

EESTI KEEL KUI FORMAALNE KEEL. MIS ON FORMAALNE SEMANTIKA JA MILLEKS TA HEA ON? *

Daniel Cohnitz
Tartu Ülikool

Käesolevas artiklis vaatame põgusalt **formaalse semantika** keeleteaduslikku uurimisprogrammi. Pärast mõningaid üldisi märkusi selle programmi põhiliste metodoloogiliste eelduste kohta (1. osa) seletame ühe näite varal formaal-semantilist lähenemist loomulikule keelele, rakendades formaalse semantika analüüsi lühikese eesti keele fragmendi peal (2. osa). Lõpetuseks kommenteerime põgusalt formaalse semantika ulatust ja piire semantilise analüüsi tarbeks üldiselt (3. osa)¹

1. Formaalne semantika

Formaalse semantika aluseks on Richard Montague arusaam, et suhet süntaksi ja semantika vahel loomulikus keeles – näiteks eesti keeles – on võimalik vaadelda suhteliselt sarnaselt süntaksi ja semantika suhtega mõnes formaalses keeles, nagu näiteks esimest järku predikaatloogika keeles.

1.1. Täendus ja loogika

Formaalse keele semantikat väljendab tavaliselt mudeliteooria, mis on matemaatiline teooria, täpsemalt öeldes matemaatilise loogika üks haru (vt Partee, ter Meulen, Wall 1990: 8.5). Millele toetub usk, et loogikast võiks loomuliku keele semantika uurimisel mingitki abi olla? Toetudes

* *Inglise keelest tõlkinud Indrek Männiste.*

¹ Hea sissejuhatusena formaalsesse semantikasse vt nt Portner 2005; Cann 1993; Chierchia, McConnell-Ginet 1990; de Swart 1998 ja Heim, Kratzer 1998.

Thomas Ede Zimmermannile ja Angelika Kratzerile (1990) on selleks vähemalt kaks head põhjust. Esiteks see, et hiljemalt alates Gottlob Frege (vt Frege 1971: 38) kujutletakse loogikat teadusena, mis uurib tõe seadusi. Niisugusel juhul ongi **tõde** ja **täendus** leidnud ilmselge ühteseotuse selles, mida Max J. Cresswell nimetas **Kõige kindlamaks printsiibiks**:

Cresswelli Kõige kindlam printsiip

Kahe lause α ja β korral kehtib, et juhul kui mingis võimalikus olukorras α on tõene ja β on väär, peavad α -l ja β -l olema erinevad tähendused²

Teiseks põhjuseks on see, et semantikateooria ülesanne (või vähemasti üks ülesandeid) on õigesti ennustada väljendite vahelisi tähendussuhteid. Kuid väidetavalt on kõik **tähendussuhted loogilised** suhted (või nende taandatavad). Seega pole näiteks **hüponüümia** midagi muud kui **skemaatiline implikatsioon**: *roheline HÜP värviline*, sest *x on roheline* implitseerib *x on värvilise*, mis iganes *x* ka poleks. **Ühtesobimatus** on samuti taandatav implikatsioonile (ja eitusele): *roheline INKOMP punane*, sest *x on roheline* implitseerib *x ei ole punane*, st *x on punane* eituse. Nii saab keeleväljendite vahelist loogilist suhet taandada loogilise (või analüütilise) implikatsiooni mõistele. Selles tähenduses võib formaalset semantikat mõista katsena kirjeldada kõiki asjakohaseid semantilisi suhteid loogiliste suhete kaudu.³

1.2. Kompositsioonilisuse dogma

Peaaegu kogu formaalse semantika alase töö⁴ keskmes paikneb veel üks montague'lik idee (tavaliselt tagasiviidav Gottlob Fregeni). See idee seisneb selles, et loomuliku keele semantika peab olema **kompositsiooniline**. Nn kompositsioonilisuse printsiipi on seetõttu sageli peetud formaalse semantika fundamentaalseks printsiibiks:

² Printsiibi selline formuleering toetub Parteele (1995).

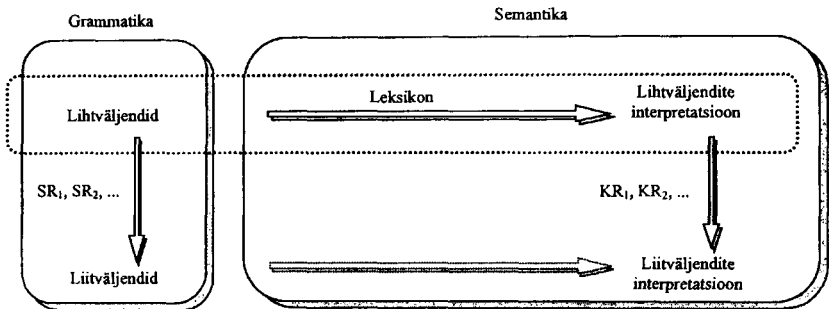
³ Formaalne semantika on seotud ka ambitsioonikama eesmärgiga, nimelt esitada loomuliku keele väljendite **tõetingimused**. Sellest lähemalt käesoleva kirjutise jaotistes 2.4–2.5.

⁴ Sellest leidub erandeid niisuguse formaalse semantika puhul, mis seda printsiipi rikub, nt **diskursuse representatsiooni teooria** (Kamp, Reyle 1993). Vt ka diskussiooni Janssen 1997.

Kompositsioonilise printsiipi

Väljend tähendus on üheselt kindlaks määratud oma osade tähenduse ning süntaktiliste seostega nende vahel.

