Laura Rituma, Baiba Saulīte, Lauma Pretkalniņa
Mākslīgā intelekta laboratorija,
LU MII

Latviešu valodas sintaktiski marķētā korpusa marķēšanas metodikas apraksts

Saturs

Ievads 3
1. Saikņu veidi 4
1.1. Atkarības (pakārtojums) 4
1.1.1. Izteicējs (*pred*) 4
1.1.2. Teikuma priekšmets (*subj)* 5
1.1.3. Papildinātājs (*obj*) 6
1.1.4. Apzīmētājs (*attr*) 6
1.1.5. Apstāklis (*adv*) 7
1.1.6. Sekundāri predikatīvs komponents (*spc*) 7
1.1.7. Situants (*sit*) 8
1.1.8. Determinants (*det*) 8
1.1.9. Iespraudums (*ins*) 8
1.1.10. Negramatisks atkārtojums (*repeat*) 8
1.2. Teikuma elementa sadalīšana sastāvdaļās 9
1.2.1. X-vārdi 10
1.2.1.1 Vārdu savienojums ar prievārdu (*xPrep*) 11
1.2.1.2 Salikts izteicējs (*xPred*) 12
1.2.1.3 Skaitļa vārdi (*xNum*) 13
1.2.1.4 Pielikumi (*xApp*) 13
1.2.1.5 Salīdzinājums (*xSimile*) 14
1.2.1.6 Vārdu savienojums ar partikulu (*xParticle*) 15
1.2.1.7 Palīgvārdu savienojums (*xFunctor*) 15
1.2.1.8 Vārdkopas analogs (*subrAnal*) 15
1.2.1.9 Vārdrindas analogs (*coordAnal*) 16
1.2.1.10 Nosaukums (*namedEnt*) 16
1.2.1.11 Frazeoloģiskās vienības (*phrasElem*) 16
1.2.1.12 Nestrukturētas vienības (*unstruct*) 16
1.2.2. Sakārtojums 17
1.2.2.1 Vienlīdzīgi teikuma locekļi (*crdParts*) 17
1.2.2.2 Ar sakārtojumu saistītas teikuma daļas (*crdClauses*) 17
1.2.3. PMC (pieturzīmju piesaiste kokam) 18
1.2.3.1 Teikums (*sent*) 19
1.2.3.2 Izteikums (*utter*) 19
1.2.3.3 Izsauksmes vārds (*interj*) 20
1.2.3.4 Patstāvīga teikuma daļa (*mainCl*) 20
1.2.3.5 Palīgteikums (*subrCl*) 21
1.2.3.6 Sekundāri predikatīva komponenta (SPK) pieturzīmes (*spcPmc*) 23
1.2.3.7 Iespraudumi un iestarpinājumi (*ins*) 24
1.2.3.8 Uzruna (*address*) 26
1.2.3.9 Tiešā runa (*dirSpPmc*) 27
1.2.3.10 Citas pēdiņas (*quot*) 31
1.2.3.11 Partikulu konstrukcijas (*particle*) 31
2. Redukcija 31
2.1. Specifiskie gadījumi 33
2.2. Negramatisku teikumu marķēšana 35
3. Specifiskie gadījumi 35

3.1. Konstrukcijas, kas attēlojamas ar vienu virsotni 35
3.2. Vairākvārdu saikļi un partikulas 36
3.3. Partikulu marķēšana 36
3.4. Sintaktisko lomu izvēle dažādos gadījumos 36
4. Virsotņu marķējums 37
4.1. x-vārdu morfoloģiskā marķēšana 40
Versijas 41

Ievads

Latviešu valodas sintaktiski marķētais korpuss (LVTB) tiek veidots, par teikuma pamata modeli izmantojot atkarību un frāžu modeļa hibrīdu. LVTB sastāv no 3 slāņiem, kas katrs satur noteikta veida marķējumu:

1. vārdu slānis — sadala tekstu rindkopās (*paragraph*) un teksta primitīvos (*token*),
2. morfoloģijas slānis — satur morfoloģisko marķējumu,
3. analītiskais slānis — satur sintaktisko marķējumu.

Katrs slānis ir neatkarīgs no slāņiem, kas atrodas „virs" viņa (piemēram, lai izmantotu 2. slānī esošo morfoloģisko informāciju, nav nepieciešams 3. slānis).

Visiem slāņiem ir XML formāti, 1. un 2. slāņa formāti ir savietojami ar atbilstošajiem *Prague Dependency Treebank* (PDT) slāņu formātiem. 3. slānis ir radīts balstoties uz PDT analītiskā slāņa pieredzi, papildinot to ar latviešu valodā nepieciešamajām konstrukcijām — analītiskajām formām, atšķirīgiem sakārtojuma attieksmju attēlojumiem u.c. 1.slāņa faila nosaukuma paplašinājums parasti ir „.w", otrā — „.m", trešā — „.a".

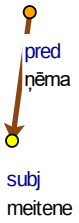
Šis dokuments satur izklāstu par to, kā LVTB jāmarķē dažādas latviešu valodas sintaktiskās konstrukcijas. Dokuments tiek papildināts.

Dokumenta 1. nodaļā ir aprakstīta dažādu sintaktisko saikņu attēlošana LVTB, 3. nodaļā ir apskatīti dažādi speciālgadījumi (ieskaitot redukciju, partikulu marķēšanu un dažas konstrukcijas, kuras *netiek* apstrādātas LVTB 3. slānī, piemēram, vairākvārdu saikļi), 4. nodaļā ir paskaidrots virsotņu marķējums.

1. Saikņu veidi

1.1. Atkarības (pakārtojums)

Atkarības attēlo pakārtojuma attieksmes starp teikuma elementiem.



1–1. attēls. „**meitene ņēma**”

TrEd-ā atkarības attēlo ar brūnu bultu. Pakārtotais loceklis (*dependant*) atrodas zemāk, bet loceklis, no kura pakārtotais ir atkarīgs (*dependency head*), — augstāk (sk. 1–1. attēlu).

Šobrīd sintakses kokos atkarības tiek izmantotas šādu teikuma locekļu piesaistei:

Nosaukums	**Angliski**	**Kam pakārtots (parasti)**
teikuma priekšmets	*subject*	izteicējs *
papildinātājs	*object*	izteicējs *
apzīmētājs	*attribute*	lietvārds
sekundāri predikatīvais komponents (spk)	*semi-predicative component*	izteicējs, teikuma priekšmets, cits spk u.c.
apstāklis	*adverbial modifier*	izteicējs, apstāklis, spk
situants	*situant*	teikuma daļa
determinants	*determinant*	teikuma daļa

* Teikuma daļas izteicējs (*predicate*) atrodas teikuma daļai atbilstošā apakškoka saknē — t.i., tas nav ne no kā atkarīgs (sk. 1.2.3.1. un 1.2.3.4. nodaļu).

Tāpat atkarības izmanto palīgteikumu piesaistei pie elementa, kuru tas paskaidro (sk. 1.2.3.5.), kā arī tiešo runu (sk. 1.2.3.9.) un iespraudumu un iestarpinājumu (sk. 1.2.3.7) piesaistei.

Atkarības izmanto arī, lai piesaistītu partikulas ar brīvu novietojumu teikumā pie vārda, uz kuru tās attiecas, piemēram, **arī Jānis**. Šajā gadījumā lieto lomu **no**. Partikulu attēlošana kokā detalizēti izskaidrota 3.3. sadaļā.

Ar atkarību tiek attēloti arī tādi teikuma elementi, kas attiecas uz visu teikumu un nav pakārtoti vienam konkrētam teikuma loceklim, piemēram, ārpusshēmas komponenti situants un determinants, pamatojuma palīgteikums, dažādi iesparudumi. Šādus visam teikumam pakārtotus elementus piesaista teikuma saknei, bet gadījumos, kad tas attiecas uz vairākām vienlīdzīgām teikuma daļām, to pakārto tuvākai apaļai virsotnei virs abām teikuma daļām (skatīt **1**–**7**. attēlu).

Turpmākās apakšnodaļās sīkāk aprakstīti teikuma locekļi, kas tiek marķēti.

1.1.1. Izteicējs (*pred*)

Izteicējs nosauc darbību, stāvokli vai pazīmi. Izteicējs atrodas teikuma daļai atbilstošā apakškoka saknē. Vienkāršs izteicējs ir verba finītā forma, kā arī dažos palīgteikumu veidos - nenoteiksme. *Bērns spēlējās ar klucīšiem. Kā viņa zina, kur mūs meklēt?* Nenoteiksme tiek marķēta kā izteicējs arī teikumos ar pavēles nozīmi un tēmas infinitīva nozīmi: *Visiem piecelties! Pagaidīt.*

Rakstīt? Nenoteiksme tiek marķēta kā izteicējs arī tad, kad tai piemīt kāda modalitāte un to teorētiski varētu marķēt par saliktu izteicēju ar modificētāju: *Kaut reizi mūžā redzēt Spāniju! Kāpēc mocīties par minimālo algu? Virtuvē neiztikt bez plīts.* Salikts izteicējs veido x-vārdu **xPred** (skatīt 1.2.1.2 nodaļu).

Piezīmes:

- vārdu savienojumos, ko veido vietniekvārds *kas* kādā locījumā vai adverbi *kur, kad, kurp* kopā ar infinitīvu, piemēram, ***Ir ko darīt. Ilgāk nav ko ņemties. Nav kad strādāt.*** , izteicēja loma **pred** tiek piešķirta verba "būt"formai patstāvīgā funkcijā, savukārt verbam nenoteiksmē tiek spiešķirta teikuma priekšmeta loma **subj.**

1.1.2. Teikuma priekšmets (*subj)*

Teikuma priekšmets tiek apzīmēts ar lomu **subj** un ir atkarīgs no izteicēja (**pred**)**.** Teikuma priekšmets nosauc runas priekšmetu, kura darbību, stāvokli vai pazīmi raksturo izteicējs, piemēram, *Jānis iet, grāmata ir interesanta.*

Ja teikumā ir salikts izteicējs, tad virsotni ar lomu **subj** pakārto tikai virsotnei **pred,** nevis kādai atsevišķai frāzes **xPred** daļai.

Parasti teikuma priekšmets ir lietvārds vai vietniekvārds nominatīvā, bet iespējami arī citi gadījumi. Vēl teikuma priekšmets var būt:

- lietvārds vai vietniekvārds ģenitīvā: 1) pie verba *nebūt, trūkt* finītās formas: *nav spēka, trūkst prāta,*2) ja lietvārds saistīts ar vārdu, kuram ir nenoteikta daudzuma nozīme, piemēram, *maz, vairāk,* vai ar nelokāmu skaitļa vārdu, kas saista pārvaldījumā vārdu ģenitīvā: *šodien atnāca vairāk klientu, tūkstoš monētu tika iebērtas maisā*.
- lietvārds vai vietniekvārds datīvā: 1) ja izteicējs ir vajadzības izteiksmē *man jāsteidzas, tam jābūt kopējam darbam,* 2) ja izteicējs ir salikts izteicējs ar modificētāju, kas piešķir nejaušības nozīmi: *Mums izdevās pastrādāt, man laimējās satikt tevi;* 3) ja izteicējs ir nenoteiksmē: *visiem nostāties mierā!* 4) pakārtots divdabim ar lomu **basElem** absolūtā datīva konstrukcijā: *Sākoties pieaugušas sievietes dzīvei, tā īpašība pazuda.*
- īpašības vārds, ja tas lietots kā īpašvārds: *Čaklais devies uz ārzemēm*;
- substantivējies īpašības vārds vai divdabis: *vecais sēdēja pie krāsns, neticamais bija noticis*;
- verba nenoteiksme: *dzīvot ir skaisti, pētīt nozīmē iedziļināties, pamosties ir veiksme, smēķēt ir aizliegts*;
- verba nenoteiksme - vārdu savienojumos, ko veido vietniekvārds *kas* kādā locījumā vai adverbi *kur, kad, kurp* kopā ar infinitīvu, piemēram, ***Ir ko darīt. Ilgāk nav ko ņemties. Nav kad strādāt.***, izteicēja loma **pred** tiek piešķirta verba "būt"formai patstāvīgā funkcijā, savukārt verbam nenoteiksmē tiek spiešķirta teikuma priekšmeta loma **subj.**

- nedalāms vārdu savienojums, kas veido x-vārdu – 1) pielikuma konstrukcija **xApp**: *dzejnieks Rainis*, 2) nosaukums **namedEntity**: *Jānis Bērziņš, Tēvzemei un Brīvībai*, 3) vārdkopas analogs **subrAnal**: *viņi abi, es pats.*

1.1.3. Papildinātājs (*obj*)

Papildinātājs ir teikuma loceklis, kas nosauc darbības tiešo vai netiešo objektu un ir pakārtots darbības vārdam, retāk lietvārdam vai īpašības vārdam. Papildinātājs ir izteikts ar lietvārda atkarīgo locījumu vai prievārdisku savienojumu (xPrep), piemēram, *Vecmāmiņa ada zeķi. Viņš priecājas par panākumiem. Zēns izbrīnīts skatījās uz grāmatu.* Papildinātāju apzīmē ar lomu **obj**, un parasti ir pakārtots izteicējam **pred.**

Piezīmes:

- kā papildinātājs **obj** tiek marķēta prievārdiska konstrukcija pie lietvārda lokatīvā šādos vārdu savienojumos: *saskaņā ar .., atkarībā no .., mierā ar .. , saistībā ar* u.tml.
- kā papildinātājs **obj** tiek marķēta prievārdiska konstrukcija pie adverba šādos vārdu savienojumos: *kopā ar .., līdz ar ..*
- kā papildinātājs **obj** tiek marķēts lietvārds datīvā pie adverba vai prievārdiska vārdu savienojuma ar adverbiālu nozīmi šādos un līdzīgos gadījumos: *paralēli līnijai, atsbilstoši laikam, par spīti tam, par godu gadadienai, speciāli gadījumam, tuvāk brīvībai, stāstam pamatā, galā caurulei, viņai zem kājām* u.c.
- ja papildinātājs saistīts ar saliktu izteicēju, tad parasti tas ir pakārtots visam xPred, bet, ja izteicējs veidots ar modificētāju **mod,** tad papildinātājs tiek pakārtots izteicēja bāzes elementam **xPred** konstrukcijā.
- prievārdiskas konstrukcijas pie deverbāliem un citiem lietvārdiem, kuras nosauc darbības objektu, tiek marķētas kā apzīmētājs, kas atkarīgs no minētā lietvārda, **(attr):** *šķiešanās ar naudu, runa par procesiem, sapnis par nākotni.*
- nomens datīvā ar nolūka nozīmi pie verba tiek marķēts kā papildinātājs (**obj**), bet pie deverbāla vai cita lietvārda, kurš nosauc darbību, - kā apzīmētājs (**attr**) *nodoms nākotnei*.

