

VILNIAUS UNIVERSITETAS
FILOLOGIJOS FAKULTETAS

Kamilė Pavydytė
Taikomosios kalbotyros MA studijų programa

**CHATGPT GENERUOJAMI ILIUSTRACINIAI LAUKAI:
TURINIO IR KOKYBĖS ANALIZĖ BEI PALYGINIMAS SU BLKŽ**
Magistro darbas

Darbo vadovė
Doc. dr. Lina Plaušinaitytė

Vilnius, 2024

Kamilė Pavydytė, *ChatGPT generuojami iliustraciniai laukai: turinio ir kokybės analizė bei palyginimas su BLKŽ*. Magistro darbas, vadovė doc. dr. Lina Plaušinaitytė, Vilniaus universitetas, 2024, 90 p.

Raktažodžiai: elektroninė leksikografija, iliustraciniai duomenų laukai, dirbtinis intelektas, ChatGPT, Bendrinės lietuvių kalbos žodynas, leksikografinių tekstų kokybė.

Anotacija

Šiame MA darbe siekiama aprašyti įrankio ChatGPT lietuviškai generuojamus žodyno iliustracinius sakinius, atliekant šių sakinių kokybinę lingvistinę analizę. Darbe taip pat siekiama įvertinti ChatGPT generuojamų iliustracinių sakinių kokybę, juos palyginant su BLKŽ esančiais iliustraciniais sakiniais. Darbe keliamas klausimas, ar ChatGPT lietuviškai sugeneruotus iliustracinius sakinius būtų galima perkelti į hipotetinį žodyną, o jei ne, kokie trūkumai tam trukdo. Darbe taip pat keliamas hipotezė, jog BLKŽ esantys iliustraciniai sakiniai šiuo metu yra kokybiškesni už iliustracinius sakinius, kuriuos generuoja ChatGPT, o norint sugeneruotus sakinius perkelti į hipotetinį žodyną, prireiktų papildomos leksikografo redakcijos. Teorinėje darbo dalyje aptariamos svarbiausios leksikografijos, dirbtinio intelekto ir su juo susijusios sąvokos, BLKŽ koncepcija, ChatGPT ypatumai ir su ChatGPT susiję leksikografiniai tyrimai. Empirinėje dalyje aprašomi sugeneruotų sakinių sintaksiniai, morfologiniai, leksiniai, pragmatiniai aspektai, aptariamas sugeneruotų sakinių turinys ir nagrinėjami apklausos, kurioje respondentų prašyta įvertinti sugeneruotų sakinių skambėjimą, rezultatai. Toliau aprašomi kiti ChatGPT sugeneruotų sakinių aspektai ir BLKŽ esantys iliustraciniai sakiniai. Empirinės dalies pabaigoje dalis ChatGPT sugeneruotų sakinių palyginami su BLKŽ esančiais iliustraciniais sakiniais ir įvertinama ChatGPT sugeneruotų sakinių kokybė. Tyrimas parodė, jog ChatGPT sugeneruotais autentiškai skambančiais sakiniais, juos poredagavus leksikografui, galima tik papildyti hipotetinį žodyną. Tai lemia sintaksinis sakinių panašumas, semantinės, leksinės, morfologinės sugeneruotų sakinių įvairovės stygius, nedidelis ChatGPT sugeneruotų sakinių kiekis. Darbas galėtų būti aktualus leksikografijos tyrėjams, leksikografams bei plačiajai visuomenei.

Turinys

Įvadas.....	5
1. Teoriniai MA darbo tyrimo pamatai	10
1.1. Spausdintiniai ir EŽ, EŽ klasifikavimas ir ypatumai.....	10
1.1.2. EŽ straipsnis, jo struktūra ir IL sąvoka	12
1.1.3. IL pavyzdžių atranka ir problematika	14
1.1.4. EŽ ir IL kokybės kriterijai.....	16
1.2. Norminamieji žodynai, jų specifika ir kritika moksle	17
1.2.1. BLKŽ koncepcija	19
1.3. DI samprata ir artimos sąvokos.....	20
1.3.1. DI (ChatGPT) pritaikymas leksikografijos tyrimuose.....	22
1.4. Įrankis ChatGPT, jo galimybės ir trūkumai	23
1.4.1. ChatGPT užklausos ir jų generavimo ypatumai	24
2. ChatGPT sugeneruotų ir BLKŽ esančių IL tyrimo ir rezultatų aprašymas	26
2.1. Tyrimo metodika.....	26
2.1.1. Lemų sąrašo sudarymas	26
2.1.2. Užklausos formulavimas ir bandomasis tyrimas	28
2.1.3. Kiti tyrime pasitelkti metodai.....	32
2.2. ChatGPT sugeneruotų ir BLKŽ esančių IL tyrimo rezultatai.....	34
2.2.1. Bendrai apie ChatGPT sugeneruotus IL	34
2.2.1.1. Autentiškai skambantys sakiniai	35
2.2.1.2. Gimtakalbių vertinimas	39
2.2.1.3. Sintaksinė autentiškai skambančių sakinių analizė.....	42
2.2.1.4. Morfologiniai autentiškai skambančių sakinių aspektai	46
2.2.1.5. Leksiniai autentiškai skambančių sakinių aspektai.....	47
2.2.1.6. Pragmatiniai autentiškai skambančių sakinių aspektai	48
2.2.1.7. Autentiškai skambančių sakinių turinys	49
2.2.1.8. Autentiškai skambančių sakinių analizės apibendrinimas	51
2.2.2. Gimtakalbių vertinimu dirbtinai skambantys sakiniai	52
2.2.2.1. Haliucinacijos.....	53
2.2.2.2. Kiti aspektai ir problemos	55
2.2.3. BLKŽ esantys IL ir jų turinys	57
2.2.4. ChatGPT sugeneruotų ir BLKŽ esančių IL palyginimas.....	59
2.2.5. ChatGPT sugeneruotų IL kokybės įvertinimas	69
Išvados.....	71
Literatūros sąrašas	73
Šaltinių sąrašas	82
Summary	84

Priedas 85

MA darbe vartojamų santrumpų sąrašas

- BLKŽ** – Bendrinės lietuvių kalbos žodynas
- ChatGPT** – Chat Generative Pre-trained Transformer
- DLKŽ** – Dabartinės lietuvių kalbos žodynas
- DI** – dirbtinis intelektas
- EL** – elektroninė leksikografija
- EŽ** – elektroninis žodynas / elektroniniai žodynai
- IL** – iliustracinis laukas / iliustraciniai laukai
- LKŽ** – Lietuvių kalbos žodynas
- VLE** - Visuotinė lietuvių enciklopedija
- ŽS** – žodyno straipsnis / žodyno straipsniai

Ivadas

Temos aktualumas ir naujumas. *Chat Generative Pre-trained Transformer* (toliau – ChatGPT) – tai novatoriškas, 2022-ųjų metų pabaigoje išleistas bei internete laisvai prieinamas dirbtiniu intelektu (toliau – DI) grįstas turinį generuojantis įrankis. Vartotojui pateikus užklausas (angl. *prompts*), įrankis geba sugeneruoti bet kokius tekstus – vertimus, tekstų santraukas, pateikia atsakymus į klausimus (Liu, Han et al. 2023: 1), dažniausiai prašomus sugeneruoti elektroninių laiškų tekstus bei programinius kodus (Lin 2023: 1) ir kita¹. Taigi, ChatGPT geba generuoti įvairaus pobūdžio tekstus. Tačiau šis įrankis turi ir trūkumų: netiksliai supranta užklausas ir jų kontekstą, pateikia faktinių klaidų, nepatikimų ar visai nesuprantamų atsakymų (Taecharungroj 2023: 2). Nepaisant to, ChatGPT susilaukė didelio dėmesio tiek viešajame diskurse², tiek ir tikslųjų (Frieder et al. 2023, Surameery, Shakor 2023), gamtos (Zhu et al. 2023), medicinos (Biswas 2023, Liu, Wang et al. 2023), taip pat ir humanitarinių (Andriansyah 2023, Leme Lopes 2023, McGee 2023, Virvou et al. 2023) mokslų tyrimuose. ChatGPT vaidmuo jau yra panagrinėtas taikomosios kalbotyros kontekste – esama straipsnių apie jo perspektyvas lingvodidaktikoje (Hong 2023, Kohnke et al. 2023).

ChatGPT geba generuoti net tik atsakymus į pateikiamas užklausas, tačiau atlikti ir sudėtingesnes užduotis, tarp kurių – žodyno straipsnių (toliau – ŽS) kūrimas. ChatGPT kaip galimas leksikografinio darbo įrankis iki šiol palyginus nedaug tirtas leksikografijos moksle. Tai yra visiškai nauja tema taikomojoje kalbotyroje. Įrankio siūlomas galimybes ir perspektyvas leksikografijoje jau aprašė Robert’as Lew’as (2024), Gilles-Maurice de Schryver’is (2023, 2023a). 2023 m. spalio mėn. išleistoje publikacijoje Schryver’is apžvelgia ir vertina dešimt kitų tyrimų, kuriuose tirtos ChatGPT panaudojimo leksikografijoje galimybės (Schryver 2023: 5). Vienas jų apima ir šiam MA darbui nelabai tolimą tyrimą, kuriame ChatGPT buvo prašyta sukurti Collins COBUILD anglų kalbos ŽS, o vėliau buvo vertinta jų kokybė (Lew 2023). Taigi, jau esama mokslinių bandymų aprašyti DI grįstų įrankių, pavyzdžiui, ChatGPT, galimą panaudojimą sudarant elektroninius žodynus (toliau – EŽ).

Šio MA darbo autorė siekia nagrinėti galimybę susieti įrankio ChatGPT generuojamus tekstus su elektroninės leksikografijos (toliau – EL) sritimi keldama tokį klausimą: **ar DI lietuviškai sugeneruotus iliustracinius sakinius būtų galima perkelti į hipotetinį žodyną? Jei ne, kokie trūkumai užkerta kelią tokiam perkėlimui?** Minėtina, jog tokio mokslinio darbo iki šiol dar nebuvo parašyta, o jo rezultatai yra aktualūs leksikografams ir lietuvių kalbos EL mokslui. Autorės nuomone, lietuvių kalbos EL yra nelabai plačiai ištirtas, tačiau mokslinė prasme itin perspektyvus laukas. Tolimesni lietuvių EL tyrimai, tarp jų ir šis MA darbas, yra svarbūs ne tik su lietuvių kalbos

¹ 2023 m. rugsėjo 25 d. bendrovė OpenAI paskelbė įrašą, jog ChatGPT versija mobiliuojuose programėlyje jau leidžia vartotojui užklausas pateikti balso, nuotraukų formatu bei gauti ChatGPT įgarsintus atsakymus. Plačiau žr.: <https://openai.com/blog/chatgpt-can-now-see-hear-and-speak>.

² Minėtinas šio dirbtiniu intelektu grįsto įrankio aktualumas žurnalistikoje, pavyzdžiui, Lietuvos (plačiau žr. <https://www.lrt.lt/tema/chatgpt>) ar JAV (plačiau žr. <https://www.nytimes.com/spotlight/chat-gpt>) žiniasklaidoje.

žodynais dirbantiems įvairių sričių profesionalams bei plačiajai visuomenei, t. y., EŽ vartotojams ir EŽ kūrėjams. Darbas, atskleidžiantis ChatGPT leksikografinių tekstų generavimo lietuvių kalba potencialą, gali būti taip pat įdomus žmonėms, naudojančioms šį įrankį.

Minėtina, jog ChatGPT ir jo generuojami tekstai nuolat tobulėja kalbine, informacine prasmėmis, todėl atliktas tyrimas fiksuos įrankio generuojamų leksikografinių tekstų kokybę ir jo siūlomas leksikografines galimybes tam tikroje laiko atkarpoje. ChatGPT programos nuolatinio tobulėjimo nulemtas sugeneruotų atsakymų nenusipėjamumas yra tiek darbo ribotumas, tiek ir galimybes vėliau plėtoti tyrimą sudarantis aspektas. Ateityje būtų galima šio tyrimo rezultatus palyginti su ChatGPT pagal užklausas generuojamais leksikografiniais tekstais. MA darbo rezultatais taip pat galės pasinaudoti kiti EL tyrėjai.

Darbo objektas – ChatGPT pagal užklausas sugeneruoti iliustraciniai laukai (toliau – IL). **Tikslas** – įvertinti ChatGPT generuojamų IL turinį ir kokybę, aprašyti BLKŽ esančius IL. Taip pat siekiama nustatyti, ar ChatGPT geba sukurti kokybiškus IL, kuriuos būtų galima perkelti į žodyną, o jei jų nebūtų galima perkelti, kokie IL trūkumai trukdo jų perkėlimą.

Tikslui įgyvendinti keliami tokie **darbo uždaviniai**:

1. Aptarti svarbiausias su darbu susijusias leksikografines sąvokas, BLKŽ koncepciją, DI sampratą ir jai artimas sąvokas, su ChatGPT panaudojimu leksikografijoje susijusius tyrimus bei įrankio ChatGPT ypatumus.
2. Sudaryti lemų sąrašą, pagal kurį bus pateiktos užklausos įrankiui ChatGPT ir kurių BLKŽ straipsniai bus apžvelgiami.
3. Suformuluoti užklausą, kuri bus pateikiama ChatGPT įrankiui.
4. Pateikti užklausą įrankiui ChatGPT ir surinkti jo sugeneruotus tekstus.
5. Sudaryti apklausą, kurioje gimtakalbiai respondentai įvertins, kurie ChatGPT sugeneruoti iliustraciniai sakiniai skamba autentiškai, o kurie – dirbtinai.
6. Aprašyti apklausos rezultatus.
7. Atlikti ChatGPT sugeneruotų autentiškai skambančių iliustracinių sakinių išsamią kokybinę lingvistinę analizę.
8. Aprašyti kitus aspektus, susijusius su neautentiškai skambančiais ChatGPT sugeneruotais sakiniais.
9. Aprašyti, kokiais bendrais bruožais pasižymi BLKŽ esantys IL ir jų turinys.
10. Palyginti dalį ChatGPT sugeneruotų ir BLKŽ esančių iliustracinių sakinių.
11. Remiantis viso MA darbo tyrimo rezultatais įvertinti ChatGPT IL kokybę.

Tiriamoji medžiaga ir metodai. Darbe pasitelkiant kokybinės lingvistinės analizės priegas tiriami ChatGPT sugeneruoti IL, dalį jų palyginant su BLKŽ esančiais iliustraciniais sakiniais. Pirmiausia iš Virginijaus Dadurkevičiaus 2020 m. sudaryto dažninio lietuvių kalbos žodžių sąrašo

atrinkta 40 dažniausių pagal vartojimą lemų, prasidedančių raide *k*. Lemų sąrašo atrankai pasirinkta būtent ši raidė, nes ji BLKŽ yra baigta ir turi daugiausia ŽS (6375 straipsnius). Toliau suformuluota ir ChatGPT įrankiui pateikta užklausa, kurioje prašyta sugeneruoti iliustracinius sakinius visoms 40 atrinktų lemų. Darbe naudota kompiuteriu pasiekiamą, nemokama GPT-3.5 versija. ChatGPT sugeneruoti IL buvo surinkti 2024 m. vasario 16 dieną – šią datą svarbu paminėti dėl nuolatinio įrankio tobulėjimo bei dėl vis skirtingai generuojamo teksto³.

Toliau sudaryta apklausa, kurioje respondentai, kuriems lietuvių kalba yra gimtoji, įvertino visų 147 ChatGPT sugeneruotų sakinių skambėjimo natūralumą / dirbtinumą. Po to apžvelgti apklausos rezultatai ir atlikta išsami autentiškai skambančių ChatGPT sugeneruotų iliustracijų⁴ analizė. Tuomet aprašyti BLKŽ esantys IL. Dalis ChatGPT sugeneruotų iliustracijų palygintos su BLKŽ esančiomis iliustracijomis. Galiausiai įvertinta ChatGPT generuojamų iliustracijų kokybė. Minėtina, jog norminamasis, aiškinamasis BLKŽ šiame darbe pasirinktas kaip atspirties taškas, palyginamasis leksikografinis veikalas, vertinant ChatGPT sugeneruotus iliustracinius sakinius.

Darbe yra keliama **hipotezė**:

1) BLKŽ esantys IL šiuo metu yra lingvistiškai kokybiškesni už ChatGPT generuojamus IL. Norint ChatGPT sugeneruotus IL perkelti į hipotetinį žodyną ir taip pagreitinti žodyno straipsnio sudarymo procesą, reikia papildomos leksikografo redakcijos.

Ši hipotezė grindžiama 2023 m. liepos-rugpjūčio mėn. darbo autorės atlikto tyrimo *ChatGPT ir žodyno straipsnių kūrimo galimybės lietuvių ir vokiečių kalbomis*⁵, kurį finansavo Lietuvos mokslo taryba, rezultatais. Atlikus tyrimą paaiškėjo, jog nors ChatGPT generuojami žodyno straipsniai pateikė gausius iliustracinius pavyzdžius, papildomą gramatinę, etimologinę informaciją apie lemas, vis dėlto straipsniuose pasitaikė nemažai faktinių klaidų, netipiškų vartosenos pavyzdžių, informacinės vertės neturinčių formuluočių. Dėl to galima manyti, jog ChatGPT sugeneruoti IL nebus tokie kokybiški, kokie yra pateikiami leksikografų sudarytame BLKŽ ir šiuo metu sugeneruotų iliustracijų nebūtų galima perkelti į hipotetinį žodyną be leksikografo redagavimo.

Darbo struktūra. Šį darbą sudaro teorinė bei empirinė dalys. Teorinėje dalyje trumpai aptariami spausdinti žodynai, EŽ, jų klasifikavimas bei ypatumai. Išsamiai pristatoma EŽ straipsnio sąvoka, jo struktūra, IL sąvoka, toliau aprašomas IL pavyzdžių atrankos procesas bei jo metu galimai kylančios problemos, apžvelgiami EŽ ir IL kokybės kriterijai. Toliau aptariama norminamojo žodyno koncepcija, žodynų normatyvumo kritika, BLKŽ koncepcija. Kitoje darbo dalyje nagrinėjama DI

³ Apie vis skirtingai sugeneruojamo teksto problematiką plačiau apžvalginame straipsnyje rašo Gilles-Maurice de Schryver'is (2023), šiam ypatumui priskirdamas atkuriamumo (angl. *reproducibility*) terminą (2023: 363).

⁴ Iliustraciniai sakiniai šiame darbe sinonimiškai dar vadinami iliustracijomis ir iliustraciniais pavyzdžiais.

⁵ Šio projekto paraiškos registracijos nr. P-SV-23-435. Jis minimas projekto vadovės doc. dr. Vilmos Zubaitienės tinklaraštyje, skiltyje *Vadovavimas studentų moksliniams tyrimams*: <https://web.vu.lt/flf/v.zubaitiene/moksline-veikla/>.

samprata ir DI artimos sąvokos, DI sietis su leksikografija. Teorinės dalies pabaigoje pristatomas įrankis ChatGPT, jo specifika ir galimybės, įrankiui pateikiamų užklausų specifika.

Empirinėje dalyje pirmiausiai išdėstoma tyrimo metodika, techniniai tyrimo medžiagos rinkimo aspektai, taip pat aprašoma užklausos formulavimo procedūra ir bandomojo tyrimo rezultatai. Po to pateikiami ChatGPT sugeneruotų iliustacinių sakinių tyrimo rezultatai: pateikiami bendri su duomenimis susiję skaičiai, specialūs autentiškai skambančių sakinių kriterijai, gimtakalbių apklausos rezultatai. Pateikiama išsami autentiškai skambančių sakinių sintaksinė analizė, aprašomi morfologiniai, leksiniai, pragmatiniai autentiškai skambančių sakinių aspektai, apžvelgiamas šių sakinių turinys ir pateikiamas šios analizės dalies apibendrinimas. Toliau aptariamos kitos dvi sugeneruotų sakinių kategorijos, papildomi su sugeneruotais sakiniais susiję aspektai. Kitoje dalyje bendrai apžvelgiami BLKŽ esantys IL ir dalis ChatGPT sugeneruotų pavyzdžių toliau palyginami su BLKŽ esančiais pavyzdžiais. Galiausiai remiantis visu tyrimu įvertinama ChatGPT generuojamų iliustracijų kokybė. Darbo pabaigoje pateikiamos išvados, 86 punktų literatūros ir 21 punkto šaltinių sąrašai, MA darbo apibendrinimas anglų kalba. Taip pat pateikiamas priedas *ChatGPT pagal užklausą sugeneruoti IL*. Darbe yra aštuoni paveikslai ir viena lentelė.

1. Teoriniai MA darbo tyrimo pamatai

1.1. Spausdintiniai ir EŽ, EŽ klasifikavimas ir ypatumai

Žodyno samprata mokslinink(i)ų yra pateikiama kiek skirtingai, išryškinant įvairius žodynų ypatumus. Henning'as Bergenholtz'as žodyną vadina leksikografiniu veikalu ir rašo, jog tokia veikalė yra su konkrečiomis temomis ar kalbos elementais susiję ŽS, o žodyne esama tokios informacijos, kuri vartotojui gali padėti teksto recepcijos, teksto produkcijos ar vertimo klausimais arba tiesiog suteikti jam informacijos apie žodį, žodžio dalį ar žodžių kombinaciją (Bergenholtz 2012: 30). Taigi, žodynas yra toks veikalas, kuris gali atlikti kelias žodyno vartotojui aktualias funkcijas vienu metu: apie tai rašo ir Sven'as Tarp'as, žodynus pavadindamas įrankiais, kurie turėtų būti sukurti tam tikriems žmogaus poreikiams patenkinti (Tarp 2009, cit. iš Müller-Spitzer 2009: 2). Vilma Zubaitienė žodynus apibūdina kaip informacijos, komunikacijos ir norminimo šaltinius (Zubaitienė 2014: 4). Kai kuriuose žodyno apibrėžimuose taip pat užsimenama apie **spausdintinių** bei kartu su kompiuterių paplitimu atsiradusių **elektroninių** žodynų skirtį. Evalda Jakaitienė žodynus įvardija knygomis, „kuriose ta pačia ar kita kalba aiškinami žodžiai ar kitokie leksiniai vienetai“ ir pabrėžia, jog tai gali būti ne tik knygos, tačiau ir „įvairios kompiuterinės laikmenos, ir elektroninės duomenų bazės bei bankai“ (Jakaitienė 2009: 308). Alan'as Kirkness'as rašo, jog žodynas yra laikomas prototipiniu leksinės informacijos šaltiniu, o ši (lingvistinė, taip pat žodyne galinti pasitaikyti ir semantinė, vartosenos, etimologinė ir kita) informacija randama spausdintinėje arba elektroninėje formoje (Kirkness 2004: 59). Toks žodynų skirstymas į spausdintinius (dar vadinamus popierinius ar knyginus) bei elektroninius (dar vadinamus kompiuterinius) yra laikomas žodynų tipologija pagal išorinę žodyno formą (Zubaitienė 2014: 56). Šiame MA darbe dėmesys yra telkiamas būtent į EŽ.

Svarbu aptarti, kaip yra klasifikuojami šiam darbui aktualūs EŽ, kurių spartų vystymąsi galima išvelgti iš jų klasifikacijų įvairovės. Prieš daugiau nei dvidešimt metų klasifikuodamas EŽ, Gilles-Maurice de Schryver'is kalbėjo apie tokius šiomis dienomis jau nebeaktuales EŽ prieigos būdus kaip, pavyzdžiui, naudojantis vienoje rankoje telpančiu kompiuteriu (angl. *palmtop*⁶), kompaktine plokštele ar intranetu⁷ (Schryver 2003: 151). Kalbant apie terpę, kurioje galima pasiekti žodyną, Evalda Jakaitienė skiria ir dar vieną kompiuterizuotą EŽ tipą, kurį apibūdina kaip jau egzistuojančių knyginių žodynų elektronines versijas (Jakaitienė 2009: 330) – tokio kompiuterizuoto žodyno pavyzdžiu lietuvių kalbos leksikografijoje būtų galima laikyti suskaitmenintą LKŽ⁸.

Schryver'is (2003: 151) pateikia internetinių EŽ kategoriją, kurių vartotojai žodynus pasiekia naudodami kompiuterius ar kitus ekranus bei interneto prieigą. Šią Schryver'io internetinių EŽ

⁶ Žodžio *palmtop* definiciją skaityti čia: <https://dictionary.cambridge.org/dictionary/english/palmtop>.

⁷ Pasak Valentinios Dagienės et al., intranetas yra „[v]ietinis tinklas, naudojantis interneto protokolus duomenims persiųsti, galintis turėti ryšį su internetu“ (Dagienė et al. 2014).

⁸ Baigtinio LKŽ kompiuterizuota versija pasiekama šiuo adresu: <https://ekalba.lt/lietuviu-kalbos-zodynas/>.

kategoriją išplėtė Annette Klosa pridėdama, jog internetinius žodynus pasiekti galima pasitelkiant ir mobiliuosius telefonus (Klosa 2009: 2). Verta paminėti, jog skiriamas ir išmaniųjų telefonų ar planšetinių kompiuterių programėlėse esančių žodynų tipas: tokioje programėlėje gali būti patalpintas vienas ar keli žodynai, gali būti ir su galimybe juos papildyti (Nomicienė, Stunžaitė 2016: 93).

EŽ paplitimas šiuolaikinėje visuomenėje neišvengiamai palietė jų klasifikavimą: rašydamas apie internetinius žodynus, Pedras A. Fuertes-Olivera's išskyrė institucijų sukurtus (angl. *Institutional Internet reference works*) bei kolektyvinius nemokamus daugiakalbius EŽ (angl. *Collective free multiple-language Internet reference works*) (Fuertes-Olivera 2009: 103-107). Pirmąjį išskiriamą EŽ tipą Fuertes-Olivera's apibūdina kaip pripažintos institucijos sudarytus, mokamus arba nemokamus žodynus (Fuertes-Olivera 2009: 103-104). Nors autorius rašo, jog dažniausiai tokie **nemokami** EŽ yra sudaromi neprofesionalių leksikografų (angl. *amateur lexicographers*), atstovaujančių tam tikrą instituciją, o ribotos prieigos EŽ pasižymi aukštesniu patikimumu dėl jų sudariusių leksikografų kompetencijos, MA darbo autorė nori pabrėžti, jog situacija per 15 metų jau pakito. Esama aukštos kokybės internete pasiekiamų anglišku nemokamų Kembridžo, Oksfordo, Kolinso, vokiškų Duden, Pons, Leo žodynų. Lietuvių kalbos EL kontekste minėtini tokie kokybiški, patikimi ir neribotos prieigos nemokami EŽ kaip BLKŽ⁹, DLKŽ¹⁰ ir kt., kurių elektronines versijas parengė Lietuvių kalbos institutas¹¹.

Rašydamas apie kolektyvinius nemokamus daugiakalbius žodynus, Fuertes-Olivera's (2009: 106) kaip pavyzdžius mini puslapius *Wiktionary*¹² ir net *Wikipedia*¹³ bei pabrėžia, jog tokie žodynai sudaro galimybę į puslapį įtraukti duomenis bet kokia kalba ir bet kokiam žmogui, tačiau vadovaujantis tam tikrais nurodymais puslapyje. Vis dėlto tokia laisvai prieinama žodynų redagavimo prieiga gali kelti problemų: minėtus nemokamus žodynus dažniausiai sudaro neprofesionalūs leksikografai, jie tiesiog randami internete ir nėra aiškaus jų autoriaus (Collins 2016: 36), o tai gali paveikti informacijos lingvistinį, semantinį, faktinį patikimumą bei kokybę.

Apibendrinant, su kompiuterių, interneto įsitvirtinimu atėję EŽ nuo spausdintinių žodynų aiškiai skiriasi pasiekiamumu elektroninėje kompiuterio ar kito išmaniojo įrenginio erdvėje, priešingai nei fizinės knygos puslapiuose. Matyti, jog EŽ mokslinink(i)ų yra klasifikuojami įvairiai: išskiriami EŽ tipai pagal žodyno pasiekimo būdus, pagal pirminį EŽ atsiradimą elektroninėje ar ne elektroninėje erdvėje (pvz., kompiuterizuoti EŽ), pagal tai, ar EŽ prieinamas programėlėje, c kas yra

⁹ BLKŽ el. versija pasiekama šiuo adresu: <https://ekalba.lt/bendrines-lietuviu-kalbos-zodynas/>.

¹⁰ DLKŽ el. versija pasiekama adresais <https://ekalba.lt/dabartines-lietuviu-kalbos-zodynas/> arba <http://www.lkz.lt/>.

¹¹ Ši informacija pasiekama paspaudus bet kurią minėtą nuorodą į žodynus portale *Ekalba* ir pasirinkus informacijos apie žodyną skiltį.

¹² *Wiktionary* (liet. *Vikižodynas*) yra daugeliu kalbų (taip pat ir lietuvių) prieinamas internetinis žodynas. Plačiau apie lietuvišką Vikižodyno versiją skaityti čia: https://lt.wiktionary.org/wiki/Viki%C5%BEodynas:Pagrindinis_puslapis.

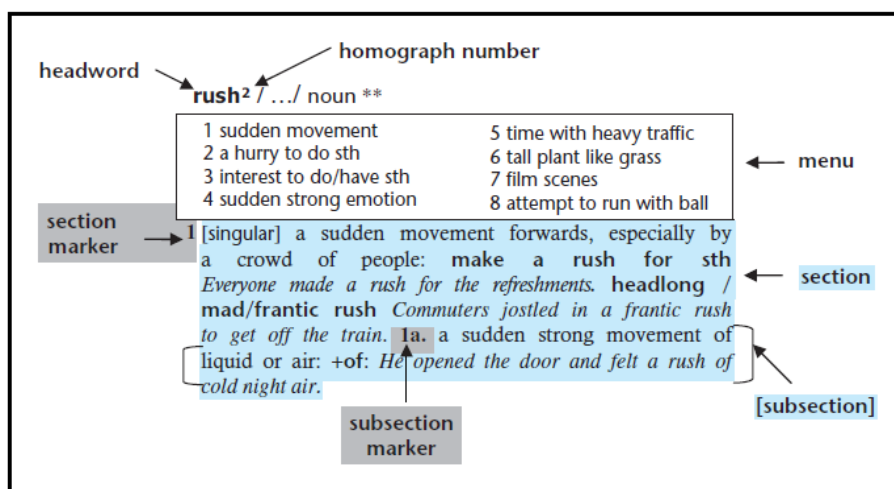
¹³ Darbo autorei netikėtai pasirodė puslapio *Wikipedia* kaip tam tikro žodyno minėjimas greta *Wiktionary* žodyno autoriaus Pedro A. Fuertes-Olivera (2009) tekste.

EŽ kūrėjas/-ai ir kt., o EŽ kokybė gali būti vertinama net pagal EŽ (ne)apmokestinimo ar (ne)prieinamumo kriterijus. Šiam darbui aktualus yra būtent internetinis autoritetingos institucijos sukurtas BLKŽ (žr. skyrių 1.2.1.).

1.1.2. EŽ straipsnis, jo struktūra ir IL sąvoka

Šiame skyriuje toliau aprašomi skirtingų autorių nurodomi EŽ mikrostruktūros principai bei pristatoma IL sąvoka.

ŽS Anetos Stojić apibūdinamas kaip dvinarė struktūra, susidedanti iš lemos ir aiškinamosios dalies (Stojić 2015: 24). Autorė rašo, jog lema yra įvardinė ir privaloma ŽS dalis, o jame lema toliau aiškinama pasitelkiant kalbinę ar faktinę informaciją. Pasak Stojić, ši apie lemą pateikiama informacija gali būti suskirstyta į skirtingus informacijos tipus (pvz., reikšmės, rašybos, kalbos dalies), o straipsnyje gali būti pateikiama sakiniiais, sutrumpinimais, parafrazėmis ir pan. Nors rašydama apie ŽS autorė nepabrėžia, jog omenyje turi būtent EŽ, tačiau šis terminas neabejotinai vartotinas ir EL srityje: vartotojui pateikus paieškos užklausą EŽ, jo ekrane atsiranda struktūruotas straipsnis, kuriame pateikiama ieškoma lema ir jos informacija¹⁴. EŽ straipsnis turi savo mikrostruktūrą (Melnikienė, Jankauskaitė 2012: 80): aprašydami ŽS struktūrinius komponentus, Sue Atkins ir Michael'as Rundell'as pabrėžia, jog kiekvienas ŽS komponentas turi fiksuotą savo vietą, paeiliui minėdami lemą, reikšmes eksplikuojantį „menu“ (dažniausiai ši skiltis atsiranda negimtakalbiams skirtuose žodynuose), lemą aiškinančius skyrelius ir poskyrius¹⁵ bei jų žymėjimus (šių ŽS struktūrinių komponentų vizualizaciją žr. 1 paveiksle) (Atkins, Rundell 2008: 204-205).



1 pav. Žodyno straipsnio komponentai (Atkins, Rundell 2008: 203).

¹⁴ Tai galima įrodyti internete patikrinus bet kokį tiek vienkaltį (žr. <https://ekalba.lt/sinonimu-zodynas/pavyzdys?paieska=pavyzdys&i=a16b8a87-4b77-4f0d-bbcf-e09a1a717332>), tiek dvikalbį / daugiakalbį EŽ (žr. <https://dict.leo.org/german-english/example>).

¹⁵ Turėtina omenyje tai, jog nebūtinai kiekviename ŽS randami poskyriai: jie būdingi tik polisemiškų lemu aiškinimui.

Pasak autorių, tokia ŽS struktūra būdinga ir spausdintiniams, ir EŽ, tačiau ŽS gali pasitaikyti ne visi išvardyti komponentai arba gali būti ir kitokių (Atkins, Rundell 2008: 202).

Svarbu atsižvelgti į tai, jog apie lemą ŽS pateikiamą informaciją galima suskirstyti į tam tikras grupes. Geert'as Booij'as aprašo ŽS neprivalomas, tačiau dažnai pateikiamos **fonologinės, morfologinės, sintaksinės** informacijos principus: autorius rašo apie fonetinę formą – žodžio tarimo transkripciją, pateiktą tarptautine fonetine abėcėle, sukirčiavimą ir suskiemenavimą (Booij 2003: 252), ŽS nurodomą žodžių kaitybos ar darybos (morfologinę) informaciją (Booij 2003: 254), sintaksinę informaciją (pvz., veiksmažodžių tranzityvumą ar tai, ar konkretus daiktavardis skaičiuotinis) (Booij 2003: 257). Fons'as Moerdijk'as rašo apie žodžio reikšmes atskleidžiančią **semantinę** informaciją: žodžio reikšmių sudėjimą į tą patį straipsnį (angl. *lumping*) ar išskyrimą į daugelį atskirų reikšmių (angl. *splitting*), konteksto svarbą žodžio reikšmei (-ėms), galimą polisemijos ar homonimijos informaciją (Moerdijk 2003: 278-282). Henk'as Verkuyl'as, Maarten'as Janssen'as ir Frank'as Jansen'as mini galimus trumpinius, nurodančius formalius žodžio aspektus (kalbos dalį, skaičių ir kt.) arba **stilistinius** duomenis (vulgarios, šnekamosios, literatūrinės ir kt. leksikos žymėjimą) (Verkuyl et al. 2003: 297-298). Nicoline van der Sijs dar mini ir galimą **etimologinę** žodžio informaciją, nors ji žodynuose nėra tokia svarbi ir daugelyje jų išvis nepasitaiko (van der Sijs 2003: 312).

Visus šiuos aspektus Evalda Jakaitienė vadina duomenų laukais, o šie yra ŽS mikrostruktūros dalys: atitinkamai pagal pateikiamą lemos informaciją esama fonetinių, morfologinių, sintaksinių, semantinių, stilistinių duomenų laukų, iliustracijų lauko (Jakaitienė 2005: 56-103). Svarbu išsamiau aptarti šiam MA darbui aktualią **IL sąvoką**: Danguolė Melnikienė ir Monika Jankauskaitė rašo, jog „[p]agrindinis iliustracinių pavyzdžių tikslas – parodyti, kaip antraštinis žodis ir jo formos funkcionuoja kalboje kartu su kitais žodžiais“ (Melnikienė, Jankauskaitė 2012: 80), o tai yra daroma per autentiškus kalbos pavyzdžius, leksikografų įprastai atrenkamus iš tekstynų (žr. 1.1.3. skyrių). Tai gali būti laisvieji ar pastovieji žodžių junginiai, kolokacijos (Melnikienė, Jankauskaitė 2012: 80-81), ištisi sakiniai, citatos, tautosaka (Zubaitienė 2014: 30). Taigi, tai gali būti įvairiausi kalbos pavyzdžiai, atitinkamai iliustruojantys lemą ir jos vartojimą bei vartojimo kontekstą, o šie iliustruojantys sakiniai turi atitikti tam tikrus kokybės kriterijus (žr. 1.1.4. skyrių). Taip pat minėtina, jog vienakalbiuose žodynuose esančių iliustracijų atliekamos funkcijos yra kiek kitokios dvikalbiuose ar daugiakalbiuose: juose iliustraciniai sakiniai patvirtina, papildo vertimą (Atkins, Rundell 2008: 509), yra susiję su kitais semantiškai artimais žodžiais (Humbly 2001: 55-56).

