

Sündmuste märgendamise juhised

Sissejuhatuse ja sündmuse definitsioon

Käesolevad juhised on suuresti koostatud inglise keele TimeML juhiste ([1], [5]) eeskujul, kohandades neid vastavalt eesti keele omapäradele.

TimeML järgi [5] on „sündmus” katusmõiste, mis hõlmab situatsioone ja seisundeid tähistavaid keeleväljendeid. Keskendutakse just sündmuste mainimisele (ingl *event mentions*, *eventualities*, eesti *sündmusviide*, *sündmusele viitav keelend*) märgendamisele tekstis, ent erinevalt näiteks ACE sündmuste mudelist¹, ei pöörata olulist tähelepanu sündmuse rollistruktuuri märgendamisele (sündmuses osalejad ning sündmuse asjaolud). Seetõttu on TimeML oluliselt üldistavam ning *ei ole sõltuv eeldefineeritud sündmusstruktuuridest* (nagu nt ACE mudel, Framenet'i mudel jms). TimeML eesmärgiks on pindmine märgendus (ingl *surface-based annotation*), mis toob esile pigem sündmusanalüüsil olulised grammatilised tunnused, kui nende lõpliku interpretatsiooni.

Sündmused võivad olla punktuaalsed (Nt *Vulkaanist purskas tuhka, gaasi ja laavat*) või kesta mingi perioodi (Nt *2000 inimest evakueeriti epitsentri lähistel*). Samuti loetakse sündmuste alla predikaadid, mis kirjeldavad *olekuid* ning *olukordi/asjaolusid*, mil mingi väide või omadus kehtib – eeldusel, et need on tekstis ajaliselt määratletud (Nt *Gaasimaskide puudus on pikka aega sõjaväele kitsaskohaks olnud*).

Sündmuse asjaoludest leiavad TimeML raamistikus märgendamist ka ajaväljendid, nende märgendamise kohta vt dokumendist „Ajaväljendite märgendamise juhised”.

Sündmuse märgendamiseks tekstis ümbritsetakse see EVENT märgenditega, mille atribuutides tuuakse välja sündmuse tunnused. Näide:

```
Vulkaanist <EVENT class="OCCURRENCE" pos="VERB" tense="IMPERFECT">
purskas</EVENT> tuhka, gaasi ja laavat.
```

Ruumi kokkuhoiu mõttes edaspidi terviklikku märgendust välja ei tooda ning märgendamisele kuuluvad sõnad tähistatakse näidetes [allajoonimise ja sinise värviga](#). NB! Nii tuuakse välja eelkõige näite seisukohalt olulised sündmused, samas võib lauses olla ka tähistamata, ent siiski märgendamisele kuuluvaid sõnu.

SündmusVerbid

Üldjuhul loetakse grammatiline predikaat (tüüpiliselt finiidne verbivorm) sündmust tähistavaks (k.a olekuid tähistavad predikaadid).

Nt *Lähis-Idas asub teadaolevalt 66% maailma naftavarudest, seejuures üksi Saudi Araabias leidub veerand varudest. Peaminister ütles, et ta ei mõista, kuidas saab riigiamet endale lubada riigi rahadega prassimist.*

Kui sündmust tähistab lause, osalause või verbifraas, märgendatakse vaid üksuse verbiline peasõna.

Nt [*Ta müüs oma vara viiele inimesele*]; [*Ostan siis, kui müüd poole hinnaga*]; [*Ta on õnnelik, et saab võla tasuda*].

Keerukamate verbifraaside märgendamine:

- **Ahverbid** ($V_1 + V_{inf}$), kus infiniitverb V_{inf} on sisuline ning V_1 väljendab modaalsust, tegevuse faase, kausatiivsust vms infot;
 - **Modaalverbid** *võima, tohtima, saama, pidama, näima, paistma, tunduma* leiavad

¹ ACE mudel (<https://www ldc.upenn.edu/collaborations/past-projects/ace> , 2014-10-09) eristab entiteedide mainimisi (ingl *entity mentions*) ning entiteete (ingl *entities*). Sündmuse mainimise all mõeldakse mingit tekstilõiku (võib olla üsna minimalistlik, nt üks sõna), mis annab edasi sündmuse kirjelduse. Sündmuse entiteedi all mõeldakse reaalse maailma sündmuseid või kogumit tekstis olevatest sündmuse mainimistest, mis kõik viitavad ühele ja samale reaalse maailma sündmusele.

prantsuse ja itaalia TimeML ([3], [4]) eeskujul märgendamist ning nende eristaatus kajastub sündmuse klassimäärangus (`class=MODAL`). Modaalverb märgendatakse lahus infiniitsest verbivormist, Nt

- Nii [saaksime](#) aasta pärast läbi [viia](#) rahvahääletuse;
- Eesti filmistuudio [peab leidma](#) uusi sponsoreid;

Infiniitverb, mille tähendust mõjutatakse modaalverbi poolt, viitab modaalverbile atribuudi `modality` abil, kus tuuakse välja modaalverbi lemma.