1.1.4. Apzīmētājs (*attr*)

Piezīmes:

- Kā apzīmētāji **attr** tiek marķēti daži noteikti apstākļa vārdi, kas pakārtoti lietvārdam un nosaka lietvārdam ģenitīva locījumu – *daudz, vairāk, visvairāk, maz, mazāk, vismazāk, cik, tik, necik, mazliet.*
- Kā apzīmētāji **attr** tiek marķēti apstākļa vārdi, kas pakārtoti lietvārdam vai vietniekvārdam un neietekmē lietvārda locījumu – *kopā būšana, blakus istaba, pilnīgi viss, šurpu turpu gājēja.*

- prievārdiskas konstrukcijas pie deverbāliem un citiem lietvārdiem, kuras nosauc darbības objektu, tiek marķētas kā apzīmētājs, kas atkarīgs no minētā lietvārda (**attr):** *šķiešanās ar naudu, runa par procesiem, sapnis par nākotni.*

1.1.5. Apstāklis (*adv*)

Piezīmes:

- par apstākli marķē lietvārdu akuzatīvā ar laika mērvienības nozīmi, kas a) nosauc darbības ilgumu: *process ilgst divas **stundas;*** b) precizē citu laika nozīmes apstākļa vārdu un ir pakārtots tam: *dažas **nedēļas** atpakaļ.*

1.1.6. Sekundāri predikatīvs komponents (*spc*)

Piezīmes:

- **spc** centrā nevar būt verba finītā (arī ne atstāstījuma, vēlējuma un vajadzības izt.) forma.
- **spc** lomu piešķir lietvārda bērnam ar apstākļa nozīmi - nomeniem lokatīvā vai prievārdiskiem savienojumiem: *māja mežā, vakars uz ezera* Tas attiecas arī uz atgriezeniskiem lietvārdiem ar izskaņu –šana, -šanās: *slīkšana jūrā, atgriešanās pēc 15 gadiem* .
- Nosaukšana vārdā veido konstrukciju ar **spc:** *Daiga* (**spc**) *viņu* (**obj**) *sauc* (**pred**). (Neitrāla vārdu secība būtu *Viņu sauc Daiga.*) Atšķirībā no konstrukcijas, kura izsaka kāda cilvēka pasaukšanu, piemēram, *Māte sauc viņu pusdienās,* kur lietvārds vai vietniekvārds nominatīvā veic teikuma priekšmeta funkciju, pirmajā piemērā lietvārds *Daiga* nominatīvā veic sekundāri predikatīvā komponenta **spc** funkciju un ir atkarīgs no izteicēja *sauc.* Tāpat gadījumos, kad nosaukšanas veids (vārds) norādīts prievārda konstrukcijā, piemēram, *Viņu sauc par Daigu*, prievārda konstrukcija veic sekundāri predikatīva komponenta funkciju un ir atkarīga no izteicēja.
- Nelokāmie divdabji, īpaši divdabis ar –am, ir pieobjekta sekundāri predikatīvi komponenti – divdabja nosauktās darbības darītājs veic objekta funkciju teikumā. Šajos gadījumos **spc** ir atkarīgs no tā verba, no kura atkarīgs objekts, nevis objekta: *Redzēt* (**subj**) *Šakiru* **(obj)** *uzstājamies* **(spc)** *uz skatuves ir neaizmirstams piedzīvojums.* (spc pakārtots teikuma priekšmetam "redzēt".) *Olivers juta lēnu, bet aukstu cietsirdības vilni* (**obj)** *paceļamies* (**spc**) *savas būtnes attālākajās malās un skalojamies uz kodolu.*
- **spc** lomu piešķir prievārdiskam savienojumam ar prievārdu **par** pie verbiem, kuru nozīme atbilst "darīt, lai kaut kas kļūst par kaut ko" vai "uzskatīt kaut ko par kaut ko" – *norīkot, iekārtot, ievēlēt, atzīt, uzskatīt, nosaukt: iecēla par direktoru, uzskatīja par gudru.*
- **"viens otr." un "cits cit."**, arī prievārdiskos gadījumos "cits ar citu", "viens pēc otra" tiek marķēti šādi:
    - pirmais vārds nominatīvā, kas tiek saskaņots ar teikuma priekšmetu dzimtē un locījumā, tiek marķēts ar lomu **spc** un tiek zīmēts kā atkarīgais no otrā vārda, kas tiek

locīts atbilstoši teikuma struktūrai (attiecīgi - *otrs* vai *cits*), vai prievārdiskas frāzes, ja otrais vārds tādā iesaistīts (*viens ar otru*);

- to dara arī gadījumā, ja vārdi nav blakus (*uzņēmēji viens informē par otru*) vai ja frāzes otrais vārds iesaistīts lietvārda frāzē (*viņi žņaudzīja viens otra rokas*).

1.1.7. Situants (*sit*)

Situants ir apstāklis, kas iesaistās teikumā neatkarīgi no vārda sakariem, tas ir, nav vārdkopu attieksmēs ne ar vienu vārdu teikumā. Parasti tas attiecas uz visu teikumu kopumā un atrodas teikuma sākumā. Tā kā latviešu valodas sintakses teorijā nav precīzi aprakstīts, kā atšķirt situantu no tāda apstākļa verbālos teikumos, kas novietots teikuma sākumā, kā arī citi kritēriji nav precīzi noteikti, sintaktiski marķētajā korpusā situants tiek marķēts tikai nepārprotamos gadījumos, kad kāds situāciju raksturojošs apstāklis ir attiecināts uz vairākām teikuma daļām vai arī pēc nozīmes tas nepārprotami attiecas uz visu teikumu. Kategorija ir izplūdusi, un korpusā tā nav marķēta viennozīmīgi.

1.1.8. Determinants (*det*)

Determinants ir teikuma paplašinātājs datīvā ar izjutēja vai īpašnieka nozīmi. Tiek uzskatīts, ka determinants ir brīvs teikuma loceklis, kas attiecas uz visu teikuma struktūru, tāpēc korpusā tas tiek marķēts kā no visa teikuma vai teikuma daļas atkarīga visotne. *Man pieder māja*.

Piezīmes:

- Determinanta lomu **det** piešķir teikuma loceklim datīvā, kas nosauc situācijas adresātu/izjutēju pie atļaušanas un aizliegšanas nozīmes verbiem: **“Meitenei ir atļauts..”**, **“Automašīnām ir aizliegts atrasties ..”**, u.tml.

1.1.9. Iespraudums (*ins*)

Atsevišķos gadījumos iesprauduma loma **ins** tiek izmantota, lai marķētu teksta iezīmētājus, kas pilda iesprauduma funkciju, bet pēc kuriem latviešu gramatiskā tradīcijā komatu neliek – paskaidrojošo vārdu grupu ievadītājvārdi *galvenokārt, (it) sevišķi, (it) īpaši, tai skaitā, to starpā, tostarp, tā sauktais, tā saucamais, u.c.* Ievadītājvārds tiek zīmēts kā atkarīgais no paskaidrojošā vārda vai visas vārdu grupas.

Piezīmes:

- iesprauduma vai paskaidrojošās vārdu grupas ievadītājvārdi *līdzīgi kā* arī pilda **ins** lomu, tikai sastāv no salīdzinājuma frāzes.

1.1.10. Negramatisks atkārtojums (*repeat*)

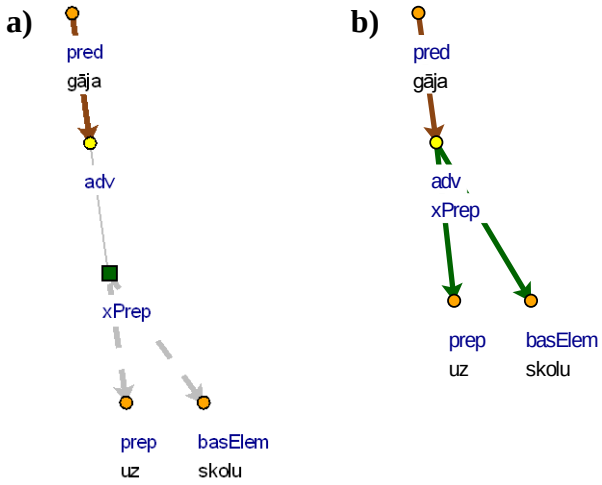
Teikumos *Viņam bija bija auksti* un *Zālē bija aizņemtas bija gandrīz visas vietas* negramatiski atkārtoto vārdu iesaista teikumā ar to lomu, kāda pēc struktūras paredzēta, savukārt pirmo minējumu, kurš tika atkārtots, apzīmē ar lomu **repeat** un pakārto tam vārdam, kas ir ticis atkārtots. Pat, ja latviešu valodā reizēm par negramatisku tiek uzskatīts otrreiz lietotais vārds, kurš teikumā attiecīgi vairs nav

nepieciešams, datu transformācijai uz Universālajām atkarībām (Universal Dependencies) ir nepieciešams, lai teikuma struktūrā vienmēr būtu iesaistīts otrais vārds un tam būtu pakārtots pirmais pieminējumu.

1.2. Teikuma elementa sadalīšana sastāvdaļās

No strukturālā viedokļa sintaktiskā koka uzbūves pamatelements ir atkarība (pakārtojuma attieksme). Katram teksta primitīvam (vārdam vai pieturzīmei) atbilst viena koka virsotne (TrEd-ā šīs virsotnes tiek apzīmētas ar aplīšiem), un atkarība ir parastā saikne, kas mēdz savienot 2 koka virsotnes.

Situācija kļūst nedaudz sarežģītāka, ja vienā atkarības galā jānovieto nevis viens vārds, bet vairāku vārdu komplekts, kas izpilda kopēju funkciju (piemēram, fragmentā „**gāja uz skolu**" jāparāda, ka **uz skolu** ir pakārtots izteicējam **gāja**). Tad tiek ieviesta virsotne, kas apvieno šos vārdus kopā vienā saliktā vienībā (TrEd-ā šīs virsotnes no pārējām var atšķirt pēc kvadrātveida formas). Šādai virsotnei pievienotās virsotnes ir saliktās vienības sastāvdaļas, nevis atkarīgie elementi. 1–2.a. attēlā redzams piemērs, kurā šī papildus virsotne ir apzīmēta ar lomu **adv**, tādejādi parādot, ka **uz skolu** teikumā izpilda apstākļa lomu. Virsotnei **adv** ir viens „īpašais" bērns — **xPrep** (pievienots ar tievu pelēku līniju) — un tas kā bērnus satur visas **adv** sastāvdaļas — piemērā tās ir **uz** un **skolu**.



1–2. attēls. „**gāja uz skolu**"

Pavisam ir trīs veidu virsotnes, kas tiek izvērstas sastāvdaļās atkarībā no tā, cik cieši un ar kāda veida lingvistiskajām saitēm sastāvdaļas ir savā starpā saistītas (TrEd-ā tās no pārējām var atšķirt pēc kvadrātveida formas). Šie veidi ir: x-vārdi, sakārtojums un pieturzīmju konstrukcijas (PMC). Par katru no veidiem tālāk ir pastāstīts sīkāk, taču šeit ir atzīmētas to kopīgās īpašības.

Pirmkārt, katra apaļā virsotne var tikt izvērsta sastāvdaļās, izmantojot x-vārdu, PMC vai sakārtojumu. Otrkārt, katrai apaļajai virsotnei var būt tikai viens veids, lai izvērstu sastāvdaļās, nevis vairāki, t.i. zem vienas apaļās virsotnes nevar atrasties divas x-vārdu virsotnes vai reizē x-vārdu un sakārtojuma virsotnes u.tml.

Treškārt, apaļajai virsotnei, kurai ir „izvērsums-sastāvdaļās" bērns, vēl varbūt arī parastie (atkarību) bērni. Tie apzīmē elementus, kas paskaidro visu salikto elementu kopumā (piemērs redzams **1**–**12**. attēlā).

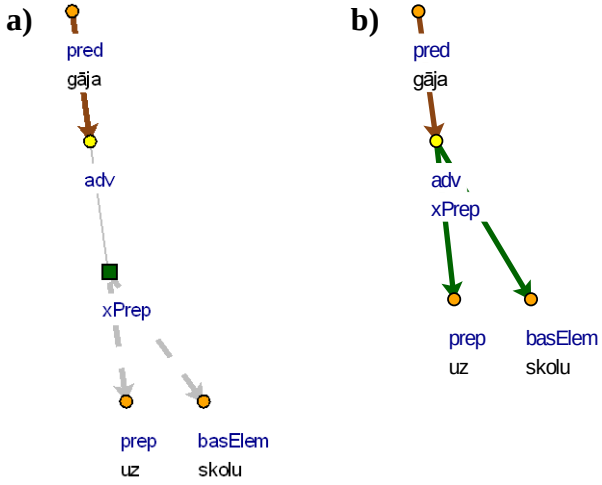
Cilvēkam, kam nepieciešams pašam veidot sintaktiskos kokus, nepieciešams izprast, kā novietojamas šīs sastāvdaļās izvērstās virsotnes, lai spētu pareizi izveidot koku. Savukārt gala lietotājiem vienkāršībai ir izveidots arī otrs attēlojuma veids, kurā sastāvdaļās izvērstās virsotnes tiek attēlotas ar atšķirīgas krāsas bultām. Šī attēlojuma paraugs redzams 1–2.b. attēlā. (Diemžēl tehnisku ierobežojumu dēļ šīs konspektīvākais attēlojuma veids nav lietojams koku rediģēšanai.)

1.2.1. X-vārdi

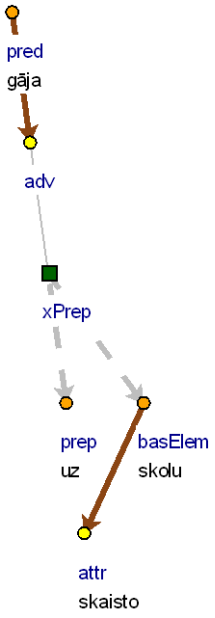
Tred-ā x-vārdus attēlo ar zaļu bultu, ja tiek izmantots kompaktais skata režīms (1–3.b. att.), vai ar zaļu papildvirsotni, ja tiek izmantots marķēšanas režīms (1–3.a. att.).

1–3. attēlā redzams tipiskais x-vārdu lietojums: izteicēju „gāja” paskaidro apstāklis, kas izteikts ar x-vārdu, konkrēti — ar x-prievārdu (**xPrep**). X-vārds sastāv no 2 vārdiem — „**uz**” un „**skolu**”.

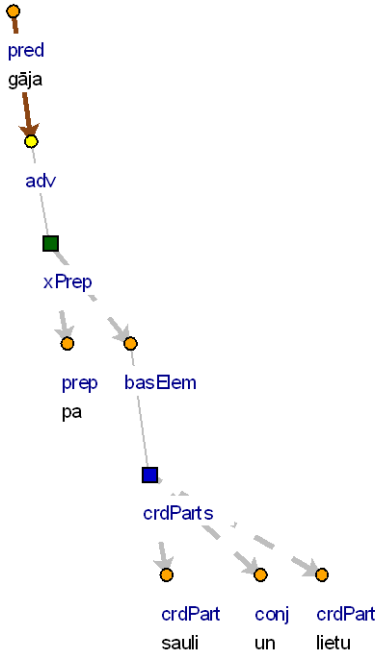
Vispārīgajā gadījumā katru no x-vārda sastāvdaļām var paskaidrot citi vārdi — piemērs 1–4. att., vai arī kāda no sastāvdaļām var būt izteikta ar kādu sarežģītāku konstrukciju — PMC, x-vārdu vai sakārtojumu — piemērs ar sakārtojumu 1–5. att.



1–3. attēls. „**gāja uz skolu**”



1–4. attēls. „**gāja uz skaisto skolu**”



1–5. attēls. „**gāja pa sauli un lietu**”

Gandrīz katram x-vārdam ir vairāki apakštipi, kurus norāda x-vārda tagā kvadrātiekavās. Šobrīd tiek izmantoti turpmāk aprakstītie x-vārdu veidi.

1.2.1.1 Vārdu savienojums ar prievārdu (*xPrep*)

TrEd virsotni apzīmē ar lomu **xPrep**. Piemēri: gāja **uz skolu**, strādāt **par direktoru**, **pa labi**, **pa ceļam**, **par vāju**.

Iespējamās sastāvdaļas:

- prep — prievārds vai prievārdisks adverbs, piemēram, **uz, blakus**;
- basElem — pamatvārds, piemēram, **skolu**.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Prievārdiskai konstrukcijai **xPrep** iespējami trīs apakštipi:

1) prepozicionālas [pre] – prievārds vienmēr novietots pirms bāzes elementa;
2) postpozicionālas [post] – prievārds vienmēr novietots pēc bāzes elementa , *manis dēļ*;
3) konstrukcijas ar prievārdiskiem apstākļa vārdiem [rel] – apstākļa vārds var būt novietots gan pirms, gan pēc bāzes elementa.

