

Vārdkopas analogi „Latviešu valodas sintaktiski marķētajā korpusā”¹

Analogue of subordinate phrase in Latvian Treebank

Laura Rituma, Gunta Nešpore-Bērzkalne, Baiba Saulīte,
Lauma Pretkalniņa

Matemātikas un informātikas institūts
Latvijas Universitāte

Raiņa bulvāris 29, Rīga, LV-1050, Latvija

E-pasts: laura.rituma@lumii.lv, gunta.nespore@lumii.lv,
baiba.valkovska@lumii.lv, lauma.pretkalnina@lumii.lv

Latvijas Universitātes Matemātikas un informātikas institūtā tiek veidots „Latviešu valodas sintaktiski marķētais korpus” (LVTB), kurā tiek marķētas gan latviešu valodas sintakses teorijā jau aprakstītās latviešu valodas sintaktiskās parādības, gan arī retākas, līdz šim gramatiskās sīkāk neanalizētas konstrukcijas. Šajā rakstā aplūkota vārdkopas analoga izpratne latviešu valodniecībā, kā arī apkopoti līdz šim LVTB marķētie vārdkopas analogi un to klasifikācija. Sīkāks vārdkopas analoga iedalījums grupās balstīts uz noteiktiem gramatiskiem (vārdšķiras piederība un sintaktiskais raksturojums) un semantiskiem kritērijiem, šajos saliktajos teikuma locekļos nedaudz analizēta arī komponentu secība. Izplatītākie vārdkopas analoga veidi ir dažādu vietniekvārdu savienojumi, kā arī konstrukcijas ar prievārda savienojumu, kas nosauc daļu no kopuma. Pārējie parādās retāk, bet piemēru daudzums ir pietiekams, lai atklātu visos konstrukciju veidos lietoto leksēmu daudzveidību un izmantošanu dažādās sintaktiskajās funkcijās. Konsekventi vārdkopas analoga noteikšanas kritēriji un vienveidīgs sintaktiskais marķējums nodrošina ērti izmantojamu valodas materiālu turpmākai vārdkopas analoga un citu sintaktisko konstrukciju analīzei.

Atslēgvārdi: latviešu valodas sintakse; sintaktiski marķēts korpus; vārdkopas analogs; salikts teikuma loceklis.

¹ Publikācija izstrādāta Valsts pētījumu programmas „Letonika latviskas un eiropiskas sabiedrības attīstībai” projektā „Mūsdienu latviešu valodas izpēte un valodas tehnoloģiju attīstība” (Nr. VPP-LETONIKA-2021/1-0006).

levads

„Latviešu valodas sintaktiski marķētais korpuss” (turpmāk tekstā – LVTB; sīkāku informāciju par korpusu sk. <https://korpuss.lv/id/LVTB>) Latvijas Universitātes Matemātikas un informātikas institūta Mākslīgā intelekta laboratorijā tiek veidots kopš 2010. gada, un šobrīd tā apjoms (aptuveni 17 000 teikumu; 282 tūkstoši tekstvienību) ir pietiekams gan dažādiem mašīnmācīšanās uzdevumiem (Pretkalniņa 2023, 55), gan dažādu latviešu valodas gramatisko parādību izpētei.

LVTB izveidē ir izmantots šī raksta autoru izstrādātais hibrīdais gramatikas modelis (sīkāk par gramatikas modeļa realizāciju mašīnlasāmā sintakses korpusā sk. Pretkalniņa 2023), kas balstīts atkarību sintaksē un papildināts ar frāzes struktūras sintakses elementiem (Rituma, Saulīte, Nešpore-Bērzkalne 2019), t. i., teikuma struktūra tiek attēlota kā atkarību sintakses koks, kurā atkarības parāda sintaktiskos sakarus teikumā starp teikuma sintaktiskajām vienībām, bet šīs vienības var būt vienkāršas (sastāv no vienas tekstvienības) vai saliktas (sastāv no vairākām tekstvienībām). Salikta vienība ir frāze, un tai kā vienotam veselumam tiek piešķirta sava sintaktiskā loma. Ar frāzi tiek attēlotas gan analītiskās formas un salikti teikuma locekļi, gan arī plašākas sintaktiskās konstrukcijas, piem., iespraudumi, tiešā runa (plašāk sk. Rituma, Saulīte, Nešpore-Bērzkalne 2019, 208–213).

Viena no frāzēm LVTB gramatikas modelī (frāzes apzīmējums modelī – subAnal) paredzēta vārdkopas analogam (Rituma, Saulīte, Nešpore-Bērzkalne 2019, 209–210), kas (Skujiņa 2007, 435) skaidrots kā „nedalāms vārdu savienojums, kas formāli atgādina vārdkopu, bet kurā starp komponentiem nav pakārtojuma sakara, piem., *kaut kas interesants, tāds skaists, viņi abi*”.

Latviešu valodniecībā vārdkopas analogi līdz šim aprakstīti pavisam nedaudz – termins pārņemts no slāvu valodniecības 20. gs. 70. gados (Kārklīņš 1976, 71), to vēlāk valodas kultūras un sintakses teorijā lieto arī Inta Freimane (1993, 244–245; 1997, 10) un Ilze Lokmane (2013a, 118), bet latviešu gramatikās (piem., Ahero et al. 1962; Lokmane 2013b; Kalnača, Lokmane 2021) tas netiek plašāk analizēts. Tomēr tas nenozīmē, ka šādu konstrukciju savdabīgums sintaktiskā aspektā nav pamanīts. Atsevišķi vārdkopas analoga konstrukcijām līdzīgi piemēri atrodami vietniekvārdu funkciju raksturojumā un saliktu teikuma locekļu aprakstos, kur tie saukti par „frazēoloģiskiem vārdu savienojumiem” (Ozols 1959, 39–40), „tradicionāliem sakopojumiem” (Ahero et al. 1959, 514), „vārdu savienojumiem” (Nītiņa 2013, 431), „sintaktiski / jēdzieniski nedalāmiem vārdu savienojumiem” (Valdmanis 1989, 32–33).

Šī pētījuma mērķis ir precizēt vārdkopas analoga izpratni, konsekventi marķēt šo grupu LVTB un raksturot iegūtos rezultātus.

LVTB tiek veidots kā „Līdzsvarotā mūsdienu latviešu valodas tekstu korpusa” (turpmāk tekstā – LVK2018; sīkāku informāciju par korpusu sk. <https://korpuss.lv/id/LVK2018>) apakškopa un joprojām tiek papildināts – divas reizes gadā tiek publicēta jauna LVTB versija (šajā rakstā izmantoti dati no LVTB 2.10 versijas). LVTB ir manuāli sintaktiski marķēts korpuss, un tā apjoms ir pietiekams, lai savāktu plašu klāstu ar vārdkopas analoga piemēriem (gandrīz 800 piemēru) un izveidotu daļījumu apakšgrupās, tomēr atsevišķos gadījumos (vārdkopas analoga komponentu novietojuma un

retāku konstrukciju analīzē) rakstā sniegti piemēri arī no daudz plašākā, automātiski morfoloģiski marķētā LVK2018.

Pētījuma gaitā vārdkopas analoga jēdziena izpratne paplašināta, balstoties gan uz vārdkopas analogu un citu saliktu teikuma locekļu aprakstiem latviešu valodniecībā (sk. 1. nodaļu), gan uz empīrisko materiālu. Vārdkopas analogi sadalīti vairākās grupās (sk. 2. nodaļu), pamatā pēc tā, no kādu vārdšķiru vārdiem tie sastāv, tomēr ir arī citi kritēriji. Nobeigumā sniegti secinājumi un izvirzīti jautājumi turpmākiem pētījumiem.

1. Vārdkopas analoga izpratne latviešu valodas sintaksē

Vārdkopas analoga jēdzienu latviešu valodniecībā ieviesis Jūlijs Kārklīšs (1976, 71–72), tieši pārņemot to no Gaļinas Zolotovas (*Galina Zolotova* 1973, 55) pētījumiem sintaksē. Viņa par vārdkopas analogu (krievu val. *analog slovosochetanija*) uzskata piekļāvumu – savienojumu, kurā starp elementiem ir nevis pakārtojuma sakars, jo vārda sakaru formālo rādītāju nav, bet jēdzieniskas attieksmes (piem., *māja mežā, skriet pa labi*). (*Aut. piez.*: mūsdienu izpratnē piekļāvums (Lokmane 2013b, 705) ir pakārtojuma veids, kurā vārdkopas atkarīgais komponents ir nelokāms vārds vai vārds nelokāmā formā.)

Inta Freimane (1993, 244–245) terminu „vārdkopas analogs” attiecina uz citu parādību loku – vārdu savienojumiem, kuri atgādina vārdkopu, bet to pamatā ir nevis pakārtojuma sakars, jo pakārtojuma sakars izriet no pakārtojuma vārdu semantikas kategoriālajām nozīmēm, bet gan jēdzieniskās attieksmes, resp., vārdkopas analogā vārdi atrodas „zināmās semantiskās koordinācijas attieksmēs”. Autore arī piebilst, ka sintaktiski vārdkopas analogs funkcionē kā viena sintaktiska vienība, to veido divi vai vairāki patstāvīgu vārdšķiru vārdi, bet leksiskās nozīmes ziņā tie ir nepilnīgi, jo nenosauc reālījas tieši, bet nosauc dažādas attieksmes. Būtiski, ka I. Freimane (op. cit.) uzskata: starp vārdkopas analoga komponentiem pastāv zināms formu salāgojums, tomēr tas nav pilntiesīgs sintaktisks sakars. Vārdkopas analogu parasti veido vietniekvārdi un skaitļa vārdi, retāk viens no komponentiem ir īpašības vārds. Tāpat I. Freimane (op. cit.) norāda, ka ir saskaņojuma analogi un pārvaldījuma analogi, tiesa, vienīgais piemērs pārvaldījumam ir vārdu savienojums *mēs ar tevi*, savukārt kā saskaņojuma piemēri sniegti *tu pats, kaut kas labs, tāds smieklīgs, nekas tāds* u. tml.

Šāds raksturojums raisa vairākus jautājumus – kas ir semantiskās koordinācijas attieksmes? Kā definēt sakaru starp vārdkopas analoga komponentiem (lai neveidotu definīcijas ar noliegumu)? Kuras no minētajām pazīmēm ir būtiskas, lai noteiktu parādības piederību pie vārdkopas analogiem? Vai vārdkopas analogi ir vienīgās konstrukcijas, kas atbilst minētajam raksturojumam?

No I. Freimanes (1993, 245) piemēriem (*mēs visi, jūs neviens, tie abi*) var secināt, ka būtiska pazīme ir vietniekvārda iesaiste vārdu savienojumā – tieši vietniekvārdu semantiskā funkcija nosaukt reāliju netieši attieksmē pret citu leksisko vienību rada ciešās semantiskās attieksmes starp vārdkopas analoga komponentiem. Tāpat būtiska ir piezīme, ka vārdu savienojums funkcionē kā viens elements sintaktiskajā struktūrā jeb, citiem vārdiem, kā salikts teikuma loceklis, tomēr ir jānoskaidro, kādos gadījumos

var runāt par saliktu teikuma locekli. Pati I. Freimane (1997, 10) ir norādījusi, ka viena teikumlocekļa funkciju spēj veikt frazeoloģismi, salikti numerāļi, arī vārdkopas analogi, ja ar tiem saprot divu vai vismaz vienas attieksmju nosaucējas vārdšķiras vārdu (t. s. deiktisko vārdu) saistījumu savā starpā vai ar kādu pilnnozīmes vārdu *dažs labs, viņi abi, mēs visi, kaut kas traks, viens pats, tāds viltīgs* u. tml.

Latviešu valodniecībā par šādiem nedalāmiem vārdu savienojumiem ir runāts vēl pirms vārdkopas analoga termina ieviešanas. 20. gs. 50. gados Arturs Ozols (1959, 39–40) jau rakstījis par vietniekvārdu savienojumiem *dažs labs, šāds tāds, viens otrs, tas pats, šis pats* un lietojis terminu „pronomiņāls frazeoloģisms”, norādot uz to veselumu jeb mūsdienu izpratnē, salikta teikuma locekļa funkciju, tomēr nav precizējis, kā raksturot sintaktisko saistījumu starp vietniekvārdiem šādā frazeoloģismā. Arī „Mūsdienu latviešu literārās valodas gramatikā” (Ahero et al. 1959, 506) personas vietniekvārdu lietošanas aprakstā minēts, ka, konkretizējot vietniekvārda saturu, tam reizēm pievienojas vai nu prievārds ar locījumu (*mēs ar tevi (viņu, viņiem), mēs ar brāļi*), vai skaitļa vārds (*mēs divi*), vai vietniekvārds (*mēs visi* – par vairākām personām; *mēs citi* – tie, kuriem runātājs pieskaita sevi), vai lietvārds resp. vārdkopa pielikuma vietā, uz ko norāda *mēs, piem., mēs, trešās klases audzēkņi*. Kā var redzēt, precīzāk netiek nosaukts ne sintaktiskais saistījums starp vārdiem, ne sintaktiskā vienība, kas izveidojas. Savukārt nodaļā par norādāmajiem vietniekvārdiem minēti „tradicionālie sakopojumi *šis tas, šo to, šāds tāds* vārdu kārtā, ko nemēdz grozīt” (Ahero et al. 1959, 514). Precīzs termins netiek lietots arī jaunākajā latviešu valodas gramatikā (Kalnača, Lokmane 2021, 183, 208), bet vietniekvārdu aprakstā norādīts, ka reizēm vietniekvārds ir vietniekvārda un adjektīva vai skaitļa vārdu kombinācija, piem., *dažs labs un viens otrs*.

LVTB vārdkopas analoga grupa veidota, ņemot vērā arī teoriju par citiem nedalāmiem vārdu savienojumiem. „Mūsdienu latviešu literārās valodas gramatikā” (Ahero et al. 1962, 16–17) vārdkopu raksturojumā skaidrots, ka ir jēdzieniski nedalāmas vārdkopas, kuras nevar sadalīt komponentos un kurām ir nostabilizējusies viena vārda nozīme. To attiecina uz frazeoloģismiem, un tiek norādīts, ka teikumā tas parasti ir viens teikuma loceklis, lai gan šajā gramatikā dotajā piemērā *mest mieru* starp komponentiem formāli ir pakārtojuma sakars. Tādējādi tiek parādīts uzskats, ka teikuma sintaktiskās struktūras noteikšanā iesaistās semantika. Teikuma locekļu aprakstā lietots termins „paplašinātais teikuma loceklis”, kuru veido viss vārdu savienojums, tiesa, te iekļauti gan atsevišķi prievārda savienojumi (*mēs ar Piču, kāds no puikām*) un vietniekvārdu savienojumi (*nekas cits, tas viss*), gan vārdu savienojumi ar kvantitatīvas nozīmes lietvārdu nominatīvā (*pulks meitu, grupiņa latviešu*), gan vārdu savienojumi, kas apzīmē vienu jēdzienu (*baltā akācija, Skroderdienas Silmačos*), un pat veseli teikumi kā paplašinātie teikuma locekļi (Ahero et al. 1962, 225–227, 243 u. c.).

Jānis Valdmanis (1989, 32–33) šķir „sintaktiski nedalāmus vārdu savienojumus” un „jēdzieniski nedalāmus vārdu savienojumus”, kas var veikt salikta teikuma locekļa funkciju. Par sintaktiski nedalāmiem vārdu savienojumiem tiek uzskatīti lietvārda, vietniekvārda vai skaitļa vārda ar daudzuma nozīmi savienojumi ar lietvārdu ģenitīvā vai datīvā (ar datīvu domāti gadījumi, kad savienojumā ir prievārda konstrukcija), piem., *grupa latviešu, daļa no teksta, maz skatītāju, daži no jauniešiem, viens no klausītājiem*,

vairāk saules, desmit cilvēku. Savukārt jēdzieniski nedalāmo vārdu savienojumu grupā J. Valdmanis iekļāvis nosaukumus, terminus, frazeoloģiskus savienojumus u. tml., bet šī raksta kontekstā aktuāli ir tādi jēdzieniski nedalāmi vārdu savienojumi kā *kāds no draugiem, nekas cits, tas viss, mēs paši*.

Arī pie adverbīāliem vārdu savienojumiem ar vietas nozīmi gramatikās ir atrodamas līdzīgas norādes uz sintaktisko un jēdzienisko veselumu. Piem., „Mūsdienu latviešu literārās valodas gramatikā” (Ahero et al. 1959, 697–698) minēts, ka tādi apstākļa vārdu savienojumi, kas cēlušies no lietvārdiem (piem., *augšā, lejā* u. tml.), bet pirmajā daļā ir salikts apstākļa vārds *kaut kur*, ir daudz ciešāk saauguši un teikumā funkcionē kā viena vienība ar vienas vietas apzīmējuma nozīmi (*kaut kur augšā*). Šo skatījumu paturējusi arī Gunta Smiltneiece (2013, 608), adverbu raksturojumā norādot, ka precīzākai vietas apzīmēšanai tiek lietoti divi vai vairāki vietas apstākļa vārdi līdzās, veidojot vienu sintaktisko vienību, piem., *tepat lejā, priekšā pa kreisi, te pie loga, tur augšā, kaut kur tālāk, kaut kur aiz meža*. Pašlaik LVTB gramatikas modelī šāda adverbīāla vārdkopas analogu apakšgrupa nav veidota, bet tiek apsvērtas iespējas tādu ieviest.

„Latviešu valodas gramatikā” vārdkopas analogi nav skaidroti. Vietniekvārdu raksturojumā (Nītiņa 2013, 431) gan minēts, ka vietniekvārdi var apvienoties un papildināt cits citu, veidojot vārdu savienojumus, piem., *viss kas, mēs visi, mēs abi, es pats, nekas cits* u. tml., tomēr sintakses daļā vārdu savienojumu iekšējā struktūra un sintaktiskās funkcijas nav plašāk aprakstītas, minēts tikai tradicionālais dalījums vārdkopās un vārdrindās (Lokmane 2013b, 693).

Šobrīd jaunākie vārdkopas analoga skaidrojumi atrodami „Valodniecības pamatterminu skaidrojošajā vārdnīcā” (Skujiņa 2007, 435) un I. Lokmanes (2013a, 118) latviešu valodas sintakses aprakstā: „vārdkopu un vārdrindu analogi – vārdu savienojumi, kuros neveidojas ne pakārtojuma, ne sakārtojuma sakars, jo vienam vai vairākiem vārdiem (vietniekvārdiem, vispārīgas nozīmes apstākļa vārdiem) nav konkrētas leksiskas nozīmes, piem., *dažs labs, viens otrs, šāds tāds, šur tur, kaut kas labs, nekas neiespējams, šad tad*”. Sintaktiskais raksturojums balstīts apgalvojumā, kā šajos savienojumos trūkst, un to būtu vēlams precizēt, konkretizējot, vai sintaktiskā sakara šādos vārdu savienojumos nav vispār vai arī ir vēl kāds cits sintaktiskā saistījuma veids, kas jādefinē. Latviešu valodas sintaksē tiek šķirts pakārtojuma un sakārtojuma sakars vārdu savienojumos un teikumā, kā arī ir definēti trīs pakārtojuma formālie paveidi (saskaņojums, pārvaldījums un piekļāvums) atkarībā no tā, vai tiek realizēti kādi sintaktisko sakaru rādītāji, savukārt predikatīvās vienībās definēts predikatīvais, sekundāri predikatīvais un determinatīvais sakars (Lokmane 2013b, 700–703). Pētījuma autore pieļauj, ka vārdkopas analogu kontekstā šo klasifikāciju var nākties precizēt. Definējot šos sakarus, jāņem vērā arī Aksela Holvūta (*Axel Holvoet* 2012, 96–105) ieteikums neuzskatīt morfosintaktisko atkarību rādītājus (locījuma pārvaldījumu un saskaņojumu) par pašiem sintaktiskajiem sakariem un pārdomas par to, vai saskaņojums ir sintaktisko vai morfosintaktisko sakaru pazīme.

Kā redzams, vārdkopas analoga aprakstos minētās konstrukcijas fragmentāri pie-minētas jau iepriekš, bet nav plašāk aprakstītas un konsekventi sintaktiski raksturotas. Vairākkārt norādīts, ka šādas konstrukcijas veic viena teikuma locekļa funkciju jeb ir

salikts teikuma loceklis. Papildus latviešu sintakses teorijā minētas arī citas nedalāmas sintaktiskas vienības, kuru sintaktiskās struktūras izpratne nākotnē būtu jāprecizē. Daļai no šīm vienībām nedalāmība ir balstīta semantikā. Jautājums, cik lielā mērā jēdzieniskā nedalāmība jāņem vērā salikta teikuma locekļa statusa noteikšanā, paliek atklāts un jārisina turpmākos pētījumos. Raksta 2. nodaļā aprakstīts, kuri no teorijā minētajiem jēdzieniski nedalāmajiem vārdu savienojumiem ir iekļauti vārdkopas analoga grupā un kuri nav.

2. Vārdkopas analogu tipi „Latviešu valodas sintaktiski marķētajā korpusā”

LVTB gramatikas modelis paredz iespēju veidot dažādus saliktus teikuma locekļus, tai skaitā vārdkopas analogus (Rituma, Saulīte, Nešpore-Bērzkalne 2019, 209). Sākotnēji šajā grupā tika iekļauti visi vārdkopas analoga definīcijā (Skujiņa 2007, 435) minētajiem piemēriem līdzīgie vārdu savienojumi ar vietniekvārdiem vai vietniekvārdiem un adjektīviem, piem., *tas pats, mēs visi, tāds priecīgs*.

Pēc tam grupa paplašināta, kā kritēriju ņemot vērā arī jēdzienisko nedalāmību, tāpēc tajā tika iekļauti arī daži citi iepriekšējos pētījumos minētie nedalāmie vārdu savienojumi, tai skaitā prievārda konstrukcija ar *no*, piem., *kāds no mums*, un salīdzinājuma konstrukcija ar *tāds kā*, piem., *tāds kā vilies*. Tomēr jēdzieniskā nedalāmība ir tikai viena no pazīmēm, tāpēc šobrīd vārdkopas analoga grupā nav iekļauti šādi nedalāmi vārdu savienojumi:

- 1) vārdrindas analogi, t. i., tādi vārdu savienojumi, kas formāli atgādina vārdrindu, bet kuros starp komponentiem nav īsta sakārtojuma sakara un formālo sakārtojuma rādītāju (saikļi vai pieturzīmes), piem., *viens otrs, šurp turp* (Skujiņa 2007, 436); vārdrindas analogiem LVTB paredzēts cits frāzes veids;
- 2) tādi prievārda savienojumi kā *mēs ar tevi*, kur prievārda savienojums nosauc pavadoni vai līdzdalībnieku un starp komponentiem saskatāms sekundāri predikatīvs sakars;
- 3) salikti nosaukumi un frazeoloģismi, piem., *Latvijas Banka* un *pūst pīlītes*, kuru iekšējā struktūra atbilst pakārtotumam un kur ir skaidri nosakāms neatkarīgais komponents; šāds lēmums saistīts ar sintakses korpusa veidotāju mērķi pēc iespējas saglabāt sintaktiskās struktūras marķējumu atsevišķi no semantikas marķējuma.

Var diskutēt, vai tādas prievārda konstrukcijas, kas nosauc daļu no veselā (piem., *viens no viņiem*), kurās ir divi pilnnozīmes vārdi, būtu iekļaujamas vārdkopas analogu grupā, jo sintaktiski skaidri nosakāms neatkarīgais komponents (*viens*), kurš teikumā var iesaistīties dažādos locījumos, un atkarīgais komponents (*no viņiem*), tāpēc šo vārdu savienojumu var attēlot arī ar atkarību. Tomēr nav skaidra atkarīgā komponenta sintaktiskā loma, un arī semantiski šīs konstrukcijas veido jēdzieniski nedalāmu savienojumu, turklāt teorētiskajos avotos tās minētas kā salikti teikuma locekļi, tāpēc vārdkopas analogs šķita atbilstošākais no saliktu teikuma locekļu veidiem.

LVTB (korpusa apjoms – 16 803 teikumi) ir marķēti 772 vārdkopas analogi un šķirtas piecas vārdkopas analoga grupas (katrai grupai norādīti piemēri un to skaits korpusā). Četras no tām veido savstarpēji saskaņotu komponentu pāri – vietniekvārdu savienojumi, piem., *mēs visi, tāds pats* (310), konstrukcijas ar skaitļa vārdu, piem., *visi četri, pirmie trīs* (28), konstrukcijas ar īpašības vārdu, piem., *dažs labs, pats labākais* (87), un salīdzinājuma konstrukcijas, piem., *tāda kā stīga* (18). Piekto, visizplatītāko, grupu veido konstrukcijas ar prievārda savienojumu, kurās var ietilpt dažādu vārdšķiru vārdi, tomēr konstrukcijas uzbūve ir nemainīga, un tā iekļauta vārdkopas analoga grupā semantisku iemeslu dēļ, piem., *daudzi no mums, viens no labākajiem* (329). LVTB tiek veidots kā līdzsvarots korpus, tāpēc ir diezgan ticams, ka šādā proporcijā konstrukcijas parādās rakstītā valodā, tomēr sarunvalodā aina varētu būt citāda.

Šeit gan jānorāda, ka šobrīd vārdkopas analogi marķēti tikai gadījumos, kad teikumā nav lietvārda, kuru šis vārdu savienojums paskaidro un kuram katru vārdkopas analoga komponentu var atsevišķi pakārtot kā apzīmētāju. Tas saistīts arī ar korpusa veidotāju mēģinājumu sintaktiskajā marķēšanā vairāk norobežoties no semantikas. Tomēr, ja kā kritēriju ņem vērā jēdzienisko nedalāmību, tad vārdkopas analogs būtu jāmarķē arī šādos piemēros:

- (1) a. *Un man bija jākārto tie paši murgainie „ideoloģiskie” eksāmeni.*
- b. *Kuru katru brīdi tāda ragana [...] varēja parādīties [...]*
- c. *Meitene, galvu nepagriezot, apveltīja mani ar slīpu, tādu kā niknu skatienu [...]*
- d. *Balsošanas dienā pirmie trīs vēlētajī ieradušies jau ap pulksten septiņiem rītā [...]*

Šādi teikumi nākamajās LVTB versijās būtu jāpārskata un, iespējams, jāpārmarķē.

Vaicājumi, ar kuriem LVTB var atrast dažādu grupu vārdkopas analogus, redzami

1. tabulā.

Meklējamā vārdkopas analoga grupa	Vaicājums
Vietniekvārdu savienojumi	a-xinfo [xtype = "subrAnal", tag ~ '\\[vv'] ;
Konstrukcija ar skaitļa vārdu	a-xinfo [xtype = "subrAnal", tag ~ '\\[skv'] ;
Konstrukcija ar īpašības vārdu	a-xinfo [xtype = "subrAnal", tag ~ '\\[ipv'] ;
Salīdzinājuma konstrukcija	a-xinfo [xtype = "subrAnal", tag ~ '\\[sal'] ;
Konstrukcija ar prievārda savienojumu	a-xinfo [xtype = "subrAnal", tag ~ '\\[set'] ;

1. tabula. Vaicājumi vārdkopas analogu meklēšanai LVTB

Turpmāk apakšnodaļās aprakstīts katras vārdkopas analogu grupas gramatiskais un semantiskais aspekts, kā arī norādīts iespējamais sastāvs. Kur iespējams, sniegts komentārs par vārdkopas analoga komponentu secību, bet sīkāk šis aspekts LVTB un citos korpusos būtu pētāms atsevišķi.

2.1. Vietniekvārdu savienojumi

LVTB kā vietniekvārdu savienojumi (saliktā teikuma locekļa morfoloģiskais marķējums – [vv]) marķēti 310 vārdkopas analogi.

Noteikšanas kritēriji

1. Komponentu vārdšķira – vietniekvārdi:
 - PRON + PRON – *tas pats*
 - PRON + PRON + PRON – *tie visi citi*
2. Sintaktiskais raksturojums – pilnīgs vai daļējs saskaņojums, salikts teikuma loceklis, savstarpējais sintaktiskais saistījums nav definēts.
3. Semantiskais raksturojums – viens vietniekvārds precizē otru, bet ne vienmēr ir skaidrs, kurš kuru precizē.

Novietojums

Vietniekvārdi atrodas blakus, bet to savstarpējais novietojums šādos savienojumos ir atsevišķa pētījuma vērts. Dažādi secības varianti ne vienmēr konstatēti LVTB, un, lai izīmētu problēmu, vairākiem vietniekvārdu savienojumiem dažādi secības varianti meklēti LVK2018, tomēr arī tas, ka kāds no secības variantiem nav konstatēts LVK2018, nenozīmē, ka tāds nav iespējams.

Dažiem savienojumiem raksturīga noteikta komponentu secība, piem., savienojumos ar attieksmes vietniekvārdu *kurš*: *kuri abi, kuri paši, kuri visi* (sk. (10a–c) piemēru tālāk tekstā), jo šie vārdkopas analogi ievada palīgteikumu. Arī noteiktā un nenoteiktā vietniekvārda savienojumā *kurš katrs* (sk. (12) piemēru) tipiska ir tieši šāda secība. Tāpat personas vietniekvārdu un noteikto vietniekvārdu savienojumos personas vietniekvārds (sk. (2a), (3a) piemēru) parasti ir pirmais, tomēr LVK2018 parādās arī pretēja secība (sk. (2b), (3b) piemēru):

- (2)
 - a. – *Ko jūs visi tā raustāties no viņa?*
 - b. *Re, man dēls, viņš visus jūs apspēlēs [...] (LVK2018)*
- (3)
 - a. *Mēs visi tā darām.*
 - b. *Visi mēs redzēsim, kā tas notiek tālāk. (LVK2018)*

No latviešu valodas kultūras viedokļa iepriekš analizēti divi vietniekvārdu savienojumi. Daina Nītiņa (2013, 431) norāda, ka vietniekvārdu savienojums *katrs mēs* (sk. (4b) piemēru) varētu tikt lietots angļu valodas ietekmē, savukārt I. Freimane (1993, 245) savienojuma *tas viss* (sk. (5a) piemēru) nominatīvā iesaka nemainīt komponentu secību, īpaši runātā tekstā:

- (4)
 - a. – *Mūsu galvenā atziņa – mēs katrs dzīvojam pats par sevi.*
 - b. *Katrs mēs par otru varam spriest tikai pēc savas pieredzes. (LVK2018)*
- (5)
 - a. *Tas viss prasa daudz dažādu zināšanu.*
 - b. *Vispirms rodas vilšanās, viss tas ir muļķīgi, bezjēdzīgi [...]*

Tomēr gan šajos divos vietniekvārdu savienojumos (*mēs katrs* (4a) piemērā un *viss tas* (5b) piemērā), gan lielākajā daļā citu savienojumu komponenti var atrasties jebkurā secībā, un, lai arī par tipiskās secības noteikšanas rādītāju varētu izmantot lietojuma

biežumu, ne vienmēr ir skaidrs, kura secība būtu uzskatāma par tipisko, kura – par netipisko, piem., kādi būtu tipiskās secības noteikšanas kritēriji savienojumā *nekas cits* (sal. (6a) un (6b) piemēru):

- (6) a. *Nekas cits neeksistē.*
 b. *[..] cits nekas neatliek, kā vien likties uz auss.* (LVK2018)

Kā jau minēts, sīkāka šādu secības variantu nozīmes analīze ir atsevišķa pētījuma uzdevums, kas šobrīd ir vieglāk risināms tāpēc, ka ir pieejams sintaktiski marķēts korpus, kurā šīs konstrukcijas ir īpaši marķētas.

Vietniekvārdu savienojumu ir daudz, tāpēc pārskatāmībai sniegts sīkāks iedalījums pēc tā, kādas semantiskās grupas vietniekvārdi veido vārdu savienojumu (šajā rakstā izmantots tradicionālais latviešu valodas vietniekvārdu dalījums deviņās grupās atkarībā no semantikas un funkcijām, sk., piem., Kalnača, Lokmane 2021, 179–181). Iepriekš minēto iemeslu dēļ sīkāk nav aplūkots komponentu savstarpējais novietojums, bet shematiskais apzīmējums, piem., PRON PERS (*es*) + PRON DEF (*pats*), nozīmē, ka šajā grupā ietilpst gan *es pats*, gan *pats es*. Tālāk nav apskatīti vārdkopas analogi ar trim komponentiem, jo LVTB ir konstatēts tikai viens šāds piemērs:

(7) *Kāds labums tiem visiem citiem?*

1. PRON PERS + PRON DEF: *es pats, tu pats, viņš pats, mēs abi, jūs abi, viņi abi, mēs visi, jūs visi, viņi visi, mēs katrs*, piem.:

- (8) a. *Ērikam izklausījās, ka mācītājs sarunājas pats ar sevi un tikai viņam pašam saprotamā valodā.*
 b. *Reiz es pats dzirdēju, kā viņa sarunājās ar mirušajiem, un man kļuva baigi.*
 c. *[..] jo visu sarunas laiku bija domājis, kā lai iesāk runāt par viņiem abiem, un par mākslu runājis tīrā autopilotā.*
 d. *Mēs abi esam precēti, mēs abi negribam pamest savas ģimenes.*
 e. *Patī tu tagad tāda ragana.*

2. PRON DEM + PRON DEF: *šis pats, tas pats, šis viss, tas viss, šie abi, tie abi, šie visi, tie visi, tāds pats*:

- (9) a. *Lai ko Sofija nedarītu, meitene spogulī darīja to pašu.*
 b. *Tie visi apklusā, apsēdās.*
 c. *Augšpus grāvja strauts ir tāds pats kā arvien – dzidrš, pelēcīgs, sevī noslēpies.*
 d. *Kad viss šis notika un ļaudis bijās, tie no sākuma mierināja sevi un citus [..]*
 e. *Ēdu visu to, kas man garšo [..]*
 f. *Dziesma ir domāta visiem tiem, kas gremdējas atmiņās [..]*

3. PRON REL + PRON DEF: *kuri abi, kuri visi, kurš pats*:

- (10) a. *[..] ko saka divi Saeimas Prezidija locekļi, kuri abi pieder pie opozīcijas [..]*
 b. *[..] lai saglabātu iespēju izmēģināt receptes arī tiem, kuriem pašiem vairs vecvecāku [..]*
 c. *Tas pārsteidz mūsu kaujas biedrus, kuri visi ir paņēmuši līdzī lielas teltis un pat saliekamās gultiņas.*

4. PRON INDF + PRON INDF: *kāds cits, [kaut] kas cits, jebkas cits, jebkurš cits:*
- (11) a. – *Tad tu salauzīsi kādu kāju vai **ko citu**, – Cvībaks paraustīja plecus.*
 b. *Tā tad arī pasaules telpai vajadzēja būt reiz tapušai no **kaut kā cita**.*
 c. *Hendrikss iesaucis mūziķi par Rendiju Kaliforniju, lai atšķirtu no **kāda cita** Rendija [..]*
 d. *[..] vari novietot grāmatas, suvenīrus, garšvielu trauciņus vai **jebko citu**, kas ienāk prātā.*
 e. *[..] man zobi nesāp, bet man sāp **cits kaut kas**.*
5. PRON INDF + PRON DEF: *kurš katrs:*
- (12) *Mums jau ir izveidojies savs muzikālais rokraksts, ko **kurš katrs** nevarētu atskaņot.*
6. PRON NEG + PRON INDF: *nekas cits, neviens cits:*
- (13) a. *Un līdz pieciem **nekā cita** nebūs, es uz autoveikalu.*
 b. ***Neviens cits** mazulim vairs nav vajadzīgs.*
7. PRON INDF + PRON DEM: *kaut kas tāds.*
- (14) *Bet vai ir **kaut kas tāds**, kas interesē visus?*
8. PRON DEF + PRON INDF: *visi citi.*
- (15) ***Visi citi** labāk par viņiem zina, kas viņiem vajadzīgs.*
9. PRON NEG + PRON DEF: *neviens pats.*
- (16) *Bet visā biežajā telefonu grāmatā nebija **nevienu paša**, kas sauktos Mellers Knāgs.*
10. PRON PERS + PRON NEG: *mēs neviens.*
- (17) ***Mēs neviens** neesam perfekts, visi esam cilvēki, un kļūdiešies ir cilvēcīgi.*
11. PRON INT + PRON DEM: *kas tas.*

Latviešu valodniecībā parādās atšķirīgi viedokļi, kā sintaktiski analizēt jautājuma un izsaukuma teikumus, kuros ir konstrukcija ar vietniekvārdiem *kas* un *tas* un prievārda savienojums ar *par*:

- (18) a. ***Kas tā** par brūkleņu zapti?*
 b. ***Kas tā** par lietu?*

Analizējot vietniekvārda *kas* sintaktisko funkciju šajās konstrukcijās, I. Freimane (1985, 59) norāda, ka vietniekvārds *kas* kopā ar nomenu vai nomena aizstājēju (šajā gadījumā – vietniekvārdu *tas*) veic teikuma priekšmeta funkciju, bet gadījumos, kad nomena vai tā aizstājēja nav, *kas* viens pats veic teikuma priekšmeta funkciju. No tā noprotams, ka vārdu savienojums *kas tas* funkcionē kā salikts teikuma loceklis. Savukārt I. Lokmane (2013b, 731) uzskata, ka *kas* ietilpst izteicējā, bet vietniekvārds ar noteiktu referenci viens pats ir teikuma priekšmets, līdz ar to izslēdzot iespēju šo vārdu savienojumu uzskatīt par saliktu teikuma loekli.

LVTB šādi vietniekvārdu savienojumi konkrētajās konstrukcijās šobrīd marķēti kā vārdkopas analogi – salikti teikuma locekļi teikuma priekšmeta funkcijā, tomēr jāņem vērā, ka LVK2018 ir konstatēti arī tādi piemēri, kur līdzīgā konstrukcijā starp abiem vietniekvārdiem atrodas arī partikula, kas nav novērojams citās vārdkopas analogu grupās.

- (19) a. **Kas tad tie par jokiem?** (LVK2018)
 b. **Kas tad tas par līgavaini, varēji jau labāku atrast!** (LVK2018)
 c. **Kas gan tie par draugiem, ar kuriem nepavadi brīvdienas [..]** (LVK2018)

Piemēri ar partikulām rāda, ka šādu vārdu savienojumu iekļaušana vārdkopas analogu grupā vēl rūpīgi jāpārdomā un konstrukcija jāanalizē kopā ar citiem līdzīgiem *kas* un *tas* lietojumiem jautājuma vai izsaukuma teikumos.

Kopumā var secināt, ka vietniekvārdu savienojumos iesaistās visas vietniekvārdu leksiski semantiskās grupas, turklāt šādās grupās ir vieglāk analizēt komponentu secību. Lielākajai daļai vietniekvārdu savienojumu iespējama abējāda komponentu secība, bet dažiem – tikai viena. Konstrukcijas komponenti vienmēr saskaņojas locījuma un parasti arī dzimtē un skaitlī, izņemot vietniekvārdu savienojumus ar personas vietniekvārdu daudzskaitlī kā *mēs katrs*, *mēs neviens*. Semantiskas atšķirības, kas izriet no komponentu leksiski semantiskās grupas un komponentu secības, pētāmas turpmāk, tāpat jāprecizē sintaktiskais saistījums kontekstā ar pielikuma konstrukcijām, kuru sintaktiskā analīze raisa līdzīgas grūtības (Lokmane 2013b, 704).

2.2. Konstrukcija ar skaitļa vārdu daudzuma izteikšanai

LVTB kā konstrukcija ar skaitļa vārdu marķēti 28 vārdkopas analogi (saliktā teikuma locekļa morfoloģiskais marķējums – [skv]).

Noteikšanas kritēriji

1. Komponentu vārdšķira:
 PRON + NUM – *mēs trīs*, *visi četri*, *tā otra*, *es viens*
 ORD + CARD – *pirmie trīs*
 ADJ + NUM – *pārējie trīs*, *pēdējie divi*
2. Sintaktiskais raksturojums – saskaņojums, salikts teikuma loceklis, savstarpējais sintaktiskais saistījums nav definēts.
3. Semantiskais raksturojums – precizē ar vietniekvārdu / īpašības vārdu / kārtas skaitļa vārdu nosauktās kopas locekļu skaitu.

Novietojums

Komponenti atrodas blakus, pamata skaitļa vārds parasti ir otrais:

- (20) a. *Diezin, vai tā viena, jau iemūrētā, to visu spēsot pavilkt.*
 b. *Tā „Disnejlendā” pie mums trijiem pieskrēja maza meitenīte: „Es jūs atpazīnu!”*
 c. *Zemgals kolēģu apvainojumus atzina par pilnībā nepamatotiem un mēģināja mierīgi izrunāties ar pārējiem trim, bet tas neizdevās [..]*

Jāpiebilst, ka vārdkopas analogā *viens pats* sastopama tikai secība ar skaitļa vārdu kā pirmo komponentu:

- (21) a. [...] *kad mājdzīvnieks ir atstāts viens pats mājās.*
 b. *Auga ticīgā ģimenē, jau septiņu gadu vecumā viena pati gāja uz četrus kilometrus attālo baznīcu.*

Arī šajā grupā īpaši būtu analizējams tipiskajam pretējs komponentu novietojums – kādas sintaktiskās attieksmes veidojas, piem., LVK2018 konstatētajā kombinācijā *trīs pēdējie* un kāda ir šādas secības maiņas nozīme:

- (22) *Trīs pēdējie ir vienatnes darbi.* (LVK2018)

Šīs vārdkopas analogu grupas kontekstā īpaši aktuāls ir jautājums par jēdzieniskās nedalāmības kritēriju salikta teikuma locekļa marķēšanā, jo šajā grupā nereti vārdkopas analogs marķēts arī tad, ja teikumā ir lietvārds, kuram var tieši pakārtot abus konstrukcijas komponentus:

- (23) *Apšaudē tika ievainoti arī divi citi policisti un trīs no pieciem noziedzniekiem, kurus visus piecus drīz pēc tam izdevās aizturēt.*

2.3. Konstrukcijas ar īpašības vārdu

LVTB kā konstrukcijas ar īpašības vārdu ir marķēti 87 vārdkopas analogi (saliktā teikuma locekļa morfoloģiskais marķējums – [ipv]).

Noteikšanas kritēriji

1. Komponentu vārdšķira – vietniekvārds un īpašības vārds (vai adjektīvisks divdabis):
 PRON + ADJ / PTCP INDF – *tāds priecīgs, nekas prātīgs, dažs labs*
 PRON + ADJ / PTCP DEF – *nekāds siltais, pats labākais, abi pēdējie*
2. Sintaktiskais raksturojums – saskaņojums, salikts teikuma loceklis, savstarpējais sintaktiskais saistījums nav definēts.
3. Semantiskais raksturojums – teikuma semantiskajā struktūrā nozīmīgāka ir ar īpašības vārdu nosauktā īpašība. Semantiski nedalāms, jo nozīme izriet no vārdkopas analoga kā veseluma.

Novietojums

Komponenti novietoti blakus, īpašības vārds parasti otrais:

- (24) a. *Acīs griezās kaut kas spožs, tas bija lukturis, kura gaisma krita tieši viņa sejā.*
 b. *Ir grūti atskatīties uz maču un nepateikt ko rupju.*
 c. *Gados jaunāki, no skata abi tādi parasti, neuzkrītoši.*
 d. *Mēs, kam jau dažs labs gads uz pleciem [...]*
 e. *Ūdens vēl nevar būt nekāds siltais.*
 f. *Arī pensionāri un jaunās māmiņas visu nepieciešamo izstaigājot no rīta [...]*

Jāpiebilst, ka pēc noteikšanas kritērijiem šajā grupā iederas vārdu savienojumi ar vietniekvārdu *pats* un īpašības vārdu pārākajā pakāpē (sk. (25a) un (25b) piemēru). Tie tiek lietoti īpašības vārdu vispārākās pakāpes formu vietā un jēdzieniskās nedalāmības dēļ būtu marķējami kā salikti teikuma locekļi arī tad, ja teikumā ir lietvārds, kuram varētu tieši pakārtot katru no vārdiem (sk. (25c) piemēru):

- (25) a. *Nekur vairs nav bijis jāsteidzas – kad viss dienišķais jau aiz muguras, bet **pats galvenais** – priekšā.*
 b. *Tas laikam ir **pats svarīgākais** – dzīvot.*
 c. *Kāpēc visam jābūt tik grūti, ja cilvēks sācis domāt par **pašiem svarīgākajiem un pašiem vispārīgākajiem jautājumiem**?*

Šajā konstrukcijā vietniekvārda funkcijā var būt arī vietniekvārdu savienojums, t. i., vārdkopas analoga sastāvā ir cits vārdkopas analogs. (26) piemērā viens vārdkopas analogs ir *ko citu* (PRON + PRON), kas kā viena vienība iesaistās nākamajā vārdkopas analogā *ko citu pretlikumīgu* (PRON + ADJ / PTCP INDF):

- (26) *Lula [...] noliedz, ka būtu darījis **ko citu pretlikumīgu**.*

2.4. Salīdzinājuma konstrukcija

Līdzīgi konstrukcijām ar īpašības vārdu (sk. (27a) piemēru) funkcionē salīdzinājuma konstrukcija ar vietniekvārdu *tāds* (sk. (27b) piemēru), tāpēc šobrīd tā iekļauta vārdkopas analoģu grupā, tomēr šo konstrukciju sintaktiskā marķēšana jāskatās kontekstā ar pārējo salīdzinājuma konstrukciju marķējumu. Abu konstrukciju līdzīgums redzams piemēros:

- (27) a. *Gribējām jau celties kājās, bet kuldīdznieki **tādi atturīgi**...*
 b. *Bet meitenes acis, **tādas kā izbidētas**, raudzīja viņa skatienu notvert un sagūstīt.*

LVTB salīdzinājuma konstrukcijas vārdkopas analogi atrodami 18 reizes (saliktā teikuma locekļa morfoloģiskais marķējums – [sal]).

Noteikšanas kritēriji

1. Komponentu vārdšķira:
 PRON + PTCL (*kā*) + ADJ / PTCP – *tāds kā vīlies*
 PRON + PTCL (*kā*) + N – *tādas kā balsis*
2. Sintaktiskais raksturojums – saskaņojums, salikts teikuma locekļis, savstarpējais sintaktiskais saistījums nav definēts.
3. Semantiskais raksturojums – konstrukcijā nozīmīgākais ir īpašība vai nosauktā parādība, ar kuru salīdzina, vārdu savienojums *tāds kā* funkcionē līdzīgi kā partikula *it kā*, piešķirot šķietamības nozīmi.

Novietojums

Komponenti parasti atrodas blakus:

- (28) a. *Tēvs kādu brīdi apsvēra Ievas teikto un pēc tam tālu atliecās atpakaļ, skatīdamies uz viņu augstprātīgi un mazliet **tāds kā vīlies**, kā caur lupu.*

- b. *Dziļāk kafejnīcā, tādā kā nišā, bariņš trīsdesmitgadīgu vīriešu sīta zolīti.*
- c. *Starp mani un mīļoto ir radusies tāda kā atsvešināšanās.*
- d. *Esmu [...] pēkšņi uz līdzenas vietas baltas aprīļa dienas vidū iekritusi tādā kā straumē, un tagad mani nes.*

Raksta autore uzskata, ka šajās salīdzinājuma konstrukcijās vērojamas semantiskās struktūras pārmaiņas un nav izmantojama tradicionālā salīdzinājuma konstrukciju divdaļīgā analīze ar sekundāri predikatīvā komponenta lomas piešķiršanu otrajam komponentam (Lokmane 2013b, 756), drīzāk vārdu savienojums *tāds kā* šajās konstrukcijās ir leksikalizējies un kā vesela sintaktiskā vienība modificē adjektīvu vai lietvārdu, norādot uz nedrošību par nosauktās pazīmes vai reālijas atbilstību īstenībai, tomēr precīzāka semantiskā analīze ir veicama turpmāk kontekstā ar pārējām salīdzinājuma konstrukcijām un nav šī pētījuma mērķis.

2.5. Konstrukcija ar prievārda savienojumu

Prievārda konstrukcijas, kas nosauc daļu no veselā, gramatiski atšķiras no citām vārdkopas analoģu grupām, bet šīs grupas piemēri konstatēti visbiežāk – LVTB kā konstrukcija ar prievārda savienojumu marķēti 329 vārdkopas analoģi (saliktā teikuma locekļa morfoloģiskais marķējums – [set]).

Kritēriji

1. Komponentu vārdšķira:
 N + PREP (*no*) + N / PRON – *daļa no latviešiem, vairums no viņiem*
 PRON + PREP (*no*) + N / PRON – *kāds no kaimiņiem, daži no viņiem*
 ADJ + PREP (*no*) + N / PRON – *vienīgais no kolēģiem, sīkākais no tiem*
 NUM + PREP (*no*) + N / PRON / ADJ / NUM – *divas no personām, trīs no viņiem, vieni no pirmajiem, viens no adekvātākajiem*
2. Sintaktiskais raksturojums – skaidri nosakāms atkarīgais un neatkarīgais komponents, kas teikumā var iesaistīties dažādos locījumos atkarībā no sintaktiskās lomas, bet prievārda savienojuma forma nemainās. Prievārda konstrukcija nevar eksistēt bez neatkarīgā komponenta. Tomēr ar šo raksturojumu nepietiek, lai noteiktu sintaktiskā sakara veidu. Vietniekvārda, īpašības vārda un skaitļa vārda savienojumos ar prievārda konstrukciju neatkarīgais komponents saskaņojas dzimtē ar prievārda konstrukcijā iesaistīto nomenu.
3. Semantiskais raksturojums – visa konstrukcija izsaka vienu jēdzienu, nosauc daļu no kopuma.

Novietojums

Komponenti parasti atrodas blakus, prievārda konstrukcija ir otrā.

- (29) a. *Tas, ka šo grāmatu, kas domāta bērniem, lasa pieaugušie, ir viena no pasaules literatūras mīklām.*
 b. *Sakarsē pusi no kausētā sviesta.*

- c. *Sāku atcerēties, kuri no maniem draugiem arī kādreiz ir pazaudējuši drēbes.*
- d. *Nevienu no tām viņš nebija pircis.*

Atsevišķos gadījumos šīs grupas vārdkopas analogi sākas ar prievārdu, t. i., sastopamas atkāpes no tipiskās secības, piem.:

- (30) a. *No aptuveni 300 projektiem, ko esam realizējuši astoņu gadu laikā, ar valsts sektorubija saistīti kādi pieci [..]*
- b. *No tiem aptuveni divi kilometri ir melnā seguma brauktuves.*
- c. *[..] Krievijā pašlaik dzīvo aptuveni 25 miljoni ārvalstu pilsoņu, no kuriem ceturtdaļa izvairās no oficiālas reģistrācijas.*

Šai vārdkopas analoga grupai netipiskā komponentu secība (sk. Valkovska 2016, 112–114) ir iespējama dažādās konstrukcijās (sk. (30a–c) piemēru), bet biežāk ir vērojama konstrukcijās ar vietniekvārdiem prievārda savienojumā (sk. (30b), (30c) piemēru). (30a) piemērā tā skaidrojama ar teikuma aktuālo dalījumu, bet (30c) piemērā šāda secība saistīta arī ar attieksmes vietniekvārda *kurš* pakārtojuma vārda funkciju un novietojumu palīgteikuma sākumā un būtu uzskatāma par tipisku savienojumos ar attieksmes vietniekvārdiem. Tomēr jāpiebilst, ka LVTB ir konstatēts arī teikums, kur attieksmes vietniekvārds savienojumā ir otrajā, t. i., no valodas kultūras viedokļa nevēlamā, pozīcijā:

- (31) *Piemēram, Leonardo da Vinči izgudrojumi, daudzi no kuriem tika izmantoti vairākus gadsimtus pēc viņa nāves [..]*

Atkāpes no tipiskās vārdu secības konstrukcijās ar prievārda savienojumu ir pētāmas plašākos latviešu valodas teksta un runas korpusos.

Secinājumi

Pētījuma mērķis aprakstīt vārdkopas analogu grupu „Latviešu valodas sintaktiski marķētā korpusa” datus un noteikt šai grupai vienotus kritērijus ir sasniegts daļēji. Apraksti latviešu valodas pētījumos un gramatikās nav pietiekami precīzi, lai būtu skaidrs, kuri vārdu savienojumi atbilst šim jēdzienam un kuri ne. Konstrukcijā parasti ir iesaistīti nomeni, sintaktiskais saistījums starp vārdkopas analoga komponentiem nav definēts, ir norādīts, ka tas nav pakārtojums, un nav norādīts, vai būtiska pazīme ir saskaņojums. Savukārt semantiskais raksturojums, ka vārdkopas analogs nenosauc reālijas tieši, bet nosauc attieksmes starp tām, arī ir pārāk vispārīgs, lai skaidri izšķirtu, kādi jēdzieniski nedalāmi vārdu savienojumi būtu iekļaujami šajā grupā. LVTB modelī, veidojot saliktas teikuma vienības jeb frāzes, netiek definēts, kāds sintaktiskais saistījums ir starp to komponentiem, tomēr, ja nav salikta teikuma locekļa noteikšanas kritēriju, nav skaidrs, kādas ir jēdzieniskās nedalāmības robežas.

Pielāgojot ļoti plašu interpretāciju, korpusa datus iekļautas strukturāli un semantiski ļoti atšķirīgas vienības, tāpēc ir veidots sīkāks dalījums piecās apakšgrupās

pamatā pēc vārdkopas analoga sastāva. Kopīgais visām grupām ir tas, ka reālija netiek nosaukta tieši, bet ar jēdzieniskām attieksmēm starp konstrukcijas komponentiem. Sintaktiskā aspektā saskaņojums ir četrās grupās, bet konstrukcijā ar prievārda savienojumu saskaņojums ir tikai dzimtē, ja kāds no komponentiem ir vietniekvārds, skaitļa vārds vai adjektīvs. Nākotnē nepieciešams pārskatīt sintaktisko sakaru raksturojumu latviešu gramatikās, ņemot vērā pētījumā novērotās grūtības raksturot šeit apskatītās vārdkopas esošajā sistēmā.

Vieni no biežākajiem vārdkopas analogiem LVTB ir dažādu vietniekvārdu savienojumi, kas ir ļoti daudzveidīgi – tos veido vietniekvārdi no visām vietniekvārdu leksiski semantiskajām grupām. Atsevišķiem vietniekvārdu savienojumiem iespējama tikai noteikta komponentu secība, bet citiem ne, un šī atšķirība ir pētāma turpmāk – aplūkojama ne tikai secības nozīme, bet arī vietniekvārdu savienojumu kā gramatiskas vienības izpratne vietniekvārda raksturojumā, pieļaujot salikta vietniekvārda ieviešanu. Daži vietniekvārdu savienojumi var tikt izmantoti kā komponenti citu vārdkopas analogu konstrukciju veidošanā.

Ļoti izplatītas LVTB ir arī konstrukcijas ar prievārda savienojumu, kas nosauc daļu no kopuma. Tajā pašā laikā korpusa apjoms un tajā esošo vārdkopas analogu piemēru daudzums ir pietiekams, lai atklātu visos konstrukciju veidos lietoto leksēmu dažādību un izmantošanu dažādās sintaktiskajās funkcijās. Ar LVTB pašlaik nepietiek, lai analizētu dažādus komponentu secības aspektus vārdkopas analogos, tomēr konsekvēnti vārdkopas analogu noteikšanas kritēriji un vienveidīgs sintaktiskais marķējums ir labs sākumpunkts arī dažādu secības variantu analīzei.

Saīsinājumi un apzīmējumi

ADJ	īpašības vārds
CARD	pamata skaitļa vārds
DEF	noteiktais (īpašības vārds, vietniekvārds)
DEM	norādāmais (vietniekvārds)
INDF	nenoteiktais (īpašības vārds, vietniekvārds)
INT	jautājamais (vietniekvārds)
LVK2018	<i>Līdzsvarotais mūsdienu latviešu valodas korpus</i>
LVTB	<i>Latviešu valodas sintaktiski marķētais korpus</i>
N	lietvārds
NEG	noliegtais (vietniekvārds)
NUM	skaitļa vārds
ORD	kārtas skaitļa vārds
PERS	personas (vietniekvārds)
PREP	prievārds
PRON	vietniekvārds
PTCL	partikula
PTCP	divdabis
REL	attieksmes (vietniekvārds)
subrAnal	vārdkopas analogs

Avoti

- Levāne-Petrova, Kristīne, Dargis, Roberts. 2018. *Līdzsvarotais mūsdienu latviešu valodas korpuss (LVK2018)*. CLARIN-LV digitālā bibliotēka. Latvijas Universitātes Matemātikas un informātikas institūts. Pieejams: <http://hdl.handle.net/20.500.12574/11>
- Rituma, Laura et al. 2022. *Latviešu valodas sintaktiski marķētais korpuss (LVTB)*, v2.10. CLARIN-LV digitālā bibliotēka. Latvijas Universitātes Matemātikas un informātikas institūts. Pieejams: <http://hdl.handle.net/20.500.12574/63>

Literatūra

- Ahero, Antonija et al. 1959. *Mūsdienu latviešu literārās valodas gramatika*. I. Rīga: LPSR Zinātņu akadēmijas izdevniecība.
- Ahero, Antonija et al. 1962. *Mūsdienu latviešu literārās valodas gramatika*. II. Rīga: LPSR Zinātņu akadēmijas izdevniecība.
- Freimane, Inta. 1985. *Vienkāršs teikums un tā paplašināšana*. Rīga: Latvijas Valsts universitāte.
- Freimane, Inta. 1993. *Valodas kultūra teorētiskā skatījumā*. Rīga: Zvaigzne.
- Freimane, Inta. 1997. Sintakses vienības latviešu valodā. *Baltu filoloģija*. VII, 6–20.
- Holvūts, Aksels. 2012. *Vispārīgās sintakses pamati*. Sirmule, Lāsma (tulk.). Rīga: Latviešu valodas aģentūra.
- Kalnača, Andra, Lokmane, Ilze. 2021. *Latvian Grammar*. Rīga: University of Latvia Press. <https://doi.org/10.22364/latgram.2021>
- Kārklīšs, Jūlijs. 1976. Vārda formas sintaktisko potenci problēma. *Valodniecības un literatūrzinātnes teorijas un prakses jautājumi*. Mieze, Marija (red.). Rīga: Zvaigzne, 54–101.
- Lokmane, Ilze. 2013a. Sintakse. *Latviešu valoda*. Veisbergs, Andrejs (red.). Rīga: Latvijas Universitātes Akadēmiskais apgāds, 109–132.
- Lokmane, Ilze. 2013b. Vienkārša teikuma formālā (strukturālā) organizācija. *Latviešu valodas gramatika*. Nītiņa, Daina, Grigorjevs, Juris (red.). Rīga: Latvijas Universitātes Latviešu valodas institūts, 710–766.
- Nītiņa, Daina. 2013. Vietniekvārds (pronomens). *Latviešu valodas gramatika*. Nītiņa, Daina, Grigorjevs, Juris (red.). Rīga: Latvijas Universitātes Latviešu valodas institūts, 428–452.
- Ozols, Arturs. 1959. Raksturīgākās latviešu tautasdziesmu nominālās vārdkopas. *Pētera Stučkas Latvijas Valsts universitātes zinātniskie raksti*. 30, 7–54.
- Pretkalniņa, Lauma. 2023. *Formāls latviešu valodas gramatikas modelis un tā realizācija mašīnlasāmā sintakses korpūsā*. Promocijas darba kopsavilkums. Rīga: Latvijas Universitāte.
- Rituma, Laura, Saulīte, Baiba, Nešpore-Bērzkalne, Gunta. 2019. Latviešu valodas sintaktiski marķētā korpusa gramatikas modelis. *Valoda: nozīme un forma*. 10, 200–216. <https://doi.org/10.22364/vnf.10.16>
- Skujiņa, Valentīna (red.). 2007. *Valodniecības pamatterminu skaidrojošā vārdnīca*. Rīga: Latvijas Universitātes Latviešu valodas institūts.
- Smiltiece, Gunta. 2013. Apstākļa vārds (adverbs). *Latviešu valodas gramatika*. Nītiņa, Daina, Grigorjevs, Juris (red.). Rīga: Latvijas Universitātes Latviešu valodas institūts, 595–618.
- Valdmanis, Jānis. 1989. Teikums. *Latviešu valodas sintakse*. Ceplītis, Laimdots, Rozenbergs, Jānis, Valdmanis, Jānis (red.). Rīga: Zvaigzne, 13–104.

- Valkovska, Baiba. 2016. *Vārdu secības komunikatīvais aspekts mūsdienu latviešu valodā*. Promocijas darbs. Rīga: Latvijas Universitāte.
- Zolotova, Galina. 1973. *Očerki funkcional'nogo sintaksisa russkogo jazyka*. Moskva: Nauka.

Summary

The Institute of Mathematics and Computer Science, University of Latvia, has developed Latvian Treebank – a syntactically annotated corpus of the Latvian language. It showcases both the syntactical phenomena of the Latvian language already described in the Latvian syntax theory, as well as the less common structures that have not yet been analysed in depth by theory. This paper considers understanding of the subordinate phrase analogue in Latvian linguistics. The paper provides an overview of the constructions labelled as such in Latvian Treebank. The principles of a further classification of the subordinate phrase analogues by certain grammatical criteria (morphological composition and syntactical characteristics) and semantic criteria are proposed. Some analysis of the order of components in these compound constituents is also provided. The two most common types of the subordinate phrase analogues in Latvian Treebank are groups of pronouns, as well as constructions with prepositional phrases. Other types appear less frequently. The amount of treebank examples is sufficient to cover the lexeme variety used for each type, as well as the range of syntactical functions fulfilled by each type. Hereby developed criteria for identifying the analogues of subordinate phrase and consistent syntactical annotation provide a convenient language material for further analysis of these constructions.

Keywords: Latvian syntax; syntactically annotated corpus; subordinate phrase analogue; compound part of sentence.



Rakstam ir Creative Commons Attiecinājuma 4.0 Starptautiskā licence (CC BY 4.0) /

This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>)