Kompositsioonilise printsiip aitab meil seletada, millisel viisil formaalne semantika tavapäraselt **funktsioneerib**. See printsiip käsitleb loomuliku keele fragmente nii, et liitväljendite põhiklassist saab keele süntaksireeglite (SR) rakendamise abil genereerida grammatiliste liitväljendite klassi (joonisel 1: SR_1, SR_2). Formaalse semantika ülesanne on niisiis esitada interpretatsioon lihtklassi väljendite tarbeks ning kompositsioonireeglite (KR) kogum, mis peegeldaks süntaksireegleid ja määraks kindlaks liitväljendite semantilised väärtused nende osade semantiliste väärtuste põhjal (joonisel 1: KR_1, KR_2).



Joonis 1. Kompositsiooniline semantika (SR – süntaksireegel, KR – kompositsioonireegel), vt Löbner 2003.

Vaatamata sellele, et kompositsioonilise printsiip annab meile hea seletusviisi formaal-semantilise analüüsi lähenemise kohta (samal moel mõistame ka meie seda siin), on kompositsioonilise printsiibi staatus formaalse semantika siseselt üsnagi ebaselge. Mõnikord (samuti alates Montague'st) on seda printsiipi kasutatud **seletusliku** kitsendusega mis tahes loomuliku keele semantika tarbeks, suutmaks seletada tööka, et inimesed võivad aru saada keeltest, kus on lõpmatult palju erineva tähendusega väljendeid. Säärases tähenduses tundub sellel printsiibil olevat seletuslik väärtus: niisiis on ta loomu poolest **empüüriline**. Sellele vaatamata väidavad mõned formaal-semantikud, et tähenduse formaalse analüüsi raames on kompositsioonilise printsiip tegelikult triviaalne

(järelkult mitte-empiriiline) ja et parim oleks seda käsitleda pelgalt metodoloogilise printsiibina. Mõistagi tuleks metodoloogilise printsiibi tarvituselevõtmiseks esitada põhjendusi, ent senimaani pole veel ühtegi esitatud.⁵

See printsiip pole mitte mingil juhul probleemidevaba, kuivõrd see juhib semantilise analüüsi – juhul, kui see ühitatakse Cresswelli Kõige kindlama printsiibi ja küllaltki enesestmõistetava oletusega, et sünonüümia on loogiline eristamatus ning seega nõrgem kui samasus – otsejoones kohta, mida nimetatakse Matesi Lõksuks (*Mate's Trap*). Lõks seisneb selles, et need kolm oletust üheskoos võetuna tunduvad vastuollu sattuvat tõsiasjaga, et mis tahes kahe erineva väljendi α ja β korral peab esimene järgnevatest lausetest olema tõene, samas kui teine tundub olevat väär:

- (1) Kellelgi on võimalik uskuda, et α on seesama mis α , ilma et ta usuks, et α on seesama mis β .
- (2) Kellelgi on võimalik uskuda, et α on seesama mis α , ilma et ta usuks, et α on seesama mis α .

Cresswelli printsiibi järgi peaksid (1) ja (2) tähenduse poolest erinema ning järelkult, lähtudes kompositsioonilisuse printsiibist, peavad ka α -l ja β -l olema erinevad tähendused.

Mõned semantikud väidavad, et see kummutab kompositsioonilisuse printsiibi (Pelletier 1994), samas väidavad teised, et see näitab meie loogilise semantika piire. Ent on võimalik ka väita, et (1) ja (2) tähenduse erinevust tuleks analüüsida pigem pragmaatiliselt kui semantiliselt.

2. Kuidas tehakse formaalset semantikat: fragment⁶

Kompositsioonilises analüüsis koosneb formaalne semantika tavaliselt kahest sammust. Esiteks tõlgitakse loomuliku keele väljendid formaalkeelde (sageli on selleks mõni predikaatloogika variantidest⁷). Seejärel

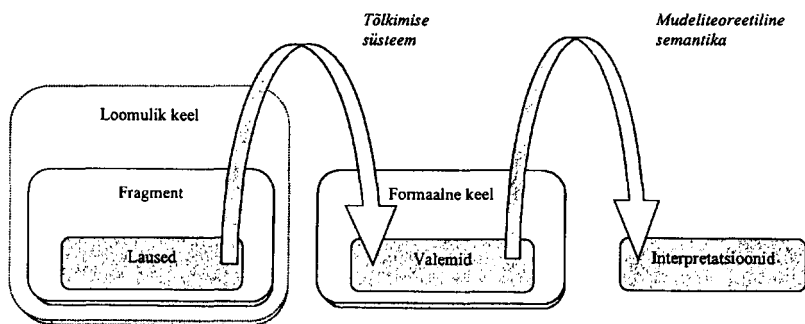
⁵ Diskussiooni sellel teemal vt Cohnitz 2005.

⁶ Alljärgnev esitus on ligilähedaselt analoogiline suurepärase esitusega teoses Löbner 2003.

⁷ Kuivõrd me ei too sisse ühtegi tuletusreeglit, vaid üksnes keele, siis oleks õigem kõnelda siin pigem formaalsest keelest kui loogikast.

esitatakse formaalkeele semantiline interpretatsioon, kasutades selleks mudeliteooriat.

Niisuguse protseduuri juures ei mängi formaalne keel kui niisugune mingisugust erilist rolli. Oleks ju võimalik (ja seda on mõnikord ka tehtud) mudeliteoreetilist tähenduse omistamist rakendada otse loomuliku keele väljenditele ilma formaalse keeleta (st kombineerida analüüsi kaks sammu üheksainsaks). See aga muudaks semantilise väärtustamise läbi- paistmatumaks ning selle kirjeldamise keerukamaks. Teisalt on formaal- keelele tähenduse omistamine formaalselt kindlaksmääratud protseduur, mille võiks isegi vahele jätta; huvitav osa praktilises analüüsis on seega tõepoolest tõlkimine järgnevas kahesammulises analüüsis (vt joonis 2):



Joonis 2. Formaalne semantika, vt Löbner 2003.

Nagu näidatud joonisel 2, tegeleb formaalne semantika seni vaid loomuliku keele fragmentidega ning on kaugel sellest, et haarata mis tahes loomulikku keelt täielikul kujul. Näitamaks, kuidas seda teha saab, vaatame üht väikest eesti keele fragmenti ainult mõne lekseemi ning mõne väga põhilise süntaksireeglga.⁸

⁸ Reeglite lihtsuse pärast oleme siia valinud üsna kunstlikud juhtumid.

2.1 Eesti keele fragment

Lekseemid, mis on selle ülesande tarbeks viidud „õigele” kujule, ning nende semantilised kategooriad on järgmised:

<i>Anna, Ekke</i>	NP (nimisõnafraas)
<i>õpetaja, ilves</i>	N (nimisõna)
<i>kena, vana</i>	A (omadussõna)
<i>teatud</i>	D (determinant)
<i>istub</i>	VP (verbifraas)
<i>tunneb, vihkab</i>	TV (sihiline tegusõna)
<i>on</i>	CV (koopula)

Iga selle fragmendi tarbeks kasutatav süntaktiline reegel peab määratlema kolme asja: (a) väljendite sisendkategooriad, mille tuletused on reegli poolt genereeritud liitväljendi koostisosad, (b) väljundväljendi tuletuse kuju ja (c) väljundväljendi kategooria. Kuivõrd eesti keel on enamike (b) juhtumite puhul üsna komplitseeritud, tuli meil lekseemide osas teha mõned kunstlikud valikud. Samuti tööme sisse determinandi, mida eesti keeles „ametlikult” ei ole.⁹ Sellele vaatamata ei tundu nendest lekseemidest moodustatud laused ebagrammatilised (kuigi ka mitte alati eriti loomulikud). Süntaksireeglid (SR) on määratletud järgnevalt:

SR _{1E}	NP VP	→ S
SR _{2E}	D N	→ NP
SR _{3E}	A N	→ N
SR _{4E}	TV NP _t	→ VP
SR _{5E}	CV NP	→ VP
SR _{6E}	CV A	→ VP

Noolest vasak pool määratleb iga reegli juures sisendkategooriad, parem aga väljundkategooria. Sisendkategooriate järjekord näitab ära väljundväljendite struktuuri (seega on kategooriate lühendid samaaegselt ka muutujad resultaadiks saadava väljendi koostisosade tarvis). Oma näites püüdsime kõiki väljundväljendeid kombineerida lihtsa **konkatenatsiooni** abil. Nii määrab esimene reegel SR_{1E}, et kui võetakse üks nimisõnafraas ning sellest paremale kirjutatakse verbifraas, saadakse lause (S). Kõik laused meie fragmendis on NP VP struktuuriga, sest meil ei

⁹ Detailset analüüsi selle kohta, mil viisil determinante igapäevases eesti keeles tegelikult (nii kõnes kui ka kirjas) kasutatakse, vt Pajusalu 2000 ja 2001.

ole ühtegi teist reeglit, mis väljundkategoriana päädiks lausega S. Reegel SR_{4E} on fragmendis ainuke reegel, mis sisaldab väljendi süntaktilist tuletust, öeldes, et kui võtta transitiivne verb ning kirjutada üks nimisõna fraas ühe lisa *t*-ga tema parema poole lõppu, saadakse verbifraas. Et näha, kuidas need reeglid töötavad, vaatame järgnevaid näidislauseid:

SR	Sisend 1	Sisend 2	Väljund
(3) <i>Anna on kena</i>			
SR _{4E}	CV	on	A <i>kena</i> VP <i>on kena</i>
SR _{1E}	NP	<i>Anna</i>	VP <i>on kena</i> S <i>Anna on kena</i>
(4) <i>Anna tunneb teatud õpetajat</i>			
SR _{2E}	D	<i>teatud</i>	N <i>õpetaja</i> NP <i>teatud õpetaja</i>
SR _{4E}	TV	<i>tunneb</i>	NP _t <i>teatud õpetajat</i> VP <i>tunneb teatud õpetajat</i>
SR _{1E}	NP	<i>Anna</i>	VP <i>tunneb teatud õpetajat</i> S <i>Anna tunneb teatud õpetajat</i>

Sellega meie eesti keele fragmendi grammatika piirduki. Fragmendi **formaalse** analüüsi esimeseks sammuks on selle loomuliku keele fragmendi tõlkimine ühte formaalsesse keelde.

2.2 Formaalne keel FK

Siin esitatava näite tarbeks kasutame ühte fragmendi predikaatloogika keelest, mille algsõnavara sisaldab ainult neid väljendeid, mida meil tõepoolest vaja läheb. Meil saab olema vähendatud hulk loogilisi konstante: indiviidkonstandid *Anna* ja *Ekke* jaoks, ühekohalised predikaatkonstandid nimisõnade, omadussõnade ja sihitute verbide jaoks (leksikaalsed VP-d) ning kahekohalised predikaatkonstandid sihiliste verbide jaoks. Lisaks nendele kasutame veel indiviidmuutujaid, mis koos indiviidkonstantidega on indiviidtermide hulgas kategoorias T. Vaatamata sellele on indiviidmuutujad samaaegselt ka omaenda kategoorias V. Et asju oleks lihtsam kirjeldada, kasutame näidetes edaspidi rasvast kirja kõigi niisuguste konstantide puhul, millele leidub otsene paariline eesti keele sõnade hulgas. Formaalkeele FK algsõnavara on järgmine:

Algväljendid

∃, ∧, =

a e

x y

Kategooria

L loogilised konstandid

T indiviidkonstandid

V, T indiviidmuutujad

õpetaja, ilves P1 ühekohaline predikaatkonstant
kena, vana
istub

tunneb, vihkab P2 kahekohaline predikaatkonstant

Kuivõrd FK on keel, milles on liitväljendeid, on seal ka süntaksireeglid selle tarbeks, kuidas moodustada lauseid algväljenditest ning liitlauseid muudest lausetest:

SR _{1P}	P1(T)	→	S
SR _{2P}	P2(T, T)	→	S
SR _{3P}	T = T	→	S
SR _{4P}	S ∧ S	→	S
SR _{5P}	∃V(S)	→	S

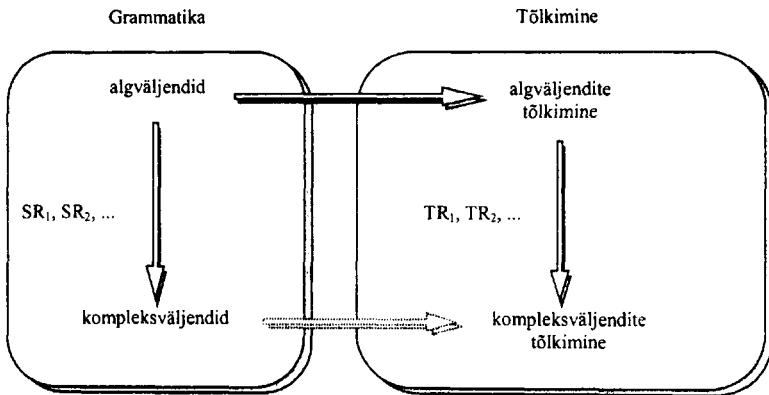
Neid süntaksireegleid võib käsitleda keele FK lausete hulka määratlevana. Näiteks ütleb reegel SR_{5P}, et juhul kui S on lause ja V on algsõnavara individumuutuja, siis kirjutades $\exists V(S)$, saame keele FK lause. Lisades ühe lisatingimuse selle kohta, et FK-s ei ole peale nende lausete, mis moodustatakse FK süntaksireeglitega, mingisuguseid teisi grammatilisi lauseid, jõuaksime FK lausete **rekursiivse määratlemiseni**. Mõistagi kehtib sama ka ülaltoodud eesti keele fragmendi puhul. Tegelikus eesti keeles leidub loomulikult rohkemgi hästimoodustatud lauseid, mida algsõnavaras esitatud lekseemidest moodustada saab, ent **fragmentis** on ainult niisugused laused, mida saab moodustada (meie esitatud) reeglite abil. Kuigi lause *Anna on õpetaja* on perfektne eesti keele lause, ei käsitleta seda fragmendis grammatilisena, samas tehakse seda kokkuleppeliselt lause *Anna on teatud õpetaja* puhul. Ehkki FK loogilistel ja mitteloogilistel konstantidel on lähedane tõlgendus ($\wedge \approx$ ja; $\exists V \approx$ Eksisteerib vähemalt üks V, mille puhul; $T = T \approx T$ on identne T-ga), ei ole keele FK lausetele veel tähendust antud. Kuid mudeliteoreetilise semantika asemel alustame hoopis loomuliku keele ja formaalse keele vahelise tõlkesüsteemi määratlemisega.

2.3 Fragmendi tõlkimine keelde FK

Tõlkesüsteemi määratledes asume endiselt süntaksi tasandil, tähendus ja semantika on veel kõrvale jäetud. Meie eesmärk on siinkohal formaalkeele ja selle süntaktiliste ressursside kujutamine loomuliku keele fragmendi süntaktilise struktuurina niisugusel moel, et fragmendis tehtud

(semantiliselt¹⁰) olulised süntaktilised eristused oleksid esindatud ka formaalkeeles. Konkreetselt tuleb meil tegeleda 1) erinevusega referentsiaalsete ja predikatiivsete nimisõnafraaside (NP) vahel; 2) erinevusega atributiivsete ja predikatiivsete adjektiivide vahel; 3) erinevusega pärisnimede ja umbmääraste nimisõnafraaside (NP) vahel ja 4) tõsiasjaga, et mõned fragmendi väljendid, juhul kui neid käsitleda FK lauseis predikaatidena, võivad evida ühesugust argumenti (nagu näiteks lauses *Teatud ilves vihkab Annat*, mille puhul tõlkesüsteemist johtuvalt tuleb *ilves(x)* ja *vihkab(x,a)* märkida sama *x*-iga).

Tõlkesüsteemis on järgitud **kompositsioonilisuse** põhimõtet. Järgnevalt määratleme tõlkefunktsiooni algväljendite jaoks ja seejärel defineerime tõlkereeglid, mis jäljendavad loomuliku keele süntaksireegleid, nii et iga süntaksireegli jaoks on üks tõlkereegel:



Joonis 3. Kompositsiooniline tõlkimine, vt Löbner 2003.

Nüüd võib alustada põhiliste väljendite tõlkimisega. Tõlgetele viidatakse edaspidi allajoonimisega, nt Anna on Anna tõlge:

T		
NP <u>Anna</u> :	<u> </u> = a	vastavalt ka <u>Ekke</u> puhul
N <u>ilves</u> :	<u>ilves</u> (<u> </u>)	vastavalt ka <u>õpetaja</u> puhul
A <u>kena</u> :	<u>kena</u> (<u> </u>)	vastavalt ka <u>vana</u> puhul
VP <u>istub</u> :	<u>istub</u> (<u> </u>)	
TV <u>vihkab</u> :	<u>vihkab</u> (<u> </u>)	vastavalt ka <u>tunneb</u> puhul

¹⁰ Tehtud eristuste õnnestumist saab hinnata alles pärast tähenduse omistamist.