- **Faasiverbid** *hakkama, asuma, saama, jääma* jms kuuluvad märgendamisele; Märgendatakse lahus käändelisest verbivormist, märkides, et tegu käändelist verbivormi aspektuaalselt piiritlevate sõnadega (`class=ASPECTUAL`);
 - Ta [hakkas sõõma](#); Noormees [jäi](#) tunnis [tukkuma](#).
 - Vigastatud Nani [jääb](#) murule [lebama](#);
- **Kausatiivverbid** *laskma, panema, ajama* jms kuuluvad märgendamisele, lahus infiniitsest verbivormist;
 - Ta [laseb](#) endale maja [ehitada](#);
 - Hispaania [pane](#)b jälle oma platsipoolet söödumängu [käima](#);
- **Koloratiivtarindis** võib koloratiivverb jääda märgendamata, märgendamisele kuulub vaid infiniitne verbivorm.
 - Tüdruk [kihistas naerda](#). Rahvas [rõkatab laulda](#);
- **Väljendverbid** (Verb + noomen) – fraseoloogiliselt terviklikud, idiomaatilised väljendid. Märgendatakse siiski vaid verb;
 - Poiss [laskis jalga](#); Pitsa [pandi nahka](#).
- **Ühendverbide** (Verb + afiksaaladverb) puhul puhul märgendatakse vaid verbiline osa:
 - Müüja [lasti](#) töölt [lahti](#); Pätt [pannakse kinni](#); Poiss [jooksis](#) sillast [üle](#);
 - NB! *olema*-verbi puhul märgendatakse siiski mõlemad, nii verbiline kui ka adverbiline osa (vt allpool „Erikonstruksioonid“);
 - modaaladverb: *Vaja* [läheb](#) vett, *puid ja paberit*. *Tarvis* [oli](#) midagi enam.
 - NB! Modaaladverb võiks kajastuda predikaadi atribuutis `modality`;
- **Sariverbid** (seriaalkonstruksioonid): võib märgendada kõik osalevad verbid, eraldiseisvalt:
 - *Mine* [vii](#) taara ära!
 - Oskar [läks võttis](#) auto ja [sõitis minema](#);Alternatiiviks on märgendada ainult kõige olulisem sündmus, nt *mine* [vii](#);
- **Perifrastilised verbivormid** – suhteliselt sisuvaese verbi (*olema, saama, jääma, pidama* jm) ühendid käändelise verbivormiga, eeldatavalt regulaarse kasutusega ning annavad edasi seda tüüpi tähendust, mida tavaliselt antakse morfoloogiliste tunnustega [6]. Kui sisutühi verb kannab modaalset või aspektuaalset tähendust, leiab see märgendamist. Muudel juhtudel peaks sisutühi verb jääma märgendamata ning sellele võib viidata käändelise verbivormi märgenduse atribuutides (`modality, aspect, tense` jms);
 - Töö [saab olema](#) põnev. Ta [pidi](#) haige [olema](#).
 - Tund oli just [lõppemas](#). Raamatud on [loetud](#).NB! Kui edasiantav tähendus pole selgelt morfoloogiline või ühendi kasutus on vähemregulaarne, võib siiski märgendada ka mõlemad vormid:
 - Ka miljonäril on [õigus käia](#) Kadaka turul odavat tráni [ostmas](#).Erandiks on piiripealsed faasiverbilised / aspektuaalsed konstruksioonid;
- **Sekundaartarindite** (moodustajad, milles puudub alus ja öeldis) **käändelised verbivormid** kuuluvad sündmustena märgendamisele, kui need on *fraasi peasõna* positsioonis;
 - Jaan [kavatses](#) [kinno [minna](#)]; [Kuju [minnes](#)] [astus](#) ta meie poolt läbi.
 - Õpetaja on [juhtunust [ärritatud](#)]; See puuvili on [söödav];Piiripealsed on siin gerundiumid, mis dubleerivad või täpsustavad lause peasündmust (on sellega sama-aegsed), ning seega on küsitav, kas nende märgendamine midagi juurde annab. Nt:
 - ... [tõusis](#) [varasema aastaga võrreldes 6,3%], [ulatudes 865 euron].

- Inglise keele märgenduses on kohata ka selliste väljendite ignoreerimist; eesti keele puhul jääb märgendamine esialgu lahtiseks küsimuseks.

Sündmusnimisõnad

Sündmuse tähistavad nimisõnad peaksid olema kasutatavad ajamääruste kontekstis (ajaliselt määratletavad) [1], nt

NIMISÕNA kestis mitmeid sekundeid/päevi/aastaid/...

NIMISÕNA oli väga kiire/hetkeline...

NIMISÕNA leiab/leidis aset AJAVÄLJENDi ajal

NIMISÕNA algas/lõppes/jätkus AJAVÄLJENDi ajal

Üldiselt märgendatakse vaid nimisõnafraasi peasõna, nt

- Tavalised NF-id: [*finantsabi maailmapangalt*]; [*paavsti külastus*];
- Nimega/nimelised sündmused: [*külm sõda*]; [*prantsuse revolutsioon*]; [*Kaie sünnipäevapidu*];

1. Täienditega nimisõnafraasid: täiendid jäetakse üldiselt märgendusest välja (nt [*kunstiteaduse võistlus*], [*karnevalide konkurss*]). Samuti jäetakse tavaliselt välja nimisõnafraasi modifitseerivad oleviku kesksõnad (nt [*keerlev hoop*], [*tagasilööv lask*]) ning nominalisatsiooni täiendina sisaldavad nimisõnafraasid (nt [*aerutamise võistlus*]); Piiripealsed on juhud, kus nii sündmuseks oleva peasõna kui ka selle täiendi võib lugeda eraldiseisvateks sündmusteks, nt:

- aspektuaalne peasõna: [*blokaadi*] [*kehtestamine*];
- perseptiivne peasõna: [*mõrva*] [*pealtnägemine*];

Sellistel juhtudel võiks mõlemad sündmustena märgendada;

2. *Sortal states* [1]: Üldiselt väljendatud 1) agenttiivsete nominaalide kaudu (nt *tööandja*, *võitja*, *tarbija*); 2) nimisõnad tähistamas rolle ja ameteid (nt *arst*, *torumees*, *õpetlane*, *saatejuht*); 3) jäigad/täpsed viitajad (nt *esimene Eesti president*). Need leiavad märgendamist vaid juhtudel, kui tegemist on predikatiivse komplemendiga² (ingl *predicative complement*). Näited predikaadi tüüpidest, kus toimub märgendamine:

- Kopulatiivsed predikaadid:
 - *Praegu on ta [Õpetajate Liidu president].*
- Inhoatiivsed (oleku tekkimisele viitavad) predikaadid:
 - *1995 sai ta [esimeseks presidendiks].*
- Aspektuaalsed (*alustama*, *jätkama*, *lõpetama*) predikaadid:
 - *Ta jätkas [uurimissuuna juhina] 1925 aastal.*
- Oleku (rolli, ameti) muutust tähistavad predikaadid (*valiti*, *määrati*, *taandus*, *astus tagasi*):
 - *Aastal 1998 ta taandus ametikohalt [emeritprofessorina].*
- Olekut (rolli, ametit) hindavad või kirjeldavad predikaadid (*loetakse*, *arvatakse*, *hinnatakse*):
 - *Teda loetakse [esimeseks naissoost rock'n'roll'i lauljaks].*

Nende märgendamine on natuke vaieldav (ka inglise keelele on *Sortal states* eristamine tulnud sisse alles 2010. aasta TimeML märgendusjuhistes), seega võib need esialgu ka välja jätta.