Piezīmes:

- x-prievārds ir arī konstrukcijas ar prievārdisku adverbu, kas piesaista nomenu datīvā, piemēram, *blakus man, cauri laikam, apkārt mājai.* Šādas konstrukcijas var veidot šādi prievārdiski adverbi: *apkārt, blakām, blakus, cauri, garām, klāt, līdz, līdzi, līdzās, līdztekus, pakaļ, pāri, pretī, pretim, priekšā, sānis, tuvu.* **NB!** Vārds *līdz* kā prievārdisks apstāklis prievārdiskā konstrukcijā ir marķēts postpozīcijā piesaistītajam nomenam, piemēram, *Nāc man līdz!,* bet pārējos gadījumos tas tiek marķēts kā prepozīcija.
- Konstrukcija **līdz ar to** marķēta kā adverbs *līdz* apstākļa vai situanta funkcijā, kuram pakārtota prievārdiska konstrukcija **xPrep** *ar to*.

1.2.1.2 Salikts izteicējs (*xPred*)

TrEd virsotni apzīmē ar lomu **xPred**. Šeit ietilpst visu veidu saliktie izteicēji — saliktā laika konstrukcijas ar palīgdarbības vārdu (**ir gājis, tika nomazgāts**), nomināli un adverbiāli izteicēji ar saitiņu (**ir muļķis**, **ir auksti, kļuva karsti**), kā arī modālie, fāzes un izpausmes izteicēji ar modificētājverbu (**sāk strādāt, grib skriet, gadījās satikt**).

Iespējamās sastāvdaļas:

- auxVerb — palīgdarbības vārdi **būt**, **tikt** un **tapt**;
- mod — modificētājs modālos, fāzes vai izpausmes saliktos izteicējos;
- basElem — pamatvārds.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Salikta izteicēja konstrukcijai **xPred** iespējami vairāki apakštipi:

1) salikti laiki [act\pass] – darāmās un ciešamās kārtas saliktie laiki;
2) nomināli izteicēji [subst/adj/num/pronom] – sastata izteicēji ar lietvārdu, īpašības vārdu, skaitāl vārdu vai vietniekvārdu;
3) adverbiāli izteicēji [adv] – sastata izteicējs ar apstākļa vārdu;
4) salikts izteicējs ar semantisku modificētāju [mod/phase/expr] – atkarībā no modificētāja tipa iedala modālos, fāzes un izpausmes izteicējos;
5) izteicējs ar nenoteiksmi [inf] – sastata izteicējs ar nenoteiksmi “dzīvot ir mīlēt”.

Piezīmes:

- Lai noteiktu, vai teikumā ir salikts izteicējs **xPred** vai arī izteicēju **Pred** paskaidro sekundāri predikatīvs komponents **spk**, jāskatās, vai var izveidot palīgteikumu ar *ka* vai *lai.* Piemēram, *solīju atdot>solīju, ka atdošu* – neveido saliktu izteicēju. Savukārt *mēģinu pacelt akmeni > *mēģinu, lai paceltu akmeni* – neveidojas dabisks palīgteikums. Otrā darbība nav mēģināšanas nolūks, bet gan tiešais objekts, tāpēc vārdu savienojums *mēģinu pacelt* veido saliktu izteicēju. Jāņem vērā arī kritēriji par nozīmi – otrajā piemērā ir zināma fāzes un modālā nozīme.
- Vārdu savienojumos ar palīgdarbības vārdu, adverbu un nenoteiksmi, piemēram, *ir labi vērot,* nenoteiksme tiek marķēta kā teikuma priekšmets **subj,** ja nenoteiksmi var pārvērst par deverbālu lietvārdu, kas veic teikuma priekšmeta funkciju pie salikta nomināla izteicēja **xPred** – *ir labi vērot > vērošana*(subj) *ir laba*(xPred)*; [ir]grūti to visu atcerēties > atcerēšanās* (subj) *ir grūta* (xPred). Gadījumos, kad šāda transformācija nav iespējama, piemēram, *ir bail iedomāties > *iedomāšanās ir bail,* nenoteiksme tiek marķēta kā sekundāri predikatīvs komponents **spk**, kas atkarīgs no salikta izteicēja **xPred***.*
- Vārdu savienojumi ar verba *būt* finīto formu, īpašības vārdu un verbu nenoteiksmē, piemēram, *ir spējīgs paciest,* tiek marķēti kā **xPred**.
- Vārdu savienojumi ar verba *būt* finīto formu, ciešamās kārtas nelokāmo divdabi ar izskaņu –ams/-ama un verbu nenoteiksmē, piemēram, *ir nepieciešams darīt,* netiek marķēts kā

viens x-vārds, bet gan kā salikts izteicējs (**xPred)** *ir nepieciešams* un tam pakārtots teikuma priekšmets (**subj**) *darīt.*

- Verbs ***kļūt*** vienmēr ir palīgdarbības vārda funkcijā. Vārdu savienojumi ar *kļūt* un prievārda *par* konstrukciju veido saliktu izteicēju **xPred**, piemēram, *kļūt par prezidentu.*
- vairāku nekā divu vārdu salikti izteicēji tiek marķēti šādi:
    - a) ja tas sastāv no vairākiem auxVerb ar lemmu *būt, kļūt, tikt, tapt* un viena basElem, tad veido vienlīmeņa **xPred** ar vairākām sastāvdaļām,: **ir bijusi sarežģīta, bija jābūt aizgājušam, veca esmu kļuvusi, ir ticis iepriecināts**,
    - b) ja tas sastāv no viena vai vairākiem auxVerb, mod, un basElem vai ja tas sastāv no vairākiem auxverb, kur kaut vienam lemma NAV *būt, tikt, tapt, kļūt,* un basElem, tad veido divlīmeņa **xPred** ar divām sastāvdaļām (auxVerb + basElem, mod+basElem), kur katru sastāvdaļu var veidot salikts izteicējs **xPred,: [ir gribējis] skriet, mēģināja [sākt strādāt], varētu [būt pagrūti], mēdz [būt ieinteresēts], [ir palicis] iespaidīgs, [šķita kļuvusi] priecīga.**

1.2.1.3 Skaitļa vārdi (*xNum*)

TrEd virsotni apzīmē ar lomu **xNum**. Piemēri: **trīsdesmit trīs**, **divi simti desmit**.

Iespējamās sastāvdaļas:

- basElem — pamatvārds.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

1.2.1.4 Pielikumi (*xApp*)

TrEd virsotni apzīmē ar lomu **xApp**. Piemēri: **dzejnieks Rainis**, laikraksts „Diena”

Iespējamās sastāvdaļas:

- basElem — pamatvārds.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Pielikuma konstrukcijai **xApp** ir divi apakštipi:

1) saskaņots pielikums [agr] - *dzejnieks Rainis*;
2) nesaskaņots pielikums [non] - *laikraksts “Diena”.*

Piezīmes:

- ir saskaņoti pielikumi un nesaskaņoti pielikumi,
- abreviatūras un saīsinājumi pirms nosaukuma, kas aizstāj sugas vārdu, arī tiek zīmētas kā pielikuma konstrukcija, piemēram, *SIA „Zandarts”, a/s Laima,*
- gadījumos, kad ir vairāki pielikuma dalībnieki, jānosaka hierarhija, kurš ir pielikums pielikumā. Piemērs: *[šurpu turpu gājēja [divgalvainā čūska amfisbena]]*

1.2.1.5 Salīdzinājums (*xSimile*)

TrEd virsotni apzīmē ar lomu **xSimile**. Teikumā tipiski izpilda **spc** lomu. Piemēri: gāja **kā vējš**, dumjš **kā zābaks**, garāks **nekā Pēteris**.

Iespējamās sastāvdaļas:

- conj — salīdzinājuma saikļi **kā, nekā** (morfoloģijā marķēts kā pakārtojuma saiklis);
- basElem — pamatvārds, piemēram, **vējš**.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Salīdzinājuma konstrukcijai ir divi apakštipi:

1) salīdzinājuma konstrukcija [comp] teikumos ar pārāko pakāpi un saikli *nekā/kā*, kas izsaka atšķirību kaut kādā mērā;
2) pielīdzinājuma konstrukcija [sim] teikumos, kas izsaka līdzību vai ekvivalenci, ar saikli "kā".

Tāpat salīdzinājuma konsturkcijām tiek šķirtas negramatizējušās un gramatizējušās salīdzinājuma konstrukcijas, ko apzīmē ar pēdējo burtu xvārda taga kvadrātiekavās – [n] vai [y].

Teikumā salīdzinājuma konstrukcija ir pakārtota īpašības vai apstākļa vārdam, kurš nosauc pazīmi, pēc kuras parādības ir salīdzinātas – ja teikumā ir pārākās pakāpes īpašības vārds vai adverbs, tad salīdzinājuma konstrukcija visticamāk būs pakārtota tam. Ja šāda pazīme ir nosaukta izteicējā, tad salīdzinājuma konstrukcija ir pakārtota visam izteicējam. Ja šāda pazīme nav nosaukta, tad salīdzinājuma konstrukcija ir pakārtota izteicējam (*skrēja kā vējš, viss notika kā vakar*) vai lietvārdam (*vīrs kā ozols*).

Piezīmes:

- gramatizējušās salīdzinājuma konstrukcijas daudzuma izteikšanai *vairāk kā/nekā, mazāk kā/nekā, tikpat kā* tiek marķētas kā **xSimile**, kurā bāzes elements ir adverbs *vairāk, mazāk vai tikpat.* Atšķirībā no ierastās salīdzinājuma konstrukciju struktūras, kurā salīdzinātājdaļa vienmēr seko aiz salīdzinājuma saikļa, gramatizējušajās salīdzinājuma konstrukcijās salīdzinājuma saiklis tiek marķēts vienā frāzē ar adverbu, kas atrodas pirms saikļa. Visu frāzi pakārto tam teikuma loceklim, kas seko aiz saikļa "kā". Daudzuma konstrukcijās gramatizējušos vienību pakārto sakitļa vārdam, ja tāds ir minēts, bet ja nē, tad lietvārdam ar mērvienības nozīmi vai adjektīvam/adverbam, kas nosauc raksturoto pazīmi: *vairāk nekā divus gadus, mazāk nekā stundu, vairāk nekā priecīgs, vairāk nekā pieklājīgi, tikpat kā jūtama;*
- gramatizējusies salīdzinājuma konstrukcija pielīdzinājuma izteikšanai ar *tāds/tāda kā, tā kā* tiek marķēta kā **xSimile,** kurā bāzes elements ir vietniekvārds *tāds* vai adverbs *tā*. Šāds xSimile var būt vārdkopas analoga **subrAnal** sastāvdaļa savienojumā ar īpašības

vārdu vai arī tas var būt pakārtots lietvārdam, piemēram, *tāda kā ēna*, vai adverbam *tā kā bēdīgi*;

- nolieguma partikula gramatizējušās konsturkcijās veido xParticle kopā ar adverbu un ietilpst xSimile sastāvā – ***[[ne vairāk] kā]***.

1.2.1.6 Vārdu savienojums ar partikulu (*xParticle*)

TrEd virsotni apzīmē ar lomu **xParticle**. Šādi attēlo vārda savienojumu ar partikulu, kad partikula ir stingri piesaistīta vienam teikuma loceklim un atrodas fiksētā pozīcijā attiecībā pret to. Piemēri: **ik vakaru**, **ne velna, kaut viens**.

Iespējamās sastāvdaļas:

- no — partikula, piemēram, **ne**, **it**, **ik**;
- basElem — pamatvārds, piemēram, **vakars**.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Partikulas konsturkcijai **xParticle** iespējami divi apakštipi:

1) ar nolieguma partikulu [neg];
2) nav ar nolieguma partikulu [aff].

1.2.1.7 Palīgvārdu savienojums (*xFunctor*)

TrEd virsotni apzīmē ar lomu **xFunctor**. Šādi attēlo vairākvārdu partikulas, vairākvārdu saikļus un vairākvārdu prievārdus. Piemēri: **lai arī**, **droši vien, līdz ko**.

Sastāv no diviem vai vairākiem **basElem**.

Visas sastāvdaļas tiek zīmētas kā zaļās virsotnes tiešie bērni.

Partikulas konstrukcijai **xFunctor** šobrīd apakštipu nav.

Piezīmes:

- Vairākvārdu saikļiem ne vienmēr visas sastāvdaļas tiek morfoloģiski marķētas kā saikļi – viena obligāti ir ar saikļa morfoloģisko tagu, bet otra var būt arī partikula, piemēram, vairākvārdu saiklis **ne tikai** sastāv no sakārtojuma saikļa **ne** un partikulas **tikai.**

1.2.1.8 Vārdkopas analogs (*subrAnal*)

Vārdkopas analogs ir nedalāms vārdu savienojums, kas formāli atgādina vārdkopu, bet kurā starp komponentiem nav pakārtojuma sakara, piemēram, ***kaut kas interesants, tāds skaists, viņi abi***. TrEd virsotni apzīmē ar lomu **subrAnal**.

Sastāv no diviem vai vairākiem **basElem**.

Vārdkopas analogam **subrAnal** ir izveidoti vairāki apakštipi:

1) vietniekvārdu konstrukcija [vv] – mēs abi, tas pats;
2) adjektīva konstrukcija [ipv] – tāds skaists, kaut kas nenosakāms, tāds kā priecīgs;
3) skaitļa vārda konstrukcija [skv] – abi divi, visi trīs, viņš viens;
4) kopuma konstrukcija [set] – viens no daudziem, labākais no labākajiem.

Piezīmes:

- Vārdkopas analogs netiek marķēts, ja seko lietvārds, kuram šos vārdus var pakārtot atributīvā funkcijā, piemēram, ***tāds skaists*** *skats*, ***visi četri*** *skolēni*, ***tāda kā*** *ēna*;
- vārdkopas analogs ar salīdzinājumu "tāds kā +adj." tiek marķēts - [xSimile "tāds kā" + "adjektīvs"], kur xSimile ir gramatizējusies frāze; vādkopas analoga tips – [ipv];
- vārdkopas analogs pagaidām netiek marķēts vārdu savienojumos ar adverbu, kaut arī tie ir ļoti līdzīgi adjketīviskajiem vārdkopas analogiem: **kaut kas vairāk, tā spēcīgi;**
- vārdkopas analogs prievārdiskā konsturkcijā xPrep iesaistās kā viens vesels bāzes elements: *pie [tā paša], no [abiem diviem], pie [viena no daudziem]*.

1.2.1.9 Vārdrindas analogs (*coordAnal*)

TrEd virsotni apzīmē ar lomu **coordAnal**. Piemēri: **trīs četri [kartupeļi]**, **mazs mazs [sunītis]**. Sastāv no diviem vai vairākiem **basElem**.

1.2.1.10 Nosaukums (*namedEnt*)

TrEd virsotni apzīmē ar lomu **namedEnt**. Šādi attēlo nosaukumus bez skaidri norādāmas iekšējās struktūras, piemēram, ***Bērziņš Investment***, un tādā gadījumā x-vārds sastāv no diviem vai vairākiem **basElem.** Visus **namedEnt** elementus zīmē tieši zem zaļās virsotnes.

Vēl šādi attēlo nosaukumus, kuriem ir skaidra iekšējā struktūra, bet kuri netiek saskaņoti ar citiem vārdiem teikumā, neļaujot tos tieši pakārtot citam teikuma loceklim, un kuri funkcionē kā salikts teikuma loceklis, piemēram, *Tēvzemei un Brīvībai.* Tādā gadījumā zem zaļās virsotnes ir viens **basElem**.

1.2.1.11 Frazeoloģiskās vienības (*phrasElem*)

TrEd virsotni apzīmē ar lomu **phrasElem**. Šādi attēlo frazeoloģiskus izteicienus, kuriem nav iespējams skaidri norādīt iekšējo struktūru vai kurus nav iespējams iesaistīt teikumā kā atkarīgu apakškoku. Piemēri: ***ka tavu almu māteri izglītību***, *[lika]* ***soli pa solim***, ***[sīkie mani izbesīja] viens un divi***.

Visbiežāk sastāv no vairākiem **basElem**.

NB! Frazeoloģiskus izteicienus, kam elementus var sakārtot struktūru parādošā apakškokā, kura sakni var iesaistīt pārējā kokā kā atkarīgu elementu, neattēlo kā x-vārdu (piemēram, **ņēma kājas pār pleciem, tikt galā**).

1.2.1.12 Nestrukturētas vienības (*unstruct*)

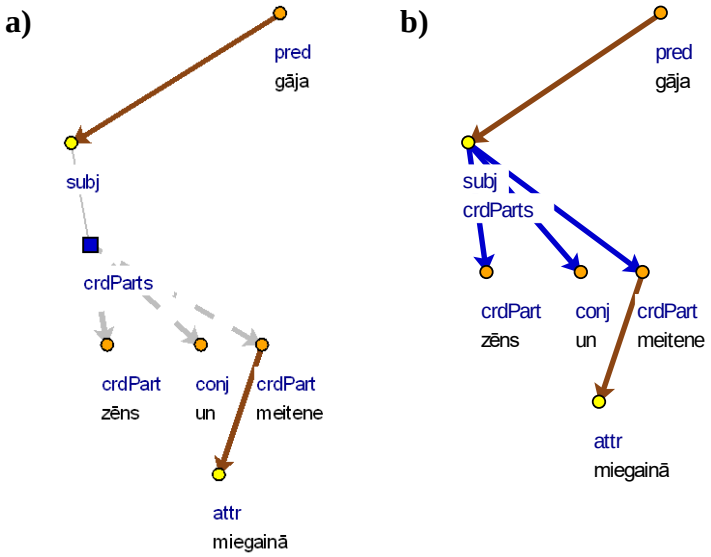
TrEd virsotni apzīmē ar lomu **unstruct**. Šādi attēlo teksta primitīvu vienības bez latviešu valodai raksturīgās gramatiskās struktūras, piemēram, izteicienus svešvalodā, matemātiskās funkcijas u.c. Piemēri: **de facto**, **per aspera ad astra, f(x) = sin x + |x| + 17**.

Sastāv no diviem vai vairākiem **basElem**.

NB! Ja vārdu virkne svešvalodā izsaka nosaukumu, to attēlo ar attiecīgu x-vārdu (*Named Entity*).

1.2.2. Sakārtojums

TrEd-ā sakārtojuma attiecības attēlo ar zilu bultu, ja tiek izmantots kompaktais skata režīms (**1**–**6**.b. att.), vai ar zilu papildvirsotni, ja tiek izmantots marķēšanas režīms (**1**–**6**.a. att.).



1–6. attēls „**zēns un miegainā meitene gāja**"

Šobrīd tiek izmantoti šādi sakārtojuma veidi.

1.2.2.1 Vienlīdzīgi teikuma locekļi (*crdParts*)

TrEd virsotni apzīmē ar lomu **crdParts**. Piemēram, **zaļš, sarkans vai zils**; **zēns un meitene**, vizuāls piemērs dots 1–6. attēlā.

Iespējamās sastāvdaļas:

- **conj** — sakārtojuma saiklis;
- **punct** — pieturzīme;
- **crdPart** — kāds no vienlīdzīgajiem teikuma locekļiem.

Gan viss sakārtojuma elements kopumā, gan katra no sastāvdaļām, kas atzīmētas ar **crdPart** var kalpot par atkarības galvu. Ar **crdPart** atzīmētās sastāvdaļas var būt izteiktas ar x-vārdu, citu sakārtojumu vai PMC.

Piezīmes:

- Saīsinājumus, kas parādās vienlīdzīgu teikuma locekļu uzskaitījuma beigās, piemēram, ***u.tml., u.c., utt.***, marķē tāpat kā partikulu – **no** un novieto zem apaļās virsotnes, kas reprezentē visu vienlīdzīgu teikumu locekļu virkni.

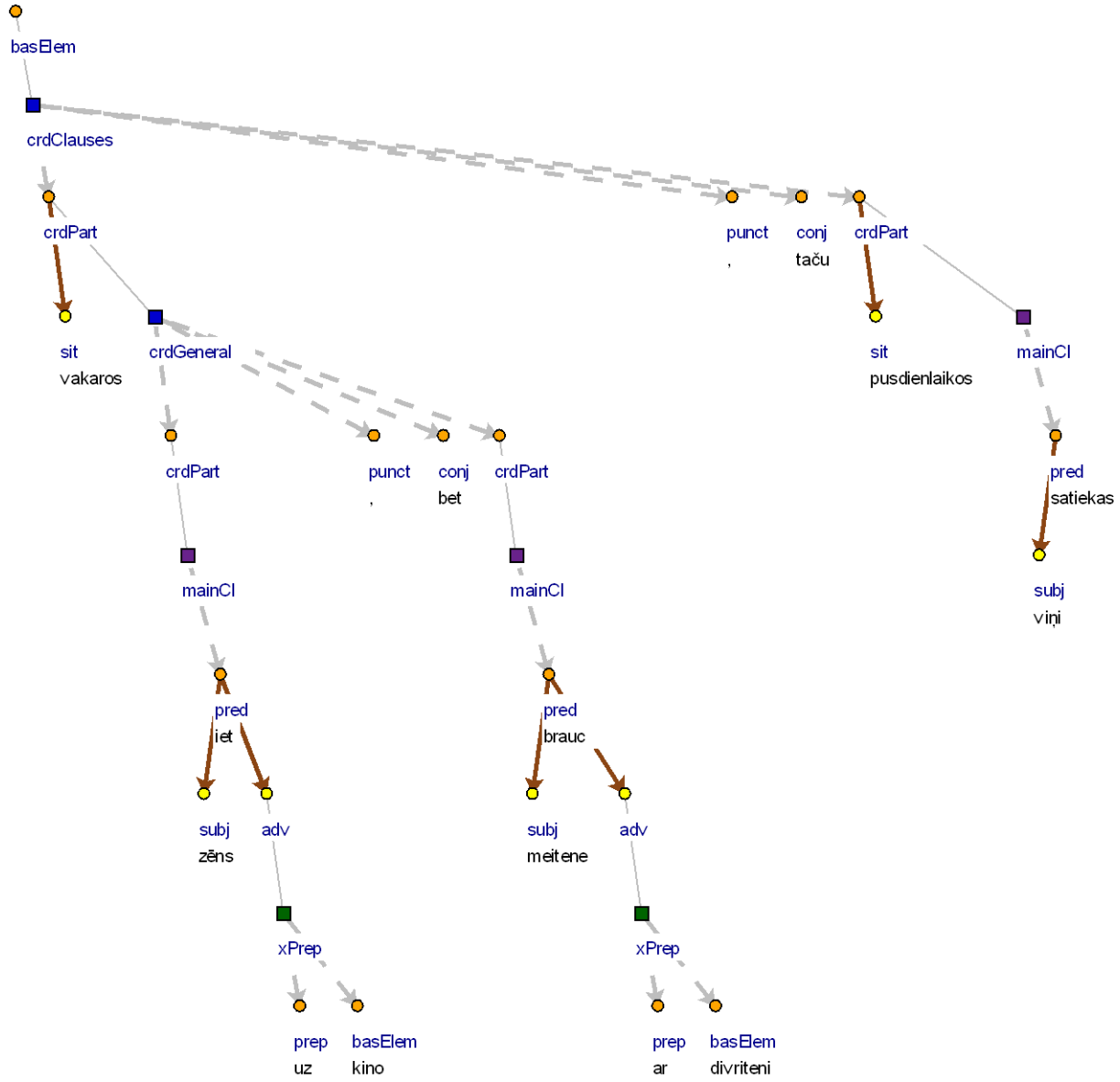
1.2.2.2 Ar sakārtojumu saistītas teikuma daļas (*crdClauses*)

Šādi attēlo gan vienlīdzīgus palīgteikumus, gan neatkarīgas teikuma daļas.

TrEd virsotni apzīmē ar lomu **crdClauses**. Vizuāli piemēri doti 1–7. attēlā (komplicēta virsteikumu kaskāde) un 1–12. attēlā (vienlīdzīgi palīgteikumi).

Iespējamās sastāvdaļas:

- **conj** — sakārtojuma saiklis;
- **punct** — pieturzīme;
- **crdPart** — teikuma daļa (palīgteikums vai neatkarīgā teikuma daļa). Piezīme: ja šādi tiek saistītas neatkarīgas daļas, tad visbiežāk sanāk, ka tālāk zem **mainCl** virsotnēm pieturzīmju vairs nav — tā, kā tas ir **1**–**7**. attēlā.

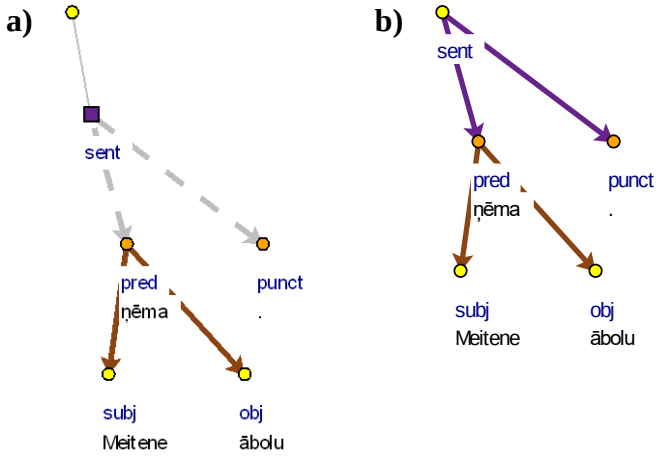


1–7. attēls. „**vakaros zēns iet uz kino, bet meitene brauc ar divriteni, taču pusdienlaikos viņi satiekas**"

1.2.3. PMC (pieturzīmju piesaiste kokam)

Ņemot vērā, ka latviešu valodā pieturzīmes tiek liktas pēc gramatiskā principa, ir nepieciešams tās piesaistīt sintakses kokam tā, lai nezaudētu informāciju, ko var izsecināt no pieturzīmju lietojuma. Šim nolūkam kalpo īpaša veida konstrukcija — pieturzīmju konstrukcija PMC (*punctuation mark construct*). No lietošanas tehniskā aspekta šī konstrukcija ir analoģiska x-vārdiem un sakārtojuma virsotnēm: arī šī saikne starp elementiem marķēšanas režīmā tiek attēlota, izmantojot papildus virsotni — šoreiz violetu (**1**–**8**.a. att.). Kompaktajā skata režīmā PMC attēlo ar violetām bultām (**1**–**8**.b. att.).

Tālāk aprakstīti izmantotie PMC veidi.



1–8. attēls. „**Meitene ņēma ābolu.**”

1.2.3.1 Teikums (*sent*)

TrEd virsotni apzīmē ar lomu **sent**. Šis ir pirmais no diviem veidiem, kā „sākt koku” (otrs veids ir izteikumi, kas aprakstīti 1.2.3.2. nodaļā). Piemērs redzams **1**–**8**. attēlā.

Iespējamās sastāvdaļas:

- **punct** — pieturzīmes (parasti teikumu noslēdzošā pieturzīme);
- **pred** — izteicējs — vienkāršiem teikumiem un saliktiem pakārtotiem teikumiem (bet ne saliktiem sakārtotiem vai jauktiem); (parcelātiem šeit var būt arī citi teikuma locekļi)
- **conj** — saiklis (šeit tiek likti tikai tie saikļi, kas **neorganizē** sakārtojumu, ja tādi ir — piemēram, **bet** teikumā „**Bet Jānis ēd.**”);
- **no** — partikulas, izsauksmes vārdi, uzrunas, iespraudumi un citi elementi, kam teikumā nav lomas — vienkāršiem teikumiem;
- **basElem** — teikuma pamatdaļa saliktiem sakārtotiem un jauktiem saliktiem teikumiem vai tiešā runa, vai piebilde.

Piezīmes:

- ja teikums ir no iepriekšējā teikuma atsevišķā teikumā nodalīts palīgteikums vai arī virsteikums varētu būt reducēts, to marķē kā **sent** konstrukciju: *Jo uzskatīja, ka citrona sula izšķīdina nejauši norītu asaku.*

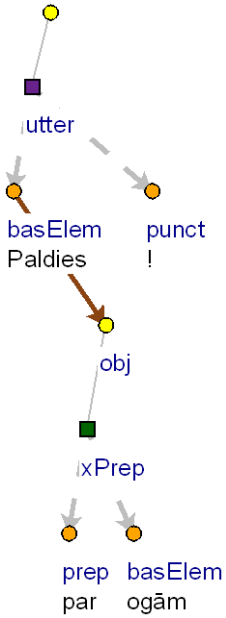
1.2.3.2 Izteikums (*utter*)

TrEd virsotni apzīmē ar lomu **utter**. Šis ir otrs no veidiem, kā „sākt koku”, — šādi tiek aprakstīti izteikumi un citas teksta daļas, kam centrā nav ne verba, ne labi atjaunojama reducēta verba (par redukcijām sk. 2. nodaļu). Piemēri: „**Jā!**”, „**Paldies, nē.**” „**Paldies par ogām!**”. Izteikuma iekšienē drīkst izmantot atkarībām lomu „**no**”, lai norādītu atkarības no izsauksmes vārdiem utt. (sk. 1–9. att.)

Tipiskās sastāvdaļas:

- **punct** — pieturzīmes;

- **no** — partikulas, izsauksmes vārdi;
- **basElem** — citi elementi, piemēram, nosaucējvienības.



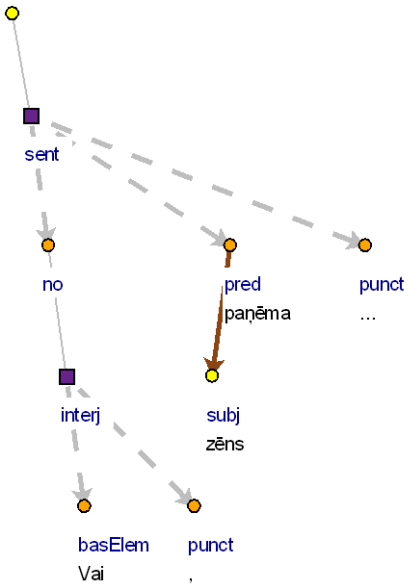
1–9. attēls. „**Paldies par ogām!**”

1.2.3.3 Izsauksmes vārds (*interj*)

TrEd virsotni apzīmē ar lomu **interj**. Šo konstrukciju lieto, lai piesaistītu teikumā esošās pieturzīmes pie izsauksmes vārdiem, kas izraisa to lietošanu — 1–10. att.

Iespējamās sastāvdaļas:

- **basElem** — izsauksmes vārds;
- **punct** — pieturzīme.



1–10. attēls. „**Vai, zēns paņēma…**”

1.2.3.4 Patstāvīga teikuma daļa (*mainCl*)

TrEd virsotni apzīmē ar lomu **mainCl**. Šādi (papildus izmantojot 1.2.2. nodaļā aprakstīto sakārtojumu) tiek aprakstītas salikta sakārtota teikuma daļas un jaukta salikta teikuma neatkarīgās daļas. Lietošanas piemērs redzams **1**–**7**. att.

Iespējamās sastāvdaļas:

- **pred** — izteicējs;
- **no** — partikulas, izsauksmes vārdi, uzrunas, iespraudumi un citi elementi, kas atrodas šajā teikuma daļā, bet kam nav sintaktiskās lomas;
- **conj** — saiklis (šeit tiek likti tikai tie saikļi, kas **neorganizē** sakārtojumu, ja tādi ir);
- **punct** — pieturzīme.