C- ja D-kategooria algväljendeid on arvesse võetud tõlkereeglite raamides. Meil polegi tarvis neile FK põhisonavaras tõlget omistada. Nagu näha, ei ole tõlgete puhul tegemist tõeliste FK väljenditega, kuivõrd nad sisaldavad tühje argumendikohti () ning täidavad argumendikoha ükskõik millega, mille tulemuseks on õiged FK väljendid. Kasutades pärisnimede puhul identsussuhet (nagu ka meie seda teeme), on võimalik välja tuua eristus predikatsioonide (mis võivad käia rohkem kui ühe objekti kohta) ja pärisnimede vahel, mis vastavas lausungikontekstis osutavad alati ainult ühele üksikule indiviidile.

Nagu juba öeldud, tegeldakse C- ja D-kategooria väljenditega tõlkereeglites. Õigupoolest on need meie süsteemi kõige lihtsamad reeglid; me omistame neile nulltõlke:

TR ₂	<u>teatud N</u>	= <u>N</u>
TR ₅	<u>on NP</u>	= <u>NP</u>
TR ₆	<u>on A</u>	= <u>A</u>

Me saame seda teha seetõttu, et FK-s ei ole tehtud **süntaktilist** eristust verbide, adjektiivide ja nimisõnade vahel, ega ka predikatiivsete ja osutuslike nimisõnafraaside (NP) vahel (kuigi asjakohane erinevus tuleb hiljem sisse).

Ülejäänud tõlkereeglid on järgmised:

TR ₁	<u>NP VP</u>	= $\exists x ((\underline{NP} \wedge \underline{VP})-x)$
TR ₃	<u>A N</u>	= $\underline{A} \wedge \underline{N}$
TR ₄	<u>TV NP_t</u>	= $\exists y ((\underline{NP} \wedge \underline{VP})-y)$

TR₁-s ja TR₄-s kasutatud märgistusviis |...|-V tähistab väljendit, mis saadakse individuumuutuja paigutamisega tühjadele argumendikohtadele (|...|). Muutuja paigutamise reeglid on sellised, et kahekojaliste predikaatide puhul täidetakse esmalt teine argumendikoht ning juhul kui meil on tegemist kobarkvantoriga (kvantorid, mis asetsevad teiste kvantorite mõjupiirkonnas) algab see muutujate paigutamisega seestpoolt väljapoole. Selle selgitamiseks olgu toodud järgnev lihtne näide:

(5) <u>teatud ilves vihkab Annat</u>	
= $\exists x(\underline{teatud\ ilves} \wedge \underline{vihkab\ Annat} -x)$	TR ₁
= $\exists x(\underline{ilves} \wedge \underline{vihkab\ Annat} -x)$	TR ₂
= $\exists x(\underline{ilves} \wedge \exists y(\underline{vihkab} \wedge \underline{Anna} -y) -x)$	TR ₄
= $\exists x(\underline{ilves}(_) \wedge \exists y(\underline{vihkab}(_, _) \wedge _ = a -y) -x)$	T

Niisiis oleme esitanud tõlke. Juhul kui me tahame nüüd muutujaid seestpoolt väljapoole asendada, alustame y -ist ning täidame seega kõigepealt kahekohalise predikaadi teise koha:

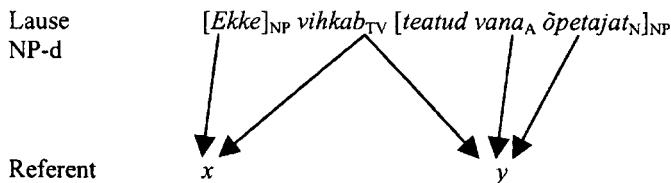
$$= \exists x(\text{ilves}(_) \wedge \exists y(\text{vihkab}(_,y) \wedge y = a|-x)$$

ja liigume seejärel x -i juurde, täites allesjäänud kohad tema kvantori mõjupiirkonnas:

$$= \exists x(\text{ilves}(x) \wedge \exists y(\text{vihkab}(x,y) \wedge y = a)$$

Ja nõnda olemegi jõudnud korrektse FK lauseni.

Nagu näeme, annavad meie tõlkereeglid tõepoolest soovitud tulemusi. 1) Referentsiaalsed NP-d on esindatud eksistentsikvantoriga (TR₁, TR₄), samas kui predikatiivsed NP-d ei ole tõlkes eksistentsikvantoriga esindatud. 2) Atributiivseid ja predikatiivseid argumente on käsitletud nii, et nad aitavad moodustada uusi predikatsioone, kas siis samale argumentile mis nimisõna (TR₃) või ka koopula konstrueerimise (TR₆) aluse osutusele. 3) Pärisnimesid ja umbmääraseid NP-sid käsitletakse ühekohaliste predikaatidena, mistõttu nendega toimitakse sarnaselt. 4) Ühised argumentid on esindatud ühiste muutujatega:



Tõlge $\exists x(x = e \wedge \exists y(\text{vihkab}(x,y) \wedge \text{vana}(y) \wedge \text{õpetaja}(y)))$

Vaatlesime ülal formaalse analüüsi esimest sammu – loomuliku keele fragmendi formaalsesse keelde tõlkimist. Selle sammu astumiseks tegime mitmeid suuri lihtsustusi: valitud eesti keele fragment oli kunstlik ning me kasutasime ebareeglipäraseid adjektiive, aga ka see, et me kasutasime determinanti (*teatud*), mida eesti keeles tavaliselt pole tarvis. Reegli oleks võinud teha ka niisuguseks:

SR_{2E}* $N \rightarrow NP$

ja determinanti mitte sisse tuua, kuivõrd sõna *teatud* saab nii või teisiti nulltõlke. Sellele vaatamata oli niisugusel moel võimalik esile tuua erinevus analüüsitava tegeliku fragmendi ja loomuliku keele vahel, mis antud lekseemide hulgast võimaldab enam süntaktilisi rekombineerimisi. Samuti oleme valinud äärmiselt puuduliku formaalkeele, kus puudub eitus ja teised standardsed konnektiivid. Vabade argumendi-kohtade asemel kasutatakse formaalses semantikas enamasti λ -operaatoreid ja λ -arvutust¹¹, mida me suurima võimaliku lihtsuse huvides siin ei kasutanud.

2.4 FK semantika

FK-le semantikat omistades omistame kaudselt semantika ka meie eesti keele fragmendi jaoks, sest selle fragmendi igale lausele leidub tõlge keeles FK. Semantiline omistus määrab ära tõeväärtused iga keeles esineva lause jaoks. Kuivõrd see toimub üldisel viisil, peaks see võimaldama anda tõetingimused kõigi selle keele väidete jaoks. Lihtsuse huvides anname tõetingimused kõikidele keele väljenditele **ühes konkreetse lausungikontekstis**. Järgnevas tuleme selle juurde veel tagasi. Seda, kas mudeliteoreetiline semantika annab tõesti seeläbi meie loomuliku keele lausete tõetingimused, on arutanud Lepore (1983), Larson ja Segal (1995) ning Ludlow (1999). Küsimust, mis juhtub siis, kui me arvestame ka teisi lausungikontekste (täpsemalt, kõiki võimalikke kontekste), arutame lühidalt jaotises 2.5.

Oma interpretatsiooni alusena kasutame üht mudelit, mis omistab tähendused kõikidele lihtväljenditele. Kõikide kompleksväljendite tähendused saab niisugusel juhul ära määrata kompositsiooniliselt, nagu lubasime jaotises 1.2.

Mudelil on kaks komponenti: esimeseks komponendiks on mingi **universum**, mis on vajalik kvantorite interpreteerimiseks. $\exists x(x = e \wedge \exists y(\text{vihkab}(x,y) \wedge \text{vana}(y) \wedge \text{õpetaja}(y)))$ ütleb seda, et eksisteerib Ekke ning eksisteerib üks vana õpetaja, keda ta vihkab. Kui see lause on tõene, siis kus nad asuvad? Nad asuvad meie arutluse universumis. Kui neid selles ei ole, on see lause väär. Teiseks komponendiks on **interpretatsioon** kõikide

¹¹ Selle leiutas 1930. aastatel Alonzo Church ning tõi lingvistikasse Richard Montague. Vt Tamme, Tammet, Prank 1997 (19.2) ja Partee, ter Meulen, Wall 1990 (13.2).

lihtväljendite jaoks. Olgu siin kõigepealt esitatud interpretatsioon kõikide lihtväljendite tarbeks ning pärast seda lühike kommentaar:

IB_{1P} Universum U

Eero *istuv õpetaja*
 Kai *kena vana ilves*
 Anna *kena õpetaja*
 Ekke *vana istuv ilves*

(Kui neid ei ole mainitud, siis nad ei istu ega ole õpetajad, ilvesed, vanad ega kenad)

IB_{2P} Nimede ja muutujate interpretatsioon

[[e]] = Ekke

[[a]] = Anna

[[x]] = Eero

[[y]] = Kai

IB_{3P} Ühekohaliste predikaatide interpretatsioon

		Eero	Ekke	Kai	Anna	
[[õpetaja]]	=	ÕPETAJA	1	0	0	1
[[ilves]]	=	ILVES	0	1	1	0
[[kena]]	=	KENA	0	0	1	1
[[vana]]	=	VANA	0	1	1	0
[[istub]]	=	ISTUB	1	1	0	0

IB_{4P} Kahekohaliste predikaatide interpretatsioon

[[vihkab]]	=	VIHKAB			
		Eero	Ekke	Kai	Anna
Eero	0	1	0	1	
Ekke	1	0	1	0	
Kai	1	1	0	1	
Anna	1	0	1	0	

[[tunneb]]	=	TUNNEB			
		Eero	Ekke	Kai	Anna
Eero	1	1	1	1	
Ekke	1	1	1	0	
Kai	1	1	1	1	
Anna	1	0	1	1	

IB_{1P}-s me lihtsalt määrasime kindlaks universumi. Esitatud näites käsitasime seda ainult neljast indiviidist koosnevana. Nagu näitab IB_{2P}, vastavad need neli indiviidi kahele nimele ja kahele muutujale. Võimalik

olnuks aga ka ühele indiviidile kaks või enam nime anda ja jätta mõned individid üldse ilma nimeta. Teisalt oli meil tarvis esitada ka x -i ja y -i kindlaksmääratud interpretatsioon, kuivõrd FK reeglid lubavad ka sidumata muutujatega lauseid. Topeltsulgudega ($\{\{\dots\}\}$) FK väljendite interpretatsioonidele osutades järgisime üldlevinud tava. Lisaks sellele tööme IB_{3P}-s ja IB_{4P}-s sisse nimed predikaatide interpreteerimiseks ja kirjutasime nad suurte tähtedega. Ühekohaliste predikaatide interpretatsioonid on funktsioonid, mis omistavad indiviididele tõeväärtusi (1 = tõene; 0 = väär). IB_{4P} näitab, et kahekohalised predikaadid on funktsioonid, mis omistavad tõeväärtused indiviidide **järjestatud paaridele**. Vertikaalselt kirjapandud individid on järjestatud paaride esimesed ning horisontaalselt märgitud individid vastavate paaride teised liikmed. Niisiis VIH_{KAB}(Eero,Kai) = 0, ent VIH_{KAB}(Kai,Eero) = 1. Pärast lihtväljendite interpretatsiooni andmist tuleb meil vaid sõnastada allesjäänud kompositsioonireeglid, mis määravad ära, mil viisil kompleksväljendite tähendused nende osade tähendustest kombineeritakse. Et kõigi FK kompleksväljendite näol on meil tegemist lausetega, ning täpsemalt väitlausestega, on kõikidel kompositsioonireeglitel alljärgnev kuju:

(6) $\{\{S\}\} = 1$ siiss (tingimused)

Vasakul pool siiss'i ('siis ja ainult siis, kui') tuleb meil arvesse võtta kõiki lausekujusid, mida on võimalik reeglite abil moodustada, mistõttu me sätestame niisuguse kujuga lausete tarvis **tõetingimused**. Alljärgnevas reeglite nimistus alustame lihtlausestest ning siirdume seejärel liitlauseste juurde, mille osadeks on teised laused:

KR _{1P}	$\{\{P1(T)\}\}$	= 1	siiss	$\{\{P1\}\}(\{T\})$	= 1
KR _{2P}	$\{\{P2(T, T^*)\}\}$	= 1	siiss	$\{\{P2\}\}(\{\{T\}\}, \{\{T^*\}\})$	= 1
KR _{3P}	$\{\{T = T^*\}\}$	= 1	siiss	$\{\{T = T^*\}\}$	= 1
KR _{4P}	$\{\{S \wedge S^*\}\}$	= 1	siiss	$\{\{S\}\}$	= 1 ja
				$\{\{S^*\}\}$	= 1
KR _{5P}	$\{\{\exists V(S)\}\}$	= 1	siiss	U-s on vähemalt üks	
				niisugune u, mille puhul	
				$\{\{V\}\} = u, \{\{S\}\}$	= 1

Nagu näha, on KR_{5P}-s vahe selles, kas eksistentsikvantor seob muutujat või mitte. Sidumata muutujad osutavad nende jaoks määratud indiviididele, seotud muutujad aga ei pruugi tingimata nendele indiviididele

osutada. Juhul kui lause on tõene, osutab seotud muutuja sellele indiviidile universumis, mis teeb selle lause tõeseks, hoolimata sellest, mille jaoks muutuja oli fikseeritud algterminite omistuse puhul. Need reeglid panevad punkti meie eesti keele fragmendi formaalsele semantikale. Vaatame, kuidas see funktsioneerib, ühe (ja viimase) näite põhjal:

(7) *teatud ilves vihkab Annat*

Analüüsi esimese sammuna tuleb see lause tõlkida keelde FK (tuletamist vt (5)).

(8) *teatud ilves vihkab Annat*

$$= \exists x(\text{ilves}(x) \wedge \exists y(\text{vihkab}(x,y) \wedge y = \mathbf{a}))$$

Teiseks sammuks on sellele tõlgitud lausele mudeliteoreetilise interpretatsiooni andmine:

- (9) $\exists x(\text{ilves}(x) \wedge \exists y(\text{vihkab}(x,y) \wedge y = \mathbf{a})) = 1$
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $[[\text{ilves}(x) \wedge \exists y(\text{vihkab}(x,y) \wedge y = \mathbf{a})]] = 1$ KR_{5P}
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $[[\text{ilves}(x)]] = 1$ ja $\exists y(\text{vihkab}(x,y) \wedge y = \mathbf{a}) = 1$ KR_{4P}
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $[[\text{ilves}(x)]] = 1$ ja
 U-s on vähemalt üks niisugune v, mille puhul $[[y]] = v$
 $[[\text{vihkab}(x,y) \wedge y = \mathbf{a}]] = 1$ KR_{5P}
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $[[\text{ilves}]]([[x]]) = 1$ ja KR_{1P}
 U-s on vähemalt üks niisugune v, mille puhul $[[y]] = v$
 $[[\text{vihkab}(x,y)]] = 1$ ja $[[y = \mathbf{a}]] = 1$ KR_{4P}
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $[[\text{ilves}]]([[x]]) = 1$ ja KR_{1P}
 U-s on vähemalt üks niisugune v, mille puhul $[[y]] = v$
 $[[\text{vihkab}]]([[x]], [[y]]) = 1$ and $[[y]] = [[\mathbf{a}]]$ KR_{2,3P}
- siis U-s on vähemalt üks niisugune u, mille puhul $[[x]] = u$
 $\text{ILVES}([[x]]) = 1$ ja IB_{3P}
 U-s on vähemalt üks niisugune v, mille puhul $[[y]] = v$
 $\text{VIHKAB}([[x]], [[y]]) = 1$ ja $[[y]] = \text{Anna}$ IB_{2,4P}
- siis U-s on vähemalt üks niisugune u, mille puhul
 $\text{ILVES}(u) = 1$ ja $[[x]] = u$
 U-s on vähemalt üks niisugune v, mille puhul
 $\text{VIHKAB}(u,v) = 1$ ja $v = \text{Anna}$ $[[y]] = v$

Kuna see tingimus on meie mudelis täidetud, on lause *teatud ilves vihkab Annat* selles tõene.

Sellega loeme oma näite lõpetatuks.

2.5 Mudeliteooria ja tähendus

Muidugi ei ole meie omistatud mudeliteoreetiline tähendus, mille meie omistasime, veel kaugeltki vastavate loomuliku keele väljendite tähendus. See on mudeliteoreetiline **semantika** ses mõttes, et ta omistab formaalkeele väljenditele keeleväliseid entiteete. **Mudel** on ta sel määral, kui palju ta hõlmab fakte, mis on relevantsed kõikide lausete väärtustamiseks. Üht niisugust mudelit (nagu ka seda, mille meie esitasime) võiks käsitleda **ühe võimaliku lausungikontekstina**. Selge on aga see, et üht niisugust mudelit ning sellest tuletatavat semantilist väärtustust ei saa samastada vastavate loomuliku keele väljendite **täheendusega**. Kui see oleks nõnda, oleks kõikidel meie mudeli sama tõeväärtusega lausetel sama tähendus. Et paremini mõista loomuliku keele väljendite kogu semantilist potentsiaali, ei tegele formaalne semantika ainult ühe lausungikontekstiga, vaid **kõigi võimalike lausungikontekstidega**. Niisugust mudeliteooriat nimetatakse **võimalike maailmade semantikaks**.¹² Võimalikud maailmad (st viisid, kuidas relevantsed faktid olla võiksid) on selles teoorias võimalikud lausungikontekstid ning lause semantiline väärtustus on funktsioon seesugustest võimalikest maailmadest tõeväärtuste hulgas.

Järelikult on **täheendusfunktsioon** võimalike maailmade semantikas määratletud järgnevalt:

Fragmendi suvaliste lausete S jaoks on $[[S]]$ funktsioon, mis omistab kõikide võimalike maailmade hulgast W igale võimalikule maailmale w S -i tõeväärtuse selles maailmas, mis on $[[S]]^w$

Selle tulemusel peaksime niisiis saama **üldised tõetingimused** kõikidele eespool vaadeldud fragmendi lausetele, sest saab väita, et võimalike lausungikontekstide hulk, kus semantiline väärtustus annab lausele väärtuse „tõene”, on ka selle lause tõetingimuste hulk.

¹² Vt ülevaadet Partee 1998. Leidub muidugi ka alternatiivseid teooriaid, nt situatsioonisemantika, aga kahjuks pole ruumpuudusel võimalik neid siin lähemalt tutvustada (vt Bremer, Cohnitz 2003 ja Barwise, Perry 1999).

3. Formaalne semantika ulatus ja piirid

Lõpetuseks kommenteerime põgusalt mõningaid formaalse semantika saavutusi ning probleeme tänapäeval.

3.1 Saavutused

Erinevalt pea kõigist teistest tänapäevastest semantikateooriatest annab formaalne semantika selge ning täpse määratluse mõnede kesksetele semantilistele terminitele. Näiteks **väljendi tähendus** (**intensioon**, funktsioon võimalikest maailmadest osutustesse); **lausungikontekst** (võimalik maailm); **väljendi osutus** (väljendi ekstensioon vastavas maailmas) ja lause **tõetingimused** (tema intensiooni kirjeldus).

Semantilist kirjeldamist võimaldava instrumendina annab formaalne semantika meile ka tähenduse esitamise võimaluse formaalses keeles, s.o lausete ja üldisemalt kompleksväljendite kompositsioonireeglite kirjelduse, samuti loogiliste omaduste ning relatsioonide kirjelduse. Üldiselt võib öelda, et formaalne semantika on tegus ning progressiivne uurimisprogramm, mis areneb endiselt väga kiiresti (mõningate uute uurimissuundade kohta vt Partee 2005 ja ter Meulen 2005).

3.2 Piirid

Nagu igal teorial, on ka formaalsel semantikal oma küsitavused; mõned neist on põhimõttelisemat laadi, teised võivad aga tulevikus lahenduse leida. Millised järgnevalt esitatavatest kuuluvad esimesse ja millised teise kategooriasse, ei ole siiski siinkohal võimalik arutada.

3.2.1 Keelelised piirid

Keelelised piirid puudutavad neid analüüsitavaid nähtusi, mida teooria piisava põhjalikkusega ei käsitle. Üheks niisuguseks probleemiks on tähenduse mitte-deskriptiivsed osad, nagu näiteks **sotsiaalne** või **ekspressiivne** tähendus. Võimalike maailmade semantikas ei ole nad tava-päraselt **üldse** esindatud, sest neil puudub roll lausete tõetingimuste määramise juures.

Hoopis keerulisem on aga see, et võimalike maailmade semantika ei näi suutvat rahuldavalt haarata isegi **deskriptiivset** tähendust. Nägime

ju artikli alul, et loogiliselt ekvivalentseid lauseid käsitletakse tähenduselt identsetena, hoolimata sellest, et nad võivad olla päris erinevad:

(10) Donald on part või Donald ei ole part.

(11) Kaks korda kaks on neli.

(10) ja (11) tähendus on erinev, aga et need mõlemad laused on loogiliselt tõesed, on nad tõesed täpselt samades maailmades (nimelt kõikides). Järelikult peab neil olema ka ühesugune tähendus.¹³

3.2.2 Võimalike maailmade semantika ja mentalistlik semantika

Teine probleem puudutab semantika seletusjõudu ning semantika tegelikku haaret. Vaadake järgnevalt kognitiivsete teooriate psühholoogilise reaalsuse kolme tasandit (Marr 1982):

1. tasand: ekstensioonis funktsiooni (*function-in-extension*) seletus, mis arvutatakse välja pakutud mooduli põhjal, viitamata sellele, kuidas seda välja arvutada saaks;
2. tasand: algoritmi seletus, mida mõistus (*mind*) / aju kasutab funktsiooni välja arvutamiseks;
3. tasand: funktsiooni psühholoogilise täideviimise seletus.

Juhul, kui me mõistame kompositsioonilisuse printsiipi empiirilise väitena (umbes nii nagu see on visandatud artiklis Cohnitz 2005), asub formaalne semantika kusagil 1. ja 2. tasandi vahepeal. