeninių išteklių kūrimą, išryškinant leksinį, struktūrinį, semantinį, sociolingvistinį, etnolingvistinį ir kt. aspektus bei slinktis.

6.1.2. Kompleksinių terminologijos dabarties ir istorijos tyrimų plėtrą, aktualizuojant lietuvių terminų kūrimo, apibrėžimo ir funkcionavimo bei raidos tendencijas.

6.2. Kurti inovatyvias šiuolaikinės bendrinės lietuvių kalbos garsyno tyrimo metodologijas ir tirti variantiškus bei probleminius lietuvių kalbos fonetikos ir fonologijos reiškinius ir procesus, išryškinant garsų akustinių ir artikuliacinių požymių niveliacijos rakursą, bendrinės tarties polinkius ir prielaidas.

6.3. Tirti bendrinės, ypač administracinės, kalbos vartoseną ir jos pokyčius, remiantis sisteminė ortografinių, leksinių ir morfosintaksinių reiškinių stebėseną ir kritine analize, nagrinėti kalbos normas ir jų raidos perspektyvas.

7. **Metodologinis tyrimų pagrindimas:**

1. Pagal 1-ą uždavinį (6.1.1) teoriniai ir taikomieji neologijos tyrimai bus atliekami siekiant nustatyti svarbiausius naujosios lietuvių kalbos leksikos atsiradimo, funkcionavimo vartosenoje ir raidos polinkius. Programoje apimami trys pagrindiniai naujažodžių tyrimo aspektai: struktūrinis, semantinis ir sociolingvistinis. Taikant sinchroninės žodžių darybos, morfeminės ir semantinės analizės ir kt. metodus bei atsižvelgiant į naujausius teorinius užsienio darybininkų darbus, bus nagrinėjamas leksinių ir semantinių naujadarų skirties klausimas, gilinamasi į hibridinės naujažodžių darybos ir morfemikos (tarp jų ir kontaminacijos) problemas. Elektroninio diskurso analizės metodu bus tiriamos kalbos vartotojų nuostatos dėl naujažodžių. Neologijos tyrimai bus atliekami analizuojant naujausius leksikos duomenis, gaunamus derinant pusiau automatinius naujažodžių surinkimo įrankius ir tikslią vartosenos stebėseną. Numatoma remtis nuolat atnaujinamu skaitmeniniu ištekliumi – Lietuvių kalbos naujažodžių duomenynu (<https://ekalba.lt/naujazodziai>), kuriame leksikografiškai aprašyta daugiau kaip 6,5 tūkst. naujažodžių (naujųjų skolinių, vertinių, naujadarų, naujų žodžių reikšmių), taip pat interneto vartosenos duomenimis.

Derinant naujus ir tradicinius metodus (integruojant modernias metodologines inovacijas), bendrinės lietuvių kalbos žodynas bus tiriamas atsižvelgiant į nuolat kintančią kalbos padėtį, naujus visuomenės bei prestižinės kalbos vartojimo poreikius, į naujausius bendrinės kalbos tyrimo, norminimo rezultatus, paskelbtą ar kitaip pareikštą kitų žodynų bei kuriamų šio žodyno tekstų kritiką, etc. Šiuolaikinė leksika ir sistemiškas dabartinės norminės leksikos vaizdas, remiantis tyrimais ir atsižvelgiant į antraštinių žodžių semantinių ryšių analizę, bus fiksuojamas elektroniniame „Bendrinės lietuvių kalbos žodyne“ (<https://ekalba.lt/bendrines-lietuviu-kalbos-zodynas>). Formuojant bei verifikuojant leksinių vienetų semantinę sistemą bus siekiama žiūrėti į reikšmę kaip į tam tikrą atvirą mentalinę struktūrą. Antraštinių žodžių apibrėžtys bus formuluojamos remiantis konceptų analizės

rezultatais (išskiriant kognityvinių kategorijų prototipinius ir periferinius narius bei kt.). Ypatingą dėmesį numatoma skirti duomenų atrankai ir klasifikacijai. Bus atsižvelgiama į tarptautinių žodynų taikomą gerąją praktiką, apibrėžiant žodžius, kur įmanoma, bus artėjama prie kognityvinės definicijos. Iliustracijose fiksuojami svarbesni prototipiniai atributai, pateikiami kuo įvairesni žodžių vartojimo kontekstai, kurie reprezentuoja atitinkamus kognityvinius modelius. Taip pat bus tiriami ir analizuojami antraštyno paradigminiai ir sintagminiai santykiai (pvz., tarpusavyje siejami sinonimai, antonimai, tiriamas ir aprašomas sintaksinis valentingumas ir kt.).

Visapusiškam dabartinės lietuvių kalbos žodyno kontekstui svarbūs šiuolaikinės bei istorinės leksikografinės medžiagos (sisteminių duomenų) tyrimai ir kognityvinės semantikos požiūriu bei derinant kognityvinį ir struktūralistinį požiūrį. Pasirinkti konceptai ir (ar) žodžių grupės semantikos sąsajos su leksiniu junglumu bus tiriami remiantis lietuvių ir kitų kalbų žodžių prasmių tinklais (pvz., <http://wordnetweb.princeton.edu/perl/webwn>, <https://ekalba.lt/zodziu-prasmiu-tinklas/>), tekstynais, aiškinamaisiais žodynais ir kitais šiuolaikiniais šaltiniais. Atliekant tyrimą bus rekonstruojami įvairių kalbos dalių (veiksmažodžių, daiktavardžių ir kt.) semantiniai tinklai, kurie remsis ne tik poliseminiais, bet ir vieno koncepto žodžių tarpusavio semantiniais ryšiais. Tokia tinklų rekonstravimo metodika pritaikoma rekonstruojant kitų konceptų žodžių semantinius tinklus, kuriant tezaurus ir ontologijas. Ši konceptų analizė tiek savo objektu, tiek metodika išplės kognityvinės semantikos tyrimų lauką ir bus aktuali plėtojant esamus ir išvestinius lietuvių kalbos išteklius (pvz., lietuvių kalbos žodžių prasmių tinklą) ir kt. Šiuolaikinės bei istorinės leksikografinės medžiagos tyrimai papildomai bus plėtojami kognityvinės etnolingvistikos aspektu, pasitelkus kognityvinės etnolingvistikos metodologiją. Bus tiriami ir aprašomi kultūriniai konceptai, vertybės ir antivertybės, stereotipai, atliekama semantinė pasirinktų sąvokų analizė, leidžianti kalbėti apie istoriją, visuomeninę patirtį, naujai ją perskaityti ir naujai interpretuoti.

Pakartotinai naudojamų atvirųjų skaitmeninių išteklių pagrindu ir paisant išteklių įvairovės sąsajų bei sąveikumo galimybių bus sukurtas modernus tekstinis skaitmeninis duomenynas, kuris apims du naujus leksikografinius duomenų rinkinius (sinonimų ir antonimų duomenų bazes) ir kelis suskaitmenintus istorinius žodynus. Numatoma, kad duomenynas pagal detaliąją paiešką generuos išsamią informaciją (metaduomenis) apie bet kurį duomenyno žodį: bus galimybė išrinkti visus žodžius pagal kalbos dalį, skiemenų skaičių, kirčiuotę, linksniuotę ir kitas morfologines bei prozodijos kategorijas, taip pat pagal dažnumą ar retumą; bus galimybė pamatyti žodžio paliudijimus istoriniuose žodynuose nuo pačių anksčiausių ir t. t. Toks modernus leksikografinis duomenynas itin aktualus ir naudingas ne tik fundamentaliems moksliniams tyrimams, bet ir siekiant užtikrinti inovatyvių kalbos technologijų, grįstų didžiais tiksliais ir išsamiais kalbos duomenimis, plėtrą ir sklaidą.

1-uoju programos uždaviniu (6.1.2) siekiama užtikrinti ir nuoseklių lietuvių terminijos dabarties bei istorijos tyrimų tęstinumą ir plėtoti bei atnaujinti terminologijos mokslinį diskursą. Tekstynų lingvistikos, kognityvinės terminologijos metodų, kalbos technologijų (*Sketch Engine*, *AntConc* ir kt.) taikymas leidžia kitaip pažiūrėti į terminų kūrimą, vartoseną, terminologijos principus ir tradicinės terminologijos postulatus bei nuostatas. Sistemine terminų analizė ir interpretacija padės atskleisti terminų kūrimo modelius ir tendencijas, motyvacijos schemas. Kognityvinės terminologijos metodais bus siekiama nustatyti konceptus, aktualizuojamus įvardijant sąvokas. Nemažai dėmesio planuojama skirti terminų vartosenai, terminologiniam variavimui – būtent tokie tyrimai užima centrinę padėtį lingvistinėje Europos terminologijoje.

Taip pat numatoma tirti ir terminų apibrėžtis, kurios iki šiol lietuvių kalbotyroje tyrinėtos minimaliai. Apibrėžčių tyrimas turi ne tik mokslinę (klasifikavimas, polinkiai, mokslinės ir teisinės terminų apibrėžties skirtybės, tipinės klaidos), bet ir praktinę vertę – padėtų sričių specialistams, rašantiems terminų apibrėžtis, aiškiai nusakyti apibrėžiamos sąvokos turinį ir atskleisti esminius požymius.

Svarbūs ir istoriniai atskirų sričių terminijos tyrimai, kaip pagrindas numatant dabartinės terminijos plėtrą ir vertinant jos kryptis. Taigi planuojama integruoti tiek dabarties, tiek istorinės terminijos tyrimus, kad galima būtų įvertinti atskirų sričių terminų sistemų ir (ar) mikrosistemų raidą, esamą padėtį ir ateities tendencijas, kreipiant dėmesį ir į terminų funkcionavimą vartosenoje, ir į jų fiksavimą. Moksliniai tyrimai bus atliekami remiantis tokiais svarbiausiais terminų šaltiniais: tekstai (mokslo straipsniai, vadovėliai, teisės aktai, interneto svetainės ir kt.), tekstynai, terminų žodynai ir terminologinių duomenų bazės (LR terminų bankas, IATE ir kt.) ir (ar) ontologijos.

2. Pagal programos 2-ą uždavinį (6.2), derinant tradicinius lietuvių fonologijos mokyklos ir naujus objektyviusius (instrumentinius, statistinius) ir subjektyviusius (audicinius, perceptivinius) metodus, iš įvairių ekstralingvistinių perspektyvų bus tiriami ir aprašomi aktualūs ir probleminiai bendrinės lietuvių kalbos fonetikos ir fonologijos relevantiniai ir nerelevantiniai reiškiniai bei procesai, pagrindinį dėmesį skiriant garsų koartikuliacijai (progresyvinei regresyvinei ir kt.), antrinei balsių ir priebalsių artikuliacijai (apima labializaciją, palatalizaciją, veliarizaciją ir kt.) ir kitiems garsų sąveikos ypatumams. Tyrimai svarbūs fiksuoti ir analizuoti įvairuojančius kirčiavimo ir tarties atvejus, nustatyti lietuvių bendrinės tarties polinkius, išryškinti akustinių ir artikuliacinių požymių niveliacijos prielaidas (pvz., kai tiriamasis segmentas, žodis, tekstas skaitomas ir kai yra tariamas natūraliame kalbos sraute) ir kt.

Šis uždavinys apima ir teorinius lietuvių kalbos garsyno tyrimo klausimus (pvz., fonologijos teorijos ir instrumentinių matavimų suderinamumą), eksperimentinių fonetikos ir fonologijos tyrimų perspektyvos pokyčius (pvz., pritaikius metodologijų inovacijas) ir naujų garsyno tyrimo metodologijų taikymą bei kritinę analizę. Naudojantis 2013–2021 m. sukaupta baltų kalbų garsyno fonoteka eksperimentiniams garsų tyrimams, šiuo uždaviniu taip pat užsibrėžta plėtoti lyginamuosius instrumentinius lietuvių ir latvių kalbų garsyno tyrimus iki šiol netirtais aspektais.

Garsinės kalbos struktūros tyrimai būtini kuriant ir tobulinant gimtosios kalbos technologijas, kurios yra svarbus kalbos visaverčio funkcionavimo garantas skaitmeniniame amžiuje. Aprašant instrumentinių ir perceptivinių bendrinės lietuvių kalbos garsyno tyrimų rezultatus bus atsižvelgiama į galimą daugiadalykį tyrimo pritaikymą ir galimybę (per)panaudoti rezultatus išvestiniams tyrimams (teisinei kalbotyrai, logopedijai, kalbų mokymo(si) ir garsų tyrimo tikslais, etc.) bei kuriant ar tobulinant kalbos sintezavimo, atpažinimo, animavimo, vertimo technologijas ir įrankius.

Instrumentiniai tyrimai bus atliekami kompiuterinėmis garsų analizės programomis (*Praat*, *WaveSurfer* ir kt.). Naujos tiriamosios medžiagos pagrindu bus sukurta garsinė duomenų bazė. Remiantis instrumentinių tyrimų rezultatais bus papildytas TFA atitikmenų bendrinės lietuvių kalbos balsinių ir priebalsinių fonemų alofonams žymėti rinkinys.

3. Pagal 3-ią uždavinį (6.3) bus atliekama sisteminė ortografinių, leksinių ir morfosintaksinių reiškinų stebėseną, kritinė analizė ir interpretacija. Tyrimai bus grindžiami šiuolaikiniais teoriniais ir metodologiniais principais ir užtikrins bendrinės kalbos, ypač jos administracinės atmainos, vartosenos bei normų tyrimų tęstinumą, siekiant atnaujinti ir išryškinti skirtingų kalbos lygmenų pokyčius lyginant su XX a. padėtimi, išsiaiškinti jų sinergiją, įvertinti globalizacijos bei Europos Sąjungos integracijos procesų nulemtą kitų kalbų poveikį lietuvių kalbai. Kalbos normų ir jų raidos tyrimai padės pagrįsti kodifikacijos tikslinimo poreikį įvertinant kalbos norminimo priemonių efektyvumą ir visuomenės nuostatų dėl kalbos tvarkybos pokyčius.

8. Numatomi rezultatai:

- 12 monografijų ir studijų (dalis bus publikuota LKI svetainėje <http://lki.lt/e-publikacijos/>),
- 80 straipsnių recenzuojamuose leidiniuose Lietuvoje ir užsienyje,
- 10 dalių (80 aut. l.) „Bendrinės lietuvių kalbos žodyno“ (bus prieinama išteklių sistemoje „E. kalba“, <https://ekalba.lt/>),
- 15 aut. l. „Lietuvių kalbos naujažodžių duomenyno“ (bus prieinama išteklių sistemoje „E. kalba“, <https://ekalba.lt/>),
- 8 aut. l. elektroninių „Kalbos patarimų“ (bus prieinama išteklių sistemoje „E. kalba“, <https://ekalba.lt/>),
- naujas elektroninis duomenynas, susidedantis iš: Sinonimų DB, Antonimų DB ir jų sąsajos su kitais skaitmeniniais žodynais (bus prieinama LKI svetainėje <http://lki.lt/skaitmeniniai-lietuviu-kalbos-istekliai/>),

Visi rezultatai atsispindės tarptautinių mokslinių konferencijų ir seminarų pranešimuose bei mokslinės sklaidos straipsniuose plačiajai visuomenei.