3. Muud sündmusele viitavad nimisõnad (nominalisatsioonid jm) leiavad alati märgendamist.

- *Moskva tõlgendas [situatsiooni] [konfliktina] „demokraatia ja reaktioonilise liikumise vahel.“*
- [*Embargo kehtestamine*] Iraagile võiks põhjustada USA majanduses [*depressiooni*] ning maailmamajanduses [*kriisi*].

² Millegipärast tundub järgnevate näidete valguses, et ingl *predicative* on natuke teises tähenduses kui eesti predikatiivne / „õeldistäiteline“ (st peamiselt *olema* verbiga-seonduv);

Mõnikord on piir tegevuse/sündmuse ning selle tegevuse asjaolusid (nt instrumenti) tähistava nimisõna vahel ka hägune. Näiteid spordireportaažidest:

- vile kui vilistamine vs vile kui tulemus / instrument:
 - Kohtuniku vile [vaikib](#) ja [mäng jätkub](#).
 - Siin on vile instrument ning ei kuulu märgendamisele;
 - ... ehkki [vilet](#) ei [kõlanudki](#).
 - ... kuid enne on [kõlanud vile](#) Hollandi kasuks.
 - Need on piiripealsed juhud, kus vile („vilistamine“) ja selle edasikandumine („kõlamine“) viitavad sisuliselt samale sündmusele. Käesolevalt on märgendatud siiski mõlemad, kuigi võib vaielda, et selline märgendus on liiane.
 - Kohtuniku vile [lõpetas mängu](#).
 - Sellise konstruktsiooni puhul on „vile“ kindlasti eraldiseisev sündmus, mis osaleb verbi „lõpetama“ poolt loodud aspektuaalses konstruktsioonis.

SündmusOmadussõnad

Atributiivsed adjektiivid (st nimisõna täiendavad adjektiivid, nt *[ilus] aed*, *[talumatu] valu*, *[õiglane] kohus*) jäetakse märgendamata [1].

Predikatiivsed adjektiivid (esinevad lauses seisundimääрусena või öeldistäitena) märgendatakse järgmistel erijuhtudel:

- Kopulatiivsed predikaadid (*on*, *oli*, *näis*):
 - *Õpilased [näisid](#) [kurnatud](#) pärast kolme nädalat klasse*.
- Inhoatiivsed (oleku tekkimisele viitavad) predikaadid:
 - *Dissident tõi lahkumise põhjuseks elu [muutumise](#) [\[talumatuks\]](#)*.
- Aspektuaalsed (*alustama*, *jätkama*, *lõpetama*) predikaadid:
 - *Koer [jõi](#) siiski [\[kurjaks\]](#)*.
- Kausatiivsed predikaadid:
 - *Saatejuht [tegi](#) publiku tõesti [\[rõõmsaks\]](#)*.
- Olekumuutmise predikaadid üldiselt;
- Tajumise predikaadid (*kuulma*, *nägema* jne);
 - *Nad [paistsid](#) kahekesi koos nii [\[õnnelikud\]](#)*.
- Hindamise ja kirjeldamise predikaadid (*arvestama*, *kirjeldama*, *esitama*):
 - *Teda [on](#) sageli kirjeldatud [\[ekstsentrilisena\]](#)*.

NB! Predikatiivsed adjektiivid loetakse märgendatavaks vaid siis, kui need kirjeldavad *mittepüsivat* omadust või olekut ning rahuldavad vähemalt üht järgnevatest kriteeriumitest:

- Omadus on kindlalt muutuv, mittepüsiv; Nt
 - *Nägu [muutus](#) vihas [\[punaseks\]](#)*.
- Omadus on esitatud kui ajaliselt piiritletud või seotud; Nt
 - *Kuninga võim [\[oli nõrk\]](#) 16ndal sajandil*.
- Omadus on esitatud kui arvamus, teadmine või kellegi usk, või on vaieldav:
 - *Valitsus [väidab](#), et mõned neist [on](#) [\[dissidentlikud\]](#)*.

Problemaatilistes kohtades tuleks omadussõna jätta *pigem märgendamata*. Analoogselt nimisõnafraaside ja verbifraasidega, märgendatakse vaid omadussõnafraasi peasõna.

Erikonstruktsioonid ja muud juhud

- Olema-verbi konstruktsioonid:
 - *olema + nimisõna(fraas) (alusena)*.
 - Kui aluseks olevat nimisõna võib käsitleda eraldiseisva sündmusena, märgendataksegi see eraldiseisvana verbist:

- Postimehe vallandatud hinnasõda on tegelikult meie kõigi ühine õnnetus.
 - Postimehe pöörase allahindluse taga on soov suretada teised lehed välja.
- Kui *olema* verb koos alusega moodustab justkui terviku (nt seisundikirjelduse), võib märgendada mõlemad ühe sündmusena:
- Ei ole kahtlust, et asjad nii lähevad.
 - Tal on hirm, et midagi juhtub.
- Kui aluseks olev nimisõna on sündmusena ebamäärane või on raske määrata, kas tegu on aluse või mingi muu moodustajaga, võib ka märgendada ainult verbi:
 - Tal on praegu leping haigekassaga.
 - Mul on hea meel, et nii arvas ka tema.
 - Riiklik töötu abiraha on praegu 240 krooni kuus.
 - *olema* + *nimisõna(fraas)* (õeldistäitena).
 - Kui nimisõnafraas viitab inimese rollile või ametile, tuleks märkida nii *olema* verb kui ka rolli tähistav sõna (fraasi peasõna) ühe sündmusena.
 - Statistikaameti teatel oli 1995. aastal iga neljas Eesti elanik rahvaraamatukogu lugeja.
 - Kui nimisõnafraas viitab tegevusele/seisundile, tuleks samuti märgendada nii *olema* verb kui ka fraasi peasõna ühe sündmusena:
 - See on normitehniline täpsustus asja selgemaks muutmiseks.
 - *olema* + *nimisõna(fraas)* (määrusena).
 - Kui nimisõnafraas ei märgi otseselt tegevust/seisundit, vaid pigem tegevuse asjaolusid (koht, aeg, valdaja, tegija, kvantiteet jms), siis märgendatakse ainult *olema* verb:
 - Metsa all on praegu veel lund.
 - Tal on aeg-ajalt hanged toas.
 - See võimalus on palju ebamäärasemas tulevikus.
 - Kui nimisõnafraas viitab tegevusele/seisundile (nt seesütlevas, saavas, olevas käändes), tuleks märgendada nii *olema* verb kui ka nimisõnafraasi peasõna ühe sündmusena.
 - Soome lapiteki harrastajate seas on moes isevärvimine.
 - Nad on nõus kahtlusalused kohtule välja andma.
 - *olema* + *adj* (õeldistäide). Märgendatakse nii *olema* verb kui ka omadussõna ühe sündmusena:
 - Suhtumine euroliitu on ebastabiilne.
 - Ennustusi selles vallas on raske teha.
 - *olema* + *adverb*. Märgendatakse nii *olema* verb kui ka adverb ühe sündmusena:
 - Ta on valmis kahjud korvama.
 - Nüüd on tuulik taas püsti.
 - Stalin oli kindel, et ...
 - Tarvis oli midagi enamat.
 - *olema* + *kvantiteet*. Märgendatakse ainult *olema* verb.
 - Üks dollar oli siis 1000 rubla.
 - Sajandi alguses oli Hiiumaal üle 500 tuuliku.
 - 1994. aastal oli lugejaid 345 000.
 - Nimistu keskmine suurus ühe arsti kohta on 1900 isikut.
 - Perifrastilised verbivormid (Verb + käändeline verb) – märgendatakse käändeline verb;
 - Tund oli just lõppemas.
 - Raamatud on loetud.
- Inglise juhised on natuke vastuolulised: [2] (ning esialgsete TimeML juhiste [5]) järgi jäetakse koopula märgendusest välja; [1] järgi aga märgendatakse nii verbaalne predikaat (koopula) kui ka predikatiivne täiend, nt
- A woman [is] the new [head] of the British Medical Association.

- *If we will not [be] [ready], we will ask USA to delay the operation.*
- *Zarei [was] [in] charge of a program to clean cities from corruption.*

- **Eitust sisaldavad konstruktsioonid:**

- Eituse abisõna ei jäetaks märgendamata, eituse olemasolu tuuakse välja verbi atribuudis *polarity*.
 - *Alonso karistuslööki omasid ei leia.*
- Sõna *pole* märgendatakse samade põhimõtete järgi, nagu *olema*-verbi konstruktsioonid.

- **Tugiverbiühendid:** predikaadi funktsioonis üldise tähendusega tugiverb (*tegema, saama, teostama, andma* jms) ja tegevust näitab substantiiv või nimisõnafraas (NF).

Kuigi tugiverb ja NF viitavad koos ühele sündmusele (ning NF annab tihtilugu sündmust paremini edasi, nt *Kohtunik annab lõpuvile*), on otsustatud parema ühilduvuse huvides (nt väljendverbidega), et märgendamisele kuuluvad nii tugiverb kui ka selle poolt kontrollitav nominalisatsioon või NF:

- *de Jong ei saa korralikku pealööki tehtud;*
- *Ronaldo teeb pika soolo;*
- *Hispaanlased teevad Ronaldo vastu vea;*
- *Kumbki meeskond pole vahetusi teinud;*
- *... saab kollase kaardi; Portugal saab karistuslöögi;*
- *Kiirel vasturünnakul annab Nani hea söödu teisele äärele Ronaldole;*
- *Kohtunik annab lõpuvile;*
- V.a väljendverbilaadsed ühendid, kus märgendatakse vaid verb ning (sisuvaene) nimisõna jääb välja;
 - *kuid Stekelenburg saab palli kätte;*
 - *Nani ei saa kaaslase pikka söötu omaks;*
 - *Nani saab läbimurdel viga;*
- Inglise keele näited – sisuvaene verb (*make, get, do, have, take, put, set, let*) ning sündmusnominalisatsioon loetakse mõlemad märgendatavateks:
 - *[Demonstrations] have [taken] place in the last week.*
 - *They will [take] this into [consideration].*

Kui tekib kahtlus, kas tegemist on tugiverbiühendiga või mitte, võib ka märgendada ainult finitise verbi.