Apibendrinant, EŽ turi savitą mikrostruktūrą, kuri ŽS išryškėja per pateikiamus informacijos apie lemą tipus ir šio pateikimo eiliškumą. ŽS randamos informacijos tipai lietuvių leksikografijos

teorijoje vadinami duomenų laukais. Jie būna fonetiniai, gramatiniai, semantiniai, stilistiniai ir iliustraciniai. Toliau MA darbe nagrinėjami su IL susiję klausimai.

1.1.3. IL pavyzdžių atranka ir problematika

Tekstynai vaidina didžiulį vaidmenį šiuolaikinėje leksikografijoje ir pasitarnauja kaip pirminis informacijos šaltinis sudarant žodynus (Enčeva, Jazbec 2014: 50). Iš tekstynų atrenkami lemas iliustruojantys sakiniai: Gilles-Maurice de Schryver'io nuomone, joks rimtas leksikografas neapsiimtų rengti žodyno be tekstyno ar be kelių tekstynų (Schryver 2003: 167). Valérie Kataros rašo, jog iliustraciniai sakiniai turi atskleisti iliustruojamos lemos reikšmę, taip pat polisemiškų lemu reikšminius atspalvius, lemos kontekstą, gramatinius kalbos dalių vartojimo ypatumus, kolokacijas (Kataros 2004: 492). Autorė taip pat mini, jog iliustraciniai sakiniai turėtų būti „gyvi“ (angl. *lively*) ir patrauklūs vartotojui (ten pat).

Nors tekstynai yra leksikografinį darbą paspartinanti priemonė ir juose galima rasti įvairių žodžio vartoseną atspindinčių pavyzdžių, tekstynai nėra be trūkumų realią kalbinę vartoseną atspindintys įrankis. John'as Sinclair'is rašo, jog nepaisant nei tekstyno dydžio, nei kruopštaus tekstyno duomenų surinkimo, **nėra** tokio tekstyno, kuris turėtų / atitiktų tokius pačius bruožus kaip ir vartojama kalba (Sinclair 2004). Esama ir vadinamojo tekstynais grįsto (angl. *corpus-based*) pavyzdžių atrankos požiūrio (Lew 2010: 5), kuomet tekstynuose rastus pavyzdžius leksikografai nežymiai pakeičia pavyzdžiui dėl to, nes šie gali būti per ilgi (Murmulaitytė 2011: 134-135, Valančiauskienė 2022: 48). Verta paminėti tai, jog kai kada leksikografai žodynuose pateiktus pavyzdžius sugalvoja patys (Lew 2010: 5), nors tokia leksikografinė praktika yra laikoma pasenusia (Karpinska 2012: 96). Pavyzdžių kūrimas yra vertinamas įvairiai: Philippe'as Humblé sugalvotus pavyzdžius vadina nepagrįstais modeliais, nes juose gali pasitaikyti negatyvių konotacijų, konstrukcijos gali būti nors ir taisyklingos, tačiau neįprastos (Humble 2001: 79). Tačiau Sue Atkins ir Michaelas Rundell'as sugalvotus sakinius laiko ekonomiškais, nes viena fraze ar sakiniu jie gali iliustruoti kelis aktualius dalykus (Atkins, Rundell 2008: 456).

Tai, kokie iliustraciniai pavyzdžiai atrenkami į žodynus priklauso nuo prie žodyno dirbančių leksikografų. Būtent dėl to verta paminėti kai kurias problemas, susijusias su iliustracijų atranka. Viena iš galimai kylančių problemų yra stereotipų atsklaida iliustracijose. Carolina Müller-Spitzer ir Jan Oliver Rüdiger'is žodyno pavyzdžiuose pasitaikančius stereotipizavimo atvejus vadina ypatinga leksikografų **atsakomybe** (Müller-Spitzer, Rüdiger 2022: 131). Jau pastebėtos lyčių vaidmenų (Müller-Spitzer, Rüdiger 2022, Iversen 2012, Hidalgo Tenorio 2000), etniškumo vaizdavimo žodynuose problemos. Dragica Žugić ir Milica Vuković-Stamatović pastebėjo, jog su etniškumu susiję terminai vienakalbiuose žodynuose dažnai yra pernelyg supaprastinami ir tai vertina

kaip klaidingą, ribojantį (angl. *reductionist*) dalyką (Žugič, Vuković-Stamatović 2021: 177), o leksikografus skatina šią praktiką keisti į įtraukesnę.

Stephen'as Coffey'is rašo apie įžeidžių (angl. *offensive*), „tabu“ žodžių vaizdavimą anglų kalbos žodynuose (Coffey 2010). Autorius pastebi, jog tirtuose anglų kalbos žodynų antraštynuose nevengiama neigiamą konotaciją turinčių, su žmogaus kūnu ir jo funkcijomis, ligomis ir pan. susijusių žodžių bei juos iliustruojančių pavyzdžių. Žodynuose vartotojui patariama, kaip būtų galima *įžeidžių* lema (angl. *offensive headword*) pasakyti neutralėsiu stiliumi (Coffey 2010: 1275). Taigi, nors įžeidūs žodžiai žodynuose pateikiami bei iliustruojami, leksikografai sąmoningai informuoja žmogų apie su lema susijusį socialinį kontekstą, o žodyno vartotojas ugdo kalbos registrų suvokimo kompetenciją (Coffey 2010: 1280).

Taigi, akivaizdu, jog nors tekstynų svarba leksikografiniame darbe yra didelė, darbas su iš tekstynų atrenkamomis iliustracijomis leksikografams kartais gali kelti įvairių iššūkių. Tai gali būti atrinktų sakinių redagavimas, iliustracijų kūrimas, siekiant kuo geriau iliustruoti lemas, taip pat leksikografų atsakomybė, susijusi su susidūrimu su stereotipais ar įžeidžia leksika iliustracijose. Šis kontekstas svarbus toliau kalbant apie IL bei jų turinį.

1.1.4. EŽ ir IL kokybės kriterijai

EŽ turinys neabejotinai priklauso nuo žodyno kūrėjų numatytos vartotojų grupės bei nuo to, ar EŽ yra vienakalbis, ar daugiakalbis (Atkins, Rundell 2008: 35). Nors MA darbo autorė pastebėjo, kad mokslinėje literatūroje nėra jokių konkrečių EŽ kokybės aprašų ar leksikografams skirtų gairių, vis dėlto esama literatūros, kurioje aprašyti esminiai EŽ bruožai. Tai būtų galima laikyti kokybiško EŽ bruožais, tačiau turint omenyje, jog ne visi literatūroje išvardijami bruožai privalo atsidurti EŽ.

Tetiana Vakaliuk et al. paeiliui vardina šiuos EŽ bruožus: lemos paieškos laukelis, nuorodų sistema (angl. *hyperlink*) (Vakaliuk et al. 2021: 44-50). Įvardijama tokia EŽ pateikiama informacija kaip fonetiniai, semantiniai, etimologiniai, stilistiniai duomenys, kolokacijos, frazeologizmai, galimai nurodoma antonimija, sinonimija, hiponimija. Autoriai taip pat mini papildomos informacijos EŽ pateikimo galimybes: paveikslėlius, vaizdo įrašus¹⁶, garso failus, tekstynų konkordanso eilutes (ten pat). EŽ pateikiamą papildomą informaciją garso ir vaizdo įrašų ar grafikų¹⁷ pavidalu Carolin Müller-Spitzer vadina **leksikografiniu tekstu** (Müller-Spitzer 2009: 2). Vakaliuk et al. bei Müller-Spitzer

¹⁶ Darbo autorei asmeniškai nėra tekę susidurti su EŽ, kurių ŽS būtų su lema susijusi video medžiaga, dėl to šis autorių išskiriamas galimas EŽ bruožas pasirodė netikėtas.

¹⁷ Darbo autorės manymu, grafiku (angl. *graph*) literatūroje yra vadinama kai kuriuose EŽ pasitaikanti papildoma funkcija, kurioje vizualiai pateikiami su lema (dažniausiai semantiškai) susiję žodžiai. Toliau pateikiamame Duden žodyno lemos *moteris* (vok. *Frau*) ŽS apačioje sugeneruota lentelė su žodžiais, kurie statistiškai dažniau pasitaiko šalia lemos *Frau*: <https://www.duden.de/rechtschreibung/Frau>.

aprašytus bruožus galima laikyti kokybiško EŽ kriterijais. Tokius EŽ bruožus kaip paieškos laukelis, su lema susijusi ŽS pateikiama informacija (semantiniai, fonetiniai, morfologiniai, iliustraciniai duomenys) galima laikyti **privalomais** kokybiškame EŽ, nes be jų funkcionalaus EŽ įsivaizduoti tiesiog neįmanoma. Tokie bruožai kaip garso failai – žmogaus įrašytas arba kalbos sintezės¹⁸ pagalba sukurtas lemos tarimo įrašas (Lew 2010: 6), vaizdinė medžiaga, nuorodų sistema ŽS, konkordanso eilutės, etimologinė informacija ir kt. gali būti laikomi **galimomis, bet nebūtinomis** papildomomis funkcijomis žodyne.

Lygiai taip pat svarbu atsižvelgti į EŽ turinį, pateikiamos informacijos taisyklingumą. Herbert'as Ernst'as Wiegand'as skiria du žodyno kokybės kriterijus: 1) leksikografinių duomenų patikimumą; 2) tinkamą leksikografinių duomenų suskirstymą (vok. *Organisation*) ir pateikimą (Wiegand 1997: 195-197). Kalbėdamas apie pirmąjį kriterijų autorius mini leksikografo prisiimamą atsakomybę už leksinių vienetų interpretaciją bei atsižvelgimą į kontekstą: pavyzdžiui, lemos *mirti* apibrėžtis gali būti nepriimtina krikščioniui¹⁹, taigi, leksikografui kai kuriais atvejais gali tekti papildomai apgalvoti definicijas (ten pat).

Plačiau aptariant IL kokybės kriterijus pirmiausia minėtina, jog iliustraciniai sakiniai neabejotinai turėtų tinkamai atskleisti aiškinamo žodžio reikšmę ar reikšmės atspalvį, tačiau žodyne negali būti iliustracijų pertekliaus, turi būti pateikiama tik lemai iliustruoti būtina informacija (Murmulaitytė 2011: 132-134). Khalid'as Ahmad'as et al. rašo, jog iliustraciniai sakiniai turėtų būti pilni (angl. *complete sentence*) ir tokie, kuriuose lema pasitaiko sakinio pradžioje, o ne pabaigoje (Ahmad et al. 1992: 146). Autoriai pastebi, jog nereikėtų lemu iliustruoti per sudėtingais, per ilgais pavyzdžiais ar tokiais sakiniiais, kuriuose pasitaiko papildomo aiškinimo reikalaujančių lemu (angl. *technical terms*) (ten pat). Pasak Philippe'o Humblé, IL turėtų būti atskleidžiami žodžio sintaksės ypatumai (Humblé 2001: 69).

Taigi, iš aptartos literatūros aiškiai matyti, jog kokybiškame EŽ privalo būti pateikta esminė su lema susijusi informacija, leidžianti vartotojui išmanyti lemos vartojimą ir vartojimo kontekstą. Kai kurios EŽ funkcijos yra laikomos privalumu, bet nebūtinybe. Kokybiškas IL privalo vartotojui visapusiškai atskleisti žodžio reikšmę, bet nebūti perkrautas.

¹⁸Pasak Ingos Bindoriūtės (2008), kalbos sintezė yra „automatinis balsinio pranešimo generavimas iš pateikto teksto ar kitos simbolių sekos, t. y. tekstu pateiktos informacijos skaitymas balsu“ (p. 17).

¹⁹ Herbert'as Ernst'as Wiegand'as rašo, jog leksikografas galėtų suformuluoti tokią neutralią apibrėžtį lemai *mirti* kaip *pereiti iš šio gyvenimo į anapus* (vok. *vom diesseitigen in das jenseitige Leben übergehen*) vietoje, pvz., *nustoti gyventi* (Wiegand 1997: 197).

1.2. Norminamieji žodynai, jų specifika ir kritika moksle

Norminamieji žodynai yra neatsiejami nuo kalbinio preskriptyvizmo sąvokos. Laima Nevinskaitė ir Giedrius Tamaševičius preskriptyvizmą supranta „pragmatiškai, kaip kalbos taisyčių visumą“ (Nevinskaitė, Tamaševičius 2019: 6). Esama ir požiūrio, jog preskriptyvizmas ir deskriptyvizmas yra persipynę tarpusavyje: kalbėdamas apie su technologijomis, programinėmis įrangomis susijusią kalbos industriją (angl. *language industry*), Jan’as Engh’as teigia, jog atitinkamas, neperdėtas kalbos aprašymas (t. y., susijęs su kalbiniu deskriptyvizmu) įmanomas tik tuomet, jei imamasi kalbinio taisyklingumo (t. y., susijusio su kalbiniu preskriptyvizmu) pozicijos (Engh 1993: 64). Esama ir preskriptyvinės leksikografijos termino: tai normatyviu požiūriu į kalbą grįsta žodynų kūrimo prieiga, telkiantis į tai, kaip kalba ar kalbinė atmaina **turėtų būti** vartojama, o ne kaip yra vartojama iš tikrųjų remiantis faktais (Hartmann, James 2002: 111). Šiuo požiūriu skirstomi ir žodynai – juos į deskriptyvinius ir preskriptyvinius galima suskirstyti pagal leksikografų asmeninį požiūrį (vok. *Einstellung des Lexikographen*) (Kühn 1989, cit. iš Buzrlová 2013: 39). Taip pat yra požiūrio, jog žodynų skirstymas į deskriptyvinius bei preskriptyvinius priklauso nuo žodyno tikslo (Čermák 1995, cit. iš Buzrlová 2013: 39).

Vis dėlto norminamasis aspektas leksikografijoje ne visada yra atvirai deklaruojamas: pasak Patrick’o Beuge’o, esama tiek norminamųjų žodynų, kuriuose kalbos norminimo pozicija nėra slepiama, tiek *paslėptų* (vok. *verdeckt*) norminamųjų žodynų, kuriuose tokia pozicija nuslepiama (Beuge 2016: 138). Kai kada normatyvinė žodyno kūrėjų pozicija paaiškėja per tam tikros leksikos, pvz., vulgarizmų, neįtraukimą į žodyno antraštinę (Bernal 2020).

Lietuvių kalbos EL dažniausiai laikomasi kalbinio preskriptyvizmo nuostatų (Melnikienė 2006: 110)²⁰. Pavyzdžiui, internete atsivertus Lietuvių žargono bazę (2010), iš jos atsirinkus 10 raide *b*²¹ prasidedančių žodžių (*babačka, babajus, babcė, babina, babinykas, babkės, babnikas, babos, bachūras, bacila*) ir juos įvedus į BLKŽ ir į DLKŽ matyti, jog abiejų žodynų antraštynuose randamas tik vienas žodis *bacila*, tačiau jo žargoninė reikšmė²² straipsniuose nenurodoma. Vis dėlto yra siekiama, jog BLKŽ neatsirastų „[n]eteiktinų ar vengtinų kalbos reiškinių“ (Liutkevičienė 2015: 1), o panašios pozicijos greičiausiai laikomasi ir kituose dabartiniuose lietuvių kalbos (elektroniniuose) žodynuose, todėl čia pateikiamas kalbinio preskriptyvizmo pavyzdys nėra netikėtas.

Apibendrinant galima teigti, jog norminamasis žodynas – tai toks (elektroninis arba spausdintinis) veikalas, preskriptyvinės leksikografijos produktas, kuriame atvirai arba užslėptai

²⁰ Darbo autorė tai taip pat pastebėjo iš asmeninės patirties akademiniams tikslams pasitelkiant bei tarpusavyje palyginant įvairiakalbius žodynus ir jų leksikografines tradicijas.

²¹ Ši raidė pasirinkta dėl to, kad pagal abėcėlę tai yra pirmoji raidė, kuria prasidedantys straipsniai jau pilnai sukelti į nebaigtinį BLKŽ ir dėl to galima palyginti BLKŽ bei DLKŽ antraštynus.

²² Lietuvių žargono bazėje pateikiama tokia žodžio *bacila* reikšmė: *sergantis, sloguojantis žmogus*.

laikomasi kalbinio preskriptyvizmo, „taisyklingos“ ir „prestižinės“ kalbos pozicijos. Norminamajame žodyne gali nebūtinai būti visapusiškai atspindėtas faktinis kalbos vartojimas, į antraštiną gali būti sąmoningai neįtraukiama tam tikra leksika (pvz., vulgarizmai, neologizmai) ar kai kurie kalbos vartojimo registrai. Norminamojo žodyno pavyzdžiu galima laikyti BLKŽ, apie kurį išsamiau rašoma kitame skyriuje.

1.2.1. BLKŽ koncepcija

BLKŽ skiltyje *Bendros žinios apie žodyną* žodyno redaktorė Danutė Liutkevičienė rašo²³, jog BLKŽ yra plačiai visuomenei skirtas „norminamasis leidinys, atspindintis realiai funkcionuojančią kalbos leksiką, taisyklingą bendrinės kalbos žodžių vartoseną“. Taip pat rašoma, jog žodyno antraštyne nėra „nenorminių, neteiktinų ar vengtinių žodžių“. Aptariant žodyno rengimo procesą užsimenama, jog tai buvo daroma „atsižvelgiant į nuolat kintančią kalbos padėtį, naujus žinių visuomenės bei prestižinės kalbos vartojimo poreikius, į naujausius bendrinės kalbos tyrimo, norminimo rezultatus, paskelbtą ar kitaip pareikštą kitų žodynų bei rengiamų šio žodyno tekstų kritiką“. Pabaigoje pabrėžiama, jog žodynas dar nėra baigtinis, yra vis atnaujinamas. Taigi, matyti, jog BLKŽ yra nebaigtinis elektroninis preskriptyvinės leksikografijos leidinys, kuriame norminamasis aspektas yra deklaruojamas atvirai (plg. Beuge 2016: 138). Tai pastebima per tam tikrų su kalbiniu preskriptyvizmu susijusių formuluočių (*taisyklinga bendrinės kalbos vartoseną; nenorminiai, neteiktini, vengtinai žodžiai; prestižinės kalbos vartojimas*) vartojimą minėtoje BLKŽ skiltyje. Vis dėlto darbo autorė taip pat pastebi kalbiniam **deskriptyvizmui** artimesnių formuluočių minėtame tekste: *realiai funkcionuojanti kalbos leksika; kintanti kalbos padėtis; nauji žinių visuomenės poreikiai*. Taigi, šioje vietoje vertėtų grįžti prie minties, jog kalbinis preskriptyvizmas gali būti glaudžiai susijęs su deskriptyvizmu (Engl 1993: 64): BLKŽ kūrėjai siekia, jog žodynas būtų norminis **ir** atspindėtų realią dabartinę vartoseną, o jame stengiamasi vaizduoti **tiesk** oficialiąją, **tiesk** šnekamąją kalbą (Liutkevičienė 2006) (šnekamosios kalbos pavyzdį BLKŽ žr. 2 paveiksle). Į BLKŽ siekiama įtraukti leksiką iš įvairių kalbos sluoksnių (Naktinienė 2000: 14), terminus, anksčiau nefiksuotus naujažodžius (Liutkevičienė 2015: 5-6).

²³ Aptariamą tekstą galima rasti šiuo adresu: <https://ekalba.lt/bendrines-lietuviu-kalbos-zodynas/apie/BLK%C5%BD%20Bendros%20%C5%BEinios%20apie%20%C5%BEodyn%C4%85>.

diplominis ¹	< > <i>diplominis</i> ²
diplominis²	diplòminis² dkt. (1)
diplominis darbas	šnek. DIPLOMINIS DARBAS: <i>Kadà parašýsi diplòminį?</i>
	Atnaujinta: 2021-05-20

2 pav. Šnekamosios kalbos registro BLKŽ pavyzdys.

Pereinant prie BLKŽ IL, pasak vienos iš BLKŽ rengėjų Daivos Murmulaitytės, BLKŽ iliustruojamos beveik visos lemos, išskyrus kai kuriuos vulgarizmus ar retai vartojamus žodžius²⁴ (Murmulaitytė 2011: 131-137). BLKŽ iliustracijos daugiausiai atrinktos naudojantis *Google* paieškos sistema, Dabartinės lietuvių kalbos tekstynu, Terminų banku, enciklopedijomis (pvz., VLE), tačiau kai kuriais atvejais „gyvas“ iliustracijas kūrė leksikografai (ten pat). Pasak BLKŽ vyr. redaktorės Danutės Liutkevičienės, BLKŽ iliustruojant lemas siekiama „praplėsti antraštinio žodžio apibrėžtį, suteikti daugiau informacijos apie tą žodį, parodyti vieno ar kito žodžio semantinę aplinką“, pateikti tokias konstrukcijas iliustracijose, kuriose gali būti daroma kalbos klaidų (Liutkevičienė 2015: 6). Taip pat BLKŽ IL yra kirčiuoti, juose pirmiausiai pateikiami lemą iliustruojantys žodžių junginiai ir tik po jų seka pilni sakiniai (Murmulaitytė 2011: 144-145).

Atsižvelgiant į aptartą BLKŽ koncepciją galima teigti, jog BLKŽ yra norminamasis leksikografinis veikalas, esantis elektroninėje erdvėje, kuriame stengiamasi ne tik atspindėti etaloninę oficialiąją vartoseną, tačiau aprėpti ir šnekamąją kalbą. BLKŽ pasižymi kūrėjų tikslingai atrinktais IL, kurie pateikiami tam tikra tvarka, yra kirčiuoti bei kuriais siekiama pavaizduoti taisyklingą vartoseną, vartojimo kontekstą. Tam tikros BLKŽ iliustracijos, kurias į žodyną atrinko leksikografų komanda, šiame MA darbe toliau lyginamos su ChatGPT, taigi, kompiuterio, sugeneruotomis iliustracijomis.

1.3. DI samprata ir artimos sąvokos

Plačiau visuomenei skirtame Lietuvos dirbtinio intelekto strategijos leidinyje rašoma, jog DI yra „sistemos, kurios demonstruoja protingą ir sumanų elgesį, analizuodamos savo aplinką ir darydamos gana savarankiškus sprendimus tikslui pasiekti“ (Ekonomikos ir inovacijų ministerija 2019: 2). Vis dėlto mokslo bendruomenė nėra sutarusi dėl vieno konkretaus DI apibrėžimo (Schuet 2019: 3, Djefal 2020: 2), tačiau straipsniuose apie DI dažnai minimas 1950-aisiais Allan'o Turing'o

²⁴ Straipsnyje nebuvo pateikta nei vulgarizmų, nei retos vartosenos lemu, įtrauktų į BLKŽ antraštinę. Verta atsižvelgti į tai, jog šis straipsnis publikuotas prieš trylika metų, dėl to antraštynas dar galėjo būti koreguotas.

sugalvotas **Tiuringo testas**. Tai yra procedūra, kai žmogus turi atskirti kompiuterio sugeneruotą tekstą nuo žmogaus sugeneruoto teksto. Jei žmogus negali atskirti kompiuterio darbo nuo žmogaus darbo (t. y., jei kompiuteris sėkmingai imituoja žmogų), tuomet laikoma, jog Tiuringo testas išlaikytas ir **tai vadintina DI** (Russell, Norvig 2010: 2-3, Schuet 2019: 3).

2007-aisiais John'as McCarthy'is DI apibrėžė kaip „**protingų mašinų, o ypač protingų kompiuterinių programų, mokslą ir jų kūrimą**²⁵“ (McCarthy 2007: 2). Kalbėdami apie DI, Stuart'as Russell'as ir Peter'is Norvig'as vartoja agento (angl. *agent*) sąvoką - autoriai ją supranta kaip veikiantį, besidarbuojantį dalyką (Russell, Norvig 2010: 4). Pavyzdžiui, racionalus agentas yra toks, kuris darbuojasi taip, kad gautų patį geriausią rezultatą (ten pat). Šie mokslininkai DI supranta kaip **mokslą, tiriantį užduotis atliekančius (sumanios) agentus** (Russell, Norvig 2010: viii). Svarbu paminėti, jog šios užduotys atliekamos **savarankiškai ir efektyviai** (Djeffal 2020: 2). Vis dėlto, nors ir kalbama apie dirbtinį **intelektą**, jo svarbu netapatinti su žmogaus kognityvinėmis galimybėmis, pažvelgti į DI iš neantropocentriško žvilgsnio. Johan'as Egbert'as Hans'as Korteling'as et al. rašo, jog DI yra tam tikri **nebiologiniai gebėjimai** savarankiškai ir efektyviai pasiekti kompleksiškus tikslus (Korteling et al. 2021: 2). Taigi, iš čia pateikiamų DI apibrėžimų matyti, jog jo konceptas yra itin kompleksiškas.

DI ir su juo susiję procesai pasižymi tam tikrais fundamentaliais matematiniais principais, pavyzdžiui, **algoritmais** – tam tikru matematiniu pagrindu (angl. *mathematical reasoning*), sąlygojančiu loginę dedukciją (Russell, Norvig 2010: 8). Kalbant apie DI pritaikymą apmokant kompiuterį atlikti užduotis, verta kalbėti apie **mašininį mokymąsi**: mašina treniruojama specialiai paruoštais duomenimis (Ertel 2017: 178) siekiant ją apmokyti atlikti konkrečias užduotis, pavyzdžiui, atpažinti balsą, atpažinti objektus nuotraukose (Vaišnoras 2020: 20) ar atlikti užduotis, susijusias su **natūralios kalbos apdorojimu** (Janiesch et al. 2021: 686). Natūralios kalbos apdorojimas grįstas matematiniais metodais, kurių pagalba apdorojami ir analizuojami skaitmeniniame formate esantys tekstai (Petkevičius 2021).

Vienas iš mašininio mokymosi pogrupių yra daugiasluoksnis **gilusis mokymasis** (Janiesch et al. 2021: 687), susidedantis iš **neuroninių tinklų**. Jais Antanas Verikas vadina tokia „informacijos apdorojimo struktūra, imituojanti kai kuriuos gyvųjų organizmų smegenyse vykstančius informacijos perdavimo procesus“ kaip mokymąsi, prisitaikymą ar adaptaciją (Verikas).

Šiais laikais DI pasitelkiamas daugelyje žmogų supančių sričių – transporte, kalbos technologijose, logistikoje, robotikoje (Russell, Norvig 2010: 28-29), medicinoje (Liu, Wang et al. 2023), lingvodidaktikoje (Kohnke et al. 2023) ir kitur. Taip pat esama įvairių plačiai visuomenei internete prieinamų DI įrankių, galinčių suprasti, versti kalbą, generuoti rišlius tekstus ir dar

²⁵ Pažodžiui iš anglų kalbos vertė MA darbo autorė. Originaliame tekste pavartotas būdvardis *intelligent*, kuri taip pat būtų galima išversti į *sumanus* (*sumanios mašinos, sumanios kompiuterinės programos*).

daugiau²⁶. Šiame MA darbe toliau domimasi, kaip vienas iš DI įrankių ChatGPT (žr. 1.4. ir 1.4.1. skyrius) lietuvių kalba generuoja žodyno IL.

1.3.1. DI (ChatGPT) pritaikymas leksikografijos tyrimuose

Lyg šiol esama nedaug mokslo darbų, kuriuose apžvelgiamos DI pritaikymo leksikografijoje galimybės: dažniausiai jie susiję su anglų kalbos leksikografija, o tokiu tyrimų lietuvių kalbos leksikografijos srityje dar nėra. Svarbu aptarti šiam MA darbui aktualaus įrankio ChatGPT pritaikymą generuojant žodyninius tekstus, kuris iki šiol jau panagrinėtas kai kurių mokslininkų. Michael'as Rundell'as remdamasis atliktu tyrimu rašo, jog vienareikšmiams žodžiams ChatGPT sugeneruoja aiškias definicijas (Rundell 2023: 5-7). Robert'o Lew'o (2023) tyrime paaiškėja, jog ChatGPT sugeneruotos apibrėžtys įvertintos labai panašiai į COBUILD žodyne esančias, leksikografų sukurtas apibrėžtis (Lew 2023). Gilles-Maurice de Schryver'is teigia, jog šiomis dienomis kuriant žodyną būtų itin atsainu nepasinaudoti įrankiu ChatGPT ir šį įrankį prilygina leksikografui (Schryver 2023: 377).

Vis dėlto, kai kurie mokslininkai ChatGPT kaip galimą leksikografinio darbo įrankį vertina ne taip gerai. Pavyzdžiui, įrankis susiduria su sunkumais generuodamas apibrėžtis polisemiškiems žodžiams: sugeneruoja per daug nerealių, išgalvotų reikšmių, apima ne visas žodžio reikšmes, pateikia sintaksiškai ir morfologiškai vienodus, neautentiškai skambančius iliustracinius sakinius (Rundell 2023: 5-7). Robert'o Lew'o tyrimas atskleidė, jog ChatGPT nepakankamai gerai generuoja iliustracijas (Lew 2023). Erin McKean ir Will'as Fitzgerald'as pastebi, jog generuojamuose leksikografiniuose tekstuose ChatGPT dažnai dubliuoja žodžius iš užklausos, į generuojamą tekstą kartais neįtraukia lemos, o pateikiamos leksikografinės informacijos taisyklingumas yra nevientisas (McKean, Fitzgerald 2023: 2-4). Kalbant apie pateikiamos informacijos taisyklingumą svarbu atsižvelgti į tai, jog ChatGPT ne taip gerai veikia kitų kalbų (Lai et al. 2023) ir ypač mažesnių, nepopuliarių kalbų kontekste (Bang et al. 2023) – tarp jų ir lietuvių. Taigi, matyti, jog DI generuojant leksikografinius tekstus neretai pasireiškia įvairūs trūkumai.

Straipsniuose taip pat keliamas klausimas, kaip galėtų atrodyti leksikografijos ateitis, leksikografo rolė, pasaulyje įsigalėjus įvairiems DI įrankiams. Nors su DI atėjimu leksikografijos laukas kinta, tampa tarpdiscipliniškesnis, vyrauja požiūris, jog ateityje vis tiek reikės profesionalių leksikografų įsitraukimo, t. y., DI jų nepakeis, o bus papildomas leksikografinio darbo įrankis (Lew 2024: 6-7, Schryver 2023: 380).

²⁶ Siekiant iliustruoti dabartinių DI įrankių galimybių daugialypiškumą, galima paminėti naują DI įrankį pavadinimu *Sora*, kuris geba generuoti hiperrealistiškus vaizdo klipus vien vartotojo pateikiamos tekstinės užklausos (angl. *text-to-video*) pagalba. Plačiau apie *Sora* skaityti čia: <https://www.euronews.com/next/2024/02/18/sora-openais-new-text-to-video-tool-is-causing-excitement-and-fears-heres-what-we-know-abo>.

Apibendrinant galima pastebėti, jog ChatGPT tikrai galima pasitelkti leksikografinio darbo procesuose: anglų kalbos kontekste įrankis neretai pateikia kokybiškas apibrėžtis vienareikšmiams žodžiams. Taip pat verta kelti klausimus apie tolimesnį šio įrankio leksikografinį potencialą. Vis dėlto dėl įvairių įrankio trūkumų ir generuojamų netikslumų tikėtina, jog greitu metu šis įrankis leksikografų žodynų sudaryme nepakeis. Toliau darbe siekiama apžvelgti, kaip ChatGPT geba generuoti leksikografinius tekstus (konkrečiai – IL) lietuvių kalba.

1.4. Įrankis *ChatGPT*, jo galimybės ir trūkumai

Dėl išpūdingai vertinamų galimybių generuoti įvairius tekstus ir dialogą imituojančios vartotojo sąsajos (angl. *interface*), įrankis ChatGPT yra vienas daugiausiai dėmesio susilaukusių didžiųjų kalbos modelių²⁷ (angl. *large language model*, LLM) (Lai et al. 2023, Bang et al. 2023). ChatGPT priklauso bendrovei *OpenAI* ir plačiai visuomenei tapo prieinamas 2022-aisiais. Šis įrankis apmokytas įvairaus pobūdžio teksta (Bang et al. 2023) iš *Common Crawl* duomenų bazės (Schryver 2023: 357). Tekstai, kuriais apmokytas ChatGPT, pagrinde yra angliški (Weber-Wulff et al. 2023: 3), o prireikus tiesiogiai verčiami į kitas didesnes kalbas (pvz., kinų, japonų) (Kohnke et al. 2023: 545). Programos apmokymo mastas itin platus: įrankio GPT-3.5. versija²⁸ apmokyta maždaug 45 terabaitais tekstinių duomenų (Lin 2023: 2), o tai lemia ChatGPT gebėjimą generuoti įvairius tekstus (Schryver 2023: 357). Be gebėjimo generuoti įvairaus pobūdžio tekstus, programinius kodus (Lin 2023: 1), įrankis geba taisyti kalbos klaidas, anotuoti tekstus, pateikti vertimus, žodynines žodžių apibrėžtis ir kt. (Kohnke et al. 2023: 539).

Šis įrankis taip pat turi tam tikrų trūkumų: generuojamuose tekstuose gali būti daugžodžiaujama, pateikiama netiksli informacija, ChatGPT nenurodo šaltinių, necituoja, taigi, neįmanoma atsekti tekstų kilmės (Kohnke et al. 2023: 544-545). Dėl programos apmokymo daugiausiai angliškais teksta, ChatGPT įrankis yra vienakultūris (angl. *monocultural*) (Alm, Watanabe 2023: 20). ChatGPT taip pat susiduria su sunkumais bandant suprasti užklausa, įvairius etikos klausimus, remiantis statistiniais principais gali pateikti absurdiškus atsakymus (Taecharungroj 2023: 2). Šioje vietoje svarbu paminėti tekstų generavimo procesuose pasitaikančias haliucinacijas (Schryver 2023: 358, Ji et al. 2023). Ziwei Ji et al. haliucinacijas²⁹ tekste supranta kaip išpūdį, jog

²⁷ Linas Petkevičius aiškina, jog „[k]albos modelis yra toks matematinis modelis, kuris užkoduoja tam tikrą tekstą į skaitinę reprezentaciją – skaičių eilutę, turinčią užkoduotą semantinę ir kontekstinę informaciją“ (Petkevičius 2021).

²⁸ GPT-3.5 versiją darbo autorė naudojo savo tyrime: ji yra nemokama, tačiau esama ir mokamų GPT-4, GPT-4o versijų, kurios apmokytos dar didesniu kiekiu duomenų, turi daugiau naudojimo funkcijų ir galimybių.

²⁹ Mokslo bendruomenėje diskutuojama dėl šio termino etiškumo: kadangi haliucinacijos yra medicininis konceptas, su kuriuo susiduria įvairiomis ligomis sergantys žmonės, keliamas klausimas, ar šis ChatGPT ypatumo įvardijimas neprisideda prie stigmos (plg. Maleki et al. 2024).

tekstas yra sklandus, nors jo turinys neatitinka tikrovės ir yra absurdiškas (angl. *nonsensical*), o teksto su haliucinacijomis iš pirmo žvilgsnio galima ir neatpažinti (Ji et al. 2023: 248:3).

Apibendrinant būtų galima sakyti, jog įrankis turi tiek įvairių tekstų generavimo privalumų, tiek ir trūkumų. Verta turėti omenyje, jog įrankis nuolat kinta ir tobulėja, atsiranda vis naujų jo versijų (žr. 28 išnašą), o galimų atlikti užduočių sąrašas vis plečiasi, dėl to šiame MA darbe apžvelgiamos ChatGPT galimybės generuoti IL lietuvių kalba ateityje neabejotinai gali kisti.

1.4.1. ChatGPT užklauso ir jų generavimo ypatumai

ChatGPT programai pateikiamos užduočių / klausimų formuluotės vadinamos užklausomis (angl. *prompts*). Užklauso generavimas (angl. *prompt engineering*) yra naujai atsiradusi metodika, kuria siekiama kuo geriau formuluoti įvairiems kalbų modeliams (pvz., ChatGPT) pateikiamas užklauso (Giray 2023: 2629). Užklauso generavimas yra neatsiejama darbo su kalbų modeliais dalis (Schryver 2023: 377). Pasak Louie'o Giray, esama skirtingų užklauso pateikimo būdų: jas įrankiui galima pateikti kaip instrukciją, kaip klausimą, prašyti užbaigti mintį ar atsižvelgti į konkretų kontekstą ir pateikti atitinkamą atsakymą ir kt. (Giray 2023: 2631). Norint gauti kuo kokybiškesnį atsakymą, reikia pateikti kuo ilgesnę (Schryver 2023: 358) užklausą, ji turėtų būti kompleksiška (Lin 2023: 2). Užklausoje reikėtų nurodyti kuo daugiau konkrečių kriterijų, kurie turėtų būti sugeneruotame atsakyme (pvz., prašoma konkrečiai paaiškinti kokio nors koncepto **privalumus** ir **trūkumus**, o ne tiesiog paaiškinti, **kas** tas dalykas yra), pateikti kuo daugiau detalių, reikiamo konteksto (Giray 2023: 2631-2632).

Vis dėlto, net jei įrankiui pateikiamos tos pačios užklauso, vartotojui bus pateikiamas vis skirtingas tekstas, taigi, šia prasme įrankis yra nenuspėjamas (angl. *non-deterministic*) (Schryver 2023: 358). Dėl to ypač svarbu kritiškai vertinti įrankio pateikiamus rezultatus (Giray 2023: 2633), net ir jam pateikus kokybišką užklausą.

Taigi, siekiant gauti kuo kokybiškesnį ir informatyvų atsakymą, vartotojui svarbu iš savo pusės užtikrinti kuo kokybiškesnę užklausą: tai nulemia žodinė jos raiška, ilgis ir kompleksškumas, priklausomai nuo įrankiui pateikiamos užduoties ar klausimo. Šio darbo 2.1.2. skyriuje autorė aprašo užklauso formulavimo procesą savo tyrime.

2. ChatGPT sugeneruotų ir BLKŽ esančių IL tyrimo ir rezultatų aprašymas

2.1. Tyrimo metodika

Siekiant atlikti tyrimą, pirmiausia sudarytas lemų sąrašas, kuriame esančioms lemoms pasitelkus specialią užklausą ChatGPT vėliau prašyta sugeneruoti iliustracinius sakinius. Toliau aprašomi šio sąrašo sudarymo žingsniai.

2.1.1. Lemų sąrašo sudarymas

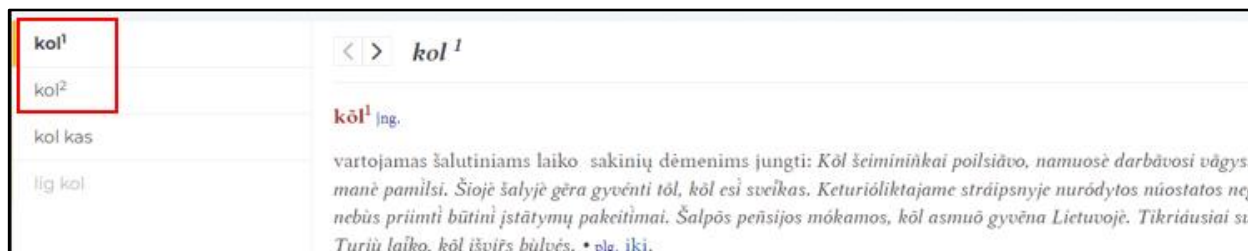
Pirmiausia autorė pasirinko surinkti lemas, prasidedančias raide *k*. BLKŽ, su kuriame esančiais IL toliau lyginami ChatGPT sugeneruoti IL, dar nėra baigtas, yra nuolat pildomas naujais straipsniais³⁰, o raidės *k* atkarpa jau yra pilnai parengta. Taigi, tolimesniam sugeneruotų IL palyginimui su BLKŽ pasirinkta tokia raidė, kurios visi straipsniai jau yra sukelti į BLKŽ.

Lemų atrankai pasitelktas Virginijaus Dadurkevičiaus 2020 m. sudarytas dažniausių 3000 lietuvių kalbos žodžių sąrašas (toliau – dažninių žodžių sąrašas). Darbo autorės žiniomis, tai yra naujausias lietuvių kalbos dažniausių žodžių sąrašas. Iš dažninių žodžių sąrašo pagal dažnumą atrinktos 40 lemų, prasidedančių raide *k*. Siekiant kiek galima labiau subalansuoti kalbos dalių lemų sąrašą skaičių ir įvairovę ir numaćius, jog į sąrašą pateks 40 lemų, pirmiausia nuspręsta atrinkti po šešias atvirų klasių kalbos dalis (daiktavardžius, būdvardžius, veiksmažodžius,rieveksmius), numaćius, jog sąrašą bus ir ribotas skaičius uždarų klasių kalbos dalių (jungtukų, dalelyčių,rieveksmių). Surinkus visas kalbos dalis (pats paskutinis pagal dažnumą į sąrašą atrinktas būdvardis), iš dažninių žodžių sąrašo parinkti trys kiti žodžiai nepaisant jų kalbos dalies tam, kad lemų sąrašo skaičius pasiektų 40 lemų. Galutinį lemų sąrašą sudaro 8 daiktavardžiai, 7 įvardžiai, 6 veiksmažodžiai, 6 būdvardžiai, 5 jungtukai, 5rieveksmiai, 2 skaitvardžiai ir dalelytė. Kaip matyti, kalbos dalių įvairovės nepavyko visiškai subalansuoti, nes kai kurių kalbos dalių dažnumo sąrašą (taigi, kalboje) pasitaiko daugiau (pvz., daiktavardžių, įvardžių), o kai kurių – žymiai mažiau (pvz., skaitvardžių, dalelyčių), turint omenyje, kad be atvirų ir uždarų žodžių klasių šio sąrašo atranka dar nulemta ir raidės *k*.

³⁰ Šią informaciją galima rasti BLKŽ internetiniame puslapyje, gelsvame lauke virš abėcėlės:

<https://ekalba.lt/bendrinės-lietuvių-kalbos-zodynai/>. Pilkai pateikiamos raidės (pvz., A, Ą, I ir t.t.) yra tokios, kurių straipsniai (lemų, prasidedančių atitinkama raide) dar nėra baigtiniai. Juodai pateikiamos raidės (pvz., B, C, Ć, tarp jų ir K ir t.t.) yra tokios, kurių straipsniai jau sukelti visi.

Taip pat pastebėta, jog lemų sąrašė du kartus atsidūrė lema *kaip*, o tai yra homonimijos atvejis: BLKŽ pirmu numeriu pažymėtame ŽS *kaip* yrarieveiksmis, o antru numeriu pažymėtame ŽS *kaip* yra jungtukas³¹. Minėtina, jog kitais homonimijos atvejais lemos kalbos dalis patikslinta remiantis BLKŽ: pavyzdžiui, pasak BLKŽ, sąrašė esantis žodis *kuo* gali būti tiek dalelytė, tiek jungtukas, tiek įvardis. Tokiais atvejais žodis suklasifikuotas pagal tą kalbos dalį, kurios straipsnis BLKŽ pažymėtas skaičiumi 1 (žr. 3 paveikslą). Kaip žinoma, homonimų straipsniai žodynuose įprastai numeruojami / jų reikšmės numeruojamos atsižvelgiant į vartojimo dažnumą.



3 pav. Homonimų numeravimas BLKŽ.

Toliau pateikiamas pagal dažninių žodžių sąrašą autorės sudarytas lemų sąrašas, kuris pasitelktas toliau aprašomame MA darbo tyrime (žr. 1 lentelę). Sąrašas pateikiamas tokia tvarka, kokia atrinktas iš dažninio žodžių sąrašo, t. y., pagal žodžių vartojimo dažnumą, o skliausteliuose nurodyta žodžio kalbos dalis, nustatyta remiantis BLKŽ:

³¹ BLKŽ yra skiriami trys lemos *kaip* straipsniai: turint omenyje tai, jog pirmu numeriu pažymėtasrieveiksmio *kaip* straipsnis, antru numeriu – jungtuko *kaip*, o trečiu – dalelytės *kaip*, galima suprasti, jog dalelytė *kaip* kalboje pasitaiko rečiau nei jungtukas *kaip* arrieveiksmis *kaip*, dėl to **dalelytė** *kaip* nepasitaikė V. Dadurkevičiaus sąrašė.

1. Kad (<i>jungtukas</i>)	21. Kartu (<i>prieveiksmis</i>)
2. Kuris (<i>įvardis</i>)	22. Kilti (<i>veiksmazodis</i>)
3. Kitas (<i>įvardis</i>)	23. Kodėl (<i>prieveiksmis</i>)
4. Kaip (<i>prieveiksmis</i>)	24. Keturi (<i>skaitvardis</i>)
5. Kas (<i>įvardis</i>)	25. Keisti (<i>veiksmazodis</i>)
6. Kai (<i>jungtukas</i>)	26. Kelti (<i>veiksmazodis</i>)
7. Koks (<i>įvardis</i>)	27. Kurti (<i>veiksmazodis</i>)
8. Komisija (<i>daiktavardis</i>)	28. Kreiptis (<i>veiksmazodis</i>)
9. Kartas (<i>daiktavardis</i>)	29. Kadangi (<i>jungtukas</i>)
10. Klausimas (<i>daiktavardis</i>)	30. Karinis (<i>būdvardis</i>)
11. Komitetas (<i>daiktavardis</i>)	31. Ketvirtas (<i>skaitvardis</i>)
12. Kalbėti (<i>veiksmazodis</i>)	32. Konkretus (<i>būdvardis</i>)
13. Kur (<i>prieveiksmis</i>)	33. Konstitucinis (<i>būdvardis</i>)
14. Kaina (<i>daiktavardis</i>)	34. Kuo (<i>dalelytė</i>)
15. Kiekvienas (<i>įvardis</i>)	35. Komercinis (<i>būdvardis</i>)
16. Keli (<i>įvardis</i>)	36. Keistas (<i>būdvardis</i>)
17. Komanda (<i>daiktavardis</i>)	37. Karštas (<i>būdvardis</i>)
18. Kiek (<i>prieveiksmis</i>)	38. Kažkoks (<i>įvardis</i>)
19. Kaip (<i>jungtukas</i>)	39. Korupcija (<i>daiktavardis</i>)
20. Kol (<i>jungtukas</i>)	40. Kanceliarija (<i>daiktavardis</i>)

I lent. Remiantis V. Dadurkevičiaus dažniniu žodžių sąrašu sudarytas raide *k* prasidedančių lemu sąrašas.

2.1.2. Užklauso formulavimas ir bandomasis tyrimas

Sudarius lemu sąrašą, tolimesnis žingsnis buvo suformuluoti užklausa, kurią pateikus programai ChatGPT, būtų gaunami kiekvienai lemai sugeneruoti vartojimo pavyzdžiai tinkami žodyno IL sudaryti. Siekdama suformuluoti užklausa, darbo autorė rėmėsi mokslininko Gilles-Maurice de Schryver'io angliškai suformuluota užklausa, kurią jis panaudojo savo žvalgomajam tyrimui su ChatGPT įrankiu, siekdamas apžvelgti įrankio ŽS generavimo anglų kalba galimybes.

Minėta užklausa internete pateikiama prie G.-M. de Schryver'io straipsnio *Generative AI and Lexicography: The Current State of the Art Using ChatGPT* (2023) priede. Autorė pasirinko būtent šią užklausą (žr. 4 paveikslą), nes ji yra viena iš nedaugelio internete skelbiamų užklausų, susijusių su leksikografijos ir ChatGPT sietimi.

PROMPT: Please give me a COBUILD-style dictionary entry for '[HEADWORD]'. Start with the frequency information at the top. Each sense should be in a numbered block. Each block then starts with the part of speech and the morphological forms of the respective sense. This is followed by a sense definition and sense examples that illustrate both the use and the meaning of each particular sense. For the example sentences, make sure to use different sentence structures, referring to different people; refer to past, present, and future situations; vary long and short example sentences; and include other elaborations so that you, dear ChatGPT, produce high-quality dictionary material for me.

4 pav. Gilles-Maurice de Schryver'io žvalgomajam tyrimui suformuluota užklausa.

Kadangi šiame MA darbe siekiama ištirti būtent ChatGPT sugeneruotus IL, autorė užklausoje toliau telkėsi į vartojimo pavyzdžius, iliustruojančius lemos reikšmę ar reikšmes (polisemijos atvejais). 2023 m. lapkričio mėn. buvo atliktas bandomasis tyrimas, kurio metu pateikta ši **pirminė** užklauskos versija dešimčiai lemuų iš minėto lemuų sąrašo:

Prašau paaiškinti žodį „X³²“ taip, kaip žodyne. Jei yra kelios šio žodžio reikšmės, prašau jas aiškiai išskirti ir apibrėžti. Prašau kiekvienai žodžio apibrėžčiai pateikti įvairaus turinio gerai reikšmę iliustruojančius pavyzdžius.

Apžvelgusi įrankio sugeneruotus IL, autorė pastebėjo, jog užklausą dar reikės patikslinti: visuose dešimt sugeneruotų atsakymų lemos buvo pateikiamos keturios arba penkios reikšmės. Patikrinus atitinkamų žodžių pateikimą BLKŽ paaiškėjo, kad toks ChatGPT reikšmių eksplikavimas yra **perteklinis**. Pavyzdžiui, BLKŽ lemai *kad* pateikiama vienintelė reikšmė, o ChatGPT jas skiria keturias su labai panašiomis arba neaiškiomis apibrėžtimis: 1) "**Kad**"³³ gali būti naudojamas išreikšti priežastinį ryšį, nurodyti priežastį arba pagrindą; 2) "**Kad**" gali reikšti tikslą arba pasekmę, nurodyti siekiamą rezultatą; 3) "**Kad**" gali būti naudojamas išreikšti prielaidą arba sąlygą; 4) "**Kad**" taip pat gali būti naudojamas nurodyti laiką, kuomet įvyksta arba įvyks tam tikras įvykis. Kitu atveju vienam žodžiui buvo išskirtos kelios reikšmės, kurioms pateikti panašūs vartojimo pavyzdžiai, pavyzdžiui, pagal užklausą žodžiui *keistas* ChatGPT sugeneravo keturias reikšmes ir pateikė keturis vartojimo pavyzdžius: 1) *Jo elgesys yra keistas; jis dažnai kalba vienas pats ir nesidomi kitų nuomone;* 2) *Jos drabužiai buvo spalvingi ir puošnūs, bet kažkaip keisti;* 3) *Jo muzikos skonis yra gana keistas, bet jis tikrai įdomus;* 4) *Savaime šviečiantis gamtos reiškinytis buvo tikrai keistas ir stebuklingas.* Neretai ChatGPT pateikdavo ir papildomą lemos „reikšmę“, pavadinimu *Kita*, o šiai pseudoreikšmei

³² Užklausoje vietoje „X“ buvo įrašomas žodis iš atrinkto lemuų sąrašo, t. y., ši užklausa ChatGPT programai pateikta dešimt kartų su vis skirtinga lema.

³³ Čia ir toliau pavyzdžiuose paryškina lema arba jos forma, kuriai iliustruoti buvo sugeneruotas sakinytis.

nepateikdavo jokio vartojimo pavyzdžio, tik pabrėždavo konteksto svarbą, pvz., *Kita: Apibrėžimas: "Ketvirtas" gali turėti įvairių nuansų³⁴, priklausomai nuo konteksto, bet pagrindinė prasmė yra numeris*. Taigi, akivaizdu, kad įrankis pateikia perteklinį tekstą generuojamuose straipsniuose. Kitas svarbus pastebėjimas yra toks, kad reikšmės eksplikacija nei vienu atveju nesutapo su atitinkamo žodžio eksplikacija BLKŽ. Pastebėtina, jog ChatGPT generuojamas reikšmių eksplikavimas tiesiogiai daro įtaką pateikiamiems iliustraciniams pavyzdžiams, kurie yra šio darbo objektas.

Grįžtant prie bandomojo tyrimo metu ChatGPT sugeneruotų vartojimo pavyzdžių, galima teigti, jog pasitaikė įvairios kokybės pavyzdžių. Kai kurie sugeneruoti vartojimo pavyzdžiai buvo negramatiški (*Ji kalba penkias kalbas ir lengvai bendrauja su įvairiatautis žmonėmis*). Taip pat pasitaikė nemažai haliucinacijų, jose pasitaikė netikri žodžiai (toliau pabraukti): *Mano automobilis pradėjo kreiptis į dešinę be mano įsikišimo; Jo veiksmai kalba geresnį negu žodžiais išsakomą dėkingumą; Šis maistas skonis kaip mamos virtinėje paruoštas; Aš ruošiuosi lauke, kad nenusinigtų, nes žinau, kad lauksime ilgai; Ji šypsosi kaip saulė, spindinčia visoje jos gerumo šviesoje*. Galima pastebėti ryškią anglų kalbos įtaką generuojamiems sakiniam. Pavyzdžiui, sakinyje *Ji mokosi kietai, kad išlaikytų egzaminą su puikiais rezultatais* žodžių junginį *mokosi kietai* būtų galima palyginti su anglų kalbos posakiu *to study hard³⁵* ir tai laikyti pateikiamu vertiniu. Kai kuriais atvejais lema išvis nepasitaikė vartojimo pavyzdyje, pavyzdžiui, lemai *kalbėti* sugeneruotame IL: *Jis nori aktyviai dalyvauti diskusijoje ir akcentuoti savo poziciją* ar lemai *klausimas* sugeneruotame IL: *Kokia tavo mėgstamiausia knyga?* – matyti, jog šiuo vartojimo pavyzdžiu aiškinamas lemos konceptas, o ne iliustruojama jo reikšmė. Taigi, tokiais atvejais būtų galima kalbėti apie nekokybiškai sugeneruotus pavyzdžius. Minėtina, jog kai kada negramatiškumo, haliucinacijų aspektai ir kiti generuojamų IL trūkumai persipina tarpusavyje – tame pačiame sakinyje gali būti tiek išgalvojami faktai, tiek sunkiai suprantama jo semantika, tiek pasitaikyti negramatiškumas (pvz., *Pardavėjas gaus 5% komisijos už kiekvieną parduotą prekę*).

Vis dėlto dalis pavyzdžių buvo kokybiškesni, pvz., *Po vakarienės mes žiūrėsime kitą filmą; Aš laikiu tavo skambučio, kad galėčiau tau pasakyti svarbią naujieną; Jei jaučiatės nesaugiai, kreipkitės į policiją arba kitus saugumo specialistus; Jei šis planas nesiseks, turime ieškoti kitų variantų* ir kt. ChatGPT sugeneravo ir tokių vartojimo pavyzdžių, kurie buvo gramatiškai taisyklingi, tačiau skambėjo ne itin natūraliai, nors reikia atsižvelgti į tai, jog generuojamos kalbos natūralumą yra sudėtinga objektyviai įvertinti³⁶. Tai padaryti sunku dėl generuojamos kalbos konteksto trūkumo,

³⁴ Pastebėtina, jog autorė nei šioje vietoje, nei kitose darbo vietose netaisė ChatGPT pateikiamų kalbinių klaidų, kurių šiam įrankiui rašant lietuviškai pasitaikė / pasitaiko neretai.

³⁵ Plačiau apie žodžių junginį skaityti Collins'o žodyne: <https://www.collinsdictionary.com/dictionary/english/study-hard>.

³⁶ Kai kurie tyrimai rodo, jog žmonėms sunku atskirti DI sugeneruotus tekstus nuo žmogaus parašytų tekstų (Tlok 2023: 3). Vis dėlto turint omenyje tai, jog ChatGPT įrankiui apmokyti 93% medžiagos pasitelkta anglų kalba (Weber-

neilgų, izoliuotų sakinių: *Jo veiksmai man kelia daug klausimų dėl jo tikslų; Jis nori kalbėti apie savo kelionę ir dalintis patirtimi su visais; Finansinė ataskaita rodo, kad įmonė gerai veikė ketvirtąjį ketvirtį; Dalijamės pyragu, ir kiekvienas gaus ketvirtą dalį* ir kt. Taigi, bandomojo tyrimo rezultatai parodė, jog ChatGPT teikiamą užklausą reikia patikslinti, suformuluoti neperkraudant jos žodžiais, norint gauti kuo aukštesnės kokybės iliustracijas.

Atsižvelgusi į bandomojo tyrimo rezultatus, autorė toliau suformulavo **galutinę užklausą**, pateiktą ChatGPT renkant MA darbo duomenis:

Paaiškink žodį „X“ taip, kaip žodyne. Paaiškink pagrindines šio žodžio reikšmes, jeigu jos yra kelios. Jeigu yra tik viena šio žodžio reikšmė, tuomet paaiškink tik ją. Kiekvienai žodžio reikšmei iliustruoti pateik kuo daugiau ir kuo įvairesnių vartojimo pavyzdžių. Vartojimo pavyzdžiuose turi būti skirtingi laikai.

Siekiant užklaustos neapkrauti nereikalingais žodžiais, atsisakyta žodžio *prašau* ir vartota liepiamoji antro asmens nuosaka. Lyginant su pirmine užklaustos versija, pirmasis užklaustos sakinytis iš esmės nepakeistas, siekiant gauti leksikografinį, aiškinamąjį tekstą. Sakinytis iš bandomosios užklaustos *Jei yra kelios šio žodžio reikšmės, prašau jas aiškiai išskirti ir apibrėžti* galutinėje užklaustos versijoje suskaidytas į du atskirus sakinius: įrankiui nurodyta paaiškinti visas žodžio reikšmes, **jeigu** jos yra kelios (pvz., remiantis BLKŽ, tai galėtų būti lemos *kuris, kitas, kaip*), o jeigu yra tik viena reikšmė (remiantis BLKŽ, tai galėtų būti *kad, kai, kartas*), tuomet paaiškinti **tik** ją. Toks nurodymas pateiktas norint užkirsti kelią pertekliniam reikšmių eksplikavimui. Paskutiniuose dviejuose užklaustos sakiniuose suformuluota užduotis pateikti IL reikalingą informaciją. Juose tiesiogiai remiamasi Gilles-Maurice de Schryver'io užklaustos pabaiga *For the example sentences, make sure to use different sentence structures, referring to different people; refer to past, present, and future situations; vary long and short example sentences ...*, tačiau ją kiek supaprastinant. Autorės užklausoje prašoma pateikti didesnę vartojimo pavyzdžių kiekį ir įvairovę. Bandydama įvairias užklausių formuluotes paleisti programoje ChatGPT autorė pastebėjo, jog įrankis nesuprato žodžių *veiksnys* ar *subjektas* užklausoje. Taip pat buvo mėginta suformuluoti prašymą vartojimo pavyzdžiuose referuoti į skirtingus asmenis, bet užklausoje šio aspekto nepavyko suformuluoti taip, jog įrankis tai suprastų. Dėl to užklaustos pabaigoje įrankio prašyta pavyzdžiuose atspindėti tik skirtingus laikus³⁷.

Wulff et al. 2023: 3), DI sugeneruotuose sakiniuose daugeliu atveju galima jausti anglų kalbos įtaką, kuri pavyzdžiams suteikia kalbinį *dirbtinumą*.

³⁷ Darbo autorei suprantama, jog galimos įvairios užklaustos formuluotės, užklausoje galima įtraukti daugiau detalių, pateikti pavyzdžių ir kt. Tai būtų galima išmėginti tolimesniuose ChatGPT kaip leksikografinio įrankio tyrimuose.

2.1.3. Kiti tyrime pasitelkti metodai

Norint suskirstyti ChatGPT sugeneruotus iliustracinius sakinius į atitinkamas kategorijas, buvo sukurta internetinė apklausa, taigi, į tyrimą įtrauktas kokybinis žiūros taškas. Visi ChatGPT sugeneruoti sakiniai buvo surašyti apklausos formoje *Google Forms* platformoje, o apklausos instrukcijoje respondentų prašyta pažymėti vieną iš dviejų po kiekvienu sugeneruotu sakiniu esančių laukų. Respondentui galėjo pažymėti laukelį **Natūraliai**, jei sakinys jam skambėjo natūraliai / įprastai, taip, kaip galėtų pasakyti pats ar įsivaizduoja, jog galėtų pasakyti kitas žmogus. Taip pat respondentas galėjo pasirinkti laukelį **Dirbtinai**, jei sakinys jam skamba nenatūraliai, kėlė abejonių (apklausos formos pavyzdį žr. 5 paveiksle). Darbo autorė kiekvienam respondentui nusiuntė apklausos nuorodą ir neatskleidė, jog visi šie sakiniai yra sugeneruoti DI, nenorėdama galimai paveikti respondentų pasirinkimų. Apklausa užpildė **devyni** respondentai: visiems respondentams lietuvių kalba yra gimtoji, beveik visi jie įgiję aukštąjį išsilavinimą (viena respondentė šio MA darbo rašymo metu dar siekia aukštojo koleginio išsilavinimo). Respondentų amžius įvairus – nuo 24-erių iki 49-erių metų.

Ar sakinys tau skamba natūraliai, ar nenatūraliai?

Žemiau pateikti 147 sakiniai lietuvių kalba. Atidžiai perskaityk kiekvieną sakinį. Prie kiekvieno sakinio pažymėk **NATŪRALIAI**, jei sakinys tavo nuomone skamba natūraliai (taip, kaip pasakytum pati / pats ar kaip pasakytų kitas žmogus) arba **DIRBTINAI**, jei jis tau skamba nenatūraliai, dirbtinai. **Galima pasirinkti tik vieną variantą.** Labai ačiū už tavo laiką!

Ji miegojo, kad pailsėtų po ilgos kelionės. *

Natūraliai

Dirbtinai

Jis nevalgė, kad numestų svorio. *

Natūraliai

Dirbtinai

Aš nusipirkau naują automobilį, kad lengvai važiuotų į darbą. *

Natūraliai

Dirbtinai

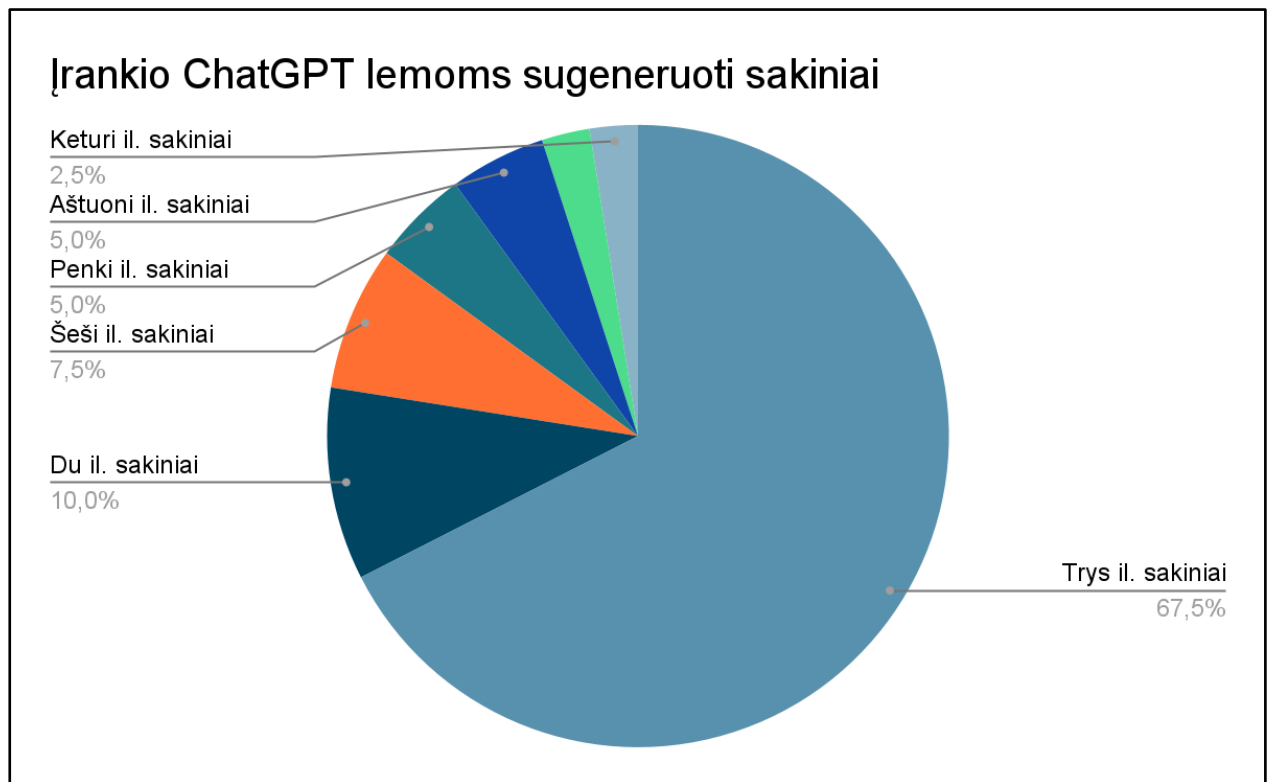
5 pav. Respondentams pateiktos apklausos instrukcijos ir pasirinkimų pavyzdys.

2.2.1.3. darbo skyriuje ChatGPT sugeneruoti sakiniai analizuojami sintaksiniu požiūriu, remiantis Vito Labučio knygoje *Lietuvių kalbos sintaksė* (1998) aprašomais įvairiais sintaksiniais kriterijais ir terminologija. 2.2.1.2. skyriaus pabaigoje ir keliose darbo vietose toliau keldama klausimą, kodėl daliai respondentų tam tikri autentiškai skambančių sakinių kategorijai priskirti pavyzdžiai kėlė abejonių dėl savo skambėjimo, autorė pasitelkia tekstynų prieigą. Sketch Engine esančiame lietuvių kalbos tekстыne ELEXIS Lithuanian Web 2021 buvo ieškota, ar kai kuriuose sugeneruotuose sakiniuose pasitaikančios formuluotės, žodžių junginiai tekстыne (taigi, pačioje kalboje) pasitaiko dažnai / retai / visai nepasitaiko ir kurios alternatyvios formuluotės pasitaiko dažniau / rečiau. Darbo autorės žiniomis, šis tekstynas yra plačiausios apimties lietuvių kalbos tekstynas Sketch Engine įrankyje (jame yra 846 563 251 žodžių), dėl to jis buvo pasirinktas.

2.2. ChatGPT sugeneruotų ir BLKŽ esančių IL tyrimo rezultatai

2.2.1. Bendrai apie ChatGPT sugeneruotus IL

Įrankis ChatGPT pagal pateiktą užklausą keturiasdešimčiai nagrinėjamų lemu iš viso sugeneravo 147 iliustracinius sakinius. Kiekvieną lemą ChatGPT iliustravo 2-10 sakiniais: 27 lemons (67,5% visų lemu) iliustruoti ChatGPT pateikė po tris sakinius, keturios lemos (*kitas, kartas, kažkoks, korupcija*) (10% lemu) iliustruotos dviem sakiniais, trys lemos (*karinis, ketvirtas, keistas*) (7,5% lemu) iliustruotos šešiais sakiniais, dvi lemos (*kilti, kadangi*) (5% lemu) iliustruotos penkiais sakiniais, kitos dvi (*komercinis, karštas*) – aštuoniais sakiniais, galiausiai lemai *kanceliarija* (2,5% lemu) pateikti keturi, o lemai *kuo* – dešimt vartojimo pavyzdžių (antroje lentelėje pažymėta ryškiai žalia spalva) (žr. 6 paveikslą):



6 pav. ChatGPT lemomis iliustruoti sugeneruotų sakinių pasiskirstymas.

Taigi, dažniausiai ChatGPT sugeneravo **tris** iliustracinius sakinius vienai lemai, o rečiausiai – keturis arba dešimt sakinių. Toks platus sugeneruotų iliustracinių sakinių kiekio svyravimas yra netikėtas, nes įrankiui pateiktoje užklausoje buvo prašoma pateikti **kuo daugiau** iliustracinių pavyzdžių kiekvienai lemai. Rezultatai rodo, jog, pavyzdžiui, keturios lemos iš sąrašo iliustruotos tik 2 sakiniiais, o tai nėra daug iliustracinių pavyzdžių, tačiau trys lemos iš sąrašo iliustruotos šešiais sakiniiais, o tai galima laikyti plačiu IL. Palyginimui galima pateikti konkretų būdvardžių pavyzdį: lemai *karinis* iš viso sugeneruoti šeši pavyzdžiai, lemai *komercinis* – aštuoni pavyzdžiai, o lemai *konstitucinis* sugeneruoti trys sakiniai³⁸. Iš pavyzdžio matyti, jog tos pačios kalbos dalies žodžiams sugeneruojamas skirtingas iliustracinių pavyzdžių kiekis. Taip pat pastebėta, jog kai kurie tą pačią lemą iliustruojantys pavyzdžiai yra labai panašūs. Pavyzdžiui, trys lemą *klausimas* iliustruojantys pavyzdžiai iliustruoja tą pačią kreipimosi reikšmę (čia ir toliau žr. nuo 1 iki 81 sunumeruotus sakinius 2.2.1.1. skyriuje: 12, 13, 14). Lemą *kadangi* iliustruojančių sakinių (3, 4, 5, 6, 7) sintaksinė struktūra yra vienoda, nes jungtukas *kadangi* sakiniuose yra toje pačioje vietoje. Todėl kyla klausimas, ar tokia generuojama pasikartojanti informacija yra naudinga ir ar kai kada nėra sugeneruojama per daug

³⁸ ELEXIS Lithuanian Web 2021 duomenimis, dažniausiai iš šių būdvardžių pasitaiko *konstitucinis* (10 419 pavyzdžių), rečiausiai – *komercinis* (3832 pavyzdžių), nors pastarajam pateikti aštuoni pavyzdžiai, o *konstitucinis* iliustruotas trimis sakiniiais. Taigi, matyti, jog ChatGPT sugeneruotas pavyzdžių skaičius nepriklauso nuo žodžių dažnumo, t. y., dažnesniems žodžiams negeneruojama daugiau pavyzdžių.

ilustracinių sakinių vienai lemai. Dėl to galima teigti, jog toks skirtingas ChatGPT pateikiamas iliustracinių sakinių skaičius lemoms nėra paaiškinamas ir yra atsitiktinis. Toliau pateikiama autentiškai skambančių ChatGPT sugeneruotų iliustracinių sakinių lingvistinė analizė.

2.2.1.1. Autentiškai skambantys sakiniai

Peržiūrėjus ChatGPT sugeneruotus IL nustatyta, jog dalis ChatGPT sugeneruotų sakinių yra tokie, kokius autorė vadina autentiškai skambančiais. Iš viso suskaičiuotas **81** autentiškai skambantis sakinytis, t. y., **~55 %** visų sugeneruotų iliustracinių sakinių. Tam, jog sakiniai patektų į autentiškai skambančių sakinių kategoriją, autorė išskyrė tokius penkis **kriterijus**: 1) sakiniai turi būti logiški ir rišlūs; 2) juose neturi trūkti esminės informacijos, leidžiančios lengvai suprasti sakinio esmę, t. y., sakinytis turi būti lengvai suprantamas; 3) sakiniai turi būti morfologiškai, sintaksiškai, leksiškai ir ortografiškai (taigi, visapusiškai kalbine prasme) taisyklingi; 4) sakiniai turi skambėti autentiškai, t. y., negalima lengvai atskirti, ar šiuos sakinius sugeneravo DI įrankis, ar tam tikrame kontekste galėjo produkuoti lietuvių kalbos gimtakalbis. Sakinių autentišką skambėjimą nustatyti padėjo devynių gimtakalbių respondentų užpildytos apklausos rezultatai, kuriais remiasi autorė (žr. skyrių 2.2.1.2.). Verta atkreipti dėmesį į tai, jog įrankio ChatGPT sugeneruoti pavyzdžiai patys savaime **negali** būti autentiški, nes juos sugeneravo DI, tačiau gimtakalbiai kalbos vartotojai be abejonės gali vertinti jų autentišką ar neautentišką skambėjimą. Taip pat minėtina, jog kai kurie autentiškai skambantys sakiniai gimtakalbių vertinimu skamba „natūraliau“ lyginant su kitais – tai galima pamatyti iš apklausos rezultatų.

Penktasis autentiškai skambančių pavyzdžių kriterijus yra toks, jog šie sakiniai tinkamai iliustruoja lemas, atskleidžia jų semantinį bei sintaksinį kontekstą. Minėtina, jog lemu sąrašė pasitaikė **13** lemu, kurioms visi sugeneruoti iliustraciniai sakiniai laikytini autentiškais, t. y., visi joms sugeneruoti sakiniai gimtakalbių įvertinti kaip skambantys natūraliai. Šios lemos buvo *komisija, klausimas, komitetas, kalbėti, kaina, kiekvienas, keli, komanda, kartu, kodėl, keisti, kreiptis, kadangi*. Taigi, hipotetiškai kalbant, visi šie minėtas lemas iliustruojantys sakiniai galėtų būti įtraukti į žodyno IL (žr. skyrių 2.2.4.).

Žemiau pateikiamas 81 ChatGPT sugeneruotas sakinytis, priskirtas autentiškai skambančių sakinių kategorijai. Sakiniuose paryškintas tas žodis (lema) iš nagrinėjamo sąrašo, kurį sakinytis iliustruoja, o sakiniai sugrupuoti pagal tos pačios lemos iliustravimą:

- 1) *Vasarą oras dažnai būna labai **karštas**, todėl svarbu gerti daug vandens.*
- 2) *Ši kava yra per **karšta** gerti, palaukite, kol atvės.*
- 3) *Ji buvo pavargusi, **kadangi** dirbo visą naktį.*
- 4) *Aš atidėjau susitikimą, **kadangi** buvau užimtas.*

- 5) *Jis nebuvo susirūpinęs dėl ateities, **kadangi** turėjo stabilų darbą.*
- 6) *Mes praleidome vakarą kartu, **kadangi** buvome geri draugai.*
- 7) *Jie nusprendė perkelti vestuves, **kadangi** oras buvo blogas.*
- 8) *Ji miegojo, **kad** pailsėtų po ilgos kelionės.*
- 9) *Jis nevalgė, **kad** numestų svorio.*
- 10) *Ta knyga, **kurią** skaitau, yra labai įdomi.*
- 11) *Pirmas autobusas buvo perpildytas, todėl laukiau **kito**.*
- 12) *Mokytojas užduoda mums daug **klausimų**.*
- 13) ***Klausimas**, kurį tu kėlei, yra labai svarbus.*
- 14) *Ar tu gali atsakyti į mano **klausimą**?*
- 15) ***Kol** vaikai miegojo, jų tėvai pasiruošė vakarienei.*
- 16) ***Karinė** operacija buvo sėkmingai įvykdyta.*
- 17) *Šiuo metu jis **kuria** naują romaną.*
- 18) *Jie planuoja **kurti** naują verslą kitąmet.*
- 19) ***Komanda** dirbo visą naktį, kad baigtų projektą laiku.*
- 20) *Futbolo **komanda** laimėjo turnyrą.*
- 21) *Rytoj **komanda** vyks į susitikimą su klientais.*
- 22) *Po **kelių** valandų jie pasiekė tikslą.*
- 23) ***Kelios** mergaitės žaidžia kieme.*
- 24) *Prieš **kelis** metus jis lankė tą pačią mokyklą.*
- 25) *Mokslo **komisija** nusprendė finansuoti šį tyrimą.*
- 26) *Valstybinė **komisija** tiria korupcijos atvejus.*
- 27) ***Komisija** patvirtino galutinį ataskaitos dokumentą.*
- 28) *Finansų **komitetas** rengia biudžetą.*
- 29) *Mokslo **komitetas** aptarinėja naujus tyrimo projektus.*
- 30) *Nusikaltimų prevencijos **komitetas** turi parengti veiksmų planą.*
- 31) *Po valandos jis pradės **kalbėti** apie savo keliones.*
- 32) *Ji **kalba** labai aiškiai ir suprantamai.*
- 33) *Anksčiau mes **kalbėjome** apie tai, ką galėtume daryti.*
- 34) ***Konstitucinis** teismas yra institucija, kuri tikrina teisės aktų suderinamumą su konstitucija.*
- 35) *Valstybės vadovo įgaliojimai yra aiškiai apibrėžti **konstituciniuose** dokumentuose.*
- 36) *Prezidento **kanceliarija** yra atsakinga už įvairius valstybės administracinius reikalus.*
- 37) *Prašome pateikti **konkretų** atsakymą į šį klausimą.*
- 38) *Šiame projekte reikia pateikti **konkretų** planą.*
- 39) *Prašau pasakyti, **kiek** laiko praleidote bibliotekoje.*

- 40) **Kiek** obuolių pirkote?
- 41) **Kiekvienas** mokinys turi baigti namų darbus iki pirmadienio.
- 42) Rytoj **kiekvienas** darbuotojas turės apmokėti mokesčius.
- 43) **Kiekvienas** svečias buvo maloniai sutiktas.
- 44) **Koks** tavo mėgstamiausias filmas?
- 45) **Koks** šios žalios spalvos atspalvis tau labiausiai patinka?
- 46) **Kodėl** jis vėluoja?
- 47) **Kodėl** taip svarbu laikytis taisyklių?
- 48) Aš nežinau, **kodėl** ji neatsiliepia į mano žinutes.
- 49) Ji **keitė** savo gyvenimo būdą, siekdama geresnės sveikatos.
- 50) Aš nusprendžiau **keisti** savo darbo vietą.
- 51) Šis kompiuteris jau senas, reikėtų jį **keisti** nauju.
- 52) Aš nežinau, **kas** tai yra.
- 53) Ar galiu **kreiptis** į tave pagalbos?
- 54) Jis **kreipėsi** į gydytoją dėl sveikatos problemų.
- 55) Jie ketina **kreiptis** į teisininką dėl teisinio patarimo.
- 56) Jis turi pranešti **kuo** greičiau.
- 57) Jie kelionėje praleido daug laiko **kartu**.
- 58) Mes valgėme pusryčius **kartu** su šeima.
- 59) Vakara praleidome **kartu** su draugais.
- 60) Jis jau buvo čia du **kartus** šią savaitę.
- 61) Rytoj susitiksime su **keturiais** draugais.
- 62) Šiame krepšyje yra **keturi** obuoliai.
- 63) Ji guli ant žolės **kaip** mieganti rožė.
- 64) Jis finišavo lenktynėse **ketvirtas**.
- 65) **Ketvirtasis** aukštas pastate yra vienas jaukiausių.
- 66) **Ketvirtas** namas į kairę nuo kryžkelės yra mano.
- 67) Saulė pradėjo **kilti**, o tai reiškė, kad artėjo rytas.
- 68) **Kai** aš atėjau į namus, ji jau miegojo.
- 69) Mokslininkai stengiasi **kuo** greičiau rasti vakciną.
- 70) Prieš keletą metų automobilio **kaina** buvo žymiai mažesnė.
- 71) Šios knygos **kaina** yra 20 dolerių.
- 72) Kokia bus būsimų metų žaliavų **kaina**?
- 73) Ji pastebėjo **keistą** objektą ant stalo ir nustebo.
- 74) Mano sapnas buvo labai **keistas**, ir aš negalėjau jo aiškiai paaiškinti.

- 75) *Keistas veidrodis buvo įspūdingas eksponatas paslaptingoje muziejaus parodoje.*
- 76) *Tai yra miestas, kur mes praleidome atostogas.*
- 77) *Jie kūrė komercinį skelbimą, norėdami pritraukti daugiau klientų.*
- 78) *Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno.*
- 79) *Jis turi komercinį projektą, kurio tikslas yra plėtoti verslą.*
- 80) *Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus.*
- 81) *Jis įsigijo komercinę nuosavybę investavimo tikslais.*

Šių sakinių analizė toliau aprašoma tokia seka: gimtakalbių apklausos rezultatai, sintaksinė sakinių analizė, morfologiniai, leksiniai, pragmatiniai sakinių aspektai, sakinių turinys ir šių sugeneruotų sakinių analizės apibendrinimas.

2.2.1.2. Gimtakalbių vertinimas

Šiame skyriuje aptariami sakiniai vadinami **autentiškai** skambančiais dėl to, jog pagal devynių respondentų vertinimą, jie skamba labiau natūraliai nei dirbtinai. Tai reiškia, jog į šią kategoriją pateko tie sakiniai, kuriuos natūraliai skambančiais įvertino arba visi respondentai, arba nuo 8 iki 5 respondentų imtinai. Pasitaikė 14 tokių sakinių (~17% visų autentiškai skambančių sakinių), kuriuos **visi respondentai** įvertino kaip skambančius natūraliai, pvz.: *Koks³⁹ tavo mėgstamiausias filmas? Prieš keletą metų automobilio kaina buvo žymiai mažesnė; Po kelių valandų jie pasiekė tikslą; Jis kreipėsi į gydytoją dėl sveikatos problemų; Aš nusprendžiau keisti savo darbo vietą.* Remiantis apklausos rezultatais galima teigti, jog šie sugeneruoti sakiniai skamba visiškai natūraliai.

Kaip ir tikėtina bet kokioje apklausoje, dėl kai kurių sakinių natūralaus / dirbtinio skambėjimo nesutapo respondentų nuomonės. Toliau aptariami sakiniai, dėl kurių respondentai sudvejojo, kurie yra morfologiškai, sintaksiškai ir semantiškai taisyklingi. 25 sakinius (~31% visų autentiškai skambančių sakinių) kaip dirbtinai skambančius įvertino tik **vienas** respondentas, pvz.: *Ta knyga, kurią skaitau, yra labai įdomi; Ji kalba labai aiškiai ir suprantamai; Šiuo metu jis kuria naują romaną; Mokslo komitetas aptarinėja naujus tyrimo projektus; Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus.* Turint omenyje tai, jog tik vienas

³⁹ Čia ir toliau sakiniuose paryškinti tas žodis, kurį tas sakinyje iliustravo.

respondentas iš devynių šiuos sakinius palaikė dirbtinai skambančiais, juos taip pat būtų galima laikyti natūraliai skambančiais lietuvių kalbos sakiniais.

20 sakinių (~25% visų autentiškai skambančių sakinių) buvo tokie, dėl kurių natūralaus skambėjimo suabejojo **du** respondentai iš devynių. Nors tai nėra didelė respondentų dalis, sakinius įvertinusi kaip skambančius dirbtinai, tačiau galima kelti klausimą, kodėl šie sakiniai dviems respondentams iš devynių taip nuskambėjo. Toliau pateikiami keli šių sakinių pavyzdžiai: *Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug vandens; Jis nevalgė, kad numestų svorio; Ar galiu kreiptis į tave pagalbos? Komisija patvirtino galutinį ataskaitos dokumentą; Prašau pasakyti, kiek laiko praleidote bibliotekoje.* Galima išskirti ypatingą atvejį, kurio natūraliu skambėjimu suabejojo du respondentai: *Ji guli ant žolės kaip mieganti rožė.* Darbo autorei pastarasis respondentų vertinimas pasirodė netikėtas dėl to, jog šis sakinytis skamba labai „literatūriškai“, „meniškai“⁴⁰, tačiau nepaisant to, absoliuti dauguma respondentų šį sakinį laikė natūraliai skambančiu.

Pasitaikė 11 sakinių (~14% visų autentiškai skambančių sakinių), ties kuriais **trys** iš devynių respondentų pažymėjo, jog jie skamba dirbtinai. Dėl to šiuos sakinius būtų galima įvardinti mažiau „natūraliai“ skambančiais: 1) *Aš nežinau, kodėl ji neatsiliepia į mano žinutes;* 2) *Jis įsigijo komercinę nuosavybę investavimo tikslais;* 3) *Ketvirtasis aukštas pastate yra vienas jaukiausių;* 4) *Ketvirtas namas į kairę nuo kryžkelės yra mano;* 5) *Saulė pradėjo kilti, o tai reiškė, kad artėjo rytas;* 6) *Kai aš atėjau į namus, ji jau miegojo;* 7) *Mokslininkai stengiasi kuo greičiau rasti vakciną;* 8) *Šios knygos kaina yra 20 dolerių;* 9) *Mes praleidome vakarą kartu, kadangi buvome geri draugai;* 10) *Valstybės vadovo įgaliojimai yra aiškiai apibrėžti konstituciniuose dokumentuose;* 11) *Rytoj kiekvienas darbuotojas turės apmokėti mokesčius.*

Į klausimą, kodėl šie sakiniai dalies (33,3%) respondentų įvertinti dirbtinai, galima pasižiūrėti patikrinant, ar tekстыne esama analogiškų formuluočių, o jei jų ten yra, ar jos pasitaiko dažnai. Pradedant pirmuoju sakiniu, kurio natūraliu skambėjimu suabejojo trys iš devynių respondentų, šnekamojoje kalboje dažnai praleidžiamas veiksnys, jei jis sutampa su kalbančiuoju, t. y., vietoje *aš nežinau* galėtų būti *nežinau*: Sketch Engine esančio tekstyno ELEXIS Lithuanian Web 2021 (toliau – tekstyno) duomenimis, *nežinau* kalboje pasitaiko žymiai dažniau (93 864 kartus) nei *aš nežinau* (5915 kartų). Žodžių junginys *investavimo tikslais* (2) tekstyने pasitaiko retai – 24 kartus. Žodžių junginys *ketvirtasis aukštas*, pasitaikantis trečiame sakinyje, tekstyने randamas 17 kartų, tuo

⁴⁰ Galima sakyti, jog šis sakinytis skamba pseudomeniška, nes mažai tikėtina, jog ChatGPT šį sakinį sugeneravo remiantis grožinės literatūros šaltiniu. Darbo autorės nuomone, tai galėtų būti vertalas ar haliucinacija (sunku pasakyti, kokių pagrindu jis sugeneruotas), tačiau remiantis apklausos rezultatais, jis atsidūrė autentiškai skambančių sakinių kategorijoje.

tarpu *ketvirtas aukštas* – 60 kartų. Žodžių junginys *saulė pradėjo kilti* (5) tekstyne nerastas, tačiau *saulė kilo* rastas 26 kartus. *Atėjau į namus* (6) tekstyne nepasitaiko, bet *grįžau į namus* randamas 50 kartų (plg. *grįžau namo* – 857 kartus). Daiktavardis *vakcina* (7) daugiausiai vartojamas su veiksmažodžiais *sukurti, kurti, pagaminti* – žodžių junginys *rasti vakciną* tekstyne pasitaikė vieną kartą. Žodžių junginys *konstituciniuose dokumentuose* (10) nėra itin įprastas, tekstyne jis randamas 3 kartus. Matyti, jog nemažai sugeneruotų žodžių junginių nėra tipiški lietuvių kalboje, o tai patvirtina minėto tekstyno duomenys. Sunku pasakyti, kas mažiau kalbiškai įprasto pasitaiko 4, 8, 9, 11 sakiniuose ir kodėl jų natūralumu suabejojo 3 respondentai.

Beveik pusė respondentų (44,4%), **keturi** iš devynių, 10 sakinių (~12% visų autentiškai skambančių sakinių) įvertino kaip skambančius dirbtinai. Šiuos sakinius taip pat galima įvardinti kaip skambančius ne taip „natūraliai“, lyginant su prieš tai skyriuje aptartais sakiniais: 1) **Keistas** *veidrodis buvo išpūdingas eksponatas paslaptingoje muziejaus parodoje*; 2) *Jis turi komercinį projektą, kurio tikslas yra plėtoti verslą*; 3) *Tai yra miestas, kur mes praleidome atostogas*; 4) *Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno*; 5) *Kokia bus būsimų metų žaliavų kaina?* 6) *Jie nusprendė perkelti vestuves, kadangi oras buvo blogas*; 7) **Klausimas**, *kurį tu kėlei, yra labai svarbus*; 8) *Po valandos jis pradės kalbėti apie savo keliones*; 9) *Prezidento kanceliarija yra atsakinga už įvairius valstybės administracinius reikalus*; 10) **Koks** *šios žalios spalvos atspalvis tau labiausiai patinka?*

Žvelgiant iš tekstynų perspektyvos, prijungiamojo sakinio dėmuo *miestas, kur* (3) yra mažiau pasitaikantis kalboje (tekstyne rastas 35 kartus) nei *miestas, kuriame* (tekstyne rastas 60 kartų). Taip pat nebūdingas yra ketvirtame sakinyje esantis pasakymas *su komerciniais tikslais* (tekstyne rastas 4 kartus), lyginant su įprastesniu žodžių junginiu *komerciniais tikslais*, kuris tekstyne pasitaikė 952 kartus. Tekstyno duomenimis, žodžių junginys *būsimų metų* (5) yra rečiausiai pasitaikantis (rastas 36 kartus), lyginant su, pvz., junginiais *kitų metų* (rastas 16 867 kartus) ar *artėjančių metų* (rastas 61 kartą). Sunkiau paaiškinti, kodėl 1, 2, 6-10 sakiniai gimtakalbiams nuskambėjo dirbtinai. Galbūt tai turi kažką bendro su pirmame sakinyje pastebimu būdvardžių pertekliumi. Antrame sakinyje pateikiama informacija yra iš dalies savaime suprantama (tai, jog komercinis projektas susijęs su verslo plėtra). *Kalbėti apie savo keliones* (8) tekstyne nepasitaiko, o *kalbėti apie keliones* pasitaiko 12 kartų. Galiausiai, *prezidento kanceliarija yra* (9) tekstyne pasitaiko septynis kartus, o *prezidento kanceliarija yra atsakinga* ar *kanceliarija yra atsakinga* tekstyne visai nerandama: reiškia, būdvardis *atsakinga* įprastai nevartojamas su daiktavardžiu *kanceliarija*. Taigi, tekstyno duomenys patvirtina, jog kai kurios ChatGPT formuluotės nėra įprastos lietuvių kalboje (tačiau jos nebūtinai neįmanomos), o tai galėjo paveikti tokį gimtakalbių vertinimą.

Be abejonės, ši apklausa turi savo ribotumą. Visada išlieka galimybė, jog atsakymų pasirinkimams įtaką gali daryti respondentų skubėjimas, sakinių neperskaitymas iki galo, neatidumas, nuovargis ir kt. Darbo autorė ruošdama apklausą suklydo ir nustatymuose nenurodė, jog respondentams būtų galima pasirinkti **tik** vieną atsakymą (tai buvo tik parašyta apklausos įvado dalyje). Todėl pasitaikė trys atvejai, kuomet vienas respondentas pažymėjo abu pasirinkimus ir šie atvejai analizei netiko. Taip pat suprantama, jog devyni respondentai nėra itin didelė gimtakalbių imtis, tačiau plataus masto apklausa ir nebuvo šio tyrimo tikslas.

2.2.1.3. Sintaksinė autentiškai skambančių sakinių analizė

Pirmiausiai aptariant autentiškai skambančių sakinių struktūrą, iš viso suskaičiuoti 53 vientisiniai sakiniai ir 28 sudėtiniai sakiniai. Taigi, autentiškai skambantys **vientisiniai** sakiniai sugeneruoti dažniausiai: iš visų autentiškai skambančių sakinių ~65 % yra vientisiniai (kitais tariant, paprastesnės sintaksinės struktūros).

Kalbant apie sudėtinius sakinius pastebėta, jog dažniausiai jie buvo **sudėtiniai prijungiamieji**: šių sakinių suskaičiuota devyniolika. Kalbant apie dažniausiai pasitaikiusius prijungiamuosius jungtukus *kadangi* (pasitaikė 5 kartus) ir *kad* (pasitaikė 3 kartus) svarbu atsižvelgti į tai, jog tam įtakos padarė tyrimo lemų sąrašas, kuriame yra abu šie žodžiai, t. y., abiem šioms žodžiams buvo prašoma sugeneruoti IL, dėl to natūralu, jog šie jungtukai sakiniuose pasitaikė dažniau. Kiti po kartą pasitaikę prijungiamieji jungtukai buvo *kuri*, *kas*, *ką*, *kodėl*, *kurio*, *kuri*, *kurią*, *kur*, *kai*, *kol*, *kiek*.

Taip pat pasitaikė 4 bejungtukiai sakiniai, kurių dėmenys daugiausiai atskirti kableliais (*Šis kompiuteris jau senas, reikėtų jį keisti nauju; Ji keitė savo gyvenimo būdą, siekdama geresnės sveikatos; Jie kūrė komercinį skelbimą, norėdami pritraukti daugiau klientų*) ir tik kartą atskirti brūkšniu (*Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno⁴¹*). Buvo rasti du daugiadėmeniai sudėtiniai sakiniai. Vitas Labutis rašo apie **mišriuosius** daugiadėmenius sakinius, „kurių vieni dėmenys susieti vienokio tipo, kiti – kitokio tipo santykiais“ (Labutis 1998: 183). Sugeneruotuose sakiniuose atpažintas sujungiamasis-prijungiamasis sakiny *Saulė pradėjo kilti, o tai reiškė, kad artėjo rytas* (*o* yra sujungiamasis, *kad* – prijungiamasis jungtukas). Taip pat rastas sakiny *Ši kava yra per karšta gerti, palaukite, kol atvės, jį galima laikyti bejungtukių-prijungiamuoju sakiniu*. Galiausiai rasti trys sudėtiniai sujungiamieji sakiniai: dviejuose iš jų dėmenys susieti jungtuku *todėl* (*Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug*

⁴¹ Šis sakiny darbo autorės vertinamas kaip vertalas, nes lietuvių kalboje būdinga sakyti *komerciniais tikslais*, o ne *su komerciniais tikslais*. Tačiau įdomu, jog dauguma respondentų jį palaikė skambančiu natūraliai.

vandens; *Pirmas autobusas buvo perpildytas, todėl laukiau kito*), o viename – jungtuku ir (*Mano sapnas buvo labai keistas, ir aš negalėjau jo aiškiai paaiškinti*).

Darbo autorė taip pat atliko sugeneruotų sakinių aplinkybių analizę. Vitas Labutis išskiria ir išsamiai aprašo šias aštuonias aplinkybes: vietos, laiko, priežasties, sąlygos, nuolaidos, tikslo, kokybės (būdo), kiekybės (Labutis 1998, 281-311). Remiantis V. Labučio aplinkybių aprašymu, toliau šia seka aptariamas jų pasitaikymas sugeneruotuose sakiniuose.

Iš viso vietos aplinkybės sugeneruotuose sakiniuose rastos 16 kartų. Didelė dalis vietos aplinkybių nurodė buvimo vietą (*šiam krepšyje; kieme; konstituciniuose dokumentuose; bibliotekoje; ant stalo; ant žolės; pastate; į kairę nuo kryžkelės; muziejaus parodoje*). Taip pat pasitaikė kryptį nurodančios aplinkybės (*į namus; į susitikimą su klientais*). Dviem atvejais buvimo vieta buvo nusakyta prieveiksmiais *čia* ir *kur*. Kai kurios vietos aplinkybės kaip *lenktynėse, kelionėje* labiau nurodė ne fizinę vietą, o su tais žodžiais susijusią įvykių erdvę (pvz., *finišavo lenktynėse⁴² ketvirtas; kelionėje praleido daug laiko*). Sakinyje *Šiame projekte reikia pateikti konkretų planą aplinkybė šiam projekte* nurodo ne fizinę erdvę, o tam tikrą abstraktą: apie semantiškai panašius abstraktumo atvejus rašo ir V. Labutis, kaip pavyzdžius įvardindamas abstrakčius asmenų pavadinimus (*mumyse*) (Labutis 1998: 283).

Laiko aplinkybės rastos 29 kartus. Jų raiška buvo gana įvairi: buvo vientisinių aplinkybių (triskart rasta *rytoj*, po du kartus – *vakarą* ir *dažnai; kitąmet; anksčiau; rytas, vasarą*). Taip pat buvo sudėtinių aplinkybių, išreikštų įvairiais žodžių junginiais (dukart rasta *visą naktį; po ilgos kelionės; šiuo metu; prieš keletą metų; po kelių valandų; prieš kelis metus; po valandos; kiek laiko; daug laiko; iki pirmadienio; kuo greičiau; du kartus šią savaitę⁴³; kuo greičiau; būsimų metų*). Kitos sudėtinės laiko aplinkybės buvo išreikštos šalutiniais sakiniais (*kol vaikai miegojo; kol atvės; kai aš atėjau į namus*). Matyti, jog kai kurios laiko aplinkybės reiškiamos konkrečiais, laikui įvardinti skirtais žodžiais (pvz., *kitąmet, rytoj, dažnai, rytas, vakaras, laikas*), o kitos aplinkybės pačios leksiškai nenurodo laiko (pvz., *kol vaikai miegojo*) ir laiko nuoroda pasidaro aiški tik konkrečiau sakinio kontekste.

Iš viso rasta 12 priežasties aplinkybių. Dažniausiai šios aplinkybės rastos sudėtiniuose sakiniuose, kurie buvo šalutiniai: *Ji buvo pavargusi, kadangi dirbo visą naktį; Aš atidėjau susitikimą, kadangi buvau užimtas; Jis nebuvo susirūpinęs dėl ateities, kadangi turėjo stabilų darbą; Mes praleidome vakarą kartu, kadangi buvome geri draugai; Jie nusprendė perkelti vestuves, kadangi oras buvo blogas; Aš nežinau, kodėl ji neatsiliepia į mano žinutes*. Taip pat pasitaikė ir priežasties raiškos pagrindiniuose sakiniuose: *Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug*

⁴² Sakinių aplinkybių analizėje čia ir toliau kai kurios teksto vietos paryškintos, siekiant iliustruoti tas aplinkybes, apie kurias kalbama.

⁴³ Nors ši laiko aplinkybė yra sudėtinė (susidedanti iš *du kartus* ir *šią savaitę*), ją darbo autorė laiko viena laiko aplinkybe, o ne dviem.

vandens; **Pirmas autobusas buvo perpildytas, todėl laukiau kito.** Šiuose sakiniuose priežasties aplinkybės buvo galima atpažinti pagalrieveiksmių *todėl, kodėl, jungtuką kadangi*. Vientisiniuose priežasties sakiniuose du po kartus rastasrieveiksmis *kodėl* ir prielinksnis *dėl*: **Kodėl jis vėluoja? Kodėl taip svarbu laikytis taisyklių?** Taip pat pasitaikė dvi prielinksninės priežasties konstrukcijos: *Jis kreipėsi į gydytoją dėl sveikatos problemų. Jie ketina kreiptis į teisininką dėl teisinio patarimo.*

Sąlygos bei nuolaidos aplinkybių nebuvo rasta. Iš viso rastos 9 tikslo aplinkybės, kurias kartais nelengva atskirti nuo priežasties aplinkybių. V. Labutis rašo, jog tikslo aplinkybė rodo „numatomą, perspektyvoje turimą priežastį“ (Labutis 1998: 300). Šiuose sakiniuose tikslas išreikštas padalyviu *siekiant*, pusdalyviais (*siekdama, norėdami*), daiktavardžiu *tikslai*, jungtuku *kad* bei kilmininko linksniu: *Komanda dirbo visą naktį, kad baigtų projektą laiku; Ji keitė savo gyvenimo būdą, siekdama geresnės sveikatos; Ji miegojo, kad pailsėtų po ilgos kelionės; Jis nevalgė, kad numestų svorio; Jie kūrė komercinį skelbimą, norėdami pritraukti daugiau klientų; Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus; Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno; Jis įsigijo komercinę nuosavybę investavimo tikslais; Ar galiu kreiptis į tave pagalbos?*

Iš viso pasitaikė 22 kokybės (būdo) aplinkybės. Iš viso 5 kartus pasitaikėrieveiksmis *labai* (*labai karštas; labai keistas; labai įdomi; labai svarbus; labai aiškiai ir suprantamai*), kartą pasitaikė šiorieveiksmio aukščiausiasis laipsnis *labiausiai*. Keturis kartus rastasrieveiksmis *kartu*, dukart – žodžių junginys *kuo greičiau*. Be kitų rastų kokybės aplinkybių (tris kartus – *aiškiai; laiku; sėkmingai; suprantamai; maloniai; žymiai; su komerciniais tikslais*), taip pat pasitaikė vienas palyginimas: *Ji guli ant žolės kaip mieganti rožė.*

Buvo rastos 9 kiekybės aplinkybės. Keturis kartus pasitaikė kiekiorieveiksmis *daug*, tarp jų ir aukštesnysis laipsnis (*daugiau klientų; daug vandens; daug klausimų; daug laiko*), dukart –rieveiksmis *kiek* (*kiek laiko; kiek obuolių*). Vieną kartą pasitaikė įvardis *kelios* (*kelios mergaitės*), dukart kiekis nusakytas skaitvardžiais (*keturi obuoliai; 20 dolerių*).

Apibendrinant, autentiškai skambančiuose sugeneruotuose sakiniuose dominuoja laiko aplinkybės, rastos 29 kartus. Taip pat rasta nemažai kokybės (būdo) (22 kartus), vietos (16 kartų) aplinkybių. Priežasties aplinkybės rastos 12 kartų. Mažiausiai, po 9 kartus rasta tikslo ir kiekybės aplinkybių, o sąlygos ir nuolaidos aplinkybių visai nebuvo rasta. Taigi, galima teigti, jog sugeneruotuose sakiniuose be abejonės pateikiamos įvairios papildomos kalbinės nuorodos, tačiau turint omenyje, jog kalbama apie 81 autentiškai skambantį sakinį, bendras rastų aplinkybių skaičius (viso suskaičiuota 97 įvairios aplinkybės) nėra išskirtinai didelis. Todėl galima kelti klausimą dėl didesnės kalbinės įvairovės sugeneruotuose sakiniuose ir ar ji išties yra.

Taip pat buvo apžvelgti funkciniai sakinių tipai. Rasti trys sakinių tipai: konstatuojamieji, skatinamieji ir klausiamieji. Toliau aptariant sugeneruotus sakinius pagal jų funkcinius sakinių tipus

taip pat remiamasi Vito Labučio (1998) knygoje *Lietuvių kalbos sintaksė* esančia terminija. Didžioji dalis sugeneruotų sakinių (68 sakiniai) yra **konstatuojamieji**, kuriuose neutraliai pateikiama, skelbiama tam tikra informacija. Toliau pateikiama keletas rastų konstatuojamųjų sakinių: *Kelios mergaitės žaidžia kieme; Kol vaikai miegojo, jų tėvai pasiruošė vakarienei; Jis turi pranešti kuo greičiau; Mokslininkai stengiasi kuo greičiau rasti vakciną; Šiame krepšyje yra keturi obuoliai.* Pasitaikė 8 **klausiamieji** sakiniai. Sakiniai *Ar tu gali atsakyti į mano klausimą? Ar galiu kreiptis į tave pagalbos?* yra **tikrinamojo klausimo**⁴⁴ sakiniai, nes prasideda klausiamąja dalelyte *ar*, o į juos galima atsakyti dalelytėmis *taip* arba *ne*. Sakiniai *Koks tavo mėgstamiausias filmas? Koks šios žalios spalvos atspalvis tau labiausiai patinka? Kiek obuolių pirkote? Kodėl jis vėluoja? Kodėl taip svarbu laikytis taisyklių? Kokia bus būsimų metų žaliavų kaina?* sakiniai yra **specialiojo (konkrečiojo) klausimo** sakiniai, nes jais klausiama kažkokio konkretaus dalyko, jie prasideda klausiamaisiais įvardžiais *koks, kokio* ir klausiamaisiaisrieveiksmiais *kiek, kodėl*. Taip pat rasti 5 **skatinamieji** sakiniai. Juos tiksliau galima įvardinti **geidžiamaisiais** sakiniiais: V. Labutis rašo, jog šiais sakiniiais reiškiamas prašymas ar kiti panašūs „pasyvesni valios aktai, kai kalbėtojas negali ar nenori stipriau, tiesiogiai paveikti to asmens ar daikto, kuriam jis reiškia savo valią“ (Labutis, 1998: 118). Geidžiamajame sakinyje *Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug vandens* yra patariama, pamokoma, sakinyje *Ši kava yra per karšta gerti, palaukite, kol atvės* išvelgiamas pasiūlymas. Sakiniuose *Prašome pateikti konkretų atsakymą į šį klausimą* ir *Prašau pasakyti, kiek laiko praleidote bibliotekoje* matyti prašymas, o sakinyje *Šis kompiuteris jau senas, reikėtų jį keisti nauju* prašoma tam tikros paslaugos, naudos sau. Funkcinių sakinių tipų analizė parodė, jog dominuoja konstatuojamieji sakiniai, o kitokio tipo (klausiamieji, skatinamieji) sakiniai pasitaikė ne taip dažnai. Šioje vietoje taip pat vertėtų kelti klausimą, ar sugeneruoti autentiškai skambantys sakiniai yra itin kalbiškai įvairūs, pavyzdžiui, kaip kokiame nors lietuvių kalbos žodyne (tarkime, BLKŽ): darbo autorė būtų linkusi sakyti, jog sakiniai yra gana panašaus pobūdžio.

2.2.1.4. Morfolginiai autentiškai skambančių sakinių aspektai

Toliau analizuojant autentiškai skambančius pavyzdžius paaiškėjo, kad sugeneruotuose sakiniuose dominavo **būstasis kartinis laikas** ir **esamasis laikai**, abu sakiniuose pasitaikė po 36 kartus. Taip pat rasti 5 sakiniai su būsimuoju laiku (*Kokia bus būsimų metų žaliavų kaina? Po valandos jis pradės kalbėti apie savo keliones; Rytoj susitiksime su keturiais draugais; Rytoj komanda vyks į susitikimą su klientais; Rytoj kiekvienas darbuotojas turės apmokėti mokesčius*). Pasitaikė 4

⁴⁴ Terminus *tikrinamojo klausimo / specialiojo (konkrečiojo) klausimo / geidžiamieji sakiniai* Vitas Labutis (1998) mini 113-118 puslapiuose.

sakiniai, kuriuose yra dvi skirtingos laiko formos: trijuose iš jų rasti esamasis ir būtasis kartinis laikai (*Klausimas, kurį tu kėlei, yra labai svarbus; Prašau pasakyti, kiek laiko praleidote bibliotekoje; Tai yra miestas, kur mes praleidome atostogas*), o viename – esamasis ir būsimasis laikai (*Ši kava yra per karšta gerti, palaukite, kol atvės*). Pastebėta, jog sudėtiniuose sakiniuose dominavo būtasis kartinis laikas, o vientisiniuose sakiniuose dažniausiai pasitaikė esamasis laikas. Taip pat minėtina, jog sakiniuose nė karto nepasitaikė būtasis dažninis laikas⁴⁵.

Analizuotuose sakiniuose rastos įvairios nuosakos. Sakiniuose dažniausiai rasta **tiesioginė nuosaka**, kuri suskaičiuota 75 pavyzdžiuose. Kai kuriuose sakiniuose pasitaikė kelios nuosakos vienu metu. Penkiuose sakiniuose rastos tiesioginė ir tariamoji nuosakos (*Ji miegojo, kad pailsėtų po ilgos kelionės; Šis kompiuteris jau senas, reikėtų jį keisti nauju; Komanda dirbo visą naktį, kad baigtų projektą laiku; Anksčiau mes kalbėjome apie tai, ką galėtume daryti; Jis nevalgė, kad numestų svorio*), o viename sakinyje – tiesioginė ir liepiamoji nuosakos (*Ši kava yra per karšta gerti, palaukite, kol atvės*).

Toliau apžvelgiami dalyviai, pusdalyviai bei padalyviai sugeneruotuose sakiniuose, kurių nerasta daug. Sakiniuose daugiausiai rasta **dalyvių** (viso devyni): *užimtas, perpildytas, sutiktas, apibrėžti, naudojami, įvykdyta, susirūpinęs, mieganti, pavargusi*. Kaip matyti, daugiausiai rasta neveikiamųjų dalyvių. Taip pat pasitaikė du pusdalyviai (*siekdama, norėdami*) ir padalyvis (*siekiant*).

Analizuojant sakinius iš giminės kategorijos perspektyvos pastebėta, jog sakiniuose dominuoja **vyriškoji** giminė. Iš viso suskaičiuota 18 vyr. g. įvardžių (12 kartų – *jis*, 6 kartus – *jie*), 3 dalyviai *užimtas, susirūpinęs, sutiktas*, pusdalyvis *norėdami*, būdvardis *geri*, skaitvardis *ketvirtas*. Taip pat rasta 12 daiktavardžių *mokytojas, vadovas, prezidentas, mokinys, darbuotojas, svečias, teisininkas, mokslininkai, klientai*, triskart – *draugai*. Buvo rasti dar du daiktavardžiai *tėvai*⁴⁶ ir *vaikai*, tačiau lietuvių kalboje jie be vyriškosios giminės neturi kito pavidalo. Kalbant apie moteriškąją giminę, iš viso 8 kartus rastas įvardis *ji*, daiktavardis *mergaitės* ir pusdalyvis *siekdama*. Taigi, matyti ženklus skirtumas, lyginant iš viso 10 kartų pasitaikiusias moteriškąją giminę žymėtas kalbos dalis su 36 kartus pasitaikiusiomis vyriškąją giminę žymėtomis kalbos dalimis. Svarbu atsižvelgti į tai, jog vyriškosios giminės įvardžiai, daiktavardžiai ir kitos kalbos dalys lietuvių kalboje neretai atlieka giminės neutralizacijos funkciją⁴⁷, tačiau akivaizdu, jog sugeneruotuose tekstuose dominuoja vyriškoji giminė ir, atitinkamai, vyriškos patirtys.

⁴⁵ Tai nėra visai netikėta, nes įprastai kalboje dažniau pasitaiko būtasis kartinis, o ne būtasis dažninis laikas. Tačiau į tai vistiek verta atkreipti dėmesį, nes ši laiko forma nepasitaikė 81 sakinyje, o tai yra nemažas kiekis sugeneruoto teksto.

⁴⁶ Kalbant apie tėvus kaip tėvo ir motinos **porą**.

⁴⁷ Apie giminės neutralizaciją savo tinklaraštyje plačiau rašo Antanas Smetona: <https://www.smetona.lt/gimine/>.

2.2.1.5. Leksiniai autentiškai skambančių sakinių aspektai

Pereinant prie leksinių autentiškai skambančių sakinių ypatumų, natūralu, jog juose pasitaikė daug įvairių bendrinių daiktavardžių⁴⁸ (pvz., *kompiuteris, kelionė, taisyklės, pagalba, reikalai, verslas, vestuvės, knyga, vakarienė, tikslas, spalva, darbas, krepšys, objektas, miestas* ir pan.). Tačiau sakiniuose visai nepasitaikė tikriniai daiktavardžiai, perkeltinės reikšmės atvejai. Taip pat nepasitaikė šnekamosios kalbos registro leksika: įdomu, jog 26 sakiniai (~32 % visų autentiškai skambančių sakinių) prasidėjo asmeniniais įvardžiais *aš / jis / ji / jie / mes* ir sutapo su sakinio veiksniu. Tačiau šnekamojoje kalboje su sakinio veiksniu sutampantys įvardžiai dažnai praleidžiami (pvz., vietoje *aš nežinau* sakoma *nežinau*⁴⁹).

Žvelgiant per leksinės semantikos prizmę, rastos kelios (viso 5) sinonimų poros (čia ir toliau paryškinta): 8) *Ji miegojo, kad pailsėtų po ilgos kelionės*; 32) *Ji kalba labai aiškiai ir suprantamai*; 45) *Koks šios žalios spalvos atspalvis tau labiausiai patinka?* 75) *Keistas veidrodis buvo įspūdingas eksponatas paslaptingoje muziejaus parodoje*; 80) *Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus*. Minėtina, jog sinonimai 75, 80 sakiniuose yra kontekstiniai, o ne pastovieji. Buvo rastos dvi antonimų poros: 37) *Prašome pateikti konkretų atsakymą į šį klausimą*; 51) *Šis kompiuteris jau senas, reiktų jį keisti nauju*. Kontekstiniais antonimais galima laikyti šias dvi poras: 15) *Kol vaikai miegojo, jų tėvai pasiruošė vakarienei*; 68) *Kai aš atėjau į namus, ji jau miegojo*. Taip pat pastebėta pora hierarchinių žodžių santykių: sakinyje *Koks šios žalios spalvos atspalvis tau labiausiai patinka?* hiperonimas yra *spalva*, o hiponimas – *žalia*. Taip pat sakinyje *Komisija patvirtino galutinį ataskaitos dokumentą* hiperonimas yra *dokumentas*, o hiponimas yra *ataskaita* (dokumento rūšis).

Tai, jog buvo prašoma iliustruoti lemas *komanda, komisija, komitetas, kanceliarija*, nulėmė dešimties metonimijos atvejų pasitaikymą, pvz., 19) *Komanda dirbo visą naktį, kad baigtų projektą laiku*; 25) *Mokslo komisija nusprendė finansuoti šį tyrimą*; 28) *Finansų komitetas rengia biudžetą*; 36) *Prezidento kanceliarija yra atsakinga už įvairius valstybės administracinius reikalus*: šiais atvejais pasitaikė sinekdocha. Metonimija taip pat pasitaikė lema *konstitucinis* iliustruojančiame sakinyje *Konstitucinis teismas yra institucija, kuri tikrina teisės aktų suderinamumą su konstitucija*. Kaip ypatingą atvejį verta paminėti personifikacijos ir palyginimo apraišką sakinyje *Ji guli ant žolės kaip mieganti rožė*.

⁴⁸ Darbo autorė čia omenyje turi bendrinius daiktavardžius kaip priešpriešą tikriniais daiktavardžiams, o ne bendrinės kalbos žodžius.

⁴⁹ Tai yra nulemta kalbos ekonomijos principo.

Apibendrinant šį poskyrį galima pasakyti, jog nors sakiniuose rastos sinonimų, antonimų poros, kelios hierarchinių žodžių santykių apraiškos, metonimijos atvejai, jų nebuvo rasta labai daug. Taip pat sugeneruotose iliustracijose pasitaikė vien bendriniai daiktavardžiai, nerasta šnekamosios kalbos leksikos, o nemaža dalis sakinių prasidėjo asmeniniais įvardžiais. Dėl to galima dar kartą kvestionuoti sugeneruotų sakinių kalbinę įvairovę, kurios neįžvelgiama itin daug.

2.2.1.6. Pragmatiniai autentiškai skambančių sakinių aspektai

Žvelgiant į sakinius iš pragmatinės perspektyvos matyti, jog juose pasitaikė keletas ilokucinių aktų, kurie artimai susiję su jau aptartais funkciniais sakinių tipais. Tai buvo patarimai (žr. 1, 2 sakinius žemiau) ir prašymai (žr. 14, 37, 39, 51, 53 sakinius žemiau). Kai kada dėl trūkstamo konteksto sunku pasakyti, ar tai tiesiog teiginiai, ar tam tikri netiesioginiai paraginimai (dėl veiksmažodžių *turėti*, *reikėti* pasitaikymo): 30) *Nusikaltimų prevencijos komitetas turi parengti⁵⁰ veiksmų planą*; 38) *Šiame projekte reikia pateikti konkretų planą*; 41) *Kiekvienas mokinys turi baigti namų darbus iki pirmadienio*; 42) *Rytoj kiekvienas darbuotojas turės apmokėti mokesčius*; 56) *Jis turi pranešti kuo greičiau*. Prašymuose du kartus pasitaikė performatyvusis veiksmažodis *prašyti*: 37) *Prašome pateikti konkretų atsakymą į šį klausimą*; 39) *Prašau pasakyti, kiek laiko praleidote bibliotekoje*. Kituose netiesioginiuose prašymuose veiksmažodžio *galėti* pagalba nurodoma į pašnekovo galėjimą padaryti tam tikrą veiksmą: 14) *Ar tu gali atsakyti į mano klausimą?* 53) *Ar galiu kreiptis į tave pagalbos?* Kitame netiesioginiame prašyme pasitaikė tariamoji nuosaka, sušvelninant adresanto pageidavimą, juolab, kad adresantas greičiausiai siekia asmeninės naudos: 51) *Šis kompiuteris jau senas, reikėtų jį keisti nauju*. Tik vieną kartą pasitaikė liepiamoji nuosaka: 2) *Ši kava yra per karšta gerti, palaukite, kol atvės* – nors tai panašu į prašymą, tai yra patarimas, nes jei adresatas nepaklausys ir nepalauks, kol kava atvės, adresantas nuo to nei nukentės, nei gaus naudos. Kitas pasitaikęs netiesioginis patarimas buvo šis: 1) *Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug vandens*. Galima matyti, jog sugeneruotų autentiškai skambančių sakinių tarpe pasitaikė penki prašymai, du patarimai, o kai kada dėl trūkstamo konteksto buvo nelengva nustatyti, koks tai sakiny (pvz., 30, 38). Šioje vietoje būtų galima kalbėti apie tam tikrą kalbinės įvairovės pragmatiniu požiūriu atosklaidą, turint omenyje, jog likusi sakinių dalis yra konstatuojamieji sakiniai.

⁵⁰ Iliustruojami pragmatiniai aspektai čia ir toliau sakiniuose paryškinti.

2.2.1.7. Autentiškai skambančių sakinių turinys

Pereinant prie sugeneruotų sakinių turinio pastebėta, jog esama tokių sakinių, kurių turinys identiškas arba labai panašus. 59 ir 6 sakiniuose kalbama apie identišką situaciją – kartu su draugais praleistą vakarą (*Vakarą praleidome kartu su draugais; Mes praleidome vakarą kartu, kadangi buvome geri draugai*). 4 ir 21 sakiniuose kalbama apie susitikimą darbo kontekste (*Aš atidėjau susitikimą, kadangi buvau užimtas; Rytoj komanda vyks į susitikimą su klientais.*), o 38 ir 19 sakiniuose – apie projektą darbo kontekste (*Šiame projekte reikia pateikti konkretų planą; Komanda dirbo visą naktį, kad baigtų projektą laiku*). Panašios sakiniuose aptariamoms situacijoms kai kada nulemtos tos pačios sakinyje iliustruojamos lemos. Lemos *komercinis* sakiniuose (78, 80) kalbama apie pelno siekį (*Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno; Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus*). Taip pat panašus kitų dviejų sakinių turinys, kuriame veiksniai yra *mokslo komisija* (25) bei *mokslo komitetas* (29), o kalbama apie su konkrečiu tyrimu susijusius dalykus (*Mokslo komisija nusprendė finansuoti šį tyrimą; Mokslo komitetas aptarinėja naujus tyrimo projektus*). Sunku pasakyti, ar toks kai kurių sakinių turinio panašumas yra atsitiktinis, ar tai yra tam tikras įrankio ChatGPT trūkumas pagal užklausas generuojant iliustracijas.

Pasitaikė sakinių, kuriuose siekiama paprastai paaiškinti tam tikrus konceptus kaip *konstitucinis teismas, konstituciniai dokumentai, prezidento kanceliarija, komercinės reklamos* ir / ar tų konceptų esminius bruožus. Toks aiškinimo stilius yra enciklopedinis: 34) *Konstitucinis teismas yra institucija, kuri tikrina teisės aktų suderinamumą su konstitucija*; 35) *Valstybės vadovo įgaliojimai yra aiškiai apibrėžti konstituciniuose dokumentuose*; 36) *Prezidento kanceliarija yra atsakinga už įvairius valstybės administracinius reikalus*; 80) *Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus*. Šioje vietoje taip pat nėra iki galo aišku, ar ChatGPT tikslingai sugeneravo būtent tokias iliustracijas, ar tiesiog klaidingai suprato užklausas, bandant paaiškinti atitinkamas lemas ir aiškino tarsi patį lemos konceptą.

Pasitaikė bendro pobūdžio nekonkrečių sakinių, pvz., 26 sakinyje *Valstybinė komisija tiria korupcijos atvejus* žodžių junginyje *valstybinė komisija* nėra papildomo įprastinio patikslinimo, nurodančio konkrečias šios komisijos funkcijas (pvz., *valstybinė finansų komisija*⁵¹ ar pan.). Neretai sakiniuose sugeneruota tik svarbiausia informacija, nepateikiant papildomo turinio, aplinkybių. 20) *Futbolo komanda laimėjo turnyrą* – neįvardijamos papildomos detalės apie turnyrą ar papildomas su

⁵¹ Tokia komisija nėra Lietuvos visuomenės realija. **Realijomis** šiame darbe vadinami įvairūs pavadinimai, būdingi konkrečiai visuomenei (pvz., įvairių institucijų, valiutų, šalių pavadinimai). Kitaip tariant, tai ne kažkokie abstraktūs įvardijimai, o konkrečios, su realybe susijusios nuorodos.

futbolo komanda susijęs pažymins, 21) *Rytoj komanda vyks į susitikimą su klientais* – netikslinama, kokia tai komanda, 28) *Finansų komitetas rengia biudžetą* – nepatikslinama, koks biudžetas rengiamas, 38) *Šiame projekte reikia pateikti konkretų planą* – nenurodomos nei detalės apie projektą, nei apie planą, 33) *Anksčiau mes kalbėjome apie tai, ką galėtume daryti* – neįvardijama nei pokalbio tema, nei numatoma situacija ir pan. Toks nekonkretumas kai kuriuose sakiniuose patvirtina, jog apskritai negalima kalbėti apie ChatGPT generuojamų iliustracijų plačią kalbinę įvairovę.

Apibendrinant autentiškai skambančių sakinių turinį matyti, jog juose vaizduojamos gyvenimiškos, kasdienės situacijos, o sakiniuose atsiskleidžia įvairi tematika: darbo aplinkos (čia ir toliau žr. 4, 21, 50 sakinius p. 35-37), verslo (18, 72, 77, 79), prekybos (70, 80), tarpasmeninių santykių (6, 53, 58, 59, 61), pomėgių (10, 31, 44), sporto (20, 64), sveikatos (9, 49, 54), mokymosi (12, 24, 41), gamtos / oro sąlygų (1, 7, 67), tačiau dėl trūkstamo konteksto ne visada galima konkrečiai įvardinti kai kurių sakinių tematiką (40, 51, 52, 66 ir kt.). Tematikos įvairovę iš dalies nulėmė lemų sąrašo leksinė sudėtis: pavyzdžiui, natūralu, jog lemai *karinis* sugeneruoti būtent su militarizmu susiję iliustraciniai pavyzdžiai ar tai, jog lema *komanda* dažnai pasitaiko darbo ar sporto diskursuose, o tai atsiskleidžia jai sugeneruotose sakiniuose.

2.2.1.8. Autentiškai skambančių sakinių analizės apibendrinimas

Apibendrinant galima teigti, jog daugiau nei pusė (~55%) visų ChatGPT sugeneruotų sakinių yra autentiškai skambantys. Tai reiškia, jog jie yra logiški, rišlūs, jų prasmė yra lengvai suprantama, jie yra taisyklingi, tinkamai iliustruoja lemas ir absoliučios daugumos apklaustų gimtakalbių įvertinti kaip skambantys autentiškai. Remiantis gimtakalbių vertinimu galima teigti, jog kai kurie sakiniai sugeneruoti sakiniai skamba natūraliau už kitus.

Kaip rodo tolimesnės autentiškai skambančių sakinių analizės rezultatai, daugiausia (65%) sugeneruoti vientisiniai sakiniai, o sudėtiniai sakiniai (viso 35%) daugiausiai yra prijungiamieji. Sakiniuose daugiausiai pasitaikė laiko aplinkybių, tačiau rasta nemažai kokybės (būdo) ir vietos aplinkybių. Vientisiniuose sakiniuose dominuoja esamasis, o sudėtinuose – būtasis kartinis laikai. Didžiojoje daugumoje sakinių rasta tiesioginė nuosaka, dominuoja vyriškoji gramatinė giminė. Sakiniuose pasitaikė vien bendriniai žodžiai, visai nerasta šnekamosios kalbos leksika. Pasitaikė keletas antonimijos, sinonimijos, metonimijos atvejų (pastaruosius nulėmė lemų sąrašo sudėtis). Sakiniai dažniausiai yra konstatuojamieji, tačiau pasitaikė ir keletas prašymų. Kai kuriuose sakiniuose trūksta papildomos informacijos, jie yra nekonkretūs. Iliustracijose dominuoja darbo, verslo, prekybos, tarpasmeninių santykių, sveikatos, pomėgių ir kt. temos, tačiau kai kada dėl trūkstamo

konteksto sunku įvardinti sakinio tematiką, kartais skirtingų sakinių turinys labai panašus, nors iliustruojamos skirtingos lemos. Taigi, galima daryti išvadą, jog DI sugeneruoti autentiškai skambantys sakiniai nepasižymi ypatinga kalbine įvairove, nes kalbama apie 81 sakinį, o ne apie žymiai mažesnį kiekį.

Palyginus šiuos rezultatus su Robert'o Lew'o (2023) tyrimo rezultatais, kur jis apžvelgia ChatGPT angliškai generuojamų iliustracinių sakinių turinį matyti, jog ten pateikiamos išvados iš esmės sutampa su šio tyrimo išvadomis. R. Lew'o tyrime ChatGPT buvo prašoma sugeneruoti ŽS penkiolikai angliškų veiksmažodžių tokiu stiliumi, kaip tai yra daroma COBUILD žodyne (angl. *COBUILD-like entries*). Tyrimas atskleidė, jog ChatGPT pateikus išsamią užklausą, sugeneruoti iliustraciniai sakiniai būna įvairūs, detalūs, skambantys autentiškai, bet juose daugiausiai dominuoja būtasis laikas (lietuvių kalboje tai atitiktų būtąjį kartinį laiką) ir kai kada sakiniai būna pernelyg šabloniški (angl. *boilerplate*⁵²) (Lew 2023). Michael'o Rundell'o (2023) tyrimas taip pat atskleidė, jog ChatGPT generuoja sintaksiškai ir morfologiškai panašius sakinius, kai kurie jų skamba neautentiškai. Šio MA darbo tyrimas parodė, jog dauguma sugeneruotų iliustracinių sakinių skamba autentiškai (ne visi, žr. skyrius 2.2.2. ir 2.2.2.1.), juose dominuoja būtasis kartinis ir esamasis laikas, tačiau ne visi jie pasižymi konkretumu, dažnai yra tarpusavyje kalbiškai panašūs, taigi, šabloniški. Taigi, šio MA darbo rezultatai sutampa su jau atliktų tyrimų rezultatais.

2.2.2. Gimtakalbių vertinimu dirbtinai skambantys sakiniai

Toliau pateikiama 12 sakinių, kurie yra sintaksiškai bei semantiškai taisyklingi, tačiau respondentų buvo įvertinti kaip skambantys dirbtinai⁵³. Šių sakinių negalima pavadinti nei autentiškai skambančiais, nei haliucinacijomis, dėl to jie sudaro atskirą sugeneruotų sakinių pogrupį. 82-88 sakinius penki iš devynių respondentų įvertino skambančiais dirbtinai, 89 sakinį – šeši iš devynių, 90-92 sakinius – septyni iš devynių, o 93 sakinį – aštuoni iš devynių:

82) *Jonas, kuris⁵⁴ sėdi šalia manęs, yra mano geriausias draugas.*

83) *Miestas patyrė žiaurų karinį konfliktą.*

84) *Menininkas kūrė paveikslą visą savo gyvenimą.*

85) *Ši teisė yra garantuota konstituciniuose dokumentuose.*

⁵² Plačiau būdvardžio *boilerplate* reikšmę žr. čia: <https://dictionary.cambridge.org/dictionary/english/boilerplate>.

⁵³ Darbo autorės nuomone, šie sakiniai galėtų pasitaikyti (būti produkuojami) tam tikruose kalbiniuose kontekstuose, pavyzdžiui, 91 sakinyje galėtų pasitaikyti grožiniame tekste.

⁵⁴ Čia ir toliau sakinyje paryškintas tas žodis, kurį sakinyje iliustravo.

- 86) *Politikų **korupcija** dažnai pasireiškia, kai jie priima dovanas ar mokesčius už savo veiksmus arba politinių sprendimų priėmimą.*
- 87) *Mano kambaryje yra stalas, **kur** aš dirbu.*
- 88) ***Kokia** orlaivių sparta⁵⁵ yra didžiausia?*
- 89) ***Karinis** lyderis vadovavo operacijai.*
- 90) ***Komercinis** užsiėmimas yra susijęs su pelno siekimu.*
- 91) ***Kas** buvo tas keistasis garsas?*
- 92) *Jo gyvenimas buvo susijęs su nuolatine **karine** treniruote.*
- 93) *Tai yra **keistas** momentas mūsų istorijoje, kai viskas tapo neaišku.*

Verta atsižvelgti į tai, jog šie sakiniai nepasirodė nei autentiškai skambantys, nei haliucinacijos būtent MA darbo autoriui. Nors ir remtasi apklausos rezultatais, negalima atmesti tikimybės, jog toks vertinimas yra subjektyvus. Dėl to keliant klausimą, ar šiuos, nors ir autorės nuomone, pilnai suprantamus sakinius, būtų galima perkelti į hipotetinį žodyną (net jei kitiems gimtakalbiams jie nepasirodė skambantys natūraliai) galima teigti, jog jų perkelti į hipotetinį žodyną nebūtų galima.

2.2.2.1. Haliucinacijos

Šiame skyriuje trumpai pristatomos ChatGPT sugeneruotos haliucinacijos⁵⁶ (termino paaiškinimą žr. 1.4. skyriuje). Iš viso buvo sugeneruotos 54 haliucinacijos (~% 37 visų ChatGPT sugeneruotų iliustracinių sakinių). Haliucinacijos išskirtos pagal šiuos kriterijus: 1) sakinių turinys netikslus, jame nėra (aiškių) loginių ryšių; 2) jie sunkiai suprantami ar visai nesuprantami, absurdiški; 3) neretai (bet nebūtinai) jie netaisyklingi morfologiškai, ortografiškai; 4) sakiniai skamba dirbtinai, galima aiškiai matyti, jog juos sugeneravo DI, o tai patvirtina gimtakalbių apklausos rezultatai. Atsižvelgiant į išvardintus sakinius ypatumus, šių sakinių giluminė lingvistinė analizė darbe nėra atliekama ir toliau dalis jų pateikiami tik kaip pavyzdžiai⁵⁷.

Dėl savo iš pažiūros taisyklingos sintaksės, haliucinacijos iš pirmo žvilgsnio gali atrodyti sklandžios (plg. Ji et al. 2023: 3), tačiau dėl migloto turinio sakinių tekstas yra sunkiai suprantamas ir nelogiškas, pvz.:

⁵⁵ Nors žodžių junginio *orlaivių sparta* nėra nei ELEXIS Lithuanian Web 2021, nei Lithuanian Web 2014 tekstynuose, nei *Google* paieškoje, teoriškai jis galėtų egzistuoti. DLKŽ *sparta* aiškinama kaip *spartumas*, *tempas*, o keturi respondentai šį sakinį įvertino kaip skambantį natūraliai, t. y., jie sakinį palaikė prasmingu.

⁵⁶ Kai kurioms lemos „ilustruoti“ sugeneruotos vien haliucinacijos, t. y., jų IL nepasitaikė autentiškai skambančių sakinių. Šios lemos yra *kaip*, *kelti* ir *kažkoks*.

⁵⁷ Savaiame suprantama, jog šie sakiniai taip pat negali tinkamai iliustruoti lemu.

- 94) *Piliečių nerimas **kėlė**⁵⁸ iškilmingas demonstracijas prieš valdžią.*
- 95) *Sodininkas uždegė savo rankas dirbdamas su **karštu** geležiniu stovu.*
- 96) *Penktadienis yra savaitės **ketvirtas** dienos darbo laikas.*
- 97) ***Kariniai** drabužiai skirti kariams kovinei aprangai.*
- 98) ***Kartą** mane aplanko paslaptis, kartą - meilė.*

Kai kurie turinio prasme neaiškūs sakiniai taip pat yra ir morfologiškai netaisyklingi. Juose trūksta reikiamų diakritinių ženklų, pasitaiko rašybos klaidų:

- 99) *Mano avinas sriubos buvo per **karšta**, ir aš netyčia sudegiau liežuvį.*
- 100) *Jis **kėlė** sunkią krepšinio kamuolį aukštyn.*
- 101) *Prieš penkerius metus jis turėjo **keturis** mašinas.*
- 102) ***Kai** būsi grįžęs, aš jau busiu miegojęs.*
- 103) *Mes ruošsimės atostogoms, **kol** oras bus geras.*

Keliose haliucinacijose pastebėti netikri žodžiai, nors pagal žodžių šaknis galima nuspėti, su kokiais tikrais žodžiais jie yra susiję:

- 104) *Jis tarnavo **karinėje** karo laivyno laivoje.*
- 105) *Aš visada valgau pusryčius **kai** saulė teksta.*
- 106) *Ji skaitys knygą, **kol** laukėjimo metas pasibaigs.*
- 107) *Jis įsipareigojo atvykti į susitikimą tam tikru, labai **konkrečiu** laiku.*

Kai kur pasitaikė pagal prasmę netinkantys žodžiai, dėl to sakiniai tampa sunkiai suprantami:

- 108) *Rask mane ten, **kur** vėjas tyčia šnypščia lapais.*
- 109) *Rytoj mes **kelsime** naują lentyną į virtuvę.*
- 110) *Jo elgesys buvo labai **keistas** – jis kalbėjo beveik savaip.*

Aštuonios haliucinacijos⁵⁹ pasitaikė dešimčia sakinių iliustruotos lemos *kuo* IL, pvz.: *Skęsta greičiau **kuo** ilgiau trukdo pagalba; Pasisakysiu prieš tave, **kuo** reikia; Tegul tai atvyksta **kuo** anksčiau; Ji nori, kad viskas būtų **kuo** geriau pasiruošę.* Toliau, nagrinėjamame lemos sąrašo du kartus esančiai lemai *kaip* sugeneruotos daugiausiai vien haliucinacijos. Dauguma šių sakinių turėjo iliustruoti jungtuko reikšmę, reiškiant palyginimą (pvz., *Ši diena yra karšta **kaip** kūno liežuvis; Šis medis yra aukštesnis **kaip** tas*), o vienas sakinyje turėjo iliustruoti prieveiksmio reikšmę reiškiant būdo

⁵⁸ Čia ir toliau pavyzdžiuose paryškinti tie žodžiai, kurie turėjo būti iliustruojami sakiniiais (yra iliustruojami netinkamai).

⁵⁹ Du sakiniai iš lemos *kuo* IL priskirti autentiškai skambančių sakinių kategorijai.

patikslinimą (*Jūs turėsite pasirinkti, kaip mums geriau veikti ateityje*). Sunku paaiškinti, kodėl kai kurioms lemoms vietoje jas iliustruojančių sakinių generuojamos beveik vien haliucinacijos.

Apibendrinant būtų galima teigti, jog ChatGPT gali pateikti įvairaus pobūdžio netinkamų sakinių. Sugeneruoti sakiniai gali būti absurdiški turinio prasme arba iš pirmo žvilgsnio atrodyti sklandūs, bet jie vistiek netinkami nei išsamesnei analizei, nei perkėlimui į hipotetinį žodyną.

2.2.2.2. Kiti aspektai ir problemos

Daug kur sugeneruotuose pavyzdžiuose pastebima anglų kalbos įtaka. Tai lemia įrankio apmokymas daugiausiai anglų kalba (plg. Kohnke et al. 2023: 545). Anglų kalbos įtaką sugeneruotiems tekstams galima atskleisti keliais žemiau pateikiamais sugeneruotais pavyzdžiais (tiek autentiškai skambančiais, tiek haliucinacijomis):

- 1) *Jie ketina kreiptis į teisininką dėl teisinio patarimo. They are going to seek legal advice from a lawyer*⁶⁰.
- 2) *Kai aš atėjau į namus, ji jau miegojo. When I returned home, she was already sleeping.*
- 3) *Jūs turėsite pasirinkti, kaip mums geriau veikti ateityje. You'll have to choose how we should better act in the future.*
- 4) *Naujas restoranas mieste yra šiuo metu labai karštas vietinis kalbų. The new restaurant in town is a hot topic.*

Kaip matyti, neretai lietuviškai sugeneruotus sakinius galima būtų laikyti vertalais. Trumpai aptariant viršuje pateiktus pavyzdžius, lietuvių kalboje labiau įprasta būtų tiesiog pasakyti (1) *ketina kreiptis į teisininką* (nes jau yra aišku, **koku** tikslu), tačiau anglų kalboje neretai patikslinama *to seek legal advice from a lawyer*. Lietuvių kalboje taip pat įprasta sakyti *kreiptis patarimo*, o ne *kreiptis dėl patarimo*⁶¹. Konstrukcija *kai aš atėjau į namus* 2 sakinyje nėra būdinga lietuvių kalbai⁶², o 3 sakinyje išvelgiama angliška sintaksė. 4 sakinyje žodžio *karštas* pasitaikymas šia prasme nėra galimas lietuvių kalboje.

Pastebima dar viena problema, susijusi su netiksliai įgyvendinta užklausa: keturiuose sugeneruotuose pavyzdžiuose visai nepasitaikė lema, kurią tie pavyzdžiai turėtų iliustruoti. Pavyzdžiui, lemą *kas „ilustruoja“ sakinyje Tai yra mano draugas, kuris visada man padeda*. Ši

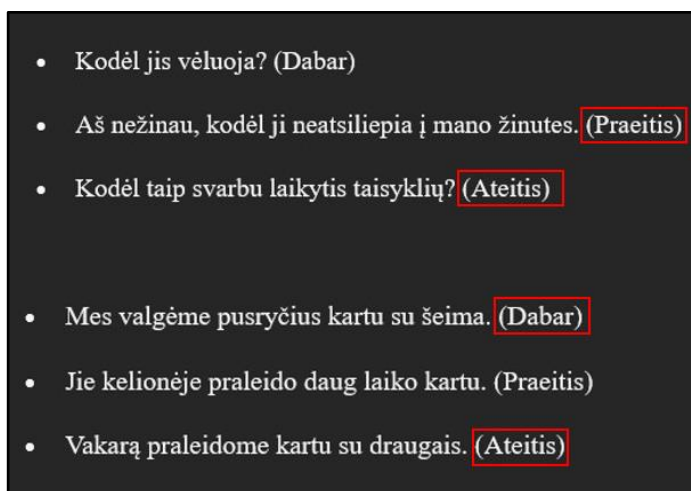
⁶⁰ Čia ir toliau sugeneruotus sakinius į anglų kalbą vertė MA darbo autorė.

⁶¹ ELEXIS Lithuanian Web 2021 tekстыne *kreiptis patarimo* pasitaiko 151 kartą, o *kreiptis dėl patarimo* – tik du kartus.

⁶² ELEXIS Lithuanian Web 2021 tekстыne konstrukcija *atėjau į namus* visai nepasitaiko. Plačiau apie šį atvejį skaityti p. 40.

problema ryški lemos *kilti* sakiniuose – 3 iš 5 jai sugeneruotų sakinių lemos visai neilustravo (*Vaikinas kėlė ranką, norėdamas atsistoti iš lovos; Po sunkaus darbo buvo sunku atsikelti iš kėdės; Piliečių nerimas kėlė iškilmingas demonstracijas prieš valdžią*). Šiais atvejais galima teigti, jog užklauso pateikimo ir atitinkamų rezultatų gavimo procesas nebuvo sėkmingas.

ChatGPT sugeneruotuose laukuose buvo pateikiama ir papildoma IL nepriklausanti informacija, kurios užklausoje nebuvo prašoma pateikti (žr. 7 paveikslą). Šioje vietoje galima kelti klausimą, ar įrankis apskritai tiksliai suprato autorės pateiktą užklausą. Papildomoje pateiktoje informacijoje neretai pateikiami išgalvoti lingvistiniai terminai, pvz., *jungiamasis junginys* (greta lemos *kad*), *rodomasis kiekinis įvardis* (greta lemos *kuris, kitas, kas*), *lyginamasis junginys* (greta lemos *kaip*), *laiko junginys* (greta lemos *kai*), *klaustukinis įvardis* (greta lemos *koks*) ir kt. Greta iliustracijų buvo pateikiami beprasmiškai reikšmės „patikslinimai“ kaip *moralinis judėjimas* (greta sakinio *Piliečių nerimas kėlė iškilmingas demonstracijas prieš valdžią*) ar *aukštytė tendencija* (greta sakinio *Palūžusi bendrovė pradeda kilti ir vėl tapti pelninga*). Pastebėtos kitos painios, sakinius papildomai „aiškinančios“ formuluotės, pvz., *Ši reikšmė padeda nurodyti arba pabrėžti veiksmo priežastį, sąlygą arba tikslą sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kad" yra naudojamas*. Tai rodo, jog haliucinacijos pasireiškia ne tik dalyje iliustracinių sakinių, bet ir pateikiamoje papildomoje informacijoje, o papildomai pateikiamos kalbinės informacijos (ne)taisyklingumas laikytinas sunkiai paaiškinamu ir atsitiktiniu.

- 
- Kodėl jis vėluoja? (Dabar)
 - Aš nežinau, kodėl ji neatsiliepia į mano žinutes. (Praeitis)
 - Kodėl taip svarbu laikytis taisyklių? (Ateitis)

 - Mes valgėme pusryčius kartu su šeima. (Dabar)
 - Jie kelionėje praleido daug laiko kartu. (Praeitis)
 - Vakara praleidome kartu su draugais. (Ateitis)

7 pav. Klaidinga papildomai sugeneruota gramatinė informacija.

2.2.3. BLKŽ esantys IL ir jų turinys

Šiame skyriuje aprašomi autorės pastebėti leksiniai BLKŽ esančių IL bruožai ir jų turinys. 40-ies lemos (iš nagrinėjamo lemos sąrašo) BLKŽ esančiuose IL suskaičiuoti viso **590** iliustraciniai

pavyzdžiai – įskaitant tiek jų lemas iliustruojančius sakinius, tiek ir žodžių junginius, priklausančius IL.

Atsivertus bet kurį BLKŽ ŽS matyti, jog ten esantys IL yra sukirčiuoti, o straipsniuose, kuriuose pasitaiko kiek daugiau iliustracinių pavyzdžių pirmiausia pateikiami žodžių junginiai, o tik po jų seka pilni sakiniai (pvz., lemu *komisija, klausimas, komitetas, kalbėti, kaina, komanda, kilti* IL). Šiuose žodžių junginiuose laužtiniuose skliaustuose neretai pateikiamos galimos alternatyvios formuluotės, pvz.: **Kilo**⁶³ *gaisras [kãras, bãdas, triũkšmas, lìgos, uždegìmas]. Archỹvo patikrinimo komisija [Komisija archỹvui patikrinti]. Komeřcinis pãstatas [transportas]. Leņgvas [pãprastas, áiškus, naivũs] klãusimas. Didmeninẽ [mažmeninẽ, valstýbẽs reguliũojama, rinkos, nãftos, dũjũ] kãina.*

Šių alternatyvių formuluočių pateikimas matomas ir kai kuriuose iliustraciniuose sakiniuose ir yra įterpiamas į sakinį skliaustuose: *Susitikę vienas [kitas] kito nebepažintume. Kamẽ [slỹpi] Jũsu sėkmẽs priežastis? Projėktã numatýta svarstýti komisijoje [Projėktã numãtẽ svarstýti komisija]. Šis klãusimas dãr svarstomas [nagrinėjamas, áiškinamas, sprėndžiamas]. Pasisvėikinau sũ kiekvienu svečiũ [sũ kiekvienu viešniã].* Kartais pačiame iliustraciniame sakinyje, skliausteliuose nepasviruoju šriftu patikslinama lemos reikšmė: *Išsiskirstẽme kãs sãu* (kiekvienas atskirai). *Kiekvienãm pirmas karãtas* (pirmieji lytiniai santykiai) *bũna skirtingas. Taiã tẽrà laĩko klãusimas* (anksčiau ar vėliau bus išspręsta). *Kalãba kaip iš knygos* (aiškiai, sklandžiai). *Kalãba kaip iš mẽdžio iškritęs* (kvailai). Taigi, neabejotinai matyti, jog pateikiami pavyzdžiai yra įvairūs, vaizdingi (paskutiniame iš minėtų iliustracinių sakinių matyti perkeltinė reikšmė), o tai pagrindžiama tuo, jog BLKŽ IL rinkti iš įvairių autentiškos kalbos šaltinių (plg. Murmulaitytė 2011).

Pereinant prie BLKŽ IL atsiskleidžiančių papildomus reikšmės aspektų, žodyne dažnai pastebima šnekamosios kalbos pavyzdžių, neretai indikuojamų trumpiniu šnek., pvz.: *Jũms tekiũba tókios žũvys, kuř pũ dẽšimt grãmũ svẽria. Nuplėšk nuũ bãtũ pãdo kãinã. Nelabaĩ smagũ, kaĩ gãuni kẽturis. Kũrk namũ!* Pasitaikė ir kitų šnekamajai kalbai būdingų pavyzdžių. Juose galima pastebėti tokias dalelytes kaip *taip, argi, dar, mat*, stilistiškai žymėtą leksikã (pvz., *tamsta, valandėlẽ*), sušukimus: *Dãlį brãškiũ virsime, kitãs taĩp suvãlgysime. Kaĩp, aĩgi jũs nepapietãusite sũ mumis? Kaĩp tamsta pasakeĩ? Dãr užtrũksiu kókiaũ valandėlẽ. Gãl dãr sulčĩũ kãm ipilti? Lėiskite dãr kiek pagalvũti. Mãt kóks gudrũs! Kókiaus gerũs órus žãda kitã savãite! Kaĩp tũ galėjai šitaip pasielgti!*

Buvo pastebėta perkeltinės reikšmės pavyzdžių: *Kilti karjẽros laiptais. Rinkoje dãrosi karšta* (pavojinga). Be abejo, perkeltinės reikšmės pavyzdžiuose rasta ir daugiau, pavyzdžiui, metonimija – *Šãlį užgróbẽ karinẽ jėgã; Kanceliãrijai pãvesta priimti, pẽrduoti ir sãugoti visũs dokumentũs; Įmonẽ dẽl netesýbũ raštũ kreĩpẽsi į teĩsmã; Vienã valdžiã keĩsdavo kitã.* Rasta metaforinių perkeltinių

⁶³ Čia ir toliau pavyzdžiuose iš BLKŽ IL paryškinami žodžiai, kuriuos iliustruoja sakiniai.

reikšmių: **Kárštas** patriotas; **Karšti** bučiniai; Populiarumo [šlovės] **káina**; *Katė* glóstoma kùpra **kēlia**.

BLKŽ iliustraciniuose pavyzdžiuose per leksinę raišką dažnai atsiskleidžia ypatingas kalbos vaizdingumas (pvz., *tvoros poškéjo*; *galvas suglaudę*; *saldzialiežuvis*; *glūdi blogio šaknys*; *auksinis smėlis*): *Užėjo tokiė šalčiai, kàd tvōros poškéjo. Ilgaĩ abù kalbėjosi gálvas suglaudę. Netikék tuō saldzialiežuviù, jìs vísq laikq kùria nebútas istōrijas. Dár jìs mùms čià áiškins, kq darýti – kám kalbét, ò kám patylét. Kamè glūdi blōgio šaknys? Kaip gēra mēgautis auksiniù smėliù ir káršta sáule.*

Tam tikros iliustracijos BLKŽ žymimos specialiais trumpiniais⁶⁴. BLKŽ pasitaiko tautosakos (trumpinys *flk.*) pavyzdžių: *Nekàsk duobės kitám – pàts įkrìsi. Nè vísa áuksas, kàs áuksu žìba. Aukštai kùlsi, žemaĩ krìsi. Žēmè kēlia žōlę, žolė kēlia rāsq.* Lemos *kelti* ŽS išskirtos 26 reikšmės, šie pavyzdžiai neretai buvo iš įvairių sričių. *Kelti* IL rasti 2 pavyzdžiai iš technikos srities (trumpinys *tech.*): *Keliamàsis krānas [mechanizmas]. Keliamóji galià.* Taip pat pasitaikė 2 pavyzdžiai iš matematikos (*mat.*) srities: *Stātmenį ĩ tiēsiąq kēlti. Kēlti skaičių ketvirtuoju láipsniu [kvadratù, kubù].* Rasta po sakinį iš religijos (*relig.*), teisės (*teis.*) sričių: *Kristus trēčiq diēnq kēlési iš numirusiųjų. Kēlti býla teismė.* Kitur rasti sakiniai iš sporto (*sport.*) srities: **Konkretieji spòrtiniai konfliktai.** Sporto, religijos ir pan. kontekstų BLKŽ esančiuose IL pasitaikė ir daugiau, bet ne visi tokie sakiniai buvo pažymėti specialiais trumpiniais⁶⁵.

Papildomai verta paminėti, jog BLKŽ ŽS pabaigoje kai kada pateikiamos papildomos antonimijos, sinonimijos nuorodos, pvz.: lemos *kartu* antonimas *skyrium* ar lemos *kas* sinonimai *kuris, kiekvienas* ir kt. ŽS apačioje neretai pateikiama konkrečios lemos frazeologizmų sritis, kurioje taip pat pateikiami šiame darbe plačiau neaptariami iliustraciniai sakiniai (žr. 8 paveikslą):

♦ **añt kójų kēlti** (kq) skatinti veikti: *Jìs vyrus añt kójų kēlė ir tēsė drqsq žyqj. (kaip) añt šakių kēlti (kq) nepelnytai, perdėtai girti: *Mótina sàvo dukterį (kaip) añt šakių kēlia. ešerį kēlti (prieš kq) didžiuotis, pūstis: *Nekėlk ešerio prieš kitus! ĩ padanges kēlti (kq) labai girti; aukštinti, garbinti: *Bè reikalo tū kelì manè ĩ padanges. ĩ puikybę kēltis* darytis išdidžiam, puikiam; puikti: *Jìs mán nepatiñka – labai ĩ puikybę kēliasi. kalbàs kēlti (apie kq) apkalbinėti, šmeižti: *Gál liaūsis ir jìe vīenq kartq apìe mùs kalbàs kēlti. kēlti aūdrą [audrās]* šėlti, triukšmauti, siausti: *Vienì pasilikę vaikaĩ aūdrq [audrās] kēlia. kēlti bálšq (prieš kq) priešintis, atsilalbinėti, priešgyniauti: *Nekėlk bálso prieš tēvus! kēlti báltq vėliavą* prisipažinti nugalėtam; sin. pasiduoti: *Krēpšininkai neketina kēlti baltōs vėliavos. kēlti bangàs bartis, konfliktuoti, siautėti: *Kò tū kelì bangàs, vīskas būs geraĩ! kēlti kóją* (iš ko, ĩ kq, pas kq) trumpam išeiti ar užėiti kur: *Jìs nekėlė kójos iš namų. Bijójau kēlti kójq ĩ jò namūs. Vísq savaitę nekėlė kójos pàs mùs. Kàd mán kójos čià daugiau nekeltum! kēlti rañkq (prieš kq) (kėsintis) mušti: *Jìs ir añksčiau rañkq prieš jq kėldavo. Nekėlk rañkos prieš tēvq! kepūrę kēlti (kam) pasisveikinti ar atsisveikinti: *Áš jam kepūrės nekėliau ir nekėsiu. nójš kēlti* didžiuotis, puikti, pūstis: *Jìs labai nójš kēlia, tóks pasipūtęs. pūgą kēlti 1. siausti, šėlti: *Diēnq miēga, ò naktį pūgq kēlia. 2. (dėl ko) piktai priekaištauti, bartis: Nevertėjo dėl tō pūgōs kēlti. skiaūterę kēlti (prieš kq) didžiuotis, pūstis, keltis: *Kò tū skiaūterę kelì prieš mùs lýg kòks pōnas. sparnūs kēlti rengtis išvykti: *Duktė kq tik grįžo iš užsienio, dabar vėl sparnūs kēlia. uodegą kēlti rengtis išvykti, išvažiuoti: *Vīenas vaikas jau išvažiavo ĩ užsienį, dabar ir kītas jau kēlia uodegq. vétras kēlti (dėl ko) siautėti, triukšmauti: *Jìs dėl meñko niēko vétras kēlia.**************

8 pav. Po lemos *kelti* ŽS pateikiama frazeologizmų skiltis.

⁶⁴ Jie paaiškinami papildomoje BLKŽ skiltyje *Ženklaĩ, šriftaĩ ir sutrumpinimai*.

⁶⁵ Skiltyje *Ženklaĩ, šriftaĩ ir sutrumpinimai* virš sutrumpinimų sąrašo rašoma, jog „[š]iuo darbo etapu daugelis sutrumpinimų nematomi“, todėl nežinoma, ar tai, jog specialūs kontekstai pažymėti ne visur yra laikinas dalykas, ar trumpiniai pažymėti tik kai kurie sakiniai.

Apibendrinant galima teigti, jog BLKŽ kalbiškai itin turtinguose IL pasitaiko įvairūs kasdieniai visuomenės kontekstai bei tematika (pvz., folkloras, technika, įvairūs mokslai, religija), o sakiniai yra tiek šnekamosios, tiek rašytinės kalbos. Iliustracijų vaizdingumas taip pat atsiskleidžia per pateikiamų pavyzdžių sintaksinę įvairovę (randami žodžių junginiai ir sakiniai), leksinių alternatyvų nuorodas, perkeltinių reikšmių atskleidimą.

2.2.4. ChatGPT sugeneruotų ir BLKŽ esančių IL palyginimas

2.2.1.1. skyriuje minima, jog pasitaikė 13 lemu (komisija, klausimas, komitetas, kalbėti, kaina, kiekvienas, keli, komanda, kartu, kodėl, keisti, kreiptis, kadangi), kurioms visas ChatGPT sugeneruotas iliustracijas galima laikyti autentiškomis (visi ChatGPT joms sugeneruoti sakiniai daugumos gimtakalbių įvertinti kaip skambantys natūraliai). Šiame skyriuje toliau bus lyginama, kaip šias lemas iliustravo ChatGPT ir kaip jos iliustruojamos BLKŽ.

Lemą komisija ChatGPT iliustravo 3 pavyzdžiais:

- Valstybinė komisija tiria korupcijos atvejus.
- Mokslo komisija nusprendė finansuoti šį tyrimą.
- Komisija patvirtino galutinį ataskaitos dokumentą.

Ši lema BLKŽ iliustruota 8 pavyzdžiais⁶⁶: *Egzaminų komisija. Valstybinė kainų ir energetikos kontrolės komisija. Įmonės likvidavimo komisija. Archyvo patikrinimo komisija [Komisija archyvu patikrinti]. Revizijos komisija. Sudaryti nuolatinės ir laikinos komisijas. Reikia nustatyti komisijos sudėtį. Projektą numatyta svarstyti komisijoje [Projektą numatė svarstyti komisija]*. Matyti, jog 3 ChatGPT sugeneruoti sakiniai yra sintaksiškai panašūs (pirmiausia pateikiamas veiksnys, tuomet – tarinys ir į klausimą *ką* atsakantis papildinys), o BLKŽ pateikiama daugiau ir sintaksiškai įvairesnių pavyzdžių. Pavyzdžiui, žodyno IL randami 5 daiktavardiniai junginiai ir 3 beasmeniai sakiniai. BLKŽ sakiniuose visuomenės realija *Valstybinė kainų ir energetikos kontrolės komisija*⁶⁷. ChatGPT ir BLKŽ sakiniuose daiktavardiniuose junginiuose dažniausiai patikslinama komisijos paskirtis (ChatGPT – *valstybinė / mokslo komisija*, BLKŽ – *egzaminų / įmonės likvidavimo / revizijos komisija* ir kt.), tačiau šie patikslinimai BLKŽ yra detalesni, paskirtį labiau patikslinantys. Taip pat matyti, jog ChatGPT lema iliustruoja tik vienaskaitos vardininko linksnyje, o BLKŽ lema dar iliustruojama daugiskaitos galininko, vienaskaitos kilmininko ir vietininko linksniuose.

Lemą klausimas ChatGPT iliustravo 3 sakiniiais:

⁶⁶ Vienu pavyzdžiu laikomas teksto vienetas (žodžių junginys ar sakiny) iki taško, pvz., *Projektą numatyta svarstyti komisijoje [Projektą numatė svarstyti komisija]* laikomas vienu iliustraciniu pavyzdžiu, o ne dviem.

⁶⁷ Tai, ar konkretus pavadinimas / žodis yra realija, patikrinama pavadinimą įvedus į internetinę paiešką.

- Klausimas, kurį tu kėlei, yra labai svarbus.
- Mokytojas užduoda mums daug klausimų.
- Ar tu gali atsakyti į mano klausimą?

BLKŽ eksplikuotos 5 sunumeruotos lemos reikšmės⁶⁸, ji iš viso iliustruota 22 pavyzdžiais: **1.** *Leñgvas [pàprastas, áiškus, naivùs] kláusimas. Kélti [dúoti] kláusimą. Atsakyti į kláusimus žodžiù [raštù]. Prášom pateikti kláusimą. Kàs turétų kláusimų? Aĩ kláusimų bús? Turiù kláusimą pranešėjui. Jám kilo daũg kláusimų. Tù mán úždavei suñkų kláusimą.* **2.** *Šis kláusimas dár svařstomas [nagrinėjamas, áiškinamas, spréndžiamas]. Palikti kláusimą neišspréstą [átvirą]. Pranešimas svařstomu kláusimu. Susirašinėjimas švietimo kláusimais. Turi búti išspréstas kláusimas dėl turto gražinimo. Tai terà laiko kláusimas (anksčiau ar vėliau bus išspręsta).* **3.** *Dėl júsų dalyvavimo konkurse nekyla jokių kláusimų. Kyla kláusimas, aĩ jì yra sąžininga. Manè kankina kláusimas, aĩ tai tiesà.* **4.** *Pranešimas teisinio švietimo kláusimais. Šis móklininkas daũg rášè bendròsios kalbòtyros kláusimais.* **5.** *Màno pašnekòvas šokinėjo nuò vieno kláusimo priè kito. Mòkytoja nuramino, kàd šiù kláusimu nereikia daũg rašyti – užtèks kelių pastráipų.* Iš ChatGPT sugeneruotų pavyzdžių matyti, jog jais iliustruojama BLKŽ skiriama pirmoji lemos reikšmė (*kreipimasis, į kurį norima gauti atsakymą*). Lyginant su lemos komisija iliustracijomis, šie ChatGPT sakiniai yra morfologiškai, sintaksiškai kiek įvairesni: lema sakiniuose pasitaiko vienaskaitos vardininke ir galininke, daugiskaitos kilmininke. ChatGPT sugeneruotas klausiamasis sakiny (jie pasitaikė ir BLKŽ IL). Trys ChatGPT su lema susijusios formuluetės (*kelti klausimą, užduoti klausimą, atsakyti į klausimą*) taip pat randamos šios lemos IL BLKŽ prie pirmosios reikšmės.

Lemą *komitetas* ChatGPT iliustravo 3 sakiniiais:

- Mokslo komitetas aptarinėja naujus tyrimo projektus.
- Finansų komitetas rengia biudžetą.
- Nusikaltimų prevencijos komitetas turi parengti veiksmų planą.

Tai pačiai lemai BLKŽ eksplikuojamos dvi reikšmės ir pateikiami iš viso 6 pavyzdžiai: **1.** *Sveikàtos [káimo] reikalų komitètas. Šis kláusimas svarstýtas Seĩmo Téisès ir teisétvarkos [Sociàliniù reikalų ir dárbo] komitetè. Parlameñtai turi nuolatiniùs komitetùs aĩ komisijas įvairiù įstàtymų projèktų išankstiniam svařstymui ir anàlizei.* **2.** *Mokinių ir tèvų komitètas. Konfereñcijos organizàcinis komitètas. Stovýklos pìrmininkui pàdeda stovýklos organizàcinis komitètas.* ChatGPT sugeneruoti sakiniai sintaksiškai labai panašūs: veiksnys *komitetas* visuomet yra vienaskaitos

⁶⁸ Šiame MA darbe lemų eksplikavimo aspektas nėra plačiau aptariamas, nes lemų eksplikavimas nėra IL dalis, tačiau tai neabejotinai yra su jais glaudžiai susiję. BLKŽ prieš eksplikuojamas lemas paaiškinama jų reikšmė ir reikšminiai atspalviai, tačiau šie tekstai šiame MA darbe pateikiami tik kartas nuo karto.

vardininke ir juo prasideda sakiniai, sekantis tarinys yra esamojo laiko, o BLKŽ viename pavyzdyje pasitaiko lemos daugiskaitos galininkas, vienaskaitos vietininkas, pateikiami 3 daiktavardiniai junginiai. Lyginant su BLKŽ matyti, jog ChatGPT sugeneravo pirmąją BLKŽ reikšmę (*institucija, vadovaujanti kuriai nors valstybinės veiklos sričiai arba vykdanči specialias užduotis*) iliustruojančius pavyzdžius. Dviejuose BLKŽ pavyzdžiuose yra vaizduojamos 4 Lietuvos realijos⁶⁹ (*sveikatos reikalų / kaimo reikalų / Seimo Teisės ir teisėtvarkos / Socialinių reikalų ir darbo komitetas*)⁷⁰. ChatGPT komiteto funkcija patikslinama daiktavardžiais *mokslo / finansų / nusikaltimų prevencijos*, o BLKŽ (be jau minėtų) – *mokinių ir tėvų / konferencijos (stovyklos)*.

Lemą *kalbėti* ChatGPT iliustravo 3 sakiniais:

- Ji kalba labai aiškiai ir suprantamai.
- Anksčiau mes kalbėjome apie tai, ką galėtume daryti.
- Po valandos jis pradės kalbėti apie savo keliones.

Šiai lemai BLKŽ eksplikuojamos 6 reikšmės ir ji iliustruojama 27 pavyzdžiais: **1.** *Balsù [týliai, pašnibždomìs, áiškiei, suñkiai, nesuprañtamai, greitai, lètai, šaltaì, karštaì] kalbèti. Kalbèkite garsiaù! Gálgi sènis tièsq mùms kalba? Jìs kalbèjo peř rãdijã. Mókytoja kalbèjo susijãudinsi, užsikirsdamà. Trumpai kalbant, reikalas svarbùs. Daùg žinók, mãža kalbék. Kàs daùg kalba, mažaì dirba. Mán skaùda gérklę, àš negaliù kalbèti. Kalba kaip iš knýgos (aiškiai, sklandžiai). Kalba kaip iš mèdžio iškritęs (kvailai).* **2.** *Ař kalbèjai sù juò dèl manęs, dèl atóstogų? Sù nepažįstamu jìs kalba nuolankiaì. Àš kalbù telefonù. Kalbèsiu visų skýriaus žmonių vardù. Kalbasi lietùviškai. Ilgaì abù kalbèjosi gálvas suglaùdę.* **3.** *Lèktorius kalbèjo apie ekologijos problemàs. Rytój apie mùdu visas káimas kalbès. Kalbék nekalbèjæs apie internetinius sukçiùs – gyvèntojai iš svetimų bédų nepasimóko. Ař dár ką nòrs tùrite kalbamam kláusimam pasakýti.* **4.** *Jìs kalbèjo lietùviškai. Vaikas jaù kalba. Nè visas jaunimas móka kalbèti tařmiškai.* **5.** *Senùtè kalbèjo póterius.* **6.** *Fãktai áiškiai kalba. Skaìčiai kalba kã kita.* Lyginant su BLKŽ reikšmių definicijomis, 3 ChatGPT pavyzdžiai atspindi tris atskiras lemos reikšmes. Pirmasis ChatGPT sakinyje atspindi pirmąją (*reikšti mintis balsu, sakyti žodžius*), antrasis – antrąją (*bendrauti kalba, šnekėti*), o trečiasis – trečiąją (*pasakoti, aptarti*) BLKŽ skiriamas lemos reikšmes. BLKŽ IL randama panašumų su ChatGPT turiniu (ChatGPT – *ji kalba labai aiškiai*, BLKŽ – *aiškiai kalbėti*). Kadangi BLKŽ pavyzdžių yra 9 kartus daugiau, natūralu, jog

⁶⁹ Atskira realija dar galima laikyti *Seimą*.

⁷⁰ Aptariant trečiajame ChatGPT sugeneruotame sakinyje pasitaikiusį žodžių junginį *nusikaltimų prevencijos komitetas*, nors ir esama Tarpministerinio nusikaltimų ir radikalizacijos prevencijos komiteto (žr. šiame dokumente - https://vrm.lrv.lt/uploads/vrm/documents/files/LT_versija/Viesasis_saugumas/Koncepcijos/Koncepcija%2020201219_g_alut.pdf), vargu, ar tai būtų galima laikyti realija ChatGPT sugeneruotame sakinyje.

ten galima įžvelgti platesnę kalbinę, iliustracinę įvairovę (pvz., žodyno IL be konstatuojamųjų dar pasitaiko šaukiamasis, klausiamieji sakiniai).

Lemą *kaina* ChatGPT iliustravo 3 sakiniais:

- Šios knygos kaina yra 20 dolerių.
- Prieš keletą metų automobilio kaina buvo žymiai mažesnė.
- Kokia bus būsimų metų žaliavų kaina?

BLKŽ išskirtos 3 lemos *kaina* reikšmės, ji iliustruota 12 pavyzdžių: **1.** *Didmeninė [mažmeninė, valstybės reguliuojama, rinkos, naftos, dujų] kaina. Didelės [māžos, palañkios] kainos. Kainos kylā [āuga, kriñta, mažėja, didėja]. Numūsti [pakėlti, užkėlti, sumāžinti, padidinti] kainą. Derėtis dėl kainos. Kokiā šių bātų kaina?* **2.** *Nuplėšk nuō bātų pādo kainą. Iš švarko apykaklės kypo kaina. Pažiūrėk, koks gamintojas añt kainos parašytas.* **3.** *Populiarumo [šlovės] kaina. Kiekvienā pėrgalē turi savo kainą. Savo pāsiekei, tačiaū kokia kaina?* ChatGPT sugeneruotuose sakiniuose iliustruojama pirmoji lemos reikšmė (*vertė pinigais*) pagal BLKŽ. *Kainos* pažyminiai sakiniuose yra leksiškai įvairūs (ChatGPT – *knygos / automobilio / žaliavų*, BLKŽ – *didmeninė, mažmeninė, naftos, dujų / batų / populiarumo* ir kt., taigi, BLKŽ randama daugiau šių pažyminių), o ChatGPT sugeneruotose iliustracijose lema pasitaiko tik vardininko linksnyje. ChatGPT ir BLKŽ pasitaiko iliustracinio turinio bendrumų (ChatGPT – *kaina buvo žymiai mažesnė*, BLKŽ – *kainos mažėja*).

Lemą *kiekvienas* ChatGPT iliustravo 3 sakiniais:

- Kiekvienas mokinys turi baigti namų darbus iki pirmadienio.
- Kiekvienas svečias buvo maloniai sutiktas.
- Rytoj kiekvienas darbuotojas turės apmokėti mokesčius.

BLKŽ skiriami du nesunumeruoti lemos reikšminiai atspalviai (*bet kuris, imant iš eilės ir visi iš eilės, skaičiuojant po vieną*). Lema iliustruojama 6 sakiniais: *Kiekvienām [kiekvienai] neįtiksi. Kiekvieno [kiekvienods] nepaklausi. / Nē kiekvienā mōteris mėgsta šeimininkāuti. Pasisvéikināu sū kiekvienu svečių [sū kiekviēna viešniā]. Kiekvienas atostogas praleidžiame kāime. Kiekvienais metais važiuoju į Pālangą. Matyti, jog visi ChatGPT sugeneruoti sakiniai atitinka BLKŽ antrąją reikšmę (*visi iš eilės, skaičiuojant po vieną*). Žodžių junginys *kiekvienas svečias* pasitaikė tiek ChatGPT sugeneruotuose sakiniuose, tiek BLKŽ (BLKŽ greta dar pateikiama leksinė alternatyva *kiekviena viešnia*). BLKŽ žodžiu *kiekvienas* kalbama ne tik apie žmones, bet ir apie abstrakčius objektus (*atostogos, metai*). ChatGPT lema referuoja tik į žmones (*kiekvienas mokinys / svečias / darbuotojas*), pasitaiko tik vardininko linksnyje.*

Lemą *keli* ChatGPT iliustravo 3 sakiniais:

- Kelios mergaitės žaidžia kieme.
- Prieš kelis metus jis lankė tą pačią mokyklą.
- Po kelių valandų jie pasiekė tikslą.

Šiai lemai BLKŽ eksplikuotos trys reikšmės, ji iliustruota 7 pavyzdžiais: **1.** *Kelių talkiniškų jūms reikia? Keliąs dieną trūks konferencija?* **2.** *Taip ir nesupratai, keli jiė teñ buvo. Pažiūrėk, keliems svečiams [kelioms viešnioms] trūksta puodėlių.* **3.** *Atėjo kėlios kaimynės. Vokiškai móku kelis žodžiūs. Keliuose būtuose nebeveikia telefónspynės.* Lyginant su BLKŽ reikšmių eksplikavimu, visi ChatGPT sugeneruoti sakiniai atitinka trečiąją BLKŽ reikšmę (*nedidelis skaičius, daugiau negu du*). Pirmame ChatGPT sugeneruotame sakinyje referuojama į žmonių grupę (*kelios mergaitės*), o kituose dviejuose – į laiką (*metus, valandas*). BLKŽ daugiausiai referuojama į žmones (*keli / keli talkininkai / svečiai / kelios viešnios / kaimynės*), bet taip pat ir į objektus (*kelios dienos / keli žodžiai / butai*).

Lemai komanda ChatGPT sugeneravo 3 sakinius:

- Futbolo komanda laimėjo turnyrą.
- Komanda dirbo visą naktį, kad baigtų projektą laiku.
- Rytoj komanda vyks į susitikimą su klientais.

Šiai lemai BLKŽ išskiriamos 6 reikšmės, ji iliustruojama 25 pavyzdžiais: **1.** *Šaudyti pagal komandą. Stotiės budėtojas davė komandą uždaryti geležinkelio pervažą. / Šuõ neklaũso komandų „Palik!“ ir „Pàs manė!. Kalýtė klaũso komandos „Ateĩk!“.* **2.** *Stipriáusios [pajėgiáusios] krepšinio komandos. Komandos kapitõnas. Ši fùtbolo komanda užėmė pirmąją vietą. Dabař šis krėpšininkas žaidžia kitojė komandoje. Mūsų komanda nenusilėido priešininkams. Añtrą kėlinį komanda žaidė dešimtiese. Pirmenybių uždarymo ceremonijoje komandos prizininkės apdovanotos taurėmis.* **3.** *Gaišrininkų [ugniagesių] komanda. Lietuvos kariuomenės komanda. Komanda paprastai atliėka kokią nors tarnybinię užduotį arba darbą.* **4.** *Komandos eilutė. Kompiuteryje pasiriñkite komandą „Spáusdinti“. Kompiuteris nemãsto, jis tik skaičiuoja ir vykdo komandas. Komanda nurõdoma, kõkias operãcijas kompiuteris turi atlikti, iš kur imti duomenis, kur pateikti skaičiãvimų rezultatus ir iš kur imti kitą komandą.* **5.** *Į laivo komandą įeina dėnio komanda, mašinu-katilinės komanda ir aptarnaujantis personãlas. Ledė įkãlinto laivo komanda sutiko Naujõsius metus. Kařtais króvinį už bõrto išmėsdavo patì laivo komanda, norėdama laivą padaryti manevringesnį arba apsãugoti nuõ apsvertimo.* **6.** *Virtuvės šėfas formuoja naują komandą. Neleñgva suburti gėrą pedagogų komandą. Kažkadà mės dirbome vienojė komandoje. Paprastai komandoje dirba nuõ dviejų iki penkiõlikos žmonių.* Remiantis BLKŽ, ChatGPT sakiniuose iliustruojamos dvi lemos reikšmės. Antroji reikšmė

pagal BLKŽ (*kartu žaidžiančių sportininkų grupė, viena iš rungtyniaujančių šalių*) iliustruojama ChatGPT sugeneruotame sakinyje *Futbolo komanda laimėjo turnyrą. Šeštoji reikšmė (žmonių grupė, darniai dirbanti tam tikrą darbą, atliekanti tam tikrą užduotį)* iliustruojama sakinių *Komanda dirbo visą naktį, kad baigtų projektą laiku; Rytoj komanda vyks į susitikimą su klientais.* ChatGPT pasitaikė tik vienas lemos pažymins futbolo komanda (šis junginys taip pat randamas BLKŽ IL), lema pasitaikė vien vienaskaitos vardininko linksnyje. BLKŽ pasitaikė daugiau daiktavardinių junginių su lema: *gaisrininkų [ugniagesių] / laivo / mašinų-katilinės komanda* ir kt. Žodyne randama realija *Lietuvos kariuomenė* (žodžių junginyje su lema *Lietuvos kariuomenės komanda*). ChatGPT sakiniai turinio prasme turi panašumų su esančiais BLKŽ (pvz., ChatGPT – *Futbolo komanda laimėjo turnyrą*, BLKŽ – *Ši futbolo komanda užėmė pirmąją vietą*), tačiau turint omenyje sugeneruotų sakinių ir žodyne pateikiamų sakinių kiekio skirtumą, tai nėra lengva palyginti.

Lemą *kartu* ChatGPT iliustravo 3 sakiniiais:

- Mes valgėme pusryčius kartu su šeima.
- Jie kelionėje praleido daug laiko kartu.
- Vakara praleidome kartu su draugais.

Šiai lemai BLKŽ eksplikuojamos 2 reikšmės, ji iliustruojama 8 pavyzdžiais: **1.** *Eimė kartū! Jiē visadà kartū – iř namiē, iř darbē, iř laisvālaikiu. Linksminkitēs kartū sū mumīs. Ji atsisākē dalyvāuti laidojē kartū sū pagarsėjusiu radikalū. Mēs kartū jaū dvīdēšimt penkeri mētai [dvīdēšimt peñkerius metūs].* **2.** *Šiltà iř kartū nerūpestinga šypsena. Tragikòmiškas – tai liūdnas iř kartū juokingas. Kaimyninėje šalyje besitęsianti krizė dāro įtaką Lietuvōs veřslui, kartū iř kapitālo riñkai.* Lyginant su BLKŽ, ChatGPT iliustruojama reikšmė atitinka pirmąją žodyno reikšmę (*viename būryje ar vienoje kompanijoje*). Dviejuose ChatGPT sugeneruotuose sakiniuose pasitaiko veiksmazodinis junginys *praleisti kartu* (šis junginys nepasitaiko BLKŽ IL). BLKŽ be konstatuojamųjų sakinių pasitaikė vienas šaukiamasis, IL prie antrosios lemos reikšmės pasitaikė realija *Lietuva (Kaimyninėje šalyje besitęsianti krizė daro įtaką Lietuvos verslui)*.

Lemai *kodėl* ChatGPT sugeneravo 3 sakinius:

- Kodėl jis vėluoja?
- Aš nežinau, kodėl ji neatsiliepia į mano žinutes.
- Kodėl taip svarbu laikytis taisyklių?

BLKŽ šią lemą taip pat iliustruoja 3 pavyzdžiais: *Kodėl tū nebuvaĩ darbē? Kodėl gi jūms neįsigijus šuñs? Išsiáiskinta, kodēl skreñdant užgēsò lēktūvo variklis.* Tiek ChatGPT, tiek BLKŽ IL

šià lemà iliustruoja dviem klausiamaisiais sakiniais ir vienu konstatuojamuoju prijungiamuoju sakiniu. BLKŽ prie prieveiksmio *kodėl* dar prisijungia dalelytė *gi* (*Kodėl gi jums neįsigijus šuns?*).

Lemà *keisti* ChatGPT iliustravo 3 sakiniais:

- Aš nusprendžiau keisti savo darbo vietą.
- Šis kompiuteris jau senas, reikėtų jį keisti nauju.
- Ji keitė savo gyvenimo būdą, siekdama geresnės sveikatos.

BLKŽ skiriamos 6 jos reikšmės, lema iliustruojama 28 pavyzdžiais: **1.** *Keičiù bûtą. Keičia senàs pãdangas naujomìs [ĩ naujàs]. Jìs nuòlat keìsdavo sàvo àpdarà. Skubiaĩ keičiame alývą. Tėvas keitė darbą [profèsiją]. Nekeìšciau ĩ niėką šito rašiklio.* **2.** *Pìnigus [valiùtà] keìsti. Eurùs keìsti dóleriais (ĩ dólerius). Stambiùs pìnigus keičia smulkiaìs.* **3.** *Viětomis keìstis. Keìstis níuomonėmis. Keìskimės laikrodžiais [knýgomis]. Sù draiégemis keìsdavomės suknėlėmis. Jiė keičiasi informàcija sù kitomis žinýbomis.* **4.** *Žmónės keičia kraštovaizdį. Žmogùs keičia sàvo pàžiūras. Keičiame įpročius, gyvenimo būdą. Laĩvas kelìs kartùs keitė krýptį.* **5.** *Plaukai keičia spalvą. Vėjas keičia krýptį. Keičiasi óras. Apliñkui viskas keičiasi.* **6.** *Trėneris dažnai keitė žaidėjus. Dieną keičia naktìs. Liėtų keičia giedrà. Vienà valdžià keìsdavo kitą. Miñtys nuòlat keičiasi. Ekranė greitai keitėsi kãdrai.* Pirmieji du ChatGPT sugeneruoti sakiniai iliustruoja pirmàjà lemos reikšmę (*vietoj vieno imti kitą tokį patį arba kitonišką*) pagal BLKŽ, o trečiasis sakinytis – ketvirtàjà reikšmę (*daryti kitokį*). BLKŽ IL randama dviejų sugeneruotų sakinių tematika (ChatGPT – *aš nusprendžiau keisti savo darbo vietą*, BLKŽ – *tėvas keitė darbą*; ChatGPT – *ji keitė savo gyvenimo būdą*, BLKŽ – *keičiame gyvenimo būdą*). Dviejuose ChatGPT sakiniuose lema *keisti* pasitaikė bendratyje.

Lemà *kreiptis* ChatGPT iliustravo 3 pavyzdžiais:

- Jis kreipėsi į gydytoją dėl sveikatos problemų.
- Ar galiu kreiptis į tave pagalbos?
- Jie ketina kreiptis į teisininką dėl teisinio patarimo.

Lema *kreiptis* randama BLKŽ ŽS *kreipti*, trečiojoje lemos sangražinėje reikšmėje⁷¹. Ji iliustruojama keturiais sakiniais: *Įmonė dėl netesybų raštù kreipėsi ĩ teismą. Žmónės kreipėsi sù pràšymu pakeìsti autobùso grãfiką. Jėi netùrite talòno [bìlieto], kreĩpkitės ĩ autobùso vairuotojà. Dėl stógo remònto mės kreĩpėmės ĩ miėsto savivaldybė.* Tiek ChatGPT, tiek BLKŽ randama po vienà su

⁷¹ Į BLKŽ įvedus lemà *kreiptis*, gaunamas daiktavardinio homonimo ŽS (*ryšys su kompiuterio įrenginiu ar programa reikalingiems veiksams atlikti*). Darbo autorė iš anksto žinojo, jog veiksmazodžio *kreiptis* reikšmės reikia ieškoti lemos *kreipti* ŽS, tačiau šioje vietoje su potencialia problema gali susidurti nepatyręs ar negimtakalbis žodyno vartotojas.

teisiniu kontekstu susijusį sakinį (*Jie ketina kreiptis į teisininką dėl teisinio patarimo; Įmonė dėl netesybų raštu kreipėsi į teismą*).

Lemai kadangi ChatGPT sugeneravo 5 sakinius:

- Ji buvo pavargusi, kadangi dirbo visą naktį.
- Aš atidėjau susitikimą, kadangi buvau užimtas.
- Jis nebuvo susirūpinęs dėl ateities, kadangi turėjo stabilų darbą.
- Mes praleidome vakarą kartu, kadangi buvome geri draugai.
- Jie nusprendė perkelti vestuves, kadangi oras buvo blogas.

Ši lema BLKŽ taip pat iliustruojama 5 sakiniiais: *Pósėdyje nedalyvavaũ, kadangi susirgaũ. Kadangi jũs neįvykdėte sutartiės sąlygų, mės jums nemokėsime. Jis turėjo pailsėti, kadangi buvo pavařges. Gyvėtojai nepageidauja atlikti vandeňs tyrimũ, kadangi ši paslauga yra mókama. Pastaruoju metu pasaulyje labai susidomėta atsinaũjinančiais energijos šaltiniais, kadangi jiė dažniausiai yra ekologiškaĩ švārũs ir garantuoja šaliės energetinę nepriklausomybę. Visuose ChatGPT pavyzdžiuose ir daugumoje BLKŽ sakinių jungtukas kadangi atsiduria po kablelio, tačiau viename BLKŽ pavyzdyje lema atsiduria sakinio pradžioje (*Kadangi jũs neįvykdėte sutarties sąlygų, mes jums nemokėsime*). Viename ChatGPT ir viename BLKŽ sakinyje kalbama apie subjekto nuovargį (*Ji buvo pavargusi, kadangi dirbo visą naktį; Jis turėjo pailsėti, kadangi buvo pavařges*).*

Šiame skyriuje palyginus ChatGPT generuojamus ir BLKŽ IL, iliustruojančius tas pačias lemas, pirmiausia paaiškėjo, jog ChatGPT generuojamų ir BLKŽ pateikiamų iliustracinių pavyzdžių skaičius ženkliai skiriasi. Šioje vietoje svarbu atkreipti dėmesį į bendrą BLKŽ esančių iliustracijų ir ChatGPT sugeneruotų iliustracijų toms pačioms lemomis kiekio skirtumą, kalbant apie visus tyrimo duomenis: BLKŽ suskaičiuota **590** iliustracijų, o ChatGPT toms pačioms lemomis sugeneravo tik **147** sakinius, taigi, maždaug keturis kartus mažiau iliustracijų. Akivaizdu, kad BLKŽ pateikiamos iliustracijos (tiek apžvelgtos šiame skyriuje, tiek 2.2.3. skyriuje) dėl savo kiekio yra kalbiškai įvairesnės, vaizdingesnės, atspindi daugiau teminių sričių, šnekamosios kalbos, nei ChatGPT generuojami sakiniai. Tai lemia ne tik faktas, jog sakiniai yra sugeneruoti DI įrankio, tačiau ir minėtas iliustracijų kiekio skirtumas: natūralu, jog įvairesni bus 28⁷² pagal lemos reikšmes išskirti iliustraciniai sakiniai nei 3 vieną lemos reikšmę iliustruojantys sakiniai⁷³.

⁷² Tiek pavyzdžių BLKŽ pateikiama lemai *keisti*.

⁷³ Taip pat svarbu turėti omenyje, jog BLKŽ iliustracijas surinko (kai kuriais atvejais sukūrė) profesionalių leksikografų ir patyrusių kalbos vartotojų komanda (taip pat visi BLKŽ IL yra sukirčiuoti, o to nerandama ChatGPT sugeneruotuose IL, tačiau nebuvo prašoma ir užklausoje).

Analizuojant autentiškai skambančius ChatGPT sugeneruotus sakinius akivaizdu, jog įrankis beveik visada sugeneruoja po 3 sakinius vienai lemai iliustruoti (išskyrus lemos *kadangi* atvejį, kuomet sugeneruoti 5 sakiniai). Nors pasitaikė pora atvejų, kuomet BLKŽ pateikiama nedaug iliustracijų, pvz., 3 ar 4 (lemos *kodėl, kreiptis*⁷⁴), BLKŽ vienai lemai pateikiamų iliustracijų vidurkis yra 12 sakinių. Nors sunku paaiškinti, kodėl ChatGPT apskritai generuoja skirtingą kiekį pavyzdžių (žr. p. 33) lemomis iliustruoti, yra aišku, jog kuo daugiau lemos reikšmių eksplikuojama BLKŽ, tuo daugiau jos iliustracinių pavyzdžių pateikiama ir atvirkščiai. Analizuojant devynias polisemiškas lemas (*klausimas, komitetas, kalbėti, kaina, kiekvienas, keli, komanda, kartu, keisti*) pastebėta, jog vidutiniškai BLKŽ eksplikuojamos keturios reikšmės, tuo tarpu pats ChatGPT visuomet nurodo tik vieną žodžio reikšmę (taip pat ir polisemiškoms lemomis). ChatGPT iliustruojami reikšminiai skirtumai pasidaro aiškūs sugeneruotus sakinius palyginus su kitu leksikografiniu veikalu kaip BLKŽ ir kaip jame eksplikuojamos reikšmės. Paaiškėjo, jog dažniausiai ChatGPT iliustruoja pirmąją (neretai ir kitas gretutines) lemos reikšmę pagal BLKŽ. Verta atkreipti dėmesį, jog tai yra atsitiktinumas ir vargu, jog ChatGPT pagal užklausą generuodamas IL remiasi BLKŽ ar koku nors kitu lietuvių kalbos žodynu.

BLKŽ pasitaiko įvairių realijų (pvz., *Lietuva, Lietuvos kariuomenė, Socialinių reikalų ir darbo komitetas, Valstybinė kainų ir energetikos kontrolės komisija*), o ChatGPT nei karto nesugeneravo Lietuvos visuomenės realijų. ChatGPT be konstatuojamųjų pasitaikė keli klausiamieji sakiniai, tuo tarpu BLKŽ be konstatuojamųjų randami klausiamieji, šaukiamieji sakiniai, kaip atskiros iliustracijos pateikiami įvairūs žodžių junginiai (pvz., *Egzaminų komisija, Valstybinė kainų ir energetikos kontrolės komisija, Įmonės likvidavimo komisija*), iliustracijos yra detalesnės, išsamesnės (tai neretai pasireiškia per detalius pažyminius). Natūralu, jog ChatGPT generuojamuose ir BLKŽ esančiuose iliustraciniuose sakiniuose neretai pasitaiko turinio panašumų (tai lemia ta pati iliustruojama lema).

Taigi, ši analizė leidžia daryti išvadą, jog BLKŽ lemas iliustruoja išsamiau, detaliau, kalbiškai įvairiau, juose pasitaiko realijų, o polisemiškas lemas ŽS paaiškina itin detaliai, lyginant su ChatGPT generuojamomis iliustracijomis, kurių kiekis dažniausiai būna nedidelis. ChatGPT generuojami pavyzdžiai yra tarpusavyje sintaksiškai ir morfologiškai panašūs, juose nėra užsimenama apie Lietuvos kontekstą, jie dažnai iliustruoja tik vieną lemos reikšmę.

⁷⁴ Tai natūralu, nes kalbama apie lemas, kurioms BLKŽ skiria tik vieną pagrindinę reikšmę.

2.2.5. ChatGPT sugeneruotų IL kokybės įvertinimas

Šiame skyriuje siekiama sugrįžti prie darbo įvade iškelto klausimo: ar ChatGPT lietuviškai sugeneruotus iliustracinius sakinius būtų galima perkelti į hipotetinį žodyną? Jei ne, kokie trūkumai užkerta kelią tokiam perkėlimui? Norint atsakyti į šį klausimą, toliau siekiama įvertinti ChatGPT sugeneruotų autentiškai skambančių sakinių kokybę, remiantis aukštos kokybės BLKŽ iliustracijomis.

Kaip jau minėta, ChatGPT iš viso sugeneravo 147 sakinius, tačiau darbe analizuotas 81 autentiškai skambantis sakiny (~55 % visų sugeneruotų iliustracinių sakinių), o siekiant įvertinti ChatGPT sugeneruotų sakinių kokybę, galima kalbėti tik tolimesnius kriterijus atitinkančius sugeneruotus sakinius. Kiekvienas jų atitiko šiuos penkis kriterijus:

1. Sakinys yra logiškas ir rišlus;
2. Sakinyje netrūksta esminės informacijos, leidžiančios lengvai suprasti sakinio prasmę;
3. Sakinys yra morfologiškai, sintaksiškai, leksiškai ir ortografiškai taisyklingas;
4. Sakinį kaip autentiškai skambantį įvertino dauguma (ar visi) gimtakalbių;
5. Sakinys tinkamai iliustruoja konkrečią lemą.

Lemoms iliustruoti generuotas labai skirtingas kiekis pavyzdžių (žr. 6 paveikslą 2.2.1. skyriuje): neretai lemai buvo sugeneruojamas vienas ar pora autentiškai skambančių sakinių, tačiau likę buvo haliucinacijos ar dirbtinai skambantys (pvz., lemų *kad*, *kuris*, *kitas* atvejai). Kaip jau rašyta, pasitaikė 13 lemų, kurioms visi sugeneruoti pavyzdžiai laikytini tinkamais ir autentiškai skambančiais: dvylikai jų sugeneruotos 3 iliustracijos, o vienai sugeneruotos 5. Šioje vietoje galima išvelgti generuojamų sakinių kokybės trūkumą: net ir polisemiškiems žodžiams, kurie BLKŽ iliustruojami keliomis dešimtimis pavyzdžių (pvz., *klausimas*, *kalbėti*, *komanda*), ChatGPT vistiek generuoja po nedidelį kiekį sakinių, nors užklausoje buvo prašoma generuoti kuo daugiau ir kuo įvairesnių vartojimo pavyzdžių. Suprantama, jog toks ribotas kiekis sakinių lemom iliustruoti nėra pakankamas ir tinkamas. Svarbu atsižvelgti ir į tai, jog ChatGPT generuojamuose sakiniuose iliustruojamos ne visos (pagal BLKŽ), o dažnai tik viena iš keleto polisemiškų lemų reikšmių (pvz., lemų *kaina*, *komanda*, *kartu* atvejai). Tokiais atvejais galima sakyti, jog nėra apimamas visas lemų vartojimo kontekstas, sakiniuose neatskleidžiami visi polisemiškos lemos reikšminiai atspalviai.

Taip pat pastebima, jog autentiškai skambančiuose sugeneruotuose sakiniuose nepasitaiko nė viena su Lietuva susijusi realija⁷⁵, nors realaus pasaulio konteksto, realijų atosklaida yra būdinga žodyniniams tekstams (apie šį aspektą žr. skyriuje 2.2.3.). Neretai ChatGPT sugeneruoti sakiniai yra vienodos sintaksinės struktūros: pavyzdžiui, jie dažnai prasideda veiksniais, įvardžiais vardininko linksnyje, po to seka tariniai, sakiniuose dominuoja dvi laiko formos, sakiniai daugiausiai vientisiniai,

⁷⁵ Vienintelė pasitaikiusi **pasaulio** realija buvo *doleris*.

konstatuojamieji ir pan. Sugeneruoti sakiniai nėra vaizdingi, ir tuo labai skiriasi nuo BLKŽ randamų iliustracijų. Svarbu paminėti įdomų paaiškėjusį faktą, jog kai kurie ChatGPT sugeneruoti sakiniai gimtakalbių kalbos vartotojų įvertinti natūraliau skambančiais (pvz., *Po kelių valandų jie pasiekė tikslą; Jis kreipėsi į gydytoją dėl sveikatos problemų; Aš nusprendžiau keisti savo darbo vietą*), o kai kurie sakiniai, remiantis apklausos rezultatais, laikytini skambančiais ne taip natūraliai (pvz., *Klausimas, kurį tu kėlei, yra labai svarbus; Po valandos jis pradės kalbėti apie savo keliones; Prezidento kanceliarija yra atsakinga už įvairius valstybės administracinius reikalus*). Suprantama, jog žodyne esantys iliustraciniai tekstai dėl savo autentiško skambėjimo neturėtų kelti abejonių gimtakalbiui žodyno vartotojui.

Atsižvelgiant į aptartus su sugeneruotų autentiškai skambančių sakinių kokybe susijusius aspektus ir grįžtant prie darbe keliamo klausimo, darytina išvada, jog autentiškai skambančiais sakiniiais (tiek aptartais 2.2.1.1., tiek išsamiau aptartais 2.2.4. skyriuose) galima papildyti hipotetinio žodyno IL, jei juos kiek redaguotų leksikografai, tačiau vien ChatGPT sugeneruotų sakinių lemu iliustravimui žodyne daugumoje atvejų nepakaktų. Tai lemia tokie ChatGPT sugeneruotų sakinių kokybės trūkumai kaip dažniausiai nedidelis ir sunkiai paaiškinamas, neprognozuojamas iliustracinių sakinių kiekis, menka polisemijos atoslaida iliustraciniuose pavyzdžiuose. Tai nulemia ne visuomet pakankamą lemos vartojimo konteksto iliustravimą pavyzdžiais. Sakiniuose nepasitaiko Lietuvos visuomenės realijos, sakiniai dažnai yra vienodos sintaksinės struktūros, nėra labai vaizdingi, kartais jiems trūksta konkretumo. Kai kurie sakiniai dėl savo autentiško skambėjimo kėlė abejonių tam tikrai daliai gimtakalbių (t. y., kai kurie skambėjo natūraliau, kai kurie mažiau natūraliai).

Išvados

Šiuo MA darbu buvo siekiama atsakyti į klausimą, ar įrankio ChatGPT lietuvių kalba sugeneruotus iliustracinius sakinius būtų galima perkelti į hipotetinį žodyną, o jei nebūtų galima, kokie generuojamų iliustracijų trūkumai trukdo tą padaryti.

Siekdama atsakyti į šį klausimą, autorė atliko išsamią lingvistinę sakinių, kuriuos pagal nagrinėtą lemų sąrašą sugeneravo įrankis, analizę. Taip pat buvo sudaryta apklausa, kurioje respondentai, kuriems lietuvių kalba yra gimtoji, įvertino ChatGPT sugeneruotų sakinių skambėjimą. Remdamasi apklausos rezultatais, darbo autorė išskyrė tris ChatGPT sugeneruotų sakinių kategorijas: autentiškai skambantys sakiniai, gimtakalbių vertinimu dirbtinai skambantys sakiniai ir haliucinacijos.

Atsižvelgiant į atlikto tyrimo rezultatus darytina išvada, jog ChatGPT sugeneruotais autentiškais sakiniiais būtų galima papildyti hipotetinio žodyno IL, tačiau vien sugeneruotų sakinių lemons žodyne iliustruoti nepakaktų. Tai lemia per mažas ChatGPT sugeneruotų iliustracijų kiekis (ypač kalbant apie polisemiškas lemas), sugeneruotų sakinių sintaksinis panašumas, nelabai platus aplinkybių, funkcinių stilių pasitaikymas, semantinės, leksinės, morfologinės sugeneruotų sakinių įvairovės stygius. Sugeneruotuose sakiniuose trūksta konkretumo, detalumo: juose visai nėra tikrinių žodžių, Lietuvos visuomenės realijų.

Apklausos rezultatai taip pat atskleidė, jog vieni ChatGPT sugeneruoti sakiniai skamba natūraliau, o kiti skamba mažiau natūraliai. Tad norint hipotetinį žodyną papildyti ChatGPT sugeneruotais sakiniiais reikėtų, kad visi konkrečiai lemai sugeneruoti sakiniai gimtakalbių kalbos vartotojų įvertinti kaip skambantys natūraliai.

Norėdama ChatGPT generuojamas iliustracijas palyginti su aukštos kokybės elektroniniu lietuvių kalbos žodynu, autorė atspirties tašku darbe pasirinko BLKŽ, o dalį autentiškai skambančių ChatGPT sugeneruotų sakinių palygino su BLKŽ esančiomis iliustracijomis. Remdamasi analize autorė BLKŽ iliustracijas įvertino kaip kokybiškesnes ir pranašesnes už ChatGPT sugeneruotas iliustracijas. Taigi, remiantis tyrimo rezultatais, darbe kelta hipotezė pasitvirtino. Nors dalį ChatGPT sugeneruotų iliustracijų ir būtų galima naudoti hipotetinio žodyno IL pildymui, tačiau leksikografams reikėtų sugeneruotus sakinius tinkamai apdoroti ir greta ChatGPT sugeneruotų sakinių į žodyno IL įdėti ir kitų, ne ChatGPT sugeneruotų iliustracijų.

Ateityje neabejotinai vertėtų kartoti šį ar panašų tyrimą ir toliau gilintis į ChatGPT lietuviškai generuojamų leksikografinių tekstų galimybes. Tam būtų galima išsamiu ar kitaip suformuluoti užklausą, įrankio paprašyti sugeneruoti iliustracinius sakinius platesniam lemų sąrašui, nurodyti daugiau iliustracijų kokybės kriterijų, dėl sugeneruotų sakinių autentiško skambėjimo apklausti daugiau respondentų. Todėl tyrimą leksikografijos tyrėjams rekomenduojama kartoti ateityje.

Šis MA darbas yra svarbus leksikografijos tyrėjams, leksikografams ir plačiajai visuomenei, naudojančiai elektroninius žodynus ir įrankį ChatGPT: šiuo darbu atskleidžiama, jog minėtą įrankį galima naudoti ir leksikografiniais tikslais.

Literatūros sąrašas

1. Ahmad, K., Fulford, H., Rogers, M. (1992). The elaboration of special language terms: The role of contextual examples, representative samples and normative requirements. *Proceedings of the 5th EURALEX International Congress*, p. 139–149. ISBN: 951-44-3111-1
2. Alm, A., Watanabe, Y. (2023). Integrating ChatGPT in Language Education: A Freirean Perspective. *Iranian Journal of Language Teaching Research*, 11(3), p. 19-30. DOI: 10.30466/ijltr.2023.121404
3. Andriansyah, Y. (2023). The Current Rise of Artificial Intelligence and Religious Studies: Some Reflections Based on ChatGPT. Editorial. *Millah: Journal of Religious Studies*, 22(1), p. ix-xviii. Prieiga internetu: <https://journal.uii.ac.id/Millah/article/download/29004/15205> [žiūrėta 2024 04 10].
4. Atkins, B. T. S., Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford University Press. ISBN: 978-0-19-927770-4
5. Bang Y., Cahyawijaya S., Lee N., Dai W., Su D., Wilie B., Lovenia H., Ji Z., Yu T., Chung W., Do Q. V., Xu Y., Fung P. (2023). *A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity*. Preprint. <https://doi.org/10.48550/ARXIV.2302.04023>
6. Bergenholtz, H. (2012). What is a Dictionary?. *Lexikos*, 22, p. 20-30. Prieiga internetu: <https://www.proquest.com/docview/1288088364?pq-origsite=gscholar&fromopenview=true&sourcetype=Scholarly%20Journals> [žiūrėta 2024 04 10].
7. Bernal, E. (2020). Normative dictionaries. Lebsanft F., Tacke F. (eds.). *Manual of Standardization in the Romance Languages*, p. 519-534. Berlin: De Gruyter. Prieiga internetu: https://repositori.upf.edu/bitstream/handle/10230/58858/Bernal_man_norm.pdf?sequence=1&isAllowed=y [žiūrėta 2024 04 10].
8. Beuge, P. (2016). Präskription oder Deskription? Zur normativen Kraft des Kodizes. Klein, W. P., Schulz, M., Staffeldt, S., Stahl, P. (Hrsg.) *Die Kodifizierung der Sprache. Strukturen, Funktionen, Konsequenzen* (17), p. 138-148. Prieiga internetu: https://opus.bibliothek.uni-wuerzburg.de/files/13808/WespA17_Kodex_Klein_Staffeldt.pdf [žiūrėta 2024 04 10].
9. Bindoriūtė, I. (2008). *Kalbos analizės ir sintezės įrankiai*. Magistro darbas. Vilnius: Vilniaus pedagoginis universitetas.
10. Biswas, S. S. (2023). Role of Chat GPT in Public Health. *Annals of Biomedical Engineering*, 51(5), p. 868-869. <https://doi.org/10.1007/s10439-023-03172-7>

11. Booij, G. (2003). The codification of phonological, morphological, and syntactic information. Sterkenburg, P. van (ed.). *A Practical Guide to Lexicography*, p. 251-259. Volume 6. John Benjamins Publishing Company.
12. Buzrllová, L. (2013). *Wörterbücher als Informationsträger und Hilfsmittel der Theorie und Praxis*. Magisterská diplomová práce. Masarykova univerzita. Prieiga internetu: https://is.muni.cz/th/b184m/Masarykova_univerzita_-_konecna_verze_1.pdf [žiūrėta 2024 04 10].
13. Coffey, S. (2010). 'Offensive' items, and less offensive alternatives, in English monolingual learners' dictionaries. *Proceedings of the 14th EURALEX International Congress*, p. 1270–1281. ISBN: 978-90-6273-850-3
14. Collins, J. B. (2016). Changes in electronic dictionary usage patterns in the age of free online dictionaries: Implications for vocabulary acquisition. *APU Journal of Language Research*, 1, p. 36-49. Prieiga internetu: https://www.researchgate.net/publication/316479856_Electronic_dictionary_usage_in_the_age_of_free_online_applications_Implications_for_Vocabulary_Acquisition [žiūrėta 2024 04 10].
15. Dagienė, V., Grigas, G., Jevsikova, T. (2014). *Enciklopedinis kompiuterijos žodynas*. IV leidimas. Vilniaus universiteto Matematikos ir informatikos fakultetas. Prieiga internetu: <http://www.ims.mii.lt/EK%C5%BD/> [žiūrėta 2024 04 10].
16. Djefal, C. (2020). Künstliche Intelligenz. Klenk, T., Nullmeier, F., Wewer, G. (eds.). *Handbuch Digitalisierung in Staat und Verwaltung*, p. 1-12. Springer Fachmedien. Prieiga internetu: <https://mediatum.ub.tum.de/doc/1584476/document.pdf> [žiūrėta 2024 04 10].
17. Enčeva, M., Jazbec, S. (2014). Wie multifunktional und wie individualisiert können und sollten Online-Lernwörterbücher sein? *Jezikoslovlje* 15.1, p. 45-66. Prieiga internetu: <https://hrcaak.srce.hr/file/175083> [žiūrėta 2024 04 01].
18. Engh, J. (1993). Linguistic Normalisation in Language Industry. Some Normative and Descriptive Aspects of Dictionary Development. *HERMES - Journal of Language and Communication in Business* (10), p. 53-64. <https://doi.org/10.7146/hjlc.v6i10.21520>
19. Ertel, W. (2017). *Introduction to Artificial Intelligence*. Second edition. Springer International Publishing. DOI: 10.1007/978-3-319-58487-4
20. Frieder, S., Pinchetti, L., Chevalier, A., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Berner, J. (2023). Mathematical Capabilities of ChatGPT. *NeurIPS 2023 Datasets and Benchmarks*, p. 1-37. <https://doi.org/10.48550/arXiv.2301.13867>
21. Fuertes-Olivera, P. A. (2009). The Function theory of lexicography and electronic dictionaries: WIKTIONARY as a Prototype of Collective Multiple-Language Internet

- Dictionary. Bergenholtz, H., Nielsen, S., Tarp, S. (eds.). *Lexicography at a Crossroads: Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*, p. 99-134. Bern: Peter Lang. Prieiga internetu https://www.researchgate.net/profile/Pedro-Fuertes-Olivera/publication/235222544_The_Function_theory_of_lexicography_and_electronic_dictionaries_WIKTIONARY_as_a_Prototype_of_Collective_Multiple-Language_Internet_Dictionary/links/57ec2cff08ae93b7fa95864f/The-Function-theory-of-lexicography-and-electronic-dictionaries-WIKTIONARY-as-a-Prototype-of-Collective-Multiple-Language-Internet-Dictionary.pdf [žiūrėta 2024 04 01].
22. Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering* 51, p. 2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>
23. Hartmann, R. R. K., James, G. (2002). *Dictionary of Lexicography*. Taylor & Francis e-Library. Prieiga internetu: https://nubip.edu.ua/sites/default/files/u5/r_r_k_hartmann_gregory_james-dictionary_of_lexicography-routledge_1998.pdf [žiūrėta 2024 04 01].
24. Hidalgo Tenorio, E. (2000): Gender, sex and stereotyping in the Collins COBUILD English language dictionary. *Australian Journal of Linguistics*, 20 (2), p. 211–230. Prieiga internetu: https://www.ugr.es/~ehidalgo/_DOCUMENTOS_PDF/Gender.pdf [žiūrėta 2024 04 01].
25. Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, p. 37-45. Prieiga internetu: <https://jeti.thewsu.org/index.php/cieti/article/download/103/64> [žiūrėta 2024 04 01].
26. Humblé, P. (2001). *Dictionaries and language learners*. Frankfurt am Main: Haag und Herchen Verlag. Prieiga internetu: https://www.researchgate.net/publication/343987260_DICTIONARIES_AND_LANGUAGE_LEARNERS [žiūrėta 2024 04 01].
27. Iversen, S. H. (2012). ‘To Teach Little Boys And Girls What It Is Proper For Them To Know’: Gendered Education and the Nineteenth-Century Children’s Dictionary. *Proceedings of the 15th EURALEX International Congress*, p. 613-618. ISBN: 978-82-303-2228-4
28. Jakaitienė, E. (2005). *Leksikografija*. Vilnius: Mokslo ir enciklopedijų leidybos institutas. ISBN: 5-420-01580-3
29. Jakaitienė, E. (2009). *Leksikologija: studijų knyga*. Vilnius: Vilniaus universiteto leidykla. ISBN: 9789955334453
30. Janiesch, C., Zschech, P., Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets* 31, p. 685-695. <https://doi.org/10.1007/s12525-021-00475-2>

31. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), p. 1-38. <https://doi.org/10.1145/3571730>
32. Karpinska, L. (2012). *Critical Analysis of English-Latvian Lexicographic Tradition*. Doctoral Thesis. University of Latvia.
33. Kataros, V. (2004). The different functions of illustrative examples in Learner's Bilingual Dictionaries. Compiling a dictionary around illustrative examples: The Experience of the Larousse School French-English, English-French Dictionary. *Proceedings of the 11th EURALEX International Congress*, p. 487-494. ISBN: 29-52245-70-3
34. Kirkness, A. (2004). Lexicography. Davies, A., Elder, C. (eds.). *The Handbook of Applied Linguistics*, p. 54-81. Prieiga internetu: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b783d2c2a5297189d62c861cfd0ec13766f3b58f> [žiūrėta 2024 04 01].
35. Klosa, A. (2009). *The lexicographical process: online dictionaries*. Preprint. Institut für deutsche Sprache Mannheim. Prieiga internetu: http://multimedia.ids-mannheim.de/mediawiki/web/images/6/6b/Klosa_LexicographicalProcess_Preprint.pdf [žiūrėta 2024 04 01].
36. Kohnke, L., Moorhouse, B. L., Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), p. 537-550. <https://doi.org/10.1177/00336882231162868>
37. Korteling, J. E. H., van de Boer-Visschedijk, G.C., Blankendaal, R. A., Boonekamp R. C., Eikelboom, A. R. (2021). Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, 4, 622364, p. 1-13. DOI: 10.3389/frai.2021.622364
38. Labutis, V. (1998). *Lietuvių kalbos sintaksė*. Antrasis pataisytas leidimas. Vilnius: Vilniaus universiteto leidykla. ISBN: 9986-19-327-3
39. Lai, V. D., Ngo, N. T., Veyseh, A. P. B., Man, H., Derroncourt, F., Bui, T., Nguyen, T. H. (2023). *ChatGPT Beyond English: Towards a Comprehensive Evaluation of Large Language Models in Multilingual Learning*. Preprint. <https://doi.org/10.48550/ARXIV.2304.05613>
40. Leme Lopes, A. P. (2023). Artificial history? Inquiring ChatGPT on historiography. *Rethinking History*, 27:4, p. 709-749. DOI: <https://doi.org/10.1080/13642529.2023.2234227>
41. Lew, R. (2010). *New ways of indicating meaning in electronic dictionaries: hope or hype?* P. 1-12. Prieiga internetu: http://wa.amu.edu.pl/~rlew/pub/Lew_New_ways_of_indicating_meaning.pdf [žiūrėta 2024 04 01].
42. Lew, R. (2023). *ChatGPT as a COBUILD lexicographer*. Preprint. Prieiga internetu: <https://osf.io/t9mbu/download> [žiūrėta 2024 04 01].

43. Lew, R. (2024). Dictionaries and lexicography in the AI era. *Humanities and Social Sciences Communications* 11, 426, p. 1-8. <https://doi.org/10.1057/s41599-024-02889-7>
44. Lin, P. (2023). ChatGPT: Friend or foe (to corpus linguists)? *Applied Corpus Linguistics*, 3 (3), p. 1-5. <https://doi.org/10.1016/j.acorp.2023.100065>
45. Liu, J., Wang, C., Liu, S. (2023). Utility of ChatGPT in clinical practice. *Journal of Medical Internet Research*. <https://doi.org/10.2196/48568>
46. Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., Ge, B. (2023). Summary of ChatGPT-Related Research and Perspective Towards the Future of Large Language Models. *Meta-Radiology* 1(2), p. 1-14. <https://doi.org/10.1016/j.metrad.2023.100017>
47. Liutkevičienė, D. (2006). "Bendrinės lietuvių kalbos žodynas". Siekiamybė ir realybė. Ritutė Petrokienė (sud.). *Leksikografija ir leksikologija. Aiškinamųjų bendrinės kalbos žodynų aktualijos*, p. 9-18. Vilnius: Lietuvių kalbos institutas. ISBN: 9789955704416
48. Liutkevičienė, D. (2015). Bendrinės kalbos normų pateikimas "Bendrinės lietuvių kalbos žodyne." *Bendrinė kalba* (88), p. 1–9. ISSN: 2351-7204
49. Maleki, N., Padmanabhan, B., Dutta, K. (2024). *AI Hallucinations: A Misnomer Worth Clarifying*. Preprint. <https://doi.org/10.48550/arXiv.2401.06796>
50. McCarthy, J. (2007). *What is Artificial Intelligence?* Stanford University. Prieiga internetu: <http://jmc.stanford.edu/articles/whatisai.html> [žiūrėta 2024 03 09].
51. McGee, R. W. (2023). *Political Philosophy and ChatGPT*. Preprint. Prieiga internetu: https://www.researchgate.net/profile/Robert-Mcgee-5/publication/369526036_Political_Philosophy_and_ChatGPT/links/641efdaa66f8522c38d2a2d2/Political-Philosophy-and-ChatGPT.pdf [žiūrėta 2024 03 09].
52. McKean, E., Fitzgerald, W. (2023). The ROI of AI in Lexicography. *Asialex 2023 Proceedings: Lexicography, Artificial Intelligence and Dictionary Users*, p. 1-10. Prieiga internetu: <https://waf-e.dubuplus.com/koreanlex.dubuplus.com/anonymous/O18CTZ8/DubuDis> [žiūrėta 2024 04 23].
53. Melnikienė, D. (2006). Žodyno antraštynas ir nenorminė leksika. *Kalbotyra* 56(3), p. 108-113. ISSN: 1392-1517
54. Melnikienė, D. Jankauskaitė, M. (2012). Iliustraciniai pavyzdžiai dvikalbiuose žodynuose: kolokacijos. *Verbum* (3), p. 79-90. ISSN: 2538-8746
55. Moerdijk, F. (2003). The codification of semantic information. Sterkenburg, P. van (red.). *A Practical Guide to Lexicography*, p. 273-293. Volume 6. John Benjamins Publishing Company.

56. Murmulaitytė, D. (2011). Iliustracinių vartosenos pavyzdžių atranka ir pateikimas “Bendrinės lietuvių kalbos žodyne.” *Kalbos kultūra* (84), p. 129–149. ISSN: 0130-2795
57. Müller-Spitzer, C. (2009). *Textual structures in electronic dictionaries compared with printed dictionaries: A short general survey*. Preprint. Institut für deutsche Sprache Mannheim. Prieiga internetu: https://www.ids-mannheim.de/fileadmin/lexik/owid/Preprint_Mueller-Spitzer_HSK5.4-Article11_01.pdf [žiūrėta 2024 03 09].
58. Müller-Spitzer, C., Rüdiger, J. O. (2022). The influence of the corpus on the representation of gender stereotypes in the dictionary. A case study of corpus-based dictionaries of German. *Dictionaries and Society. Proceedings of the XX EURALEX International Congress*, p. 129-141. Prieiga internetu: https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202022/EURALEX2022_Pr_p129-141_Mueller-Spitzer-Ruediger.pdf [žiūrėta 2024 03 09].
59. Naktinienė, G. (2000). Bendrinės kalbos žodynas ir leksikos norminimo problemos. *Kalbos kultūra* (73), p. 12–24. ISSN: 0130-2795
60. Nevinskaitė, L., Tamaševičius, G. (2019). Ar preskriptyvizmas veikia? Taisomi žodžiai Lietuvos radijo ir TV eteryje 1960–2010. *Taikomoji kalbotyra*, 13, p. 1-118. <https://doi.org/10.15388/TK.2019.19172>
61. Nomicienė, D. E., Stunžaitė, A. (2016). Skaitmeninė leksikografija: Įvairių žodynų tipų naudojimo tyrimas Vilniaus Gedimino technikos universitete. *Santalka: Filologija, edukologija* (24)2, p. 90–100. DOI: 10.3846/cpe.2016.293
62. Petkevičius, L. (2021). *Natūralios kalbos analizė: nuo tekstinių duomenų rūšiavimo iki netikrų tekstų generavimo*. Vilniaus universiteto Matematikos ir informatikos fakultetas. Prieiga internetu: <https://mif.vu.lt/lt3/kas-vyksta-fakultete/naujienos/fakulteto-naujienos/3350-nat%C5%ABralios-kalbos-analiz%C4%97s-modeliai> [žiūrėta 2024 03 09].
63. Rundell, M. (2023.) Automating the creation of dictionaries: are we nearly there? *Asialex 2023 Proceedings: Lexicography, Artificial Intelligence and Dictionary Users*, p. 1-9. Prieiga internetu: <https://waf-e.dubuplus.com/koreanlex.dubuplus.com/anonymous/O18CTZ8/DubuDisk/public/ASIALEX2023-Proceedings.pdf> [žiūrėta 2024 04 23].
64. Russell, S. J., Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Third edition. London: Pearson Education.
65. Schryver, G. M. de (2003). Lexicographers' dreams in the electronic-dictionary age. *International Journal of Lexicography*, 16(2), p. 143-199. Prieiga internetu: <https://doi.org/10.1093/ijl/16.2.143> [žiūrėta 2024 04 23].

66. Schryver, G. M. de (2023). Generative AI and Lexicography: The Current State of the Art Using ChatGPT. *International Journal of Lexicography*, 36, p. 355-387. <https://doi.org/10.1093/ijl/ecad021>
67. Schryver, G. M. de, Joffe, D. (2023a). *The end of lexicography, welcome to the machine: On how ChatGPT can already take over all of the dictionary maker's tasks*. Paskaita. Prieiga internetu: <https://biblio.ugent.be/publication/01GWSGV0M8HYZVNXV8SSZ0VXBG> [žiūrėta 2024 03 09].
68. Schuett, J. (2019). A legal definition of AI. *SSRN Electronic Journal*, p. 1-14. Prieiga internetu: https://www.researchgate.net/profile/Jonas-Schuett/publication/336198524_A_Legal_Definition_of_AI/links/5e20599a458515ba208b9e4c/A-Legal-Definition-of-AI.pdf [žiūrėta 2024 03 09].
69. Smetona, A. *Giminė*. Prieiga internetu: <https://www.smetona.lt/gimine/> [žiūrėta 2024 04 07].
70. Sijs, N. van der. (2003). The codification of etymological information. Sterkenburg, P. van (ed.). *A Practical Guide to Lexicography*, p. 312-321. Volume 6. John Benjamins Publishing Company.
71. Sinclair, J. (2004). *Developing Linguistic Corpora: a Guide to Good Practice*. Prieiga internetu: <https://users.ox.ac.uk/~martinw/dlc/chapter1.htm> [žiūrėta 2024 03 09].
72. Stojić, A. (2015). *Lexikologie und Lexikographie*. Lehr- und Übungsbuch. Sveučilište u Rijeci. Prieiga internetu: <http://izdavastvo.ffri.hr/wp-content/uploads/2016/07/Udzbenik-Leksikologija-i-leksikografija.pdf> [žiūrėta 2024 03 09].
73. Surameery, N. M. S., Shakor, M. Y. (2023). Use Chat GPT to Solve Programming Bugs. *International Journal of Information Technology & Computer Engineering*, 3(1), p. 17-22. <https://doi.org/10.55529/ijitc.31.17.22>
74. Taecharungroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, 7, 35, p. 1-10. <https://doi.org/10.3390/bdcc7010035>
75. Tlok, T. (2023). *Robuste Erkennung von KI-generierten Texten in deutscher Sprache*. Master Thesis. Fachhochschule Wedel. Prieiga internetu: https://ki.fh-wedel.de/files/TomTlok_MA.pdf [žiūrėta 2024 03 09].
76. Vaišnoras, Ž. (2020). *Mašininio mokymosi metodų vystymas įsilaužimams aptikti kompiuterių tinkluose*. Mokslinė ataskaita. Vilnius: Vilniaus universiteto Duomenų mokslo ir skaitmeninių technologijų institutas. Prieiga internetu: https://www.mii.lt/files/doc/lt/doktorantura/ataskaitine_konferencija/n009_ma_2020_vaisno_ras.pdf [žiūrėta 2024 03 09].

77. Vakaliuk, T., Chernysh, O., Figen, K., Ferrer-Conill, R. (2021). Examining the Context of E-Dictionary Use in Language Studies. *International Conference on New Trends in Languages, Literature and Social Communications (ICNTLLSC 2021)*, p. 42-52. Atlantis Press. Prieiga internetu: <https://www.atlantis-press.com/article/125957104.pdf> [žiūrėta 2024 03 09].
78. Valančiauskienė, A. 2022. *Enciklopedinės žinios dvikalbiuose lingvistiniuose žodynuose*. Humanitarinių mokslų daktaro disertacija. Vilnius: Vytauto Didžiojo universitetas, Lietuvių kalbos institutas.
79. Verikas, A. *Neuroninis tinklas*. VLE. Prieiga internetu: <https://www.vle.lt/straipsnis/neuroninis-tinklas/> [žiūrėta 2024 03 08].
80. Verkuyl, H., Janssen, M., Jansen, F. (2003). The codification of usage by labels. Sterkenburg, P. van (ed.). *A Practical Guide to Lexicography*, p. 297-311. Volume 6. John Benjamins Publishing Company.
81. Virvou, M., Tsihrintzis, G. A., Sotiropoulos, D. N., Chrysafiadi, K., Sakkopoulos, E., Tsihrintzi, E. A. (2023). ChatGPT in Artificial Intelligence-Empowered E-Learning for Cultural Heritage: The case of Lyrics and Poems. *14th International Conference on Information, Intelligence, Systems & Applications (IISA)*, p. 1-9. <https://doi.org/10.1109/IISA59645.2023.10345878>
82. Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity* 19, 26, p. 1-39. <https://doi.org/10.1007/s40979-023-00146-z>
83. Wiegand, H. E. (1997). Über die gesellschaftliche Verantwortung der wissenschaftlichen Lexikographie. *HERMES - Journal of Language and Communication in Business* (18), p. 177-202. Prieiga internetu: <https://tidsskrift.dk/her/article/download/25416/22337> [žiūrėta 2024 03 08].
84. Zhu, J. J., Jiang, J., Yang, M., Ren, Z. J. (2023). ChatGPT and environmental research. *Environmental Science & Technology* 57(46), p. 17667-17670. DOI: 10.1021/acs.est.3c01818.
85. Zubaitienė, V. (2014). *Lietuvių leksikografija: istorija ir dabartis*. Vilnius: Vilniaus universiteto leidykla.
86. Žugić, D., Vuković-Stamatović, M. (2021). Problems in Defining Ethnicity Terms in Dictionaries. *Lexikos*, 31(1), p. 177-194. <https://doi.org/10.5788/31-1-1635>

Šaltinių sąrašas

1. ChatGPT. OpenAI. Prieiga internetu: <https://chat.openai.com/> [žiūrėta 2024 05 11].
2. Dadurkevičius, V. 2020. *Wordlist of Lemmas from the Joint Corpus of Lithuanian*. CLARIN-LT digital library in the Republic of Lithuania. Prieiga internetu: https://www.lu.lv/fileadmin/user_upload/LU.LV/Apaksvietnes/Projekti/perlsi/3000_zodziu.pdf [žiūrėta 2024 05 11].
3. Dudeno žodynas. Prieiga internetu: <https://www.duden.de/> [žiūrėta 2024 05 11].
4. Keinys, S. (vyr. red.). (2021). *Dabartinės lietuvių kalbos žodynas*. 8-as patais. ir papild. leid. Vilnius: Lietuvių kalbos institutas. Prieiga internetu: <https://ekalba.lt/dabartines-lietuviu-kalbos-zodynas/> [žiūrėta 2024 05 11].
5. Kembridžo žodynas. Prieiga internetu: <https://dictionary.cambridge.org/> [žiūrėta 2024 05 11].
6. Kolinso žodynas. Prieiga internetu: <https://www.collinsdictionary.com/> [žiūrėta 2024 05 11].
7. LEO žodynas. Prieiga internetu: <https://www.leo.org/german-english/> [žiūrėta 2024 05 11].
8. Lietuvių kalbos žodynas. Prieiga internetu: <https://ekalba.lt/lietuviu-kalbos-zodynas/> [žiūrėta 2024 05 11].
9. Lietuvos dirbtinio intelekto strategija. (2019). Ekonomikos ir inovacijų ministerija. Prieiga internetu: [https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT\(1\).pdf](https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT(1).pdf) [žiūrėta 2024 05 11].
10. Lietuvių žargono bazė. (2010). Vilniaus universiteto Filologijos fakultetas. Prieiga internetu: <https://www.flf.vu.lt/mokslas/elektroniniai-istekliai#duomenu-bazes> [žiūrėta 2024 05 11].
11. Liutkevičienė, D., Naktinienė, G., Petrokienė, R., Svetikienė, D., Vosylytė, K., Zabarskaitė, J. (sud.) *Bendrinės lietuvių kalbos žodynas*. Vilnius: Lietuvių kalbos institutas. Prieiga internetu: <https://ekalba.lt/bendrines-lietuviu-kalbos-zodynas/> [žiūrėta 2024 05 11].
12. Lyberis, A. (2015). *Sinonimų žodynas*. Vilnius: Lietuvių kalbos institutas. Prieiga internetu: <https://ekalba.lt/sinonimu-zodynas/> [žiūrėta 2024 05 11].
13. *Nusikaltimų prevencijos ir kontrolės sistemos modelio Lietuvoje koncepcija*. (2020). Nusikalstamumo prevencijos Lietuvoje centras. Prieiga internetu: https://vrm.lrv.lt/uploads/vrm/documents/files/LT_versija/Viesasis_saugumas/Koncepcijos/Koncepcija%2020201219_galut.pdf [žiūrėta 2024 05 11].
14. Portalas Euronews. Prieiga internetu: <https://www.euronews.com/next/2024/02/18/sora-openais-new-text-to-video-tool-is-causing-excitement-and-fears-heres-what-we-know-abo> [žiūrėta 2024 05 11].
15. Portalas LRT. Prieiga internetu: <https://www.lrt.lt/> [žiūrėta 2024 05 11].
16. Portalas OpenAI. Prieiga internetu: <https://openai.com/blog/chatgpt-can-now-see-hear-and-speak> [žiūrėta 2024 05 11].

17. Portalas The New York Times. Prieiga internetu: <https://www.nytimes.com/> [žiūrėta 2024 05 11].
18. Skiltis *Vadovavimas studentų moksliniams darbams* doc. dr. Vilmos Zubaitienės puslapyje (nuoroda į darbo autorės 2023 m. atliktą tyrimą). Prieiga internetu: <https://web.vu.lt/flf/v.zubaitiene/moksline-veikla/> [žiūrėta 2024 05 21].
19. Tekstynų įrankis Sketch Engine. Prieiga internetu: <https://www.sketchengine.eu/> [žiūrėta 2024 05 11].
20. Užklauso pavyzdys (pagal G. M. de Schryver, docx formatu), priedas prie straipsnio. Prieiga internetu: <https://academic.oup.com/ijl/article/36/4/355/7288213#supplementary-data> [žiūrėta 2024 05 11].
21. Vikižodynas. Prieiga internetu: https://lt.wiktionary.org/wiki/Viki%C5%BEodinas:Pagrindinis_puslapis [žiūrėta 2024 05 11].

Illustrative Sentences Generated by ChatGPT: The Analysis of Its Content, Quality, and Comparison with the Dictionary of the Standard Lithuanian Language

Summary

This Master's thesis aims to describe illustrative sentences generated by ChatGPT in Lithuanian by conducting a qualitative linguistic analysis of said sentences. Another goal of this thesis is to evaluate the quality of illustrative sentences generated by ChatGPT and compare them to illustrative sentences found in the Dictionary of the Standard Lithuanian Language. This research questions the possibility of transferring illustrative sentences generated by AI to a hypothetical dictionary. The suggested hypothesis states that illustrative sentences found in the Dictionary of the Standard Lithuanian Language are at the moment of higher quality than the ones generated by ChatGPT. The transfer of illustrative sentences generated by ChatGPT is currently impossible without additional lexicographer's editing. The theoretical part of this thesis includes most important lexicographical and AI-related terms, main principles of the Dictionary of the Standard Lithuanian Language, description of ChatGPT and its ties with lexicographical research. The empirical part encompasses syntactic, morphologic, lexical and pragmatic aspects of illustrative sentences generated by ChatGPT and the results of a survey where respondents were asked to evaluate if generated sentences sound authentic or not. The description of other aspects of AI-generated illustrative sentences and of sentences found in the Dictionary of the Standard Lithuanian Language are also included in this part. The end of the empirical part contains the comparison of illustrative sentences generated by ChatGPT with illustrative sentences found in the Dictionary of the Standard Lithuanian Language. The study shows that illustrative sentences generated by ChatGPT are syntactically similar to each other and not much semantic, lexical, morphological variety is found amongst them. The number of generated sentences is also low compared to the Dictionary of the Standard Lithuanian Language. Due to these findings one can conclude that it is possible to supplement a dictionary with lexicographically edited illustrative sentences generated by ChatGPT but not replace the illustrations usually found in a dictionary. This Master's thesis is relevant to the researchers of lexicography, lexicographers and the general public.

Priedas

ChatGPT pagal užklausą sugeneruoti IL

Žodis "kad" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Jungiamasis junginys, nurodantis priežastį, sąlygą arba tikslą. Ši reikšmė naudojama siekiant paryškinti, kodėl tam tikras veiksmas atliekamas arba turi būti atliktas, kokia yra priežastis ar sąlyga, lemianti kitą veiksmą. Pavyzdžiai:
 - Ji miegojo, kad pailsėtų po ilgos kelionės. (Čia "kad" nurodo priežastį - ji miegojo tam, kad pailsėtų.)
 - Jis nevalgė, kad numestų svorio. (Čia "kad" nurodo tikslą - jis nevalgė tam, kad numestų svorio.)
 - Aš nusipirkau naują automobilį, kad lengviau važiuotų į darbą. (Čia "kad" nurodo priežastį - aš nusipirkau naują automobilį tam, kad lengviau važiuotų į darbą.)

Ši reikšmė padeda nurodyti arba pabrėžti veiksmo priežastį, sąlygą arba tikslą sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kad" yra naudojamas.

Žodis "kuris" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Rodybinis įvardis, rodomasis kiekinis įvardis, nurodantis į konkretų asmenį, daiktą ar reiškini. Ši reikšmė naudojama aprašant arba nurodant apie ką kalbama. Pavyzdžiai:
 - Jonas, kuris sėdi šalia manęs, yra mano geriausias draugas. (Šiuo atveju "kuris" nurodo konkretų asmenį - Joną.)
 - Ta knyga, kurią skaitau, yra labai įdomi. (Čia "kuris" nurodo konkretų daiktą - knygą.)
 - Susitikome su žmogumi, kuris moka kalbėti septyniais skirtingais kalbomis. (Šiuo atveju "kuris" nurodo konkretų asmenį - žmogų.)

Ši reikšmė padeda tiksliai nurodyti arba pabrėžti tam tikrą sąvoką sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kuris" yra naudojamas.

Žodis "kitas" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Rodybinis įvardis, rodomasis kiekinis įvardis, naudojamas nurodyti į kitą asmenį, daiktą ar reiškini, kuris skiriasi nuo to, apie ką jau buvo kalbėta arba kuris buvo žinomas. Pavyzdžiai:
 - Pirmas kavos puodelis buvo mano, o antrasis buvo kito svecio. (Šiuo atveju "kitas" nurodo į kito asmenį arba daiktą, skirtingą nuo jau žinomo.)
 - Pirmas autobusas buvo perpildytas, todėl laukiau kito. (Čia "kitas" nurodo į kitą autobusą, skirtingą nuo to, kuris buvo pirmasis.)

Ši reikšmė padeda nurodyti arba pabrėžti tam tikrą sąvoką sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kitas" yra naudojamas.

Žodis "kaip" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Lyginamasis junginys, naudojamas nurodyti panašumą arba lyginimą tarp dviejų ar daugiau dalykų, veikslių ar reiškinių. Pavyzdžiai:
 - Ji guli ant žolės kaip mieganti rožė. (Lyginama jos poilsio būseną su miegančia rože.)
 - Jis skaito knygą taip, kaip mašina rašo. (Lyginami skirtingi veiksmai: skaitymas ir rašymas.)
 - Ši diena yra karšta kaip kūno liežuvis. (Lyginama dienos temperatūra su kūno liežuviu.)

Ši reikšmė padeda perteikti analogiją, panašumą arba lygumą tarp dalykų arba veikslių, kuriuos norime pabrėžti sakinio kontekste.

Žodis "kas" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Rodybinis įvardis, rodomasis kiekinis įvardis, naudojamas nurodyti konkrečią asmenį, daiktą ar reiškinį. Pavyzdžiai:
 - Tai yra mano draugas, kuris visada man padeda. (Šiuo atveju "kas" nurodo konkretų asmenį - draugą.)
 - Kas buvo tas keistasis garsas? (Čia "kas" nurodo konkrečią sąvoką - keistasis garsas.)
 - Aš nežinau, kas tai yra. (Šiuo atveju "kas" nurodo nežinomą reiškinį arba daiktą.)

Ši reikšmė padeda tiksliai nurodyti arba klausti apie konkrečią sąvoką sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kas" yra naudojamas.

Žodis "kai" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Laiko junginys, nurodantis tam tikrą laiko momentą arba intervalą. Ši reikšmė naudojama siekiant nurodyti arba apibrėžti, kada vyksta arba vyko tam tikras veiksmas arba reiškinys. Pavyzdžiai:
 - Kai aš atėjau į namus, ji jau miegojo. (Čia "kai" nurodo laiko momentą - tą akimirką, kai aš atėjau į namus.)
 - Aš visada valgau pusryčius kai saulė teksta. (Čia "kai" nurodo laiko momentą - kada saulė teksta.)
 - Kai būsi grįžęs, aš jau busiu miegojęs. (Čia "kai" nurodo laiko momentą - tą akimirką, kai asmuo grįš.)

Ši reikšmė padeda nurodyti konkretų laiko momentą arba intervalą sakinio kontekste, kuris yra būtinas siekiant suprasti, kada vyksta arba vyko tam tikras veiksmas arba reiškinys.

Žodis "koks" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Klaustukinis įvardis, rodybinis kiekinis įvardis, naudojamas klausti apie kokybę, pobūdį, savybes arba rūšį. Pavyzdžiai:
 - Koks tavo mėgstamiausias filmas? (Čia "koks" klauso apie mėgstamiausio filmo pobūdį ar savybes.)
 - Kokia orlaivių sparta yra didžiausia? (Čia "koks" klauso apie didžiausios spartos rūšį.)
 - Koks šios žalios spalvos atspalvis tau labiausiai patinka? (Čia "koks" klauso apie spalvos atspalvio savybes.)

Ši reikšmė padeda nurodyti arba klausti apie tam tikrą kokybę, pobūdį, savybes arba rūšį sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "koks" yra naudojamas.

Žodis "komisija" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Grupė asmenų, įpareigota ištirti, aptarti, spręsti tam tikrus klausimus arba priimti sprendimus konkrečiais klausimais. Pavyzdžiai:
 - Valstybinė komisija tiria korupcijos atvejus. (Šiuo atveju komisija yra įsteigta tyrimui atlikti.)
 - Mokslo komisija nusprendė finansuoti šį tyrimą. (Komisija priima sprendimą dėl finansavimo.)
 - Komisija patvirtino galutinį ataskaitos dokumentą. (Komisija peržiūri ir patvirtina dokumentą.)

Ši reikšmė padeda nurodyti arba įvertinti specialią grupę arba instituciją, įpareigotą atlikti tam tikrą užduotį arba priimti sprendimus, sakinio kontekste, kuri yra būtinas siekiant suprasti, kaip žodis "komisija" yra naudojamas.

Žodis "kartas" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Skaitvardis, reiškiantis veiksmo, būsenos ar reiškinio pasikartojimą tam tikru laiko momentu ar laikotarpiu. Ši reikšmė naudojama norint pabrėžti, kad veiksmas arba būseną pasikartojo arba gali pasikartoti. Pavyzdžiai:
 - Jis jau buvo čia du kartus šią savaitę. (Šiuo atveju "kartas" nurodo, kad jis buvo čia du kartus, pabrėžiant veiksmo pasikartojimą.)
 - Kartą mane aplanko paslaptis, kartą - meilė. (Šiuo atveju "kartas" nurodo pasikartojimą, kai aptariamos skirtingos būsenos.)

Ši reikšmė padeda nurodyti arba pabrėžti veiksmo arba būsenos pasikartojimą sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "kartas" yra naudojamas.

Žodis "klausimas" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Rodybinis daiktavardis, reiškiantis išsakytą arba parašytą frazę arba sakinių dalį, kuriam reikia atsakymo arba dėmesio. Pavyzdžiai:
 - Klausimas, kurį tu kėlei, yra labai svarbus. (Šiuo atveju "klausimas" nurodo išsakytą frazę arba idėją, kuriai reikia atsakymo.)
 - Mokytojas užduoda mums daug klausimų. (Čia "klausimai" reiškia užduotas frazes arba sakinių dalis, kurias mokiniai turi atsakyti.)
 - Ar tu gali atsakyti į mano klausimą? (Čia "klausimas" rodo išsakytą frazę, kuriai laukiama atsakymo.)

Ši reikšmė padeda nurodyti arba klausti apie tam tikrą išsakytą idėją arba frazę, kuri reikalauja atsakymo arba dėmesio, sakinio kontekste, kuris yra būtinas siekiant suprasti, kaip žodis "klausimas" yra naudojamas.

Žodis "komitetas" dažniausiai naudojamas su viena pagrindine reikšme:

1. Grupė asmenų, įpareigota ištirti, aptarti ar spręsti tam tikrą klausimą, rengti sprendimus ar parengti pasiūlymus dėl konkrečios srities veiklos. Pavyzdžiai:
 - Mokslo komitetas aptarinėja naujus tyrimo projektus.
 - Finansų komitetas rengia biudžetą.
 - Nusikaltimų prevencijos komitetas turi parengti veiksmų planą.

Ši reikšmė padeda nurodyti arba įvertinti specialią grupę asmenų, įpareigotą atlikti tam tikrą užduotį arba spręsti tam tikrus klausimus sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "komitetas" yra naudojamas.

Žodis "kalbėti" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Perduoti mintis, idėjas arba informaciją žodžiais, bendraudami su kitais asmenimis arba auditorija. Pavyzdžiai:
 - Ji kalba labai aiškiai ir suprantamai. (Dabarties laikas)
 - Anksčiau mes kalbėjome apie tai, ką galėtume daryti. (Praeitį)
 - Po valandos jis pradės kalbėti apie savo keliones. (Ateitį)

Ši reikšmė padeda nurodyti veiklą, kurios metu asmuo naudoja savo balso arba žodžių ryšį, kad perteiktų mintis, idėjas arba informaciją. Tai yra būtina norint suprasti, kaip žodis "kalbėti" yra naudojamas.

Žodis "kur" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Vietos nurodymo prievoksmis, naudojamas nurodyti vietą arba padėtį. Pavyzdžiai:
 - Mano kambaryje yra stalas, kur aš dirbu. (Dabarties laikas)
 - Tai yra miestas, kur mes praleidome atostogas. (Praeitį)
 - Rask mane ten, kur vėjas tyčia šnypščia lapais. (Ateitį)

Ši reikšmė padeda nurodyti vietą arba padėtį konkrečiame kontekste. Tai yra būtina norint suprasti, kaip žodis "kur" yra naudojamas.

Žodis "kaina" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Piniginė vertė, už kurią galima įsigyti arba parduoti prekę ar paslaugą. Pavyzdžiai:
 - Šios knygos kaina yra 20 dolerių. (Dabar)
 - Prieš keletą metų automobilio kaina buvo žymiai mažesnė. (Praeitis)
 - Kokia bus būsimų metų žaliavų kaina? (Ateitis)

Ši reikšmė padeda nurodyti piniginę sumą, kuriai reikia mokėti už prekę ar paslaugą, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kaina" yra naudojamas.

Žodis "kiekvienas" dažniausiai naudojamas su viena pagrindine reikšme:

1. Pabrėžti, kad kalbama apie visus individualius atvejus arba visus asmenis ar daiktus, priklausančius tam tikrai grupei arba kategorijai. Pavyzdžiai:
 - Kiekvienas mokinys turi baigti namų darbus iki pirmadienio. (Dabar)
 - Kiekvienas svečias buvo maloniai sutiktas. (Praeitis)
 - Rytoj kiekvienas darbuotojas turės apmokėti mokesčius. (Ateitis)

Ši reikšmė padeda nurodyti, kad kalbama apie kiekvieną individualų atvejį arba asmenį, priklausančią tam tikrai grupės arba kategorijos sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kiekvienas" yra naudojamas.

Žodis "keli" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Daugiau nei vienas, bet ne daugybė. Pavyzdžiai:
 - Kelios mergaitės žaidžia kieme. (Dabar)
 - Prieš kelis metus jis lankė tą pačią mokyklą. (Praeitis)
 - Po kelių valandų jie pasiekė tikslą. (Ateitis)

Ši reikšmė padeda nurodyti mažą, bet ne vieną kiekį, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "keli" yra naudojamas.

Žodis "komanda" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Grupė asmenų, dirbančių kartu tam tikrą užduotį atlikti arba bendrą tikslą pasiekti. Pavyzdžiai:
 - Futbolo komanda laimėjo turnyrą. (Dabar)
 - Komanda dirbo visą naktį, kad baigtų projektą laiku. (Praeitis)

- Rytoj komanda vyks į susitikimą su klientais. (Ateitis)

Ši reikšmė padeda nurodyti grupę asmenų, bendradarbiaujančių tam tikro tikslo ar užduoties vykdymui. Tai yra būtina norint suprasti, kaip žodis "komanda" yra naudojamas.

Žodis "kiek" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Kiekio arba dydžio nurodymas. Tai gali būti tiek kiekybinis, tiek kokybinis dydis. Pavyzdžiai:
 - Kiek obuolių pirkote? (Dabar)
 - Prašau pasakyti, kiek laiko praleidote bibliotekoje. (Praeitį)
 - Ateikite, kiek greičiau galite. (Ateitis)

Ši reikšmė padeda nurodyti kiekį arba dydį sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kiek" yra naudojamas.

Žodis "kaip" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Lyginamoji prasmė, išreiškianti panašumą arba skirtumą tarp dviejų dalykų ar reiškinių. Pavyzdžiai:
 - Šis medis yra aukštesnis kaip tas. (Panašumas, dabartis)
 - Ji mažiau kalba kaip anksčiau. (Skirtumas, praeitis)
 - Jūs turėsite pasirinkti, kaip mums geriau veikti ateityje. (Lyginamoji, ateitis)

Ši reikšmė padeda palyginti dvi ar daugiau dalykų ar reiškinių, nurodant jų panašumą arba skirtumą sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kaip" yra naudojamas.

Žodis "kol" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Laiko tarpas, pasireiškiantis veiksniu arba įvykiu vykdymo viena po kito tvarka ribose. Šis laiko tarpas turi būti nurodytas, kai įvykis arba veiksmas, pasireiškiantis po "kol", yra baigtinas arba įvykęs. Pavyzdžiai:
 - Ji skaitys knygą, kol laukėjimo metas pasibaigs. (Dabar)
 - Kol vaikai miegojo, jų tėvai pasiruošė vakarienei. (Praeitį)
 - Mes ruošsimės atostogoms, kol oras bus geras. (Ateitis)

Ši reikšmė padeda nurodyti laiko tarpą, per kurį vyksta arba vyko tam tikri veiksmai arba įvykiai, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kol" yra naudojamas.

Žodis "kartu" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Bendras veiksmas ar būseną, vykstanti su kita asmeniu ar daiktu tuo pačiu metu arba tuo pačiu viename kontekste. Pavyzdžiai:
 - Mes valgėme pusryčius kartu su šeima. (Dabar)
 - Jie kelionėje praleido daug laiko kartu. (Praeitį)
 - Vakara praleidome kartu su draugais. (Ateitis)

Ši reikšmė padeda nurodyti, kad veiksmas ar būseną vyksta vienu metu su kita asmeniu ar daiktu, bendrame kontekste. Tai yra būtina norint suprasti, kaip žodis "kartu" yra naudojamas.

Žodis „kilti“ gali turėti kelias reikšmes, tačiau čia pateiksiu pagrindinę:

Kilti - pakilti į viršų, atsistoti, iškilti virš paviršiaus, pasipriešinti žemyninei jėgai.

Pavyzdžiai:

1. *Fizinis judėjimas*: Vaikinas kėlė ranką, norėdamas atsistoti iš lovos.
2. *Atsistatymas*: Po sunkaus darbo buvo sunku atsikelti iš kėdės.
3. *Pasisukimas*: Saulė pradėjo kilti, o tai reiškė, kad artėjo rytas.
4. *Moralinis judėjimas*: Piliečių nerimas kėlė iškilmingas demonstracijas prieš valdžią.
5. *Aukštinė tendencija*: Palūžusi bendrovė pradeda kilti ir vėl tapti pelninga.

Šioje reikšmėje „kilti“ dažniausiai naudojama pažymėti judėjimą į viršų arba pakilimą, tiek fiziškai, tiek metaforiškai.

Žodis "kodėl" gali turėti kelias reikšmes, tačiau viena iš pagrindinių yra:

1. Priežastis arba motyvo išreikšimas, kuris paaiškina, kodėl įvyko arba vyksta tam tikras veiksmas arba įvykis. Pavyzdžiai:
 - Kodėl jis vėluoja? (Dabar)
 - Aš nežinau, kodėl ji neatsiliepia į mano žinutes. (Praeitį)
 - Kodėl taip svarbu laikytis taisyklių? (Ateitis)

Ši reikšmė padeda nurodyti priežastį arba motyvą tam tikram veiksmui arba įvykiui, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kodėl" yra naudojamas.

Žodis "keturi" turi vieną pagrindinę reikšmę:

1. Skaičius, nurodantis kiekį, lygų keturiems. Pavyzdžiai:
 - Šiame krepšyje yra keturi obuoliai. (Dabar)
 - Prieš penkerius metus jis turėjo keturis mašinas. (Praeitį)
 - Rytoj susitikime su keturiais draugais. (Ateitis)

Ši reikšmė padeda nurodyti skaičių keturiems, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "keturi" yra naudojamas.

Žodis "keisti" gali turėti kelias reikšmes, tačiau čia paaiškinsiu pagrindinę reikšmę:

1. Pakeisti vieną dalyką arba būklę į kitą, dažniausiai siekiant gerėjimo arba tobulėjimo.
Pavyzdžiai:
 - Aš nusprendžiau keisti savo darbo vietą. (Dabar)
 - Šis kompiuteris jau senas, reikėtų jį keisti nauju. (Ateitis)
 - Ji keitė savo gyvenimo būdą, siekdama geresnės sveikatos. (Praeitis)

Ši reikšmė padeda nurodyti veiksmą, kuomet kažkas yra pakeičiama į kitą, dažniausiai siekiant gerėjimo arba tobulėjimo, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "keisti" yra naudojamas.

Žodis "kelti" gali turėti kelias reikšmes, tačiau čia paaiškinsiu pagrindinę reikšmę:

1. Pakelti arba perkelti ką nors iš vienos vietos į kitą, dažniausiai vertikalčiai arba aukštyn.
Pavyzdžiai:
 - Jis kėlė sunkią krepšinio kamuolį aukštyn. (Dabar)
 - Juokaujant, jie kėlė mane ant pečių. (Praeitis)
 - Rytoj mes kelsime naują lentyną į virtuvę. (Ateitis)

Ši reikšmė padeda nurodyti veiksmą, kuomet kažkas yra pakeliamas arba perkeliamas iš vienos vietos į kitą, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kelti" yra naudojamas.

Žodis "kurti" gali turėti kelias reikšmes, tačiau čia paaiškinsiu pagrindinę reikšmę:

1. Kūrybinis procesas, kurio metu yra sukuriamas, kuriamas arba įgyvendinamas kažkas naujo arba unikalaus. Pavyzdžiai:
 - Menininkas kūrė paveikslą visą savo gyvenimą. (Praeitis)
 - Šiuo metu jis kuria naują romaną. (Dabar)
 - Jie planuoja kurti naują verslą kitąmet. (Ateitis)

Ši reikšmė padeda nurodyti veiksmą, kuomet kažkas naujo arba unikalaus yra kuriamas, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kurti" yra naudojamas.

Žodis "kreiptis" gali turėti kelias reikšmes, tačiau čia paaiškinsiu pagrindinę reikšmę:

1. Kreiptis į kažką arba kažką dėl pagalbos, patarimo arba informacijos. Tai reiškia ieškoti arba prašyti paramos, patarimo arba informacijos iš tam tikros institucijos, asmenų arba šaltinių.

Pavyzdžiai:

- Jis kreipėsi į gydytoją dėl sveikatos problemų. (Praeitis)
- Ar galiu kreiptis į tave pagalbos? (Dabar)
- Jie ketina kreiptis į teisininką dėl teisinio patarimo. (Ateitis)

Ši reikšmė padeda nurodyti veiksmą, kuomet asmuo arba grupė pritampa informacijos, pagalbos arba patarimo ieškotojais iš tam tikro šaltinio, sakinio kontekste. Tai yra būtina norint suprasti, kaip žodis "kreiptis" yra naudojamas.

Žodis "kadangi" naudojamas siekiant pateikti priežastį ar paaiškinimą tam, ką pasakoma sakinio dalyje. Šis junginys rodo, kad sakinio dalis, prie kurios jis pridedamas, yra priežastinė ir tiesiogiai lemia tai, kas pasakoma kitame sakinio dalyje.

Pavyzdžiai:

1. Ji buvo pavargusi, kadangi dirbo visą naktį. (Praeitis)
2. Aš atidėjau susitikimą, kadangi buvau užimtas. (Dabar)
3. Jis nebuvo susirūpinęs dėl ateities, kadangi turėjo stabilų darbą. (Praeitis)
4. Mes praleidome vakarą kartu, kadangi buvome geri draugai. (Dabar)
5. Jie nusprendė perkelti vestuves, kadangi oras buvo blogas. (Ateitis)

Šie pavyzdžiai iliustruoja, kaip žodis "kadangi" naudojamas siekiant nurodyti priežastinį ryšį tarp dviejų ar daugiau sakinio dalių. Tai padeda aiškiau suprasti, kodėl tam tikras veiksmas ar situacija įvyko ar yra svarbi.

Žodis "karinis" gali turėti kelias pagrindines reikšmes:

1. Susijęs su karu arba kariniais veiksmais, taikomas kariuomenės veiklai arba militariniams reikalams.
2. Būdingas karui, militarizmui, kariuomenei arba ją sudarančioms struktūroms.

Pavyzdžiai:

1. Pirmoji reikšmė - susijusi su karu arba kariniais veiksmais:
 - Karinė operacija buvo sėkmingai įvykdyta.
 - Jis tarnavo karinėje karo laivyno laivoje.
 - Miestas patyrė žiaurų karinį konfliktą.
2. Antroji reikšmė - būdingas karui arba susijęs su kariuomenės veikla:
 - Kariniai drabužiai skirti kariams kovinei aprangai.
 - Karinis lyderis vadovavo operacijai.
 - Jo gyvenimas buvo susijęs su nuolatine karine treniruote.

Šie pavyzdžiai aiškiai parodo, kaip žodis "karinis" gali būti naudojamas tiek tiesiogiai susijus su karu ir kariniais veiksmais, tiek ir kaip būdingas arba susijęs su kariuomene bei jos veikla.

Žodis „ketvirtas“ gali turėti kelias reikšmes, tačiau pagrindinė jo reikšmė yra susijusi su skaičiavimu ir nurodo ketvirtą poziciją arba laipsnį sekos tvarka.

Pavyzdžiai:

1. Skaičiavimo reikšmė:
 - Šiandien yra ketvirtas spalio mėnesio diena.
 - Penktadienis yra savaitės ketvirtas dienos darbo laikas.
 - Jis finišavo lenktynėse ketvirtas.
2. Vieta arba pozicijos reikšmė:
 - Ketvirtas namas į kairę nuo kryžkelės yra mano.
 - Jis buvo ketvirtasis, kuris pateko į tikslą.
 - Ketvirtasis aukštas pastate yra vienas jaukiausių.

Šie pavyzdžiai aiškiai parodo, kaip žodis „ketvirtas“ gali būti naudojamas tiek skaičiavimo, tiek vieta arba pozicijos prasme.

Žodis „konkretus“ turi kelias pagrindines reikšmes, tačiau pagrindinė iš jų yra susijusi su aiškumu, nurodytu arba apibrėžtu.

1. Aiškumo, nurodyto arba apibrėžto reikšmė:
 - Šiame projekte reikia pateikti konkretų planą.
 - Jis įsipareigojo atvykti į susitikimą tam tikru, labai konkrečiu laiku.
 - Prašome pateikti konkretų atsakymą į šį klausimą.

Šie pavyzdžiai rodo, kaip žodis „konkretus“ naudojamas siekiant pabrėžti aiškumą, nurodytumą arba apibrėžtumą. Tai padeda aiškiai nurodyti, kas arba kas tiksliai yra turima omenyje, be dviprasmybių ar neapibrėžtumo.

Žodis „konstitucinis“ turi vieną pagrindinę reikšmę:

1. Susijęs su konstitucija, pagrindinių įstatymų arba valstybės pagrindinių teisės aktų nuostatomis:
 - Konstitucinis teismas yra institucija, kuri tikrina teisės aktų suderinamumą su konstitucija.
 - Valstybės vadovo įgaliojimai yra aiškiai apibrėžti konstituciniuose dokumentuose.
 - Ši teisė yra garantuota konstituciniuose dokumentuose.

Šie pavyzdžiai parodo, kad žodis „konstitucinis“ naudojamas aprašant ką nors, kas yra susijęs arba grindžiamas pagrindiniais valstybės įstatymais arba konstitucija. Tai gali būti taikoma institucijoms, teisės aktams, įgaliojimams ar teisėms.

Žodis „kuo“ turi keletą pagrindinių reikšmių:

1. Palyginimo būdas arba priemonė:
 - Jis yra aukštesnis kuo auga.
 - Skęsta greičiau kuo ilgiau trukdo pagalba.
2. Priežastis, tikslinės paskirties arba siekio pabrėžimas:
 - Mokslininkai stengiasi kuo greičiau rasti vakciną.
 - Ji nori, kad viskas būtų kuo geriau pasiruošę.
3. Lyginimo prielaida, sąlyga:
 - Pasisakysiu prieš tave, kuo reikia.
 - Atvykstu, kuo jau bus baigęsi susirinkimas.
4. Pasirinkimo riba, aukščiausias laipsnis:
 - Žaisti įmanoma kuo geriau.
 - Pasisstengė, kad tik jo rezultatai būtų kuo didesni.
5. Laikotarpio riba, terminas:
 - Jis turi pranešti kuo greičiau.
 - Tegul tai atvyksta kuo anksčiau.

Pateikti pavyzdžiai demonstruoja, kaip žodis „kuo“ gali būti naudojamas skirtingose situacijose, pabrėžiant palyginimus, priežastis, sąlygas, pasirinkimo ribas arba laiko tarpus.

Žodis „komercinis“ turi keletą pagrindinių reikšmių:

1. Susijęs su prekyba arba verslu:
 - Komercinis užsiėmimas yra susijęs su pelno siekimu.
 - Jis turi komercinį projektą, kurio tikslas yra plėtoti verslą.
2. Skirtas reklamai arba rinkodarai:
 - Jie kūrė komercinį skelbimą, norėdami pritraukti daugiau klientų.
 - Komercinės reklamos dažnai naudojamos siekiant didinti produkto ar paslaugos pardavimus.
3. Atitinkantis rinkos poreikius arba reikalavimus:
 - Jų prekės turi būti komercinio paskirties, kad atitiktų klientų lūkesčius.
 - Komercinis automobilis yra pritaikytas verslo reikmėms, turintis tam tikrų funkcijų ir patogumų.
4. Susijęs su pelno siekimu arba finansiniais interesais:
 - Jis įsigijo komercinę nuosavybę investavimo tikslais.
 - Verslo planas buvo parengtas su komerciniais tikslais – siekiant didelio pelno.

Šie pavyzdžiai parodo, kaip žodis „komercinis“ gali būti naudojamas aprašant veiklą ar objektus, susijusius su prekyba, verslu, reklama, rinkodara ar finansiniais interesais.

Žodis „keistas“ turi keletą pagrindinių reikšmių:

1. Netipinis, neįprastas arba neįprastai elgiantis:
 - Ji pastebėjo keistą objektą ant stalo ir nustebo.
 - Jo elgesys buvo labai keistas – jis kalbėjo beveik savaip.
2. Neaiškus, painus arba nesuprantamas:
 - Mano sapnas buvo labai keistas, ir aš negalėjau jo aiškiai paaiškinti.
 - Tai yra keistas momentas mūsų istorijoje, kai viskas tapo neaišku.
3. Susijęs su magija, mistika arba paslaptinguumu:
 - Senovės rūsyje jie aptiko keistą artefaktą, kuris kėlė daugybę klausimų.
 - Keistas veidrodis buvo įspūdingas eksponatas paslaptingoje muziejaus parodoje.

Šie pavyzdžiai padeda suprasti, kaip žodis „keistas“ gali būti naudojamas aprašant neįprastą, painų arba paslaptinę dalyką, elgesį arba situaciją.

Žodis „karštas“ turi kelias pagrindines reikšmes:

1. Aukšta temperatūra arba šiluma:
 - Ši kava yra per karšta gerti, palaukite, kol atvės.
 - Vasarą oras dažnai būna labai karštas, todėl svarbu gerti daug vandens.
2. Aštrus, degus arba skaudrus dėl aukštos temperatūros:
 - Mano avinas sriubos buvo per karšta, ir aš netyčia sudegiau liežuvį.
 - Sodininkas uždegė savo rankas dirbdamas su karštu geležiniu stovu.
3. Labai populiarius, intensyvus arba svarbus:
 - Naujas restoranas mieste yra šiuo metu labai karštas vietinis kalbų.
 - Šis metų filmas yra karštas kandidatas apdovanojimams, visi jau kalba apie jį.
4. Emociškai aštrus, įkaitęs arba intensyvus:
 - Prieš kovą tarp komandų susidaro karštos diskusijos.
 - Politinės debatai gali tapti labai karšti ir įtempti.

Šie pavyzdžiai padeda suprasti, kaip žodis „karštas“ gali būti naudojamas skirtingose situacijose, kuriose kalbama apie temperatūrą, populiarumą, aštrumą arba emocinę intensyvumą.

Žodis „kažkoks“ gali turėti keletą reikšmių, tačiau dažniausiai jis naudojamas siekiant nurodyti nežinomą arba neapibrėžtą dalyką ar asmenį.

1. Nurodo nežinomą arba neapibrėžtą dalyką ar asmenį:
 - Pirkau kažkokį daiktą rinkinį, bet dar nežinau, kas jame yra.
 - Šiame mieste susitiksiu su kažkokiais draugais iš praeities.

Šis žodis gali būti naudojamas norint pabrėžti neaiškumą arba neapibrėžtumą tam tikroje situacijoje.

Papildomai, „kažkoks“ gali turėti ir kitų reikšmių kontekste, tačiau dažniausiai jis naudojamas kaip aukščiau paminėta: nurodyti nežinomą arba neapibrėžtą dalyką ar asmenį.

Žodis „korupcija“ paprastai žymimas kaip neteisėtas veiksmas, kai asmuo ar organizacija piktnaudžiauja savo įgaliojimais arba galiomis siekdami neteisėtos naudos. Pagrindinės šio žodžio reikšmės yra:

1. Neteisėtas veiksmas, siekiant asmeninės naudos: Tai yra veiksmas, kai asmuo ar grupė išnaudoja savo oficialią padėtį ar galimybes pasiekti neteisėtą naudą, dažniausiai finansinę, paliekant šalį ar instituciją žalai.

Pavyzdžiai:

- Politikų korupcija dažnai pasireiškia, kai jie priima dovanas ar mokesčius už savo veiksmus arba politinių sprendimų priėmimą.
- Korupcija teismuose gali reikšti teisėjų ar advokatų priėmimą kyšių už sprendimų priėmimą ar teisinės pagalbos teikimą.

Ši pagrindinė reikšmė „korupcija“ yra plačiai naudojama ir aprėpia įvairias situacijas, kuriose vyksta neteisėti veiksmai siekiant asmeninės arba organizacinės naudos.

Žodis „kanceliarija“ gali turėti kelias reikšmes, tačiau čia pateiksiu pagrindinę:

Kanceliarija - tai patalpos arba vietos, skirtos administraciniams darbams atlikti, dokumentų tvarkymui, susitikimams su klientais ir pan.

Pavyzdžiai:

1. *Administracinė patalpa*: Direktorius kanceliarija yra įrengta antro aukšto ir turi puikų vaizdą į miesto centrą.
2. *Verslo vieta*: Advokato kanceliarija suteikia paslaugas įvairioms teisinėms problemoms spręsti.
3. *Valstybinė institucija*: Prezidento kanceliarija yra atsakinga už įvairius valstybės administracinius reikalus.
4. *Asmeninė darbo vieta*: Mažame miestelio kanceliarijoje vietinio notaro tarnybininkas priima dokumentų teikimą.

Šioje reikšmėje „kanceliarija“ dažniausiai naudojama verslo, administracijos ar institucijų kontekste, žymint administracinę patalpą ar vietą, skirtą darbui ir dokumentų tvarkymui.