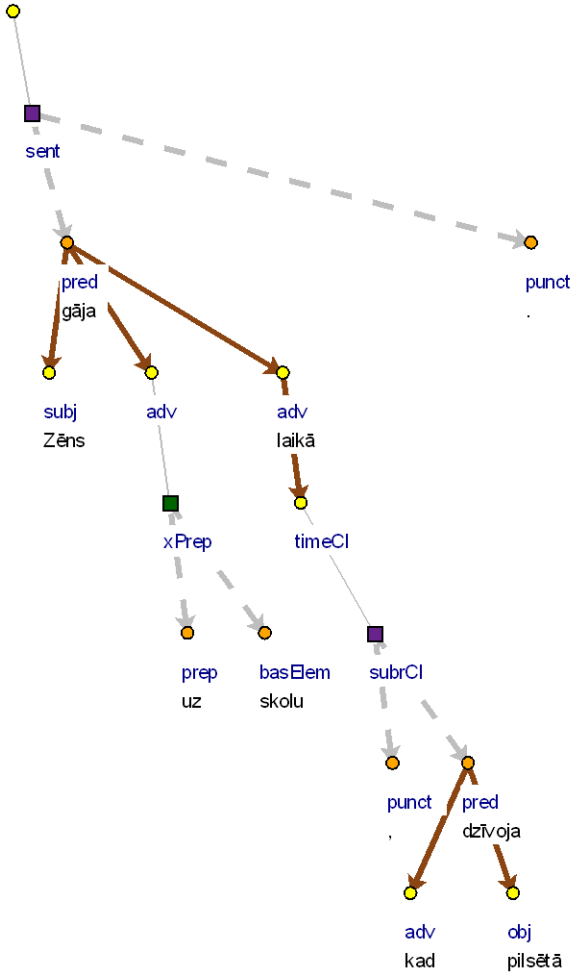
1.2.3.5 Palīgteikums (*subrCl*)

TrEd virsotni apzīmē ar lomu **subrCl**. Lietošanas piemērs redzams 1–11. attēlā. Palīgteikums tiek iesaistīts kokā šādi: elementam (vārdam, x-vārdam vai pat teikuma daļai), kuru palīgteikums paskaidro, tiek uztaisīts atkarīgais bērns, kuram kā lomu norāda palīgteikuma veidu, un tad zem šī bērna tiek novietota violetā **subrCl** virsotne.

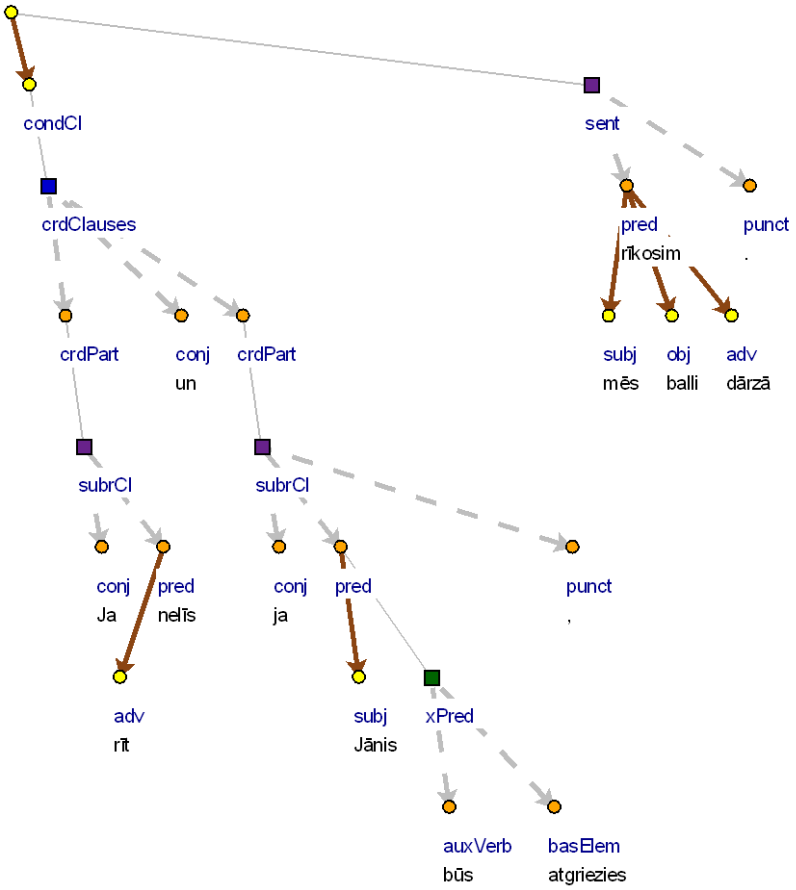
Vienlīdzīgus palīgteikumus norāda, izmantojot sakārtojuma virsotnes — piemērs 1–12. attēlā. Ja saliktā pakārtotā teikumā palīgteikums paskaidro visu virsteikumu, tad virsotni, kas norāda palīgteikuma veidu, novieto kā atkarīgu zem koka saknes virsotnes (1–12. att.). Ja jauktā saliktā teikumā palīgteikums paskaidro kādu no virsteikumiem visā pilnībā, nevis atsevišķu tā locekli, tad palīgteikumu pievieno pie virsotnes, kas ir atbilstošā virsteikuma **mainCl** tiešais vecāks (1–13. att.).

Iespējamās sastāvdaļas:

- **pred** — izteicējs;



1–11. attēls. „**Zēns gāja uz skolu laikā, kad dzīvoja pilsētā.**”

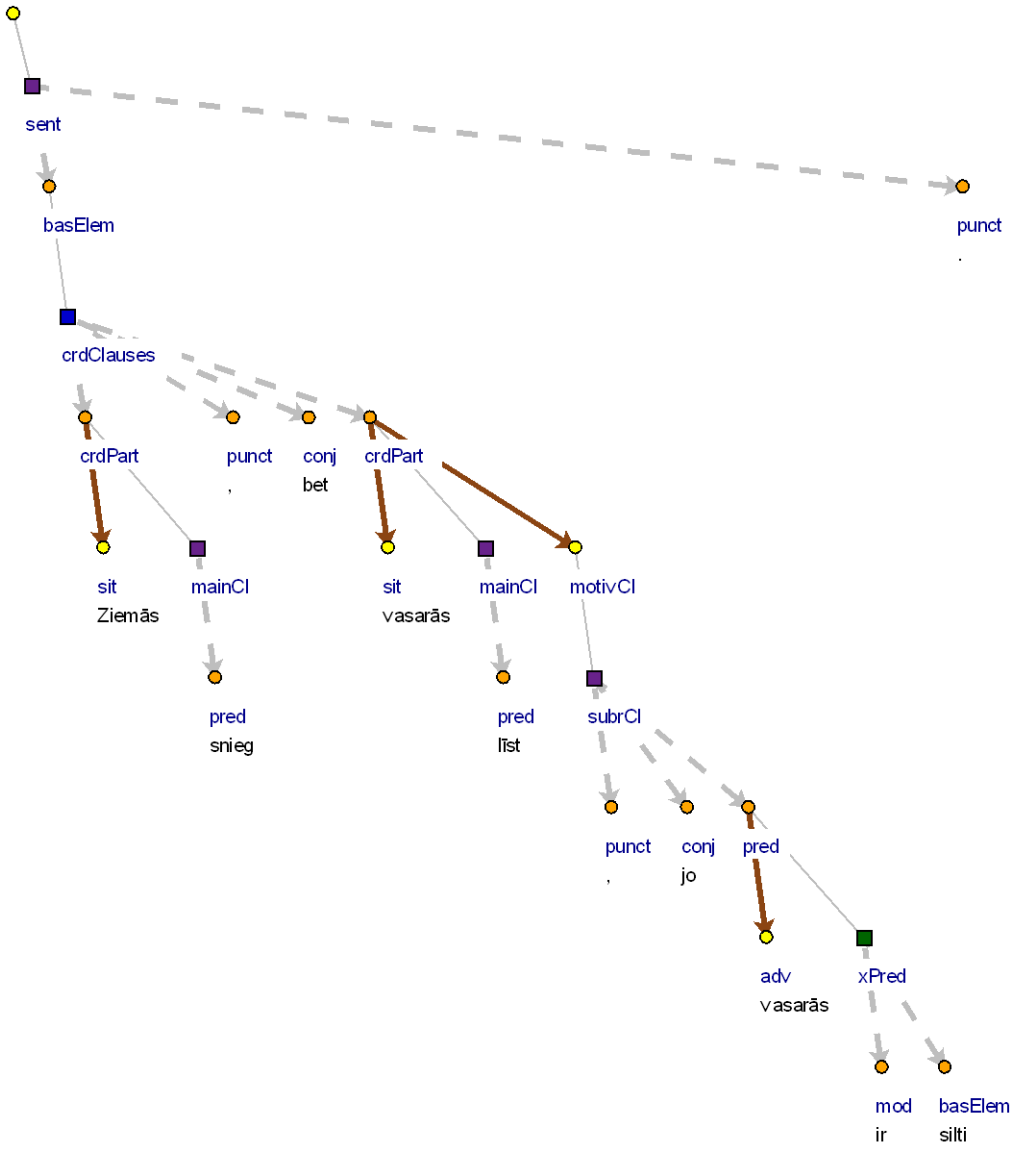


1–12. attēls. „**Ja rīt nelīs un ja Jānis būs atgriezies, mēs rīkosim balli dārzā.**”

- **no** — partikulas, izsauksmes vārdi, uzrunas, iespraudumi un citi elementi, kas atrodas šajā teikuma daļā, bet kam nav sintaktiskās lomas;
- **conj** — saiklis;
- **punct** — pieturzīme.

Piezīmes:

- ja teikums ir no iepriekšējā teikuma atdalīts palīgteikums, to zīmē kā teikuma (**sent**) konstrukciju: *Jo uzskatīja, ka citrona sula izšķīdina nejauši norītu asaku.*



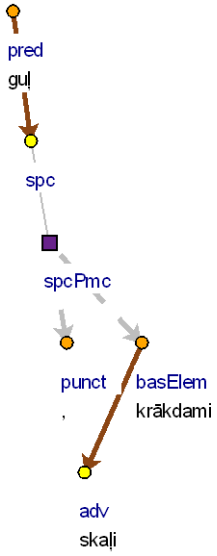
1–13. attēls. „**Ziemās snieg, bet vasarās līst, jo vasarās ir silti.**"

1.2.3.6 Sekundāri predikatīva komponenta (SPK) pieturzīmes (*spcPmc*)

TrEd virsotni apzīmē ar lomu **spcPmc**. Šādi kokam piesaista SPK pieturzīmes, piemēram, komatus, kas tiek lietoti divdabja teicienu un citu savrupinājumu dēļ (1–14. att.).

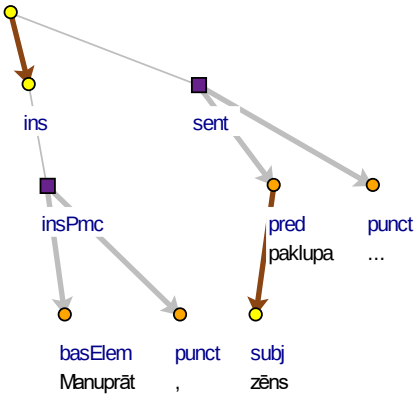
Iespējamās sastāvdaļas:

- **basElem** — spk;
- **punct** — pieturzīme.

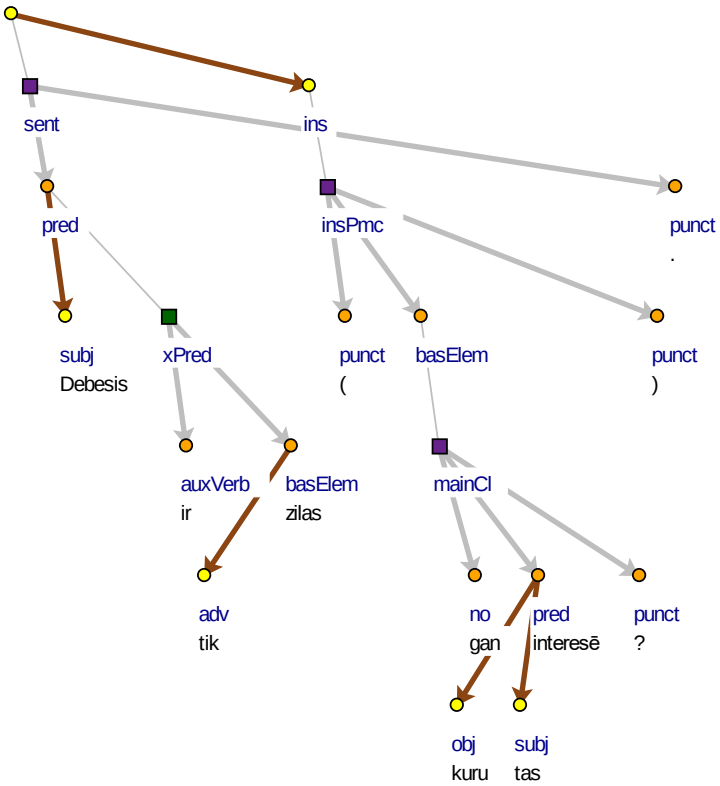


1–14. attēls. „**guļ, skaļi krākdami**”

1.2.3.7 Iespraudumi un iestarpinājumi (*ins*)



1–15. attēls. „**Manuprāt, zēns paklupa…**”



1–16. attēls. „**Debesis ir tik zilas (kuru gan tas interesē?).**”

Iespraudumi un iestarpinājumi marķējot netiek nošķirti. Ar **ins** lomu apzīmē visas konstrukcijas, kuras ievada tipiskais ievadītājvārds *kā* un kuras norāda informācijas avotu, tāpat konstrukcijas, kas norāda uz autora attieksmi pret teksta saturu, kā arī konstrukcijas iekavās, domuzīmēs vai ar īpašiem ievadītājvārdiem, kuras sniedz dažāda rakstura papildinformāciju. Šīs konstrukcijas TrEd-ā apstrādā šādi. Pirmkārt, ar PMC virsotni ar lomu **insPmc** veic iespraudumа norobežojošo pieturzīmju piesaistīšanu iespraudumа galvenajam vārdam. Otrkārt, apaļo virsotni virs **insPmc** marķē ar lomu **ins** un ar atkarību piesaista teikuma daļai vai teikuma loceklim uz kuru iespraudums attiecas. Piemērs, kur iespraudums piesaistāms noteiktam teikuma loceklim, dots 1–17. att., piemēri, kur iespraudums piesaistāms teikumam — 1–15. att., 1–16. att.

Tipiskās **insPmc** sastāvdaļas:

- **basElem** vai **pred** — iespraudumа galvenais vārds (**pred** lieto, ja iespraudums ir predikatīva vienība, savukārt **basElem** lieto, ja iespraudums ir sekundāri predikatīva vai nepredikatīva vienība, kā arī gadījumos, kad iespraudumu veido predikatīva vienība ar savu teikuma beigu pieturzīmi, kādēļ nepieciešams veidot pieturzīmes konstrukciju **mainCl**);
- **punct** — pieturzīme.

Predikatīva vienība kā iespraudums veidojas:

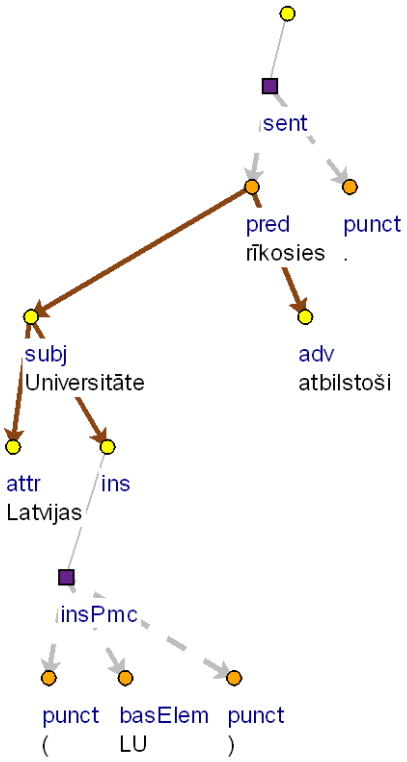
- ja predikatīvs iespraudums norāda autora attieksmi: *jāatzīst, kā likās, izrādās*;
- ja iespraudums norāda informācijas avotu un sākas ar tipisko ievadītājvārdu *kā*: .. *kā uzsver japāņu medicīna, kāju pēdās ir izvietoti nervu gali, kuri ir savienoti ar visiem iekšējiem orgāniem.*
- ja predikatīvs iespraudums atdalīts ar iekavām, domuzīmēm: .. *bez diviem tehniskajiem darbiniekiem vēl būs vajadzīgs sporta organizators – noares entuziasts (pašlaik šos pienākumus ar savējiem apvieno kultūras organizatore).*;
- ja to ievada tipiskie ievadītājvārdi (piem., *proti, tas ir*) un nevar noteikt pakārtojuma vai sakārtojuma attieksmes starp teikuma daļām;
- ja tiešajā runā piebilde ar verbu finītā formā ir novietota teikumam pa vidu, tad to marķē kā predikatīvu iespraudumu, kas atkarīgs no visas tiešās runas.

Piezīmes:

- Teikuma sākumā esoša nominăla frāze, kas nosauc tālāk teikumā paustās informācijas avotu un kurai seko kols, tiek marķēta kā iespraudums, kuram piesaistīta pieturzīme kols un kurš attiecas uz visu teikumu (skatīt 1–15. att.) tad, ja sekojošā teikuma daļa pēc kola netiek marķēta kā tiešā runa. Piemēram, *Lipmans: Notiek sarunas par jaunas „Grindeks" rūpnīcas būvniecību Krievijā.* Savukārt gadījumos, kad teikuma pirmajā daļā pirms kola bez nominālas frāzes ir arī vēl citi paplašinātāji, teikums tiek attēlots kā salikts pakārtots teikums ar reducētu izteicēju pirmajā daļā, piemēram, *Satiksmes ministrs Aivis Ronis intervijā Inārai Eglei: „SM jomās mēs vēl veiksim būtiskas reformas"* (par redukciju skatīt 2. nodaļu). Šajā gadījumā salikts pakārtots

teikums tiek zīmēts arī tāpēc, ka pēc kola attēlota tiešā runa ar pēdiņām (par tiešās runas attēlojumu skatīt 1.2.3.9. nodaļu).

- atsevišķos gadījumos loma **ins** tiek lietota arī vienam vārdam, kas neveido konstrukciju ar pieturzīmi (sk. 1.1.9. nodaļu).



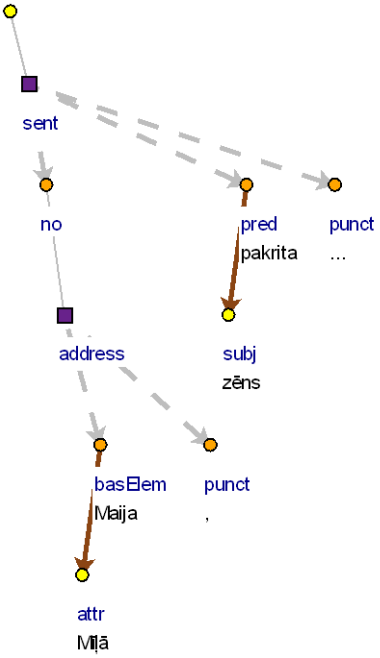
1–17. attēls. „**Latvijas Universitāte (LU) rīkosies atbilstoši.**"

1.2.3.8 Uzruna (*address*)

TrEd virsotni apzīmē ar lomu **address**. Šo konstrukciju lieto, lai piesaistītu teikumā esošās pieturzīmes pie uzrunas, kas izraisa to lietošanu. Lietošanas piemērs — 1–18. att.

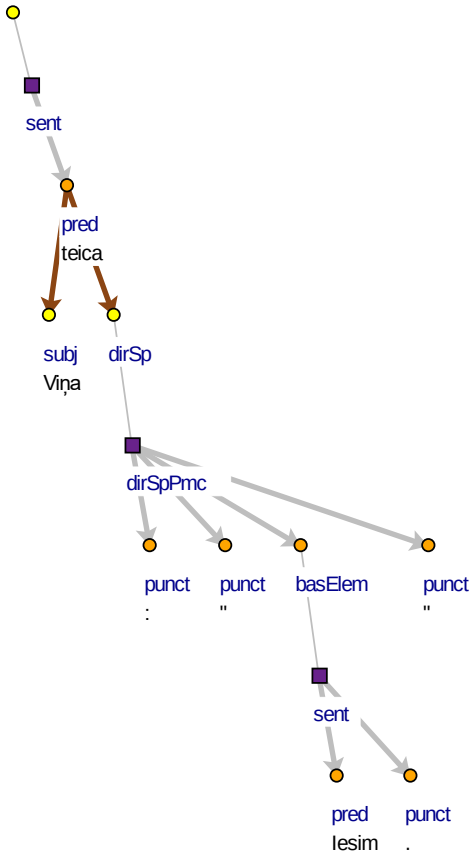
Iespējamās sastāvdaļas:

- **basElem** — uzrunas galvenais vārds;
- **punct** — pieturzīme.



1–18. attēls. „**Mīļā Maija, zēns pakrita…**”

1.2.3.9 Tiešā runa (*dirSpPmc*)



1–19. attēls. „**Viņa teica: „Iesim.”**”

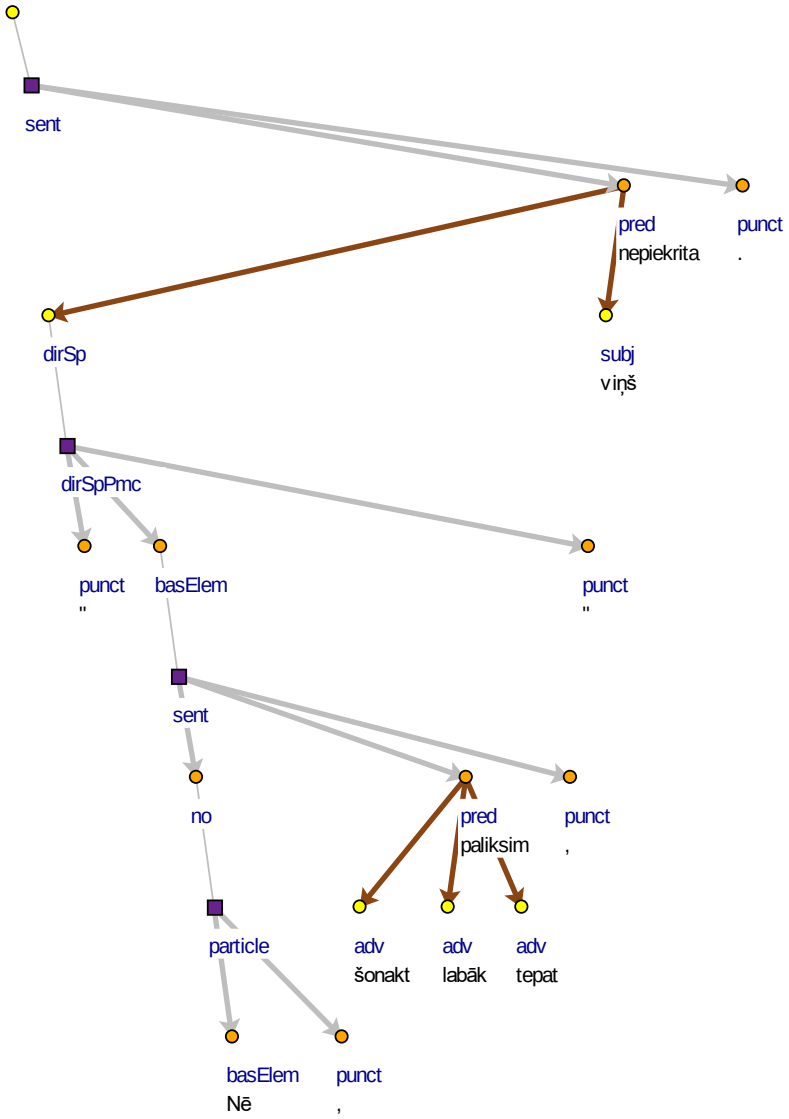
Tiešās runas apstrādei TrEd-ā tiek izmantota atkarība un PMC konstrukcija: pirmkārt, lai atzīmētu pašu tiešo runu un piesaistītu tiešo runu norobežojošās pieturzīmes — kolu, pēdiņas, domuzīmes, tiek izmantota virsotne ar lomu **dirSpPmc**, zem kuras caur **basElem** tiek pievienota **sent** vai **utter** konstrukcija. Otrkārt, apaļo virsotni virs **dirSpPmc** marķē ar lomu **dirSp** un ar atkarību piesaista atbilstošajam piebildes izteicējam. Ja tiešā runa atrodas starp piebildes vārdiem un satur

vairākus izteikumus, katrs no tiem tiek attēlots kā atsevišķs **basElem** zem **dirSpPmc**. Ja tiešā runa atrodas pirms vai pēc piebildes un satur vairākus izteikumus, tad kopā ar piebildi vienā kokā tiek zīmēts tikai viens piebildei tuvākais tiešās runas izteikums; pārējie tiek zīmēti atsevišķos kokos, turklāt tiešās runas virsotni izmanto tikai tam runas izteikumam, kuram ir attiecīgas tiešās runas pieturzīmes. Gadījumos, kad tiešās runas teikumiem tiešās runas pieturzīme ir vienā pusē un kad tiešā runa lietota bez piebildes, zem saknes atrodas PMC virsotne ar lomu **dirSpPmc**, zem tās apaļā virsotne ar lomu **basElem**, savukārt zem kuras atrodas PMC virsotne ar lomu **sent**.

Piemēri ar piebildes novietojumu pirms tiešās runas un pēc tiešās doti attiecīgi **1**–**19**. un **1**–**20**. att.

Iespējamās **dirSpPmc** sastāvdaļas:

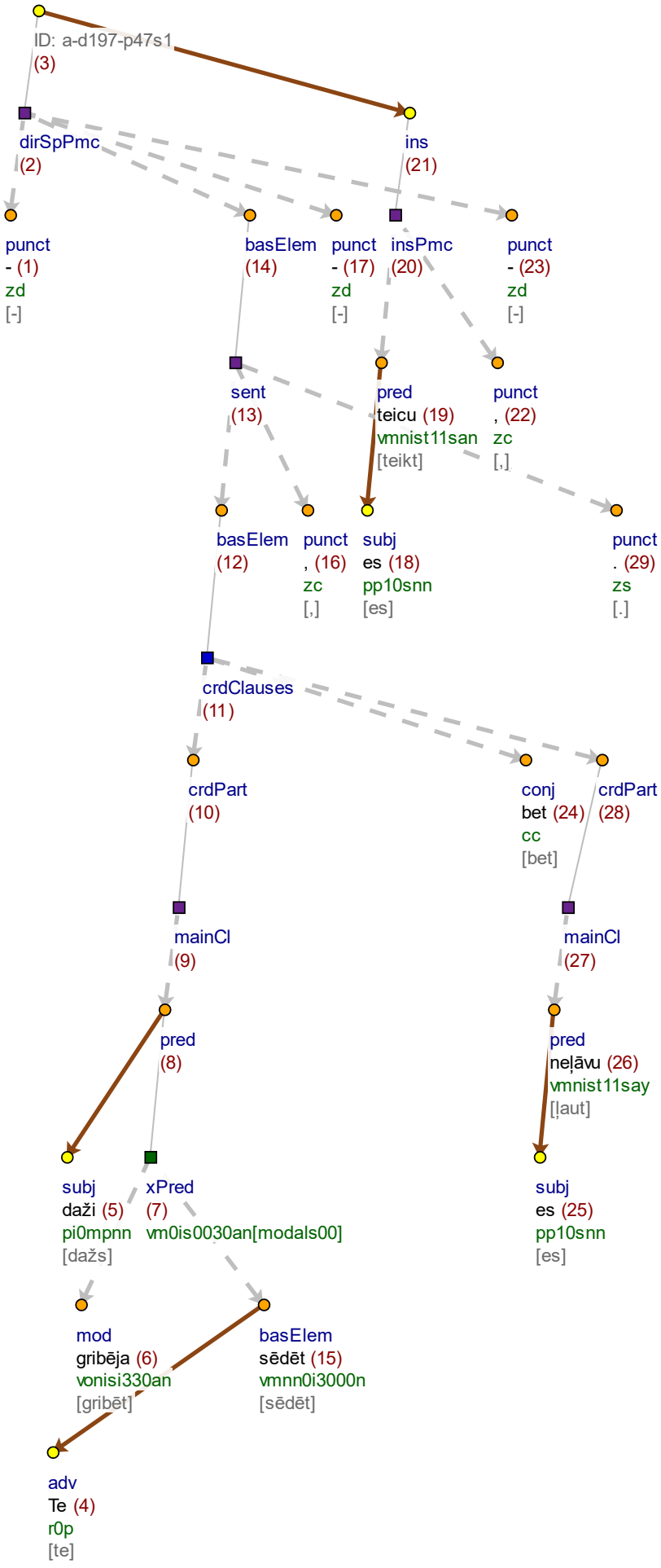
- **basElem** — tiešā runas izteikums;
- **punct** — pieturzīme.



1–20. attēls. „**„Nē, šonakt labāk paliksim tepat," viņš nepiekrita.**"

Piezīmes:

- Teikuma sākumā esoša nominăla frāze, kas nosauc tālāk teikumā paustās informācijas avotu un kurai seko kols un tiešās runas pieturzīme (pēdiņas vai domu zīme), tiek marķēta kā teikums ar reducētu izteicēju, kuram pakārtota tiešā runa. Piemēram: ***Inita Krūma: - Jā, tā ir lieta.*** Savukārt, ja pēc kola neseko tiešās runas pieturzīme, tad nominālā frāze pirms kola tiek attēlota kā iespraudums, kas pakārtots teikuma saknei.
- ja tiešajā runā piebilde ar verbu finītā formā ir novietota teikumam pa vidu, tad to marķē kā predikatīvu iespraudumu, kas atkarīgs no visas tiešās runas, skatīt 1-21. attēlu.



1-21. attēls. Teikums ar piebildi vidū: "– Te daži gribēja sēdēt, es teicu, - bet es neļāvu."

1.2.3.10 Citas pēdiņas (*quot*)

TrEd virsotni apzīmē ar lomu **quot**. Šādi teikuma struktūrai pievieno visas pārējās pēdiņas — tās, kas nav saistītas ar tiešo runu, piemēram, pēdiņās ieliktus nosaukumus, terminus, pārnestā nozīmē lietotus vārdus. Vēl šādas pēdiņas lieto gadījumos, kad teikumā teikuma locekļa funkcijā ir iesaistīts citēts vārds vai izteikums, vai uzraksts (netiek veidota tiešajai runai raksturīgā konstrukcija ar piebildi un kolu vai domu zīmi), piemēram: **„Sofijai Amundsenai" stāvēja rakstīts uz nelielās aploksnes**. Citāti, kas ir vesels teikums vai izteikums, piemēram, *Gluži tāpat kā vakarnakt skanēja "Es atceros gan mūsu pirmo, gan pēdējo tikšanos".* arī tiek apzīmēti kā konstrukcija pēdiņās, tikai konstrukcijas saknē tiek lietota loma **pred,** nevis **basElem**.

Iespējamās sastāvdaļas:

- **basElem** vai **pred** — pēdiņās ieliktā vienība, **pred** liek tad, ja citāts vai nosaukums ir predikatīva vienība;
- **punct** — pieturzīme.

1.2.3.11 Partikulu konstrukcijas (*particle*)

TrEd virsotni apzīmē ar lomu **particle**. Šo konstrukciju lieto, lai piesaistītu teikumā esošās pieturzīmes pie partikulas vārdiem, kas izraisa to lietošanu. Piemērs: „**Jā, es nāku.**"

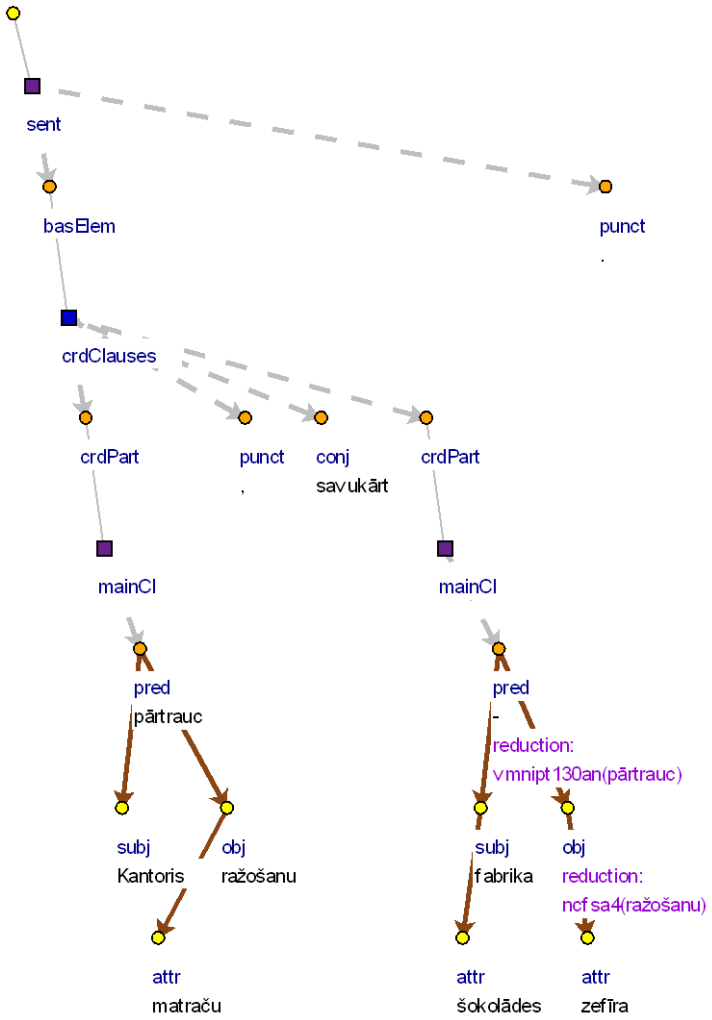
Iespējamās sastāvdaļas:

- **basElem** — partikula;
- **punct** — pieturzīme.

2. Redukcija

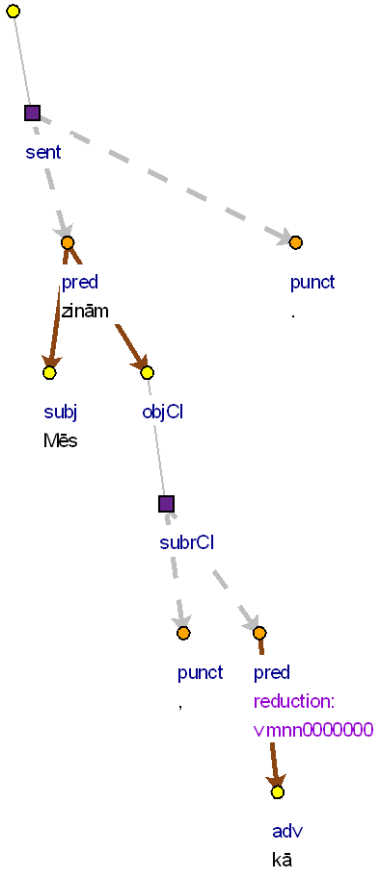
Ja teikums satur kādus reducētus vārdus, tad tie tiek atjaunoti, izmantojot tikai attiecīgajā teikumā esošo informāciju. Atjaunošana tiek veikta, ja:

- reducēto vārdu paskaidro kāds teikuma loceklis, kurš citādi nevar tikt piesaistīts kokam, jo neveido semantiskās attieksmes un sintaktisko sakaru ne ar vienu no konkrētajā teikuma daļā realizētajiem locekļiem (piemēram, vārds **zefīra** 2–1. attēlā,). Tipiskākie gadījumi šādai redukcijai:
    - reducētas teikuma daļas, kurās teikumā ir realizēti tikai teikuma priekšmets un apstāklis ar laika nozīmi – šādi apstākļi neveido semantiskas attieksmes tieši ar realizēto teikuma priekšmetu un nevar būt tieši pakārtoti, piemēram, *Aiz ratiem nāca Albertīns, tad [nāca] Helēna un tad [nāca] visi septiņi vāķētāji*.
- reducētais vārds ir saitiņa vai palīgdarbības vārds, piemēram **[ir] dzīvs**, **[ir] jādod**. Atsevišķos gadījumos tiek atjaunots arī semantiskais modificētājs saliktā izteicējā, ja tas teikumā reducēts, jo atkārtojas: *Tāpēc sēdāmies ar sievu pie galda, paņēmām divas baltas papīra lapas, pārdalījām katru uz pusēm, la vienā var rakstīt to, ko gribam, bet otrā – **[var rakstīt]**ko kategoriski negribam*.



2–1. attēls. „**Kantoris pārtrauc matraču ražošanu, savukārt šokolādes fabrika — zefīra.**”

Atjaunošana notiek šādi. Ja vārds reducēts par domuzīmi vai citu pieturzīmi, tad šīs pieturzīmes virsotne kokā tiek novietota vietā, kur vajadzētu atrasties reducētajam vārdam. Ja vārds reducēts, to pilnībā izlaižot, tad to atjauno, izveidojot jaunu virsotni, kuru novieto kokā atbilstošajā vietā. Ja par domuzīmi reducēts nevis vārds, bet vesels apakškoks, tad domuzīmi novieto reducētā apakškoka saknē, bet pārējo apakškoku izveido no jaunām virsotnēm.



2–2. attēls. „**Mēs zinām, kā.**”

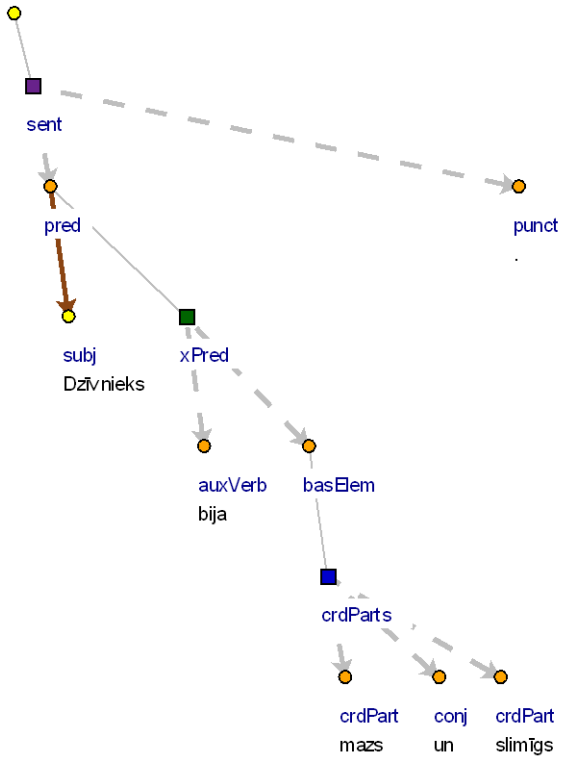
Pēc tam redukcijā iesaistītajām pieturzīmju virsotnēm un visām virsotnēm, kas atjaunošanas dēļ izveidotas no jauna, aizpilda „**reduction**” lauku(par virsotņu marķēšanu precīzāk sk. 4. nodaļu). Ja no teikuma ir labi saprotams, kāds vārds ir reducēts (piemēram, reducēts kāds atkārtots vārds vai darbības vārds **būt**), tad **reduction** laukā ieraksta reducētā vārda pilnu morfoloģisko tagu un aiz tā iekavās pašu reducēto vārdu atbilstošajā formā (piemērs 2–1. attēlā).

Ja reducētais vārds nav viennozīmīgi atjaunojams (piemēram, reducētais palīgteikuma izteicējs teikumā „**Mēs zinām, kā.**” 2–2. attēlā), tad **reduction** laukā ieraksta tikai morfoloģisko tagu, nenosakāmās pozīcijas aizstājot ar nullēm. Darbības vārdiem šādas neatjaunojamas pazīmes ir atgriezeniskums, transitivitāte, konjugācija, noliegums. Citu pazīmju atjaunojamība atkarīga no konteksta.

Nomeniem (lietvārdiem, vietniekvārdiem) redukcijas tagu raksta pēc lietvārdu morfoloģiskā taga parauga. Neatjaunojamās pazīmes (ja nav nosakāms precīzais vārds) ir tips un deklinācija, pārējo atjaunojamība atkarīga no konteksta.

2.1. Specifiskie gadījumi

Vārdu savienojumi, kuros viens palīgvārds tiek attiecināts uz vienlīdzīgiem teikuma locekļiem, netiek uzskatīti par redukcijas gadījumu, tāpēc piemēros „**ceļot pa mežiem un siliem**”, „**bija mazs un slimīgs**” redukcijas nav, t.i. tos nezīmē kā „**ceļot pa mežiem un [pa] siliem**”, „**bija mazs un [bija] slimīgs**”. Pareiza koka piemērs dots 2–3. attēlā.

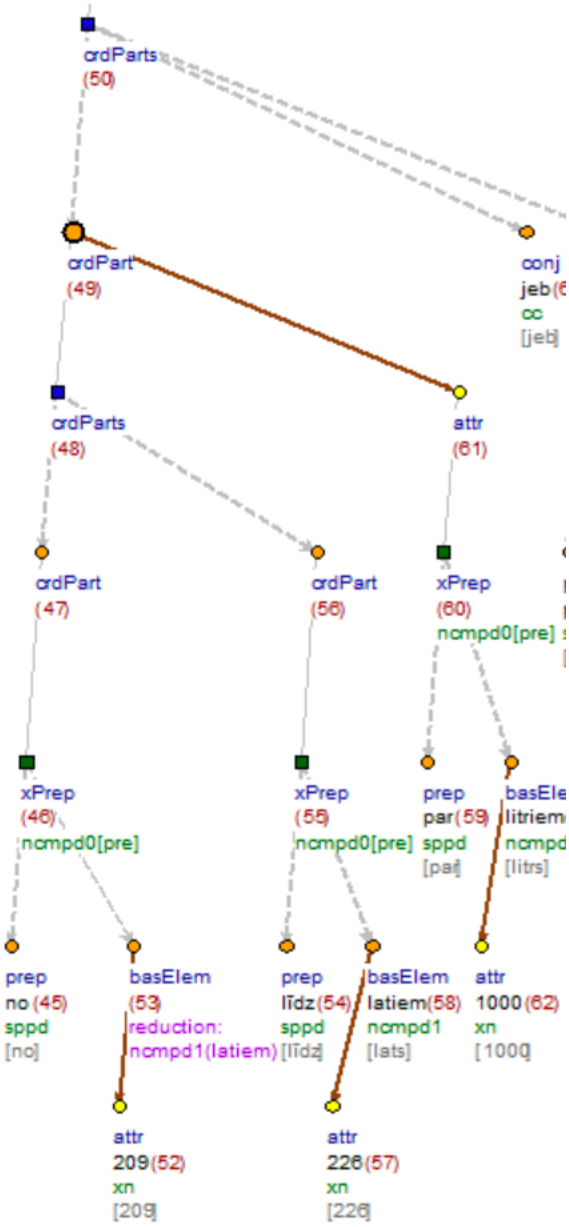


2–3. attēls. „**Dzīvnieks bija mazs un slimīgs.**”

Palīgdarbības vārda redukcija nav gadījumā, kad izteicēja funkciju veic noliegts lokāms darāmās kārtas pagātnes divdabis ar izskaņu –is/-usi ar priedēklis „ne-”: **„Nezinājis, kā naudu tērēt.”** Kaut arī formāli iespējams izveidot analītisku formu „ir nezinājis”, tiek ņemts vērā, ka latviešu valodā šādām formām noliegumu veido ar palīgdarbības vārda „nebūt” finītām formām”, piemēram, „nav zinājis”. Tāpēc šajā gadījumā divdabis veic izteicēja funkciju un iegūst lomu **pred.**

Tāpat netiek attēlota tāda lietvārda redukcija, kuru raksturo skaitļa vārds, piemēram, **“Trīs Satversmes tiesas tiesnešus apstiprina pēc ne mazāk kā desmit Saeimas deputātu priekšlikuma, divus – pēc ministru kabineta priekšlikuma un vēl divus – pēc Augstākās tiesas plēnuma priekšlikuma”**. Kaut arī šajā piemērā var skaidri noteikt reducēto lietvārdu, citos tā var nebūt, jo reducētais lietvārds atrodams kontekstā ārpus teikuma. Abos gadījumos tiek pieņemts, ka skaitļa vārds reducēta lietvārda gadījumā var veikt lietvārda funkciju.

Atšķirīgi tiek vērtēti gadījumi ar lietvārda redukciju frāzēs ar prievārdiem **“no – līdz”**, piemēram **“no diviem līdz pieciem kilometriem stundā”**. Reducētais lietvārds tiek atjaunots, bet kopējais atribūts abām prievārda konstrukcijām tiek pakārtots visai vienlīdzīgu teikuma locekļu konstrukcijai, kā redzams attēlā 2 – 4.



2-4. attēls. Reducēts lietvārds konstrukcijā "no 209 līdz 226 latiem par 1000 litriem"

2.2. Negramatisku teikumu marķēšana

3. Specifiskie gadījumi

3.1. Konstrukcijas, kas attēlojamas ar vienu virsotni

Tāpat ar vienu virsotni tiek attēloti saīsinājumi — **utt.**, **u.tml.**, **u.c.**, **utjp.**, **funkc.**, **k-gs** — un skaitļi, piemēram, kārtas skaitļi, daļskaitļi, pulksteņlaiki un skaitļi, kas kā tūkstošu atdalītāju izmanto atstarpi — **13.**, **14,57**, **15:34:02**, **114 701**.

Ar vienu virsotni tiek attēloti arī tādi saīsinājumi, starp kuru daļām iespējama atstarpe:

- skaitļi un telefona numuri (xn un xx), piemēram, 2 358 000 un +371 20001234;

saīsinājumi *u.t.jpr.*, *u.c.*, *u.tml.*, *v.tml.*, *u.t.t.*, *N.B.*, *P.S.* un *P.P.S.* ar jebkuru *P.* daudzumu, ja tajos ir atstarpes.

3.2. Vairākvārdu saikļi un partikulas

Vairākvārdu saikļi un partikulas tiek marķēti kā xvārds **xFunctor** (skatīt 1.2.1.7. nodaļu). Vairākvārdu saikļiem morfoloģiski vismaz viens no vārdiem tiek marķēts kā saiklis, bet otrs var būt arī partikula.

NB! **Kaut kas**, **kaut cik**, **kaut kur** un citus vārdu savienojumus, kam izpildās abi nosacījumi — pirmkārt, vārdu savienojumus satur vismaz vienu patstāvīgu vārdu, un, otrkārt, vārdu savienojums nav partikula — zīmē kā **xParticle** (sk. 1.2.1.6).

3.3. Partikulu marķēšana

Partikulu novietojums kokā atkarīgs no tā, vai partikula ir piesaistīta konkrētam vārdam, vai attiecināma uz visu teikumu.

Partikulu marķēšanas principi:

- partikulas, kas, līdzīgi kā prievārdi, piesaistās pie konkrēta vārda fiksētā pozīcijā, piemēram, **ik**, **it**, **ne**, zīmējamas ar **xParticle** (sk. 1.2.1.6);
- partikulas, kas attiecināmas uz konkrētu vārdu/teikuma locekli, bet var būt brīvi novietotas pirm vai pēc vārda, zīmējamas ar „atkarību" un lomu **no** (sk. 1.1);
- partikulas, kas nav attiecināmas uz konkrētu teikuma elementu, zīmējamas kā tās PMC konstrukcijas, kas ietver attiecīgo teikumu vai teikuma daļu (piemēram, **sent**, **utter**, **mainCl**, **subrCl**, **spcPmc**), sastāvdaļu ar lomu **no** (sk. 1.2.3).

3.4. Sintaktisko lomu izvēle dažādos gadījumos

Konstrukcijas tips	Nozīme	Sint. loma	Piemēri
lietvārda ar -šana/-šanās pakārtotie	objekta	attr	runāšana **par lietu**
	apstākļa	spc	skriešana **pa lietu;** ārstēšanās **slimnīcā (arī pie -ums –** grozījumi civillikumā, ierobežojumi darījumos)
	subjekta	attr	**mātes** čalošana
divdabja -ts pakārtotie	objekta	obj	**par** noteiktu **tēmu** rakstītais [raksts]
	apstākļu	adv	**vakar** runātais
	subjekta	spc	**mātes** teiktais
prievārdiska konstrukcija pie lietvārda	objekta	attr	sapnis **par māju**
	apstākļa	spc	māja **uz kalna**
	piederības	spc	lapa **ar sarakstu**
	līdzbiedrs	spc	mēs **ar tēvu**
pārākā pakāpe + "par + X"	objekts, ar ko salīdzina	obj	lielāks **par mani**
datīvs pie verba	mērķis	obj	šis pasākums paredzēts **labdarībai**
	izjutējs, īpašnieks	det	**Man** ir darbs, **Tev** pieder māja, **Man** ir bail?

	adresāts	obj	kā var sist **sunim**!
	subjekts	subj	vajadzības izteiksme "**man** ir jādara", ar modificētājiem "**man** gribas atpūsties", "**man** laimējās noķert"
Datīvs pie divdabja	Izjutējs, īpašnieks	??	**Nemierniekiem** simpatizējošs radio (Pie finītās formas vārdam būtu determinanta loma?)
divdabja -dams pakārtotie	subjekts	spc	pārbaudes **modeļu** izpildāmie faili
ģenitīvs pie verbālnomena	objekts	attr	**vadlīniju** izstrāde
datīvs pie lietvārda	objekts	attr	adresāts (vēstule **mātei**), mērķis (kompanjone **ceļojumam**, prognoze **nākotnei**)
Datīvs pie lietvārda apstākļa nozīmē		??	**Pasaulei** otrā pusē viss ir savādāk (vai atšķirtos lomas no "pasaules otrā pusē"?)
lietvārds pie īpašības vārda	objekts	obj	pilna **gaismas**, pilns **āboliem**
lietvārds pie adverba	Adverbs nosauc attieksmi, kādā atrodas pret objektu	obj	kopā **ar kādu**, atbilstoši **restorānam**
Nenoteiksme pie lietvārda		spc	Tiesības **pretendēt**, iespēja **apprecēties**
Lokatīvs ar objekta nozīmi		obj	Dalīties **jaunumos** (dalīties ar jaunumiem, nav tas pats, kas dalīties gabalos, kas būtu tipiskā lokatīva nozīme).

4. Virsotņu marķējums

Katrai virsotnei var norādīt noteikta veida marķējumu atkarībā no virsotnes veida. Šajā nodaļā ir izskaidrots, kādi marķējuma elementi (lauki) katra veida virsotnei ir un ko tie nozīmē.

Visām **apaļajām virsotnēm** (izņemot koka sakni) ir šādi lauki:

- **ord** — kārtas numurs (obligāts tām virsotnēm, kam atbilst kādi teksta primitīvi). Šis numurs norāda virsotņu secību kokā. Kārtas numuri drīkst nebūt pēc kārtas, taču ir ļoti svarīgi, lai kārtas numurs virsotnēm, kas atbilst teksta primitīviem (vārdiem vai pieturzīmēm), būtu augošā secībā atbilstoši teksta primitīvu secībai tekstā. Tas ir, pirmā vārda numuram jābūt mazākam par otrā vārda numuru, otrā vārda numuram — par trešā, utt.
- **role** — loma — teikuma loceklis vai palīgteikuma veids virsotnēm, kas kokā veido atkarību no vecāka, vai kāda no lomām **basElem**, **punct**, **auxVerb** u.c, ja šī virsotne attēlo vecāka sastāvdaļu. Obligāts lauks.
- **reduction** — šeit ieraksta redukcijas tagu (sk. 2. nodaļu) un aiz tā — apaļajās iekavās — reducēto vārdformu, ja to var noteikt. Šis lauks ir jāaizpilda visām redukcijas virsotnēm un jāatstāj tukšs visām citām, jo šī lauka aizpildījums ir tā pazīme, kas identificē, vai šeit ir redukcija.

Virsotnēm, kurām **atbilst kāds teksta primitīvs** (vārds vai pieturzīme), ir vēl citi lauki — tām ir viens **m** lauku bloks (apraksta morfoloģiju) un patvaļīgs skaits **w** bloku (katrs bloks apraksta vienu teksta primitīvu — piemēram vārdam „**3,14**" var būt trīs **w** bloki).

- **m/w/token** — tekstā lietotā vārdforma
- **m/w/no_space_after** — 1, ja starp doto teksta primitīvu un nākamo netiek lietota atstarpe (piemēram, ja šis ir „**(**" vai nākamais ir „**.**").
- **m/form** — pareizā vārdforma. Šis lauks var atšķirties no **m/w/token**, ja tekstā ir bijusi drukas kļūda vai ja vārds sastāv no vairākiem teksta primitīviem (piemēram, ja **m/w/token** ir **zabaks**, tad **m/form** ir **zābaks**) (obligāts).
- **m/form_change** — ja **m/form** atšķirās no **m/w/token**, šeit tiek norādīts veiktā labojuma veids. (TRED logā vietā, kur ir redzams teikums, parādās dzeltenā krāsā iekrāsots ievietots vai labots elements, izņemot gadījumu, kad laboti kopā sarakstīti vārdi (kopā sarakstīts "kaut kas")). Iespējamās vērtības (vajag likt visas, kas atbilst):
    - **spell** — izlabota drukas kļūda;
    - **punct** — ievietota vai izņemta pieturzīme (neattiecas uz situāciju, kad pieturzīmi aizstāj ar citu pieturzīmi);
    - **insert** — ielikta jauna tekstvienība, kas tekstā bija izlaista (šobrīd ievietojam tikai trūkstošu pieturzīmi, ne vārdus);
    - **spacing** — ielikta oriģinältekstā izlaista atstarpe;
    - **union** — apvienotas divas vai vairākas tekstvienības vienā vārdā;
    - **num_normalization** — mainīts skaitļu pieraksts, piemēram, no AM/PM uz Latvijā lietoto, utt. — *vai šāds vispār datos ir lietots?*

  Piemēri:
    - ielikts trūkstošs komats — **punct** un **insert**;
    - izdzēsts lieks komats, to apvienojot ar iepriekšējo vārdu — **punct** un **union**;
    - vārdam pielikta trūkstoša garumzīme — **spell**;
    - atsevišķi rakstāmi vārdi ("**kaut kas")** sarakstīti kopā ("**kautkas**") — **spacing** pie abām daļām (atdalīt prasa Laumai);
    - kopā rakstāms vārds ("**jādara**") uzrakstīts atsevišķi ("**jā dara"**) — **spacing** un **union**;
    - nepareizi satokenizēts, bet ar atstarpēm viss ir kārtībā ("**1.**" rakstīts bez atstarpēm, bet sadalīts divos w blokos "1" un "."; **"nozare.lv**" ir rakstīts bez atstarpēm, bet sadalīts trīs **w** blokos — "**nozare**", "**.**", "**lv**"; **"kaķ's"**) — tikai **union**;
    - kopā rakstāms vārds uzrakstīts atsevišķi un vēl kļūdaini ("ne specigi" >> "nespēcīgi") – **union** , **spell** un **spacing** (apvieno virsotnes Tredā marķētājs pats).

NB! Ja tiek atdalīts defissavienojums, piemēram, "prasību-pārbaužu", "pieņemšanas-nodošanas", tad **m/form_change** lauks nav jāaizpilda (atdalīšanu veic Lauma, jāprasa atsevišķā github biļetē).

NB! Ja tiek atdalīta pieturzīme no skaitļa, tad **m/form_change** lauks nav jāaizpilda (atdalīšanu veic Lauma).

- **m/lemma** — vārda pamatforma vai nenoteiksme (obligāts).
- **m/tag** — vārda morfoloģiskais marķējums (obligāts).
- **m/src.rf** — morfoloģiskā marķējuma avots (neobligāts).

PMC virsotnēm ir šādi lauki:

- **pmctype** — konstrukcijas tips (**sent**, **insPMC** utt.). Obligāts lauks.
- **ord** — neobligāts lauks, kas var tikt aizpildīts, lai panāktu, ka TrEd uzskatāmāk attēlo koku. Par sintaktisko/morfoloģisko marķēšanu šis lauks informāciju nenes. Tipiski tiek aizpildīts ar automātiskiem līdzekļiem.

Sakārtojuma virsotnēm ir šādi lauki:

- **tag** — morfoloģiskais marķējums, kas parāda, kādu lomu šis sakārtojums izpilda teikumā. Šis lauks ir jāaizpilda tad, ja vienlīdzīgie elementi ir vārdi vai x-vārdi, bet jāatstāj neaizpildīts tad, ja vienlīdzīgie elementi ir palīgteikumi vai virsteikumi.
  Ja vienlīdzīgie elementi visi ir vienas vārdšķiras, marķējumu nosaka pēc morfoloģiskās tabulas nodaļas, kas atbilst attiecīgajai vārdšķirai.
  Ja vienlīdzīgie elementi ir dažādu vārdšķiru:
  - lietvārds + vietniekvārds + … → x-vārdu marķē kā lietvārdu;
  - īpašības vārds + divdabis + … → marķē kā īpašības vārdu.
- **ord** — neobligāts lauks, kas var tikt aizpildīts, lai panāktu, ka TrEd uzskatāmāk attēlo koku. Par sintaktisko/morfoloģisko marķēšanu šis lauks informāciju nenes. Tipiski tiek aizpildīts ar automātiskiem līdzekļiem.

X-vārdu virsotnēm ir šādi lauki:

- **xtype** — konstrukcijas tips (**xPrep**, **namedEnt** utt.). Obligāts lauks.
- **tag** — morfoloģiskais marķējums, kas parāda kādu lomu šis x-vārds izpilda teikumā. Lauks plānots obligāts, bet šobrīd koncepcija, kā veikt marķēšanu, izstrādāta tikai daļēji.

X-prievārdi tiek marķēti tā, kā aprakstīts morfoloģiskā marķējuma tabulā — formā *marķējums-pēc-lietvārdu-tabulas*[*marķējums-no-x-prievārdekļa-tabulas*] (ja tas ir vārdu savienojums ar lietvārdu) vai *marķējums-pēc-vietniekvārda-tabulas*[*marķējums-no-x-prievārdekļa-tabulas*] (ja tas ir savienojums ar vietniekvārdu), bet netiek norādīta 7.3.pazīme (t.i. pašu prievārdu marķējumā nerakstām).

X-izteicēji tiek marķēti tā, kā aprakstīts morfoloģiskā marķējuma tabulā — formā *marķējums-pēc-darbības-vārdu-tabulas*[*marķējums-no-x-izteicēja-tabulas*]. Tipus **modal_red** („es gribu saldējumu") un **phase_red** („es sāku darbu") nelietojam, jo redukcijas fakts parādās jau pašā kokā.

X-skaitļi tiek marķēti tā, kā morfoloģiskā marķējuma tabulā aprakstīta skaitļa vārdu marķēšana.

Pārejo x-vārdu marķēšanas principi vēl ir jāapzina.

- **ord** — neobligāts lauks, kas var tikt aizpildīts, lai panāktu, ka TrEd uzskatāmāk attēlo koku. Par sintaktisko/morfoloģisko marķēšanu šis lauks informāciju nenes. Tipiski tiek aizpildīts ar automātiskiem līdzekļiem.

Koka saknei ir tikai **ord** lauks, kas tāpat kā saliktajām un tukšajām virsotnēm var tikt aizpildīts, lai panāktu, ka TrEd uzskatāmāk attēlo koku, un šis lauks sintaktisko/morfoloģisko informāciju nenes.

4.1. x-vārdu morfoloģiskā marķēšana

Piezīmes:

- Ja **xPred** sastāvā ir prievārda konstrukcija, tad lietvārds ir akuzatīvā (var būt dsk. dat. pie tā prievārda), piemēram, *kļūt par prezidentu* — **xPred** morfoloģiskā marķējuma pēdējā pozīcija iekavās iespējama vēl cita, ne tikai nominatīva apzīmējums ***n*** — *vm0is003san[substsf**a**]*. Līdzīgi sastata nominālos izteicējos morfoloģiskā marķējuma pēdējā pozīcija ir datīvā, jo saliktā izteicēja bāzes elements ir datīvā — *būt sportistam — vm0nn000san[substsm**d**]*, *būt skaistam — vm0nn000san[adjsm**d**]*.

Versijas

Versija	Moduļa versija	Publiskās relīzes versija	Apraksts	Datums	Autors
0.1			Ar kaut ko ir jāsāk		Lauma
0.2			Drukas kļūdu labojumi		Lauma
0.3	2.4		Izteikumi		Lauma
0.4	2.5		Izteikumi un fragmenti svešvalodās	2011–05–31	Lauma
0.5	2.6		Pielikums; vispārinošais vārds	2011–06–07	Lauma
0.6	2.6, 2.7		„Vairāk kā”; iestarpinājumi	2011–07–26	Lauma
0.7	2.7, 2.8		Attēli ar pareizajiem sakārtojuma tipiem „uz siliem un mežiem”	2011–11–15	Lauma
0.8	2.9		Determinants	2011–12–07	Lauma
0.9	2.9		Komentārs par **mainCl** bez pieturzīmēm. Kaut kas, lai gan zīmēšana.	2012–01–23	Lauma
0.10	2.10		**xParticle**, **xSimile** „mazie vārdiņi”. 26. janv. seminārā ar Lokmani izrunātais.	2012–01–30	Lauma
0.11	2.11		Valodas uzlabojumi, kļūdu labojumi, tiešās runas.	2012–02–08	Lauma, Laura
0.12	2.12		Jaunākās bildes, iespraudumi un tiešās runas ar atkarībām, vārdgrupu un vārdrindu analogi. Izmests **abbr** un **numPmc**.	2012–07–30	Lauma
0.13	2.13		**dirSpPmc** var atrasties saknē.	2012–09–07	Lauma
0.14	2.13		Daudzi nelieli metodoloģiskie labojumi, „ne vairāk kā” u.c.	2013–04–04	Laura
0.15	2.14, 2.15		Ieviests **unstruct** x-vārds.	2013–04–29	Laura
0.16	2.14, 2.15		Precizēts, ko attēlo ar x-vārdiem **NamedEntity** un **phrasElem**	2013–07–23	Laura
0.17	2.16		Likvidēta sakārtojuma konstrukcija ar vispārinošo vārdu **crdGeneral** (attiecīgi arī lomas **Gen** un **genList**)	2014–02–25	Laura
			Ja vārds marķēts kā partikula, tas nevar būt adv funkcijā: teikumā *Šeit jau ir arī žurnāli... jau* pārmarķēts no adv uz no.	2016–04–28	Laura
	4.0		Detalizēts apraksts par **form_change**	2017–12–01	Lauma
	4.2		Ieviests gramatizējušos salīdzinājuma kosntrukciju termins, pieliktas piezīmes par to marķēšanu, izvākta nodaļa 2.3 par vairāk kā/nekā un vairāk par konstrukciju marķēšanu	2018–02–15	Laura
	4.3	UDv2.2	Pēc būtības mainīta vairākvārdu saikļu un partikulu marķēšanas nodaļa, pielikti xvārdu iespējamie apakštipi, tiešās runas nodaļā izvākta bilde par vairākiem teikumiem tiešajā runā. Ieviests ins pie pakārtojuma lomām (līdz šim bija tikai pie PMC).	2018–09–03	Laura
	5.0	Būs uz UDv2.4	Pievienota loma **repeat** un x-vārds **xFunctor**	2018–11–29	Lauma

	5.0	UDv2.4	Precizētas lomas **repeat** marķēšana (1.1.10.), kas ir virsotne, pielikta piezīme par "nav ko darīt"konstrukciju marķēšanu(1.1.1. un 1.1.2.), izņemts piemērs par piebildi tiešās runas vidū, ieliktas piezīmes, ka tajos gadījumos piebilde kļūst par iespraudumu (1.1.9. un 1.2.3.9.). Precizēta informācija par vairākvārdu saikļu un partikulu marķēšanu (3.2.).	2019–04–30	Laura
	5.1	UDv2.4	Redukcija kā atsevišķa nodaļa, sīki teksta precizējumi.	2019–05–21	Lauma