- **Aspektuaalsed konstruktsioonid:** märgendatakse nii aspektuaalne predikaat (nt verb *alustama, jätkama* või nimisõna *alustamine, lõpetamine*) kui ka selle täiend/komplement:

- *Ronaldo [jätkab] [mängu].*
- *[Sõja alustamine] kujunes kulukaks.*

Inglise keele näited:

- *They probably would have [began] [responding] to the plan with new proposals.*
- *The banks must wait at least 30 days before [closing the purchase].*

- **Inhoatiivsed konstruktsioonid:** väljendavad oleku tekkimist; märgendatakse nii inhoatiivne predikaat kui ka selle täiend / komplement:

- *Kurikael hakkas^{ASPECTUAL} plaanima^{L-STATE} pahandust.*

Inglise keele näited:

- *They could [become] [hostages].*
- *They just [got] [scared].*

- **Kausatiivsed konstruktsioonid:** väljendavad mingi sündmuse põhjustamist; Nt

- *Muutunud aastaaeg põhjustas korratusi liikluses.*

Inglise keele näited:

- *The [rains] [caused] the [flooding].*
- *John [caused] the [fire].*
- **Konstruksioonid funktsionaalsete nimisõnadega** (*temperatuur, suurus, kaal, intensiivsus jms*), mis väljendavad entiteedi mingit konkreetset või suhtelist mõõdet; Märgeandakse nii funktsionaalnimisõna kui osalause peaverb; Üldistavad tüübid:
 - *NF on X:*
 - *Venemaa [rahvaary] [on] umbkaudu 142,9 miljonit;*
 - *(ingl) The current USA [population] [is] above 300 million.*
 - *(ingl) The market [value] of the house [is] \$150000.*
 - *(ingl) Catalonia [has] a [population] of around 7 million people.*
 - *(ingl) The substance [had] a [temperature] of 3 K.*
 - *NF muutub / tõuseb / langeb ... (X-ilt) (Y-le):*
 - *Eelmise aasta neljandas kvartalis [tõusis] keskmine [brutokuupalk] aasta varasema võrreldes 6,3%, ulatudes 865 euron.*
 - *[Temperatuur] [langes] -21,4 kraadini.*
 - *(ingl) In the nine months, [net] [rose] 4.3% to \$525.8 million.*

Selline märgendusviis on suhteliselt uus, veel inglise TempEval-1 korpus on märgendusviis teistsugune: märgendatud on väärtused (summad), ent verb ja entiteet on jäetud märgendamata;

NB! Teatud juhtudel on väljendus minimalistlik, tuues esile vaid kõige olulisema – muutunud (arvulise) väärtuse. Nt mänguseisu edasiandmine jalgpallisündmuste kirjelduses:

- *1:1 Ronaldo viigistab!*
- *2:1 Ronaldo!*

Käesolevalt on märgendatud vaid verbiline osa; Alternatiiv oleks märgendada olekuna (*STATE*) ka mänguseis (eelnevates näidetes mustalt allajoonitud), mis vastaks varasemale TimeML märgendusjuhisele.

- **Mitmele sündmusele viitavad väljendid:** märgendatakse nagu tavalised sündmused, atribuudis *cardinality* tuuakse välja sündmuste hulk (määramatu hulga korral *cardinality*=“plural”);
 - *John [õpetas] esmaspäeviti ja teisipäeviti.*
 - *[Paduvihmad] põhjustasid [üleujutusi].*
- *Adverbid* on üldiselt sündmuse asjaolusid täpsustavad ning ei kuulu märgendamisele. Inglisekeelsetes juhistes [1] on toodud mõned piirijuhud, mis siiski on märgendatavad – teatud lokatiivsed adverbid:
 - *Me oleme siin, et jääda üsna pikaks perioodiks.*
 - *Me oleme siin, kuna siin otsustatakse edasine asjade käik.*
 Nende märgendamine eesti keele tekstides on praegu lahtine;
- **Mittesündmused.** Sündmusteks ei loeta nt üldistavaid/üldkehtivaid väiteid ja definitsioone (mis pole konkreetset ajastatavad):
 - *Eesti keel on soome-ugri keel.*
 - NB! Alternatiiv on sellised sündmused siiski märgendada, ent eristada neid tavalistest, ajastatavatest sündmustest sellega, et määrata *genericity*=“GENERIC”;

Sündmuse atribuutide määramine

Sisulised atribuudid

Üldiselt peaks sündmuse atribuutide määramine tuginema pindmistele tunnustele (ingl *surface linguistic cues*), st mittenõudma süvendatud analüüsi. Erandiks on sündmuse klassifitseeriv atribuut *class* [2].

Atribuut **class** – määrab sündmuse liigi. Võimalikud liigid:

- **REPORTING** – teate, kommentaari, arvamuse edastamise või vahendamise sündmus:
 - Obama [sõnutsi kahekordistab](#) see kava USA tootevõimet.
 - Politsei [teatel](#) nädalavahetusel rikkumisi [polnud](#).
- **PERCEPTION** – mingi teise sündmuse füüsiline tajumine:
 - Tunnistaja [nägi](#) meest jooksmas.
 - Võis [kuulda](#) sadu väikseid plahvatusi, ütles tunnistaja.
- **ASPECTUAL** – konstruktsioon, milles sündmus viitab mingi teise sündmuse algusele, jätkuvusele / keskosale või lõpule;
 - Initsieeriv: *alustama, algatama (kampaaniat), välja saatma (missiooni), käivitama (juurdluse);*
 - Reinitseeriv: *taaskäivitama (kampaania), taasalustama (tuumakõnelusi);*
 - Termineeriv: *katkestama (kohtumise), blokeerima (abisaatmise), murdma välja (vangistusest);*
 - Kulmineeriv: *lõpule viima (õpingud), lõpetama (miitingu);*
 - Jätkav: *jätkama, viima edasi (analüüsi), säilitama (tasakaalu);*
 - Näiteid:
 - Nani ... [jääh](#)^{ASPECTUAL} murule [lebama](#)^{STATE}.
 - Aga kui teist sündmust pole: ... kaitsja [jääh](#)^{OCCURRENCE} murule pikali.
 - Aga:
 - Üksikmängus [lõppes](#)^{OCCURRENCE} Reino ja Saluste jaoks [turniir](#)^{OCCURRENCE} avaringis.
 - („turniir“ kui selline kestis edasi, kuigi lõppes Reino ja Saluste jaoks)
- **I_ACTION** – tahtlik / intentsionaalne tegevus (aga mitte ainult), mille argumendiks peab olema samuti sündmus, mis on tekstis ilmutatud kujul. Mõned representatiivsed verbid:
 - *üritama, püüdma; (ingl attempt, try)*
 - Mõned globaalsed firmad [püüavad monopoliseerida](#) ligipääsu Internetile.
 - *uurima; süvenema; (ingl investigate; look at; delve)*
 - Politsei [uurib](#) juhtumit.
 - Inspektor [süvenes](#) mõrvade uurimisse.
 - *takistama; viivitama; pidurdama; edasi lükkama; (ingl delay, postpone; set back)*
 - Vaidlused [lükkasid](#) miljoniprojekti aastateks edasi.
 - Kas puhkusetasu [maksmisega](#) võib [viivitada](#)?
 - *vältima, ära hoidma, tühistama; (ingl avoid, prevent, cancel)*
 - Serbia [vältis](#) kaotust;
 - Kohus [tühistas](#) politsei trahviotsuse.
 - *paluma; käskima; veenma; nõudma; (ingl ask; order; request; command)*
 - Holland [nõudis](#) võimude kiiret [sekkumist](#) konflikti.
 - Hesburger [palus](#) töötajail salatist ussikesed välja [korjata](#).
 - *lubama; pakkuma; nõustuma; otsustama; (ingl promise; offer; assure; agree; decide)*
 - Värske Euroopa meister [nõustus](#) arsenaliga [liituma](#).
 - Merko [otsustas](#) dividende mitte [maksta](#).
 - Suurepärane vindiga [kauglöökk](#) tagumise posti kõrvale ei [jätta](#) Rui Patriciole

- vähimatki võimalust;
 - Hollandlaste jäme valesööt lubab minna Portugalil otse värava peale;
 - laskma (käest);
 - Uskumatu, millise võimaluse Nani käest laseb
 - vanduma; tõotama; (ingl swear, vow)
 - Taliban tõotas kättemaksu;
 - nimetama; määrama; teadustama; (ingl name; nominate; declare; proclaim)
 - Ja mees määratigi presidendi ametikohale.
 - väitma; kinnitama; vihjama; (ingl claim, allege, suggest)
 - Putin kinnitas Venemaa WTOsse astumise;
- **I_STATE** – tahtlik / intentsionaalne olek või oleku säilitamine, mille argument on samuti märgendatav sündmus. Mõned representatiivsed verbigrupid:
 - uskuma; arvama; kahtlustama; kindel olema; (ingl believe; think; imagine; suspect; be sure);
 - David kahtlustas mind Victoria võrgutamises;
 - Saddam oli kindel, et elab sõja üle;
 - tahtma; ihaldama; janunema; (ingl want, love, like, desire);
 - Aga väiksena tahtis ta saada hoopis ninjakilpkonnaks.
 - lootma, ootama, plaanima; (ingl hope; expect; aspire)
 - Ema lootis, et poeg tuleb koju.
 - Portugal plaanib vahetust;
 - kartma; muretsema; pelgama; (ingl fear; hate; worry; be afraid)
 - Lumememm pelgas, et tuleb kevad.
 - Ka Kati Murutar kartis esinemist!
 - vajama; nõudma; (ingl need, require; demand)
 - Auto vajas põhjalikku remonti.
 - Ühest väravast Hollandile aga veel ei piisa, edasipääsuks vajatakse ka teist;
 - Holland vajab kahte väravat.
 - olema valmis; olema valmistunud; (ingl be ready; be eager; be prepared)
 - Sukles oli valmis palgamõrva eest maksma;
 - olema suutlik; suutlik mitte olema; (ingl be able; be unable)
 - Õpilane on suutlik iseseisvalt õppima.
 - Moutinho ei suuda samuti väravat lüüa.
 - Portugal ei suuda kõiki oma võimalusi ära kasutada.
 - NB! ingl could ei kuulu märgendamisele, ent eesti keele suutma on praeguse seisuga ikkagi märgendatav.
 - valitsema
 - valitseb liiga pragmaatiline suhtumine
- **STATE** – kirjeldab tingimusi/asjaolusid, milles midagi kehtib või on tõene.
 - Pole teada, kas -21,4 jääb tänaseks külmarekordiks.
 - Mured lääne kodanike turvalisuse pärast kasvasid.
 - Iraagis on 17 miljonit kodanikku.
 - Collins on ka varem olnud kaaspiloot.
- **MODAL** – eriklass, kuhu kuuluvad modaalsust väljendavad verbid võima, tohtima, saama, pidama, näima, paistma, tunduma.
 - Nii saaksime läbi viia rahvahääletuse.
 Siia alla kuulub ka sõna tulema kasutus modaalses tähenduses: Meil tuleb hakata tegutsema.
- **OCCURRENCE** – klass, kuhu kuuluvad kõik ülejäänud märgendatavad sündmused.

NB! Kui ei suuda klassi otsustada, võib ka märkida mitu varianti, eraldatuna püstkriipsuga, nt OCCURRENCE | STATE.

Andre Bittar'i [3] soovitatud lisaklassid (prantsuse keele põhjal):

- **EVENT_CONTAINER** – verb, mille subjektiks (aluseks) on sündmus (nominalisatsioon).
Nt
 - *Õnnetus*^{OCCURRENCE} *toimus*^{EVENT_CONTAINER} *täna öösel kella 01.30 paiku.*Esialgu on nende märgendamine küsitav.
- **CAUSE** – kausatiivverbid, mis märgivad sündmuse põhjustamist. Kui sündmus A põhjustas sündmuse B, märgendatakse mõlemad ning lisatakse kausaalsuse seosed. Kui põhjustav sündmus A pole tekstis sündmusena märgendatav, lisatakse selle kohta ilma tekstilise sisuta EVENT märgend.
 - *Muutunud aasta*^{CAUSE} *põhjustas* *korratust*^{OCCURRENCE} *liikluses.*
 - Siin jääb „*Muutunud aasta*“ sündmusena märgendamata, selle kohta lisatakse ilma tekstilise sisuta EVENT märgend.

Järgnevast atribuutide loetelust tuleks käsitsi märgendada vaid *polarity* ja *modality*.

Atribuut **pos** – määrab sündmuse sõnaliigi: *VERB*, *NOUN*, *ADJECTIVE* või *OTHER*.

Atribuut **tense** – määrab pöördelise verbi grammatilise aja. Käändelise verbi puhul täpsustatakse siin verbi tüüp (vanas TimeML spetsifikatsioonis atribuut *vform*). Morfoloogia tasemel määratakse väärtused:

- **INFINITIVE** – Siia alla kuuluvad *da-* ja *ma-*infiniitivid ning käändelised verbivormid *maks*, *mas*, *mast*, *mata*.
- **PASTPART** – Siia alla kuuluvad infiniitsed mineviku kesksõnad *-nud* ja *-tud*.
- **PRESPART** – Siia alla kuuluvad infiniitsed oleviku kesksõnad *-tav* ja *-v*.
- **GERUNDIVE** – Siia alla kuulub infiniitne *des-*vorm.
- **PAST** – Siia alla kuuluvad pöördelised tingiva ja kaudse kõneviisi minevikuvormis verbid.
- **IMPERFECT** – Siia alla kuuluvad pöördelised kindla kõneviisi lihtminevikuvormis verbid.
- **PRESENT** – Siia alla kuuluvad pöördelised kindla, tingiva ja kaudse kõneviisi olevikuvormis verbid ning kõik käskiva kõneviisi verbivormid.

Atribuut **mood** – määrab sündmusverbi kõneviisi (*INDICATIVE*, *CONDITIONAL*, *IMPERATIVE*, *QUOTATIVE*, *NONE*); Kasutusel ainult pöördeliste sündmusverbide puhul (verbi morfoloogiline tunnus), muude sündmuste puhul võib jääda määramata või saada väärtuse *NONE*.

- **INDICATIVE** (kindel): *Rolando väravad kukutasid Hollandi.*
- **CONDITIONAL** (tingimuslik): *Paljud investorid tahaksid aktsiaid osta.*
- **IMPERATIVE** (käskiv): *Praegu valitseb „oota ja vaata“ meeleolu.*
 - NB! Eitus siia alla ei kuulu: *Nad ei oota meid.*
- **QUOTATIVE** (kaudne): *See ei kõlbavat.*

Atribuut **polarity** – märgib, kas tegu on sündmuse jaatamise või eitamisega; vaikeväärtuseks on *positive*, seega tuleb märkida vaid negatiivsete juhtude korral; Nt

- *Nad ei lõpetanud*^{NEG} *ettekannet.*
- *Nii ei jõua*^{NEG} *me kunagi valmis.*

NB! Eituse skoop: eitus peaks hõlmama kõiki eituse markeri alla kuuluvaid sündmuseid antud fraasis või osalauses, Nt

- See ei ole^{NEG} ainuüksi rumal^{NEG}, vaid ka ohulik^{POS}
- Ei saa keelduda^{NEG} alandlikult uut elamist oodanud^{POS} inimesele korterit andmast^{NEG}
- Nad ei ole isiklikult nõusolekut lahkuda^{NEG} andnud^{NEG}

Osalausepiiridest kaugemal eitust siiski ei märgita:

- Ma ei usu^{NEG}, et poliitikud loovutaksid^{POS} võimaluse riigipea koha ümber kombineerida^{POS} ja intrigeerida^{POS}.

Atribuut **modality** – kasutusel kui sündmust modifitseerib modaalne sõna (nt ingl *may, might, must, would, should, could, can, have to, shall*); Väärtuseks modaalse sõna algvorm;

- Modaalverb:
 - John peaks lahkuma (modality=pidama) juba täna õhtul.
 - Nad võivad kohe siia jõuda (modality=võima).
 - Käsitsimärgendamisel: kui modaalne sõna on juba märgendatud (EVENT class="MODAL"), siis võib *modality* ka välja jätta.
- Mõni muu modaaltähendust kandev sõna:
 - Võib-olla nad jõuavad (modality=võib-olla) siia.

Atribuut **aspect** – grammatiline aspekt. Kuna grammatiline aspekt ei avaldu eesti keeles selgete tunnuste kaudu, siis jääb see esialgu märgendusest välja. Aspekti märgendamise võimalused vajavad edasist uurimist.

Atribuut **type** – nn leksikaalne aspekt ehk sündmuste liigitamine Z.Vendleri (1967) klassifikatsiooni eeskujul, eristades kolme üldist sündmustüüpi:

- **PROCESS** – protsess/tegevus ehk sündmus, millel puudub kulminatsioon (Nt *Mari jooksis*). Kehtib alamintervalli omadus: kui *Mari jooksis pool tundi*, siis selle *poole tunni* sees suvalise ajalõigu kohta võib öelda, et *Mari jooksis* selle ajalõigu vältel. Kuna puudub kulminatsioon, pole võimalik loendada, *mitu korda* protsess-sündmus toimus.
- **TRANSITION** – saavutused/sooritused ehk sündmused, millel on kulminatsioon (Nt *Jack ehitas maja valmis*). Alamintervalli omadus siin üldiselt ei kehti.
- **STATE** – staatilised sündmused; kirjeldavad kestvaid situatsiooni või situatsiooni algust (inseptiivsed). Alamintervalli omadus kehtib.

Esialgu on ka leksikaalse aspekti määramine väikese prioriteediga – selle võib määramata jätta.

Atribuut **comment** – kasutatakse vaid siis, kui märgendus on mingil määral problemaatiline või vaieldav – sisuks lühikommentaar probleemi kohta.

Tehnilised atribuudid – *eiid* ja *eid*

Et oleks võimalik märgendada sündmustevahelisi ajalisi seoseid, peab igal sündmusel olema küljes teksti piires unikaalne identifikaator, mille kaudu saab sündmusele viidata. Seda atribuuti nimetatakse *eiid* on ning selle väärtused on kujul *eiN*, kus *N* on sündmuse unikaalne järjekorranumber. Kui üks sündmus koosneb mitmest eraldiseisvast sõnast, lubatakse *eiid* väärtusel korduda – sama väärtust kasutatakse kõikide sündmuse alla kuuluvate sõnade juures. Järgnevalt on toodud lause „Pärast väikest seisakut on Nani valmis jätkama“ märgendus:

```
Pärast väikest <EVENT eiid="ei1" class="STATE">seisakut</EVENT> <EVENT  
eiid="ei2" class="STATE">on</EVENT> Nani <EVENT eiid="ei2" class="STATE">  
valmis</EVENT> <EVENT eiid="ei4" class="OCCURRENCE">jätkama</EVENT>.
```

Eelnevas näites on mitmesõnaline sündmus „on valmis“, mille mõlemad märgendused kuuluvad ühe *eiid* alla – *ei2*.

Probleemsed juhud

Järgnevates näidetes on võimalikud sündmused tähistatud allajoonimisega, tegelikult märgendatud sündmused aga tähistatud sinise värviga.

- **Sündmuste olulisuse probleem:** milline on **vähem-aktiivne** sündmus tekstis, mille võiks märkimata jätta?
 - Võitja selgitavad penaltid!
 - Kuna tekstis järgnes penaltide detailne kirjeldus, tundus, et *penaltid* märgendamine sündmusena võib olla ka üleliigne.
 - Võib ka tõlgendada nii, et penaltid (sündmused) põhjustavad (selgitavad) sündmuse võit.
 - Kriteerium, mida võiks katsetada oluliste sündmuste valimisel:
 - kui sündmus-sõna ei ole verb, aga on mõne märgendatud sündmusverbi vahetu argument
ning
 - kui seda sündmus-sõna mainitakse tekstis rohkem kui üks kord (selle järgi on justkui tegemist on olulise sündmusega)
 - Sellisel juhul tuleks sündmus märgendada. Näide:
Seejärel siirduvad kuus meest Tour de Skile, ent Saarepuud nende hulgas pole.
"Loomulikult tahaksin Tour de Skil kaasa lüüa ja loodan, et mind lülitatakse sõitjate sekka, " teatas Saarepuu .
 - Kuna sündmust (*Tour de*) *Ski* mainitakse eelnevas tekstis rohkem kui üks kord, ning mõlemal korral on see märgendatud sündmuse (sündmusverbi) vahetu argument, on see ka märgendatud.
- **Mitmene tõlgendus võimalik** (spordireportaažid):
 - Ronaldo karistuslöök lendab rikošetist nurgalöögiks;
 - Portugali löök lendab napilt üle;
 - Iseenesest tundub „löök lendab“ olevat keelekomistus (ikkagi „*pall lendab*“, mitte „*löök lendab*“); „löök“ on siin märgendatud sündmusena, kuna see on antud tekstis väga sage; Kui sõna oleks „*pall*“, tuleks see märgendamata jätta.
 - Esimeses lauses: lahtine küsimus on, kas „*lendab*“ tuleks kuidagi eriliselt märkida (ühe sündmuse üleminek teiseks)? Tegemist pole CAUSE-tüüpi sündmusega (pigem on *rikošett* see, mis tegelikult põhjustab nurgalöögi.)
 - Patricio püüab van der Vaarti tsenderduse;
 - Analoogselt tundub keelekomistusena (*püüab palli, mitte tsenderduse?*);
 - Busquets teenib kollase kaardi.
 - Kaks sündmust: Busquets teeb midagi ning Busquets saab selle eest kollase kaardi; Teatud mõõndustega võiks ka ainult kollase kaardi sündmuseks lugeda, kuna see on olulisem (pealegi – infostruktuur: lõpus on uus info). Vrdl nt juhuga, kui lauses on eksplitsiitselt teenimise põhjus olemas:
 - Pepe teeb Xabi Alonso vastu vea ja teenib kollase kaardi.
 - Hispaania vahetab: Silva puhkama, Jesus Navas mängu.
 - Alternatiiv oleks märkida sündmustena ka „*puhkama*“ ja „*mängu*“;

Kirjandus

[1] *Annotating Events in English: TimeML Annotation Guidelines*. Roser Saur¹, Lotus Goldberg, Marc Verhagen, James Pustejovsky. 2009.

<http://www.timeml.org/tempeval2/tempeval2-trial/guidelines/EventGuidelines-050409.pdf>

[2] *Evita: A Robust Event Recognizer For QA Systems*. Roser Saur¹, Robert Knippen, Marc Verhagen, James Pustejovsky. 2005.

[3] *French TimeBank: an ISO-TimeML annotated reference corpus*. André Bittar, Pascal Amsili, Pascal Denis, and Laurence Danlos. 2011. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2 (HLT '11), Vol. 2. Association for Computational Linguistics, Stroudsburg, PA, USA, 130-134.

[4] *Annotating Events, Temporal Expressions and Relations in Italian: the It-TimeML Experience for the Ita-TimeBank*. Caselli, T., Lenzi, V.B., Sprugnoli, R., Pianta, E., Prodanof, I. 2011. In Linguistic Annotation Workshop. 143-151.

[5] *TimeML Annotation Guidelines, Version 1.2.1*. Roser Saur¹, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. 2006.

http://timeml.org/site/publications/timeMLdocs/annguide_1.2.1.pdf

[6] Mati Erelt; Reet Kasik; Helle Metslang; Henno Rajandi; Kristiina Ross; Henn Saari; Kaja Tael; Silvi Vare. *Eesti keele grammatika II: Süntaks*. Eesti Teaduste Akadeemia, Keele ja Kirjanduse Instituut, Tallinn. 1993.