

Bendrinės lietuvių leksikos duomenų bazė – ne elektroninis Bendrinės lietuvių kalbos žodynas

Daiva Murmulaitytė

Anotacija. Trumpai apžvelgus aplinkybes, kurioms susiklosčius gimė idėja kurti *Bendrinės lietuvių leksikos duomenų bazę*, priminus jos paskirtį ir uždavinius bei patikslinus anksčiau skelbtą *Bazės* duomenų struktūrą, straipsnyje nagrinėjamas jos santykis su šiuo metu baigiamu rengti *Bendrinės lietuvių kalbos žodynu*. Neretai *Bazė* klaidingai laikoma elektroniniu šio žodyno variantu, nors pagal *Bazės* koncepciją, ji su *Žodynu* būtų susijusi tik pradiniais savo duomenimis. Iš tiesų šie dabartinės lietuvių leksikos šaltiniai skirtųsi savo duomenų struktūra, informacijos apimtimi ir pateikimu, paieškos apimtimi ir sudėtingumu, jos rezultatų vaizdavimo būdais ir galimybėmis, požiūriu į savarankiškų leksikografinių vienetų pateikimą ir kitkuo. Straipsnyje keliama mintis siekiant tikslumo pakeisti *Bazės* pavadinimą.

Reikšminiai žodžiai: leksikos duomenų bazė; elektroninis žodynas; žodyninis straipsnis; (savarankiškas) leksikografinis vienetas; semema; rodinys.

Įvadas

Mintis kurti *Bendrinės lietuvių leksikos duomenų bazę* (toliau – *Bazę*) gimė nepavykus suderinti tradicinio (popierinio) *Bendrinės lietuvių kalbos žodyno* (toliau – *Žodyno*) rašymo ir jo kompiuterizavimo. Pradžioje norėta, kad vienu metu būtų parengti tradicinis ir kompiuterinis žodyno variantai, ir kad pastarasis būtų pritaikytas ne tik skaityti jį kompiuterio ekrane, bet ir rimtesniems moksliniams tyrimams – su įvairialypės paieškos galimybėmis (ne vien pagal antraštinį žodį) ir pan. (Mikelionienė, 2003, p.81; Poliakovas, 2003, p.98). Tam reikėjo žodyną iš karto rašyti gerokai struktūriškiau, sistemingiau nei įprasta, nes „net ir tokia iš esmės pernelyg popierinio žodyno vertės nemenkinanti bėda, kaip ne visur vienodas formalus duomenų pateikimas, elektroniniame žodyne sumažina automatinės paieškos galimybes“ (Jakaitienė, 2005, p.296). Siūlymo pirma sukurti duomenų bazę, o ją naudojantis rašyti žodyną irgi buvo atsisakyta, nes atrodė, kad tai užtruks, o *Žodynas* buvo reikalingas, kaip sakoma, „jau vakar“. Taigi pagal modernesnę metodiką buvo rašomas tradicinis žodynas, o jo kompiuterizavimas paliktas ateičiai. Taip susiklosčius aplinkybėms pagalvota apie originalios struktūros duomenų bazę, kuri būtų kuriama neprisirišant prie šio *Žodyno* ar kokio kito leksikografijos šaltinio, o su *Žodynu* būtų susijusi tik tiek, kad iš jo būtų įvedami pirminiai šios bazės duomenys. Vėliau bazė būtų pildoma duomenimis iš kitų leksikografinių, tradicinės ir elektroninės žiniasklaidos šaltinių, fiksuojant vartoseną internete ir pan. Tokios bazės koncepcija buvo sukurta ir 2006 m. pristatyta Leksikografijos centro seminare, Lietuvių kalbos institute. Įvairiose konferencijose ir publikacijose (Murmulaitytė, 1999, 2002, 2006a, 2006b, 2007a, 2007b, 2008, 2009) kalbėta ir apie jos duomenų struktūrą, kūrimo eigą, nuveiktus ir ateities darbus bei problemas, išskylančias juos dirbant. 2008 m. LKI Leksikografijos centre organizuotas mokslinis seminaras „*Bendrinės lietuvių kalbos leksikos duomenų bazė* – nuveikti ir ateities darbai“, kuriame perskaitytas pranešimas „*Bendrinės lietuvių kalbos leksikos duomenų bazės* kūrimo eiga, tarpiniai rezultatai ir problemos“. Seminare įsibėgėjančius programavimo darbus pristatė *Bazę*

kartu kuriančių Matematikos ir informatikos instituto programuotojų vadovė Snieguolė Meškauskienė, taip pat dalyvavo programuotojai Antanas Grincevičius ir Remigijus Andžius.

Ilgainiui įsitikinta, kad kalbininkų visuomenė, ypač nesusipažinusi su *Bazės* koncepcija, dažnai tapatina ją su *Žodynu*, tikisi ją būsiant *Žodyno* rengimą spartinančia priemone ar pan. Taigi šio straipsnio tikslas – priminti kai kuriuos pamatinius, ankstesnėse autorės publikacijose aptartus *Bazės* koncepcijos dalykus, patikslinti tai, kas dirbant buvo koreguota, ir tiksliau nurodyti šių dviejų lietuvių leksikos šaltinių riboženklis bei sąlyčio taškus. Straipsnis parengtas pranešimo, 2009 m. balandžio 3 d. skaityto Vytauto Didžiojo universitete vykusioje konferencijoje-seminare „Lietuvių kalba ir naujosios technologijos“, pagrindu.

Bazės paskirtis ir duomenų struktūra

Primintina, kad *Bazę* sumanyta siekiant:

- kaupti dabartinės lietuvių kalbos leksiką;
- preliminariai ją apdoroti leksikografiškai;
- moksliskai tirti sukauptus duomenis;
- nuolat juos tikslinti, redaguoti, atnaujinti;
- fiksuoti probleminius atvejus, ppr. nepatenkančius į žodynus;
- jos duomenų pagrindu rengti įvairius žodynus, mokslo leidinius.

Bazėje būtų kaupiami keturių rūšių leksikografiniai vienetai:

- žodžiai;
- sudėtiniai terminai ir sudėtiniai pavadinimai (pvz., *namų šeimininkė, greitoji pagalba, Grįžulo ratai, ekonomikos perkaitimas, Jo Ekscelencija* ir kt.);
- frazeologizmai;

- žodžių dalys (turima galvoje tarptautinės žodžių dalys – t. y. pradiniai tarptautinių žodžių ir hibridų dėmenys bei priešdėliai (pvz., *avia-*, *bio-*, *pseudo-*, *anti-* ir kt.), taip pat lietuvių kalbos priešdėliai, aptariamajame ir kituose žodynuose pateikiami atskirais straipsniais, aiškinami ir iliustruojami. Ilgainiui *Bazėje* analogiškai galėtų būti pateikiami ir tarptautinių žodžių baigmenys (pvz., *-logija*, *-grafija*, *-skopas* ir kt.), taip pat lietuvių kalbos morfemos (priesagos ar baigmenys).

Bazė įsivaizduojama kaip išsami, nuolat pildoma kartoteka, kurią sudaro tokių leksikografinių vienetų sememų kortelės, t. y. vienoje kortelėje būtų kaupiami vieno leksikografinio vieneto kuria nors viena jo reikšmė duomenys. Kiekviena kortelė, savo ruožtu, būtų sudaryta iš 85 laukų, kuriuose būtų įrašoma ir kaupiama visa žodyniniame straipsniui būdinga informacija bei papildomi, redagavimui ar tyrinėjimui svarbūs duomenys (žr. 1-ąją lentelę; tai dirbant pataisytas *Bazės* duomenų struktūros variantas, todėl lentelė kiek skiriasi nuo anksčiau skelbtų pradinių *Bazės* duomenų struktūros lentelių):

1 lentelė. BLDB duomenų struktūra

1. Antraštė:
1.1. leksikografinis vienetas;
1.2. homonimo/ homografo ženklas.
2. Paastraštė:
2.1. leksikografinis vienetas;
2.2. homonimo/ homografo ženklas.
3. Pagrindinės formos:
3.1. II pagr. forma;
3.2. III pagr. forma.
4. Paradigma:
4.1. Linksniavimas:
4.1.1. linksniuotė;
4.1.2. paradigma (nr.);
4.2. Asmenavimas:
4.2.1. asmenuotė (nr.);
5. Kirčiavimas:
5.1. kirčiuotė (nr.);
5.2. kirčiuotos formos;
5.3. indeksai.
6. Sememos eilės numeris.
7. Semantinių grupių pažymos.
8. Gramatinės pažymos:
8.1. kalbos dalis;
8.2. giminė;
8.3. skaičius;
8.4. linksnis;
8.5. nuosaka;
8.6. laikas;
8.7. asmuo;
8.8. rūšis;
8.9. laipsnis;
8.10. veiksmazodžio forma;
8.11. kitkas.
9. Sintaksinės pažymos:

9.1. junglumas;
9.2. kitkas.
10. Leksinės pažymos.
11. Vartosenos ypatumų pažymos.
12. Vartojimo srities pažymos.
13. Stilistinės pažymos.
14. Tarminės pažymos.
15. Daryba:
15.1. pamatinis žodis;
15.2. pamat. žodžio homonimo ženklas;
15.3. pamat. žodžio sememos numeris;
15.4. darybos kategorija.
16. Kilmė:
16.1. etimologija;
16.2. etimono kalbos pažyma;
16.3. indeksai.
17. Aprašomasis reikšmės aiškinimas:
17.1. apibrėžtis
17.2. vartosenos nusakymas;
17.2.1. bendroji dalis;
17.2.2. atskiroji dalis;
17.3. lotyniškas pavadinimas;
17.4. komentarai.
18. Norminamosios pažymos.
19. Nuorodos:
19.1. variantas:
19.1.1. vertinimo laipsnio ženklas,
19.1.2. leksikografinis vienetas;
19.1.3. homonimo ženklas;
19.1.4. sememos numeris;
19.1.5. indeksai.
19.2. žr.
19.2.1. leksikografinis vienetas;
19.2.2. homonimo ženklas;
19.2.3. sememos numeris;
19.2.4. indeksai.
19.3. kitkas.
20. Sinonimas:
20.1. leksikografinis vienetas;
20.2. homonimo ženklas;
20.3. sememos numeris;
20.4. komentarai
20.5. indeksai.
21. Antonimas:
21.1. leksikografinis vienetas;
21.2. homonimo ženklas;
21.3. sememos numeris;
21.4. komentarai
21.5. indeksai.
22. Susijusi sąvoka (plg.):
22.1. leksikografinis vienetas;
22.2. homonimo ženklas;
22.3. sememos numeris.
23. Iliustracija:

23.1. vartosenos pavyzdys;
23.2. paaiškinimas;
23.3. vartojimo srities pažymos;
23.4. stilistinės pažymos;
23.5. šaltinio santrumpa (klasifikuota);
23.6. komentarai;
23.7. indeksai.
24. Vartojimo dažnis:
24.1. rodiklis (skaičius);
24.2. šaltinio santrumpa;
24.3. indeksai.
25. Indeksai.
26. Žodyninė sąsaja:
26.1. antraštinis leksikografinis vienetas;
26.2. homonimo/ homografo ženklas;
26.3. sememos numeris.
26.4. komentarai.
27. Kortelės šaltinis.
28. Autorinė redaktoriaus žyma.

Bazė iš esmės būtų skirta moksliniam (ypač leksikografiniam) darbui. Plačioji visuomenė galėtų naudotis darbo, atlikto naudojantis ja, rezultatais, kurie, tikėtina, greičiau atsirastų ir būtų tobulesni, turint tokią darbo priemonę. Ateityje, įgijus didesnės darbo su šia baze patirties ir sukaupus tikslesnius duomenis, būtų galima pagalvoti ir apie internetinį jos variantą, turintį siauresnes paieškos galimybes, pateikiantį tik būtiniausią informaciją peržiūrint duomenų paieškos rezultatus, nesuteikiantį teisės ir galimybių redaguoti, pildyti ar kitaip keisti tą informaciją pačioje bazėje.

Bazės santykis su Žodynu

Neretai klaidingai įsivaizduojama, kad *Bazė* kuriama kaip elektroninis *Žodyno* variantas.

Pradiniai *Bazės* duomenys, kaip jau minėta, ištis būtų įvedami iš pirmojo *Žodyno* leidimo, pabaigus jį rengti ir pagal tam tikrą metodiką struktūriškai šiek tiek pertvarkius žodyninius straipsnius. Tačiau *Bazė* nebūtų elektroninis šio žodyno variantas. Pagrindiniai šių dviejų dabartinės lietuvių kalbos leksikos rinkinių skirtumai (turint galvoje, kad *Žodynas* taip pat turėtų būti išleistas elektronine forma, antraip lyginimas apskritai vargiai įmanomas) būtų šie:

1. *Bazėje* apskritai būtų gerokai daugiau informacijos, nei jos katinama pateikti *Žodyne*. Tarkim, joje numatyta:

- ne tik pateikti žymiai daugiau tarptautinių žodžių, bet ir nurodyti jų etimologijas;
- pažymėti semantines leksikos grupes – tai ypač aktualu redaguojant daiktavardžius, nes šiuo požiūriu jie *Žodyne* niekaip nepažymėti (išskyrus deminutyvus);
- pateikti gerokai daugiau naudingos gramatinės informacijos (pvz., nurodyti kiekvieno kaitomo žodžio linksniuotę ar asmenuotę – tokia informacija įprastinė mokomiesiems nei bendriesiems žodynams);

- diferencijuoti ir indeksuoti leksikografinius vienetus iliustruojančius vartosenos pavyzdžius (atskirti žodžių junginius ir sakinius);
- leksikografinius vienetus žymėti norminamuoju aspektu;
- iš įvairių šaltinių pateikti leksikografinių vienetų vartojimo dažnio rodiklius;
- įvairius semantinius, gramatinius, vartosenos ir kitokius leksikos tyrėją ar kitą *Bazės* naudotoją dominančius dalykus žymėti atitinkamais indeksais, pagal kuriuos jis galėtų klasifikuoti *Bazės* duomenis, formuluodamas paieškos užklausą pagal savo poreikius.

2. *Bazėje* būtų kaupiama ne tik bendrinės lietuvių kalbos (norminė) leksika ir lygiaverčiai jos variantai, kaip kad yra *Žodyno* atveju, bet ir šalutiniai normos variantai, žargonas, neteiktini dalykai ir kt. Visa tai būtų atitinkamai pažymima, nustačius tokio reiškinio vietą bendroje leksikografinio vieneto struktūroje ir įvertinus jį vartosenos teiktumo požiūriu. Turint galvoje šį leksikos norminimo aspektą teisybės dėlei reiktų pasakyti, kad aptariamasis bazės pavadinimas – *Bendrinės lietuvių leksikos duomenų bazė* – nėra tikslus. *Bazės* ištakos, kaip minėta, susijusios su *Bendrinės lietuvių kalbos žodyno* kompiuterizavimu, kompiuteriniu, kaip tada vadinta, jo variantu, tad natūralu, kad tuo metu kuriamas dalykas buvo atitinkamai pavadinamas. Bet vėliau nuo šio žodyno atitolus struktūros, paskirties ir kitomis prasmėmis bei nutarus bazėje kaupti ir nenorminę leksiką, toks iš senų laikų užsilikęs pavadinimas akivaizdžiai nebeatitinka iš principo naujos bazės turinio.

3. Iš esmės skirtųsi *Žodyno* ir *Bazės* duomenų struktūros. Pagrindinis *Bazės* struktūrinis vienetas yra semema, t. y. viena leksikografinio vieneto reikšmė. Visi jos duomenys rašomi į vieną kortelę, išskyrus:

- keleriopo kirčiavimo atvejus, kai kiekvienam kirčiavimo variantui pildoma atskira kortelė, ir
- antraštines giminėmis kaitomų žodžių formas, kurioms atskiros kortelės pildomos dėl skirtingų gramatinių ir vartosenos duomenų (ateityje tai turėtų palengvinti jų vartosenos tyrimą, taip pat galėtų labai pasitarnauti kuriant morfologinių formų žodyną).

Pagrindiniu *Žodyno* vienetu tradiciškai laikytinas antraštinis žodis, t. y. visos jo sememos (reikšmės), kartu su kita būtina informacija pateikiamos žodyniniame straipsnyje. Tarkim, *Žodyno* straipsnis *racionalūs* (pavyzdį žr. toliau; tai dar nėra galutinė jo redakcija) *Bazėje* būtų išskaidytas į 6 korteles, nes negimininė forma ir frazeologizmas laikytini atskirais leksikografiniais vienetais; be to, kadangi gramatinė ir kt. moteriškosios ir vyriškosios būdvardžio giminų formų informacija (šiuo atveju – linksniuotė, paradigma) skiriasi, šios formos irgi būtų pateikiamos atskiromis kortelėmis (*Žodyne* linksniuotės nežymimos, bet *Bazėje* numatyta jas žymėti ir tai būtų daroma vėliau, jau pačioje *Bazėje* tvarkant duomenis). Šiuo atveju vyriškosios giminės būdvardžio forma *racionalūs* yra (i)u linksniuotės ir linksniuojama pagal 5-ąją paradigmą, o i linksniuotės moteriškosios giminės forma *racionali* – pagal 8-ąją paradigmą, taigi analogiškuose skirtingų kortelių laukuose, skirtuose šios rūšies gramatinei informacijai žymėti, turėtų

būti įrašyti skirtingi duomenys. Taigi *Bazėje* būtų tokios 6 žodžio *racionalūs* kortelės (pilka spalva pavyzdyje pateikiami įvairūs „nematomi“ indeksai ir kt. duomenys; jie reikalingi duomenų paieškai, bet nebūtinai (specialiai pasirinkus) rodomi pateikiant jos rezultatus):

Žodynas	Bazė
racionalūs, racional bdv. (4) 1. protingas, tikslingas, tinkamas, pagrįstas: <i>Racionalūs póžiūris.</i> <i>Racionalūs šteklių naudójimas. Racional</i> mityba. racionalù bev.: <i>Ta ganà racionalù.</i> 2. mat. bendramatis su sveikuoju skaičiumi, vienetu: <i>Racionalùsis dýdis.</i> <i>Racionalióji fúncija.</i> <i>Racionalùsis skačius</i> (sveikasis teigiamas skaičius). racionalùs gr das tai, kas atitinka tikrovę, tiesa, teisybė: <i>Šiamè teiginyjè yrà racionalaùs gr do.</i>	racionalus racionalùs bdv. v. ((i)u/5) (4) 1. protingas, tikslingas, tinkamas, pagrįstas • <i>racionalùs póžiūris žj</i> • <i>racionalùs šteklių naudójimas žj</i> racionalus racionalùs bdv. v. ((i)u/5) (4) 2. mat. bendramatis su sveikuoju skaičiumi, vienetu: • <i>racionalùsis dýdis žj Δ?</i> • <i>racionalùsis skačius</i> (sveikasis teigiamas skaičius) <i>žj Δ?</i> racionalus racional bdv. m. (i/8) (4) 1. protinga, tikslinga, tinkama, pagrįsta • <i>racional mityba žj</i> racionalus racional bdv. m. (i/8) (4) 2. mat. bendramatė su sveikuoju skaičiumi, vienetu • <i>racionalióji fúncija Δ?</i> <i>apbr?</i> racionalus racionalù bdv. bev. • <i>Ta ganà racionalù.</i> ↑ racionalus 1 racionalus grūdas racionalùs gr das tai, kas atitinka tikrovę, tiesa, teisybė • <i>Šiamè teiginyjè yrà racionalaùs gr do.</i> ↑ racionalus BŽ

4. *Bazės* duomenų paieškos rezultatus numatoma vaizduoti pasirenkant vieną iš keturių pagrindinių, o kirčiavimo variantų turinčių gimininių žodžių atveju – iš dar dviejų papildomų vaizdavimo būdų. Taigi iš viso numatyti 6 rodinių tipai.

4.1. *Kortelė*. *Bazės* naudotojui būtų rodoma vienos sememos kortelė su visais duomenimis, įrašytais jiems skirtuose laukuose (žr. 1 pav.). Šis būdas nuo kitų penkių skiriasi tuo, kad naudotojui ne tik pateikiami patys duomenys, bet akivaizdžiai parodoma ir jų vieta bazės duomenų struktūroje. Toks struktūriškas vaizdavimas daugiausia būtų reikalingas įvedant naujus bei tikrinant ir koreguojant jau įvestus duomenis. Toliau pateikiamuose pavyzdžiuose matyti, kaip vienas *Žodyne* pateikiamas straipsnis (būdvardis *adat* *nis*, *adat* *nė*) prieš perkeltas duomenis į *Bazę*

būtų išskaidomas į sememas, taip pat vaizduojamas vienos iš jų (*adat* *nė* 2) kortelės rodinys:

Žodynas	Bazė
adat <i>nis</i> , adat <i>nė</i> bdv. (2) 1. susijęs su adatomis (1 r.): <i>Adat nė pagalv lė</i> (adatoms susmeigti). 2. spec. susijęs su adatomis (2 r.), adatėlėmis: <i>Adat nis spausdintūvas</i> (komp.).	adatinis adat <i>nis</i> bdv. v. nlpsn. ((i)a/4) (2) 1. susijęs su adatomis (1 r.) il adatinis adat <i>nis</i> bdv. v. nlpsn. ((i)a/4) (2) 2. spec. susijęs su adatomis (2 r.), adatėlėmis: <i>adat nis spausdintūvas žj</i> (komp.). adatinis adat <i>nė</i> bdv. m. nlpsn. (è/9) (2) 1. susijusi su adatomis (1 r.): <i>adat nė pagalv lė žj</i> (adatoms susmeigti) adatinis adat <i>nė</i> bdv. m. nlpsn. (è/9) (2) 2. spec. susijusi su adatomis (2 r.), adatėlėmis il

4.2. *Semema*. Vienos sememos duomenys (viena paaiškinta, gramatiškai ir kitaip apiforminta bei iliustruota leksikografinio vieneto reikšmė ir kita informacija iš vienos kortelės) būtų pateikiami žodyninio straipsnio ar panašiu pavidalu, t. y. tam tikra tvarka išdėstant juos paileiui.

Sememos rodinių pavyzdžiai:

įžvalgus įžvalgùs bdv. v. ((i)u/5) (4) 1. pasižymintis įžvalga • <i>įžvalgùs pol tikas žj</i> įžvalgus įžvalgùs bdv. v. ((i)u/5) (4) 2. numatantis sin. apdairus • <i>įžvalgùs pirk jas žj</i> įžvalgus įžvalg bdv. m. (i/8) (4) 1. pasižyminti įžvalga įžvalgus įžvalg bdv. m. (i/8) (4) 2. numatanti sin. apdairi	gerbiamasis gerbiamàsis įvr. bdv. v. (K: <i>ge biamąj</i>) 1. pagarbos žodis, vartojamas kreipiantis • <i>Gerbiám eji, nóriu jùms pranèsti gèrą ž niq.</i> • <i>Gerbiámàsis, gál gal tumètè atidaryti dur s?</i> gerbiamasis gerbiámóji įvr. bdv. m. (K: <i>ge biamąj</i>) 1. pagarbos žodis, vartojamas kreipiantis • <i>Gerbiámóji, J s manè ne ta p suprātote.</i>
--	---

4.3. *Postraipsnis*. Visų sememų, turinčių vienodą paant-raštę, duomenys, pateikiami paileiui. Čia plačiau paaiškin-tinas ANTRAŠTĖS (pavyzdžiuose ji pateikiama pusjuo-džiu šriftu, pabraukta) ir PAANTRAŠTĖS (pavyzdžiuose šie duomenys rašomi pusjuodžiu šriftu) skyrimo reikalas.

ANTRAŠTĖS laukai *Bazėje* būtų skirti tiems atvejams, kai tas pats leksikografinis vienetas turi kelias antraštines for-mas – pvz., būdvardžių, įvardžių, skaitvardžių, dalyvių vyriškosios ir moteriškosios gimininių bei negimininės for-mos, kirčiavimo variantai ir pan. *Bazėje* jos būtų pateikia-

mos atskirai – t. y. vyriškosios giminės formai būtų pildoma viena kortelė, kurios PAANTRAŠTĖS 2.1. lauke būtų rašoma vyriškosios giminės forma, o moteriškosios giminės – kita kortelė, kurios PAANTRAŠTĖS 2.1. lauke atitinkamai būtų rašoma moteriškosios giminės forma. Abiejų šių kortelių ANTRAŠTĖS 1.1. lauke būtų patei-

kiamas nežymėtasis šių formų grupės narys – vyriškosios giminės forma, kad automatiškai formuojant straipsnį, abiejų formų informacija būtų surinkta ir pateikiama kartu. Kirčiavimo variantų atveju ANTRAŠTĖS lauke būtų rašoma nekirčiuota leksikografinio vieneto forma.

1 pav. Kortelės rodinio pavyzdys

Taigi pasirinkus aptariamąjį, postraipsnio, vaizdavimo būdą, būtų rodomi tik vienos iš vardažodžių giminių (vyriškosios arba moteriškosios) ar negimininės formos, vieno

kirčiavimo varianto ir pan. duomenys, kadangi kitų giminių formų, kirčiavimo variantų ir pan. duomenys būtų kaupiami kitose kortelėse ir turėtų kitą paantraštę. Kitaip sa-

kant, postraipsnio rodinyje būtų pateikiama visų kortelių, turinčių vienodą paantraštę, duomenys.

Postrapsnių pavyzdžiai:

ižvalgus

ižvalgūs bdv. **v.** ((i)u/5) (4)

1. pasižymintis įžvalga

• ižvalgūs poltikas žj

2. numatantis

sin. apdairus

• ižvalgūs pirkjas žj

ižvalgus

ižvalg bdv. **m.** (i/8) (4)

1. pasižyminti įžvalga

2. numatanti

sin. apdairi

4.4. Straipsnis. Naudotojui būtų rodomi visi vieno leksikografinio vieneto, tradiciniame žodyne sudarančio straipsnio antraštę (t. y. visų jo sememų), duomenys, pateikiami paeiliui. Šiuo būdu vaizduojant paieškos rezultatus į vieną žodyninį straipsnį būtų jungiami vardažodžių abiejų gimininių formų bei negimininės formos, visų kirčiavimo variantų duomenys ir kita informacija, kuri tradiciniuose žodynuose pateikiama vienu straipsniu, turinčiu kelias antrašines formas, o *Bazėje* dėl jau minėtų priežasčių būtų išskaidyta į atskiras korteles. Kitaip tariant, straipsnio rodinyje būtų pateikiami visų kortelių, turinčių vienodą antraštę, duomenys. Moteriškosios giminės formai sukurta apibrėžtis ir kt. šiai formai pritaikyti dalykai, matomi sememos ir postraipsnio rodiniuose, čia būtų ignoruojami, taigi vaizdas prilygtų įprastam žodyniniui.

Straipsnių pavyzdžiai:

ižvalgus

ižvalgūs, ižvalg bdv. ((i)u, i/ 5, 8) (4)

1. pasižymintis įžvalga

• ižvalgūs poltikas žj

2. numatantis

sin. apdairus

• ižvalgūs pirkjas žj

gerbiamasis

gerbiamasis, gerbiamoji įvt.

bdv. (K: ge biamąjį,

ge biamąją)

1. pagarbos žodis, vartojamas

kreipiantis

• Gerbiamoji, nōriu jūms

pranēšti gēra žōniq. •

Gerbiamasis, gāl

gal tumēte atidaryti

duris? • Gerbiamoji,

Jūs manē ne tap

suprātote.

4.5. morfologinis straipsnelis ir 4.6. kirčiavimo straipsnelis – minėtieji du papildomi rodinių tipai. Jų prireiktų vaizduojant kirčiavimo variantų turinčius gimininius žodžius – būdvardžius, dalyvius, kai kuriuos skaitvardžius ir kai kuriuos įvardžius.

Jei būdvardis labiau rūpėtų morfologiniu aspektu, būtų galima rinktis morfologinio straipsnelio rodinį ir vienoje vietoje matyti atitinkamai sugrupuotus abiejų būdvardžio gimininių formų bei negimininės formos duomenis. Atkreiptinas dėmesys, kad jie būtų pateikiami laikantis tokios pačios metodikos, kaip ir straipsnio rodinyje – apibrėžtis, sinonimai, antonimai ir pan. duomenys būtų imami tik iš vyriškosios giminės formos kortelių, o atitinkami moteriškosios giminės ir negimininės formų duomenys morfologiniame straipsnelyje, kaip ir straipsnyje, būtų ignoruojami. Viena morfologiniame straipsnelyje būtų rodomi vieno kirčiavimo varianto duomenys, taigi morfologinių straipsnelių būtų tiek, kiek būtų kirčiavimo variantų. Tuo tarpu tiriant kirčiavimo dalykus, matyt, patogiau būtų rinktis kirčiavimo straipsnelio rodinį, kuriame vienoje vietoje būtų pateikiama informacija apie visus vienos kurios morfolo-

ginės formos (giminės) kirčiavimo variantus. Šiuo atveju apibrėžtys, sinonimai, antonimai būtų pritaikyti kiekvienai konkrečiai formai, t. y. moteriškosios giminės formos būtų atitinkamai apibrėžiamos ir turėtų moteriškosios giminės sinonimus bei antonimus, analogiškai būtų pateikiamos ir negimininės formos duomenys – plg. toliau pateikiamus pavyzdžius.

Morfologinių straipsnelių pavyzdžiai:

raudonas

raudonas, raudona lpsn.

kokyb. bdv. ((ia)/ 1,

(i)o/ 6) (1 kv1:

raudonam, raudonu,

raudoname, raudoni,

raudonų, raudonais,

raudonuose; raudonos,

raudonų, raudonomis,

raudonose)

1. kuris kraujo spalvos

• raudoni dažai žj

• raudonas vėidas žj

• raudonos lpos

[aguonos] žj

• raudonos g lės žj

2. susijęs su revoliuciniu

komunistiniu veikimu

raudonas

raudonas, raudonà lpsn.

kokyb. bdv. ((ia)/1,

(i)o/6) (3 kv2:

raudonam, raudonamē,

raudon, raudon,

raudona s,

raudonuosē; raudonōs,

raudonojē, raudon,

raudonom s,

raudonosē)

1. kuris kraujo spalvos

2. susijęs su revoliuciniu

komunistiniu veikimu

Kirčiavimo straipsnelių pavyzdžiai:

raudonas

raudonas lpsn. kokyb. bdv. **v.**

((ia)/1) (1 kv1: raudonam,

raudonu, raudoname,

raudoni, raudonų,

raudonais, raudonuose),

raudonas (3 kv2:

raudonam, raudonamē,

raudon, raudon,

raudona s, raudonuosē)

1. kuris kraujo spalvos

• raudoni dažai žj

• raudonas vėidas žj

2. susijęs su revoliuciniu

komunistiniu veikimu

raudonas

raudonà lpsn. kokyb. bdv. **m.**

((i)o/6) (1 kv1: raudonos,

raudonų, raudonomis,

raudonose), **raudonà** (3

kv2 raudonōs, raudonojē,

raudon, raudonom s,

raudonosē)

1. kuri kraujo spalvos

• raudonos lpos

[aguonos] žj • raudonos

g lės žj

2. susijusi su revoliuciniu

komunistiniu veikimu

5. Kuriant *Bazės* koncepciją galvota apie įvairialypę duomenų paiešką. Remtasi geru *Dabartinės lietuvių kalbos žodyno* 3-ojo leidimo kompiuterinio varianto (toliau – DŽ(k)) pavyzdžiu, atsižvelgiant į kai kuriuos tobulintinus dalykus ir numatant papildomų, kuriuos sąlygotų kitokia, gerokai sudėtingesnė *Bazės* struktūra. Pavyzdžiui, numatyta paieška pagal vieną ar kelis požymius, įvairius jų derinius, požymius jungiant keliomis loginėmis jungtimis (IR, ARBA), taip pat paieška pagal nesantį duomenį; numatytas neribotas paieškos kartojimas ankstesnės paieškos rezultatuose (DŽ(k) paiešką kartoti buvo galima tik vieną kartą) ir t. t. Tokia įvairialypė paieška įmanoma tik gerokai sistemiškiau, nei tai daroma *Žodyne*, kaupiant ir tvarkant šio žodyno ir kt. šaltinių duomenis. Taigi ir šiuo, duomenų paieškos, aspektu klaidinga būtų tapatinti aptariamąją *Bazę* su elektroniniu *Žodynu*. Apskritai kol kas sudėtinga net lyginti *Bazės* ir elektroninio *Žodyno* duomenų paieškas, nes pastarosios klausimas šiuo metu, kiek žinoma, dar nėra svarstomas.

6. Dar vienas svarbus dalykas – savarankiškų leksikografinių vienetų pateikimas kitų, pamatinių ar pagrindinių, leksikografinių vienetų straipsniuose (Svetikienė, 2006; Šepetytė, 2006). Palyginus su pateikimu *Dabartinės lietuvių kalbos žodyne* pasakytina, kad *Bendrinės lietuvių kalbos žodyne* iš pamatinių žodžių straipsnių į atskirus straipsnius bus iškelti:

- veiksmazodžių ir vardažodžių abstraktai su priesagomis *-imas/-ymas, -umas, -ybė*;
- veikėjų pavadinimai su priesagomis *-tojas, -a, -ėjas, -a, -ikas, -ė*;
- rieveiksmiai su priesagomis *-(i)ai, -yn*;
- taip pat sudėtiniai pavadinimai, *Dabartinės lietuvių kalbos žodyne* buvę pateikti po trikampio ženklo, ir analogiški nauji sudėtiniai terminai ir pavadinimai.

Tačiau kai kurie savarankiški leksikografiniai vienetai aptariamajame *Žodyne* dar lieka susieti su pagrindiniais ar pamatiniais žodžiais ir pateikiami jų straipsniuose. Tai:

- sangražiniai veiksmazodžiai;
- frazeologizmai;
- deminutyvai;
- veiksmazodžiai su neiginiais (*ne-, nebe-*);
- dalis sudėtinių teminių ir pavadinimų (tie, kurie nėra aiškinami).

Kai kurie jų (tarkim, sangražiniai veiksmazodžiai) formaliai pateikiami kaip pamatinių ar pagrindinių žodžių (šiuo atveju – nesangražinių veiksmazodžių) atskiros reikšmės ar jų atspalviai; frazeologizmai dedami pagrindinio žodžio straipsnio gale po rombo ženklo; kiti minėti dalykai kol kas buvo pateikiami kaip pamatinio žodžio iliustracijos. Pažymėtina, kad šiuo metu dar rašoma *Žodyno* redagavimo instrukcija, taigi kai kurie dalykai dar gali keistis. Vis dėlto akivaizdu, kad *Žodyne* nebus laikomasi principo kiekvienam savarankiškam leksikografiniam vienetai skirti atskirą žodyninį straipsnį. Atsižvelgiant į tai *Žodyno* duomenys, prieš srautiniu būdu įvedant juos į *Bazę*, turėtų būti struktūriškai pertvarkyti, t. y. pritaikyti jos struktūrai, kurioje atskiriems leksikografiniams vienetais skirtos atskiros kortelės, o sąsajas su pagrindiniais ar pamatiniais žodžiais, kurių straipsniuose jie buvo *Žodyne*, numatyta pažymėti tam tikslui skirtuose atitinkamuose *Bazės* laukuose. Tarkim, įprasta manyti, kad lizdais pateikiami frazeologizmai lengviau randami *žodyne* (antra vertus, neretai jų eilės tvarkos taisyklės gana painios ir jei lizdas milžiniškas, orientuotis jame dar sudėtingiau, nei skaityti viską pagal abėcėlę). Taigi dėl šios priežasties *Bazėje* paliekama galimybė grupuoti frazeologizmus pagal pagrindinį žodį, kurio straipsnyje jie bus įrašyti *Žodyne* (tam skirta *Bazės* laukų grupė ŽODYNINĖ ŠĄSAJA). Tas pats pasakytina ir apie kitus leksikografinius vienetus, kurie *Žodyne* atskirų straipsnių neturės.

7. Gerai parengto tradicinio žodyno duomenys neturėtų kelti abejonių. Todėl kai kurie mokslinė prasme diskutuotini, įvairiai interpretuojami kalbotyros dalykai, siekiant neklaidinti eilinio naudotojo, kartais iš viso nepateikiami

arba, jei be jų išsiversti neįmanoma, pateikiami taip, kad labai nekristų į akis – tarp kitų iliustracinių pavyzdžių ir pan. Tarkim, tam tikru metu ima sparčiai kisti kokios nors srities žodynas (kaip kad atsitiko su kompiuterijos leksika), randasi naujų žodžių ir jų junginių, įvairūs geresni ar prastesni naujadarai ieško savo vietos, kovodami su svetimybėmis ir pan., todėl tokią žodyno sritį tuo metu itin sunku tinkamai parašyti ir suredaguoti. Kol knyga išeina iš spaudos, dalis terminų jau gali būti pasenę arba taip ir neprigiję. Tad mažiausiai rizikinga būtų apskritai dėti kuo mažiau šios srities leksikos. Naudotojas nebus patenkintas, neradęs vieno ar kito jam rūpimo žodžio ir informacijos apie jį, bet nebus ir suklaidintas (ar bent jau bus mažiau suklaidintas, nei gavęs apie rūpimą žodį netikslią informaciją). Tokia praktika, kad abejotini, nenusistovėję dalykai tradiciniuose žodynuose vis dėlto būtų pateikiami, atitinkamai pažymint jų problemškumą, nėra įprasta, bent jau lietuvių leksikografijoje. Tuo tarpu *Bazėje* atitinkamai pažymėti rastų savo vietą ir probleminiai, dar nenusistovėję, diskutuojami atvejai.

Išvados

Taigi aptartieji *Bendrinės lietuvių leksikos duomenų bazės* ir *Bendrinės lietuvių kalbos žodyno* struktūros, informacijos pateikimo ir kt. dalykai akivaizdžiai parodo, kad šių dviejų leksikografijos šaltinių jokių būdu nereikėtų tapatinti. Juos siėtų tik bendri duomenys, ir tai tik pradiniu *Bazės* kūrimo etapu, nes vėliau ji būtų pildoma kitų leksikografijos šaltinių (pradedant, tarkim, tarptautinių žodžių žodynu) duomenimis. Kiti *Žodyno* leidimai, be abejonės, galėtų būti kokybiškesni ir greičiau parengiami, dirbant su šia *Baze*, bet į juos tikrai patektų ne visa tai, ką tik *Bazė* būtų sukaupusi. Taigi ji, be kita ko, būtų ir naujas, modernus kalbininkų darbo įrankis, gerokai palengvinantis žodynininkų ir kitų tyrėjų darbą, o ne elektroninis tradicinio *Žodyno* variantas, nuo pastarojo besiskiriantis tik pritaiikymu naudotis elektroninėje terpėje ir kiek platesnėmis tradicinės paieškos galimybėmis. Jau nekalbant apie nuolat atnaujinamų duomenų gausą, kuria ši *Bazė* tiek nuo minėto, tiek nuo kitų žodynų turėtų skirtis pirmiausia.

Taip pat pasakytina, kad jei *Bazės* kūrimas būtų tęsiamas, reikėtų pakeisti jos pavadinimą, galbūt pavadinant *Mokslinė dabartinės lietuvių leksikos duomenų bazė* ar pan., nes iš tiesų joje būtų kaupiami ne tik bendrinės kalbos duomenys. Beje, pakeitus šios bazės pavadinimą mažiau būtų ir pagundų painioti ją su elektroniniu *Bendrinės lietuvių kalbos žodynu*.

Literatūra

1. Jakaitienė, E., 2005. *Leksikografija*. Vilnius: Mokslo ir enciklopedijų leidybos institutas.
2. Murmulaitytė, D., 1999. „Bendrinės lietuvių kalbos žodyno“ kompiuterizavimo klausimu. *Lituanistica*, 3 (39), pp.92-98.
3. Murmulaitytė, D., 2002. Bendrinės lietuvių kalbos leksikos kompiuterizavimo naujovės. *Lituanistica*, 3 (51), pp.64-73.
4. Murmulaitytė, D., 2006a. „Bendrinės lietuvių kalbos leksikos duomenų bazė“ – paskirtis ir struktūra. *Leksikografija ir leksikologija 1*, Aktualūs aiškinamųjų lietuvių kalbos žodynų klausimai. Vilnius: LKI, pp.247-296.
5. Murmulaitytė, D., 2006b. Naujas bendrinės lietuvių leksikos tyrimo įrankis. *Baltų filologija*. Rīga, p.15, pp.41-57.

6. Murmulaitytė, D., 2007a. Vedinys – savarankiškas leksikografinis vienetas. *Valoda – 2007. Valoda dažadu kultūru kontekstā, Zinātnisko rakstu krājums 17*. Daugavpils: Daugavpils Universitātes Akadēmiskais apgāds „Saule“, pp.462-471.
7. Murmulaitytė, D., 2007b. Leksikografinis vienetas tradiciniame žodyne ir elektroninėje duomenų bazėje. *Prace bałtystyczne*. Warszawa, p.4.
8. Murmulaitytė, D., 2008. Bandomasis „Bendrinės lietuvių leksikos duomenų bazės“ duomenų rengimas (kartu su E. Polekaite). *Studentų mokslinė praktika. 2006–2007 m. konferencijos medžiaga*. Vilnius: Vilniaus pedagoginio universiteto leidykla, pp.60-63.
9. Murmulaitytė, D., 2009. Semema tradiciniame žodyne ir elektroninėje duomenų bazėje (priimtas spausdinti mokslinių straipsnių rinkinyje *Leksikografija ir leksikologija 2*, Vilnius: LKI).
10. Mikelionienė, J., 2003. Tradicinės ir kompiuterinės leksikografijos sąlyčiai ir skirtybės. *Leksikografijos ir leksikologijos problemos*. Vilnius: Lietuvių kalbos instituto leidykla, pp.80-89.
11. Poliakovas, O., 2003. Lietuviškas žodis tradicinės ir elektroninės leksikografijos epochų sandūroje. *Leksikografijos ir leksikologijos problemos*. Vilnius: Lietuvių kalbos instituto leidykla, pp.89-106.
12. Svetikienė, D., 2006. Kai kurios išvestinių būdvardžių ir iš jų padarytų prieveiksmių pateikimo aiškinamuosiuose lietuvių kalbos žodynuose problemos. *Leksikografija ir leksikologija 1*, Aktualūs aiškinamųjų lietuvių kalbos žodynų klausimai. Vilnius: LKI, pp.181-191.
13. Šepetytė, R., 2006. Dėl žodžių iliustravimo jų vediniais aiškinamuosiuose lietuvių kalbos žodynuose. *Leksikografija ir leksikologija 1*, Aktualūs aiškinamųjų lietuvių kalbos žodynų klausimai. Vilnius: LKI, pp.209-220.

Daiva Murmulaitytė

The Database of Standard Lithuanian Lexicon is not an Electronic Version of the Dictionary of Standard Lithuanian

Summary

The idea to set up the *Database of Standard Lithuanian Lexicon* was launched after having rejected the idea to set up an electronic version of the *Dictionary of Standard Lithuanian* meant for the wider public as well as research purposes. The *Database* started with devising an original structure for it; as a result, the *Database* is related to the *Dictionary* exclusively through its core data. However, the *Database* is often treated as an electronic version of the *Dictionary*. In fact, they would have to differ in the structure of the data, the scope of information and its presentation, the scope and complexity of search, modes and possibilities of presenting the results of the search, approaches to the presentation of individual lexicographic units etc. There is also a suggestion made to modify the title of the *Database*.

Straipsnis įteiktas 2009 04
Parengtas spaudai 2009 11

Apie autorę

Daiva Murmulaitytė, hum. m. (kalbotyra) daktarė, LKI Leksikografijos centro vyresnioji mokslo darbuotoja.

Mokslinės veiklos sritys: leksikologija ir leksikografija, žodžių daryba, morfemika, bendrinės kalbos žodynų rengimas bei kompiuterizavimas, elektroninių kalbos duomenų bazių kūrimas.

Adresas: Lietuvių kalbos institutas, P. Vileišio g. 5, 10308 Vilnius

El. paštas: daminga@takas.lt

DOI: 10.5755/j01.sal.1.15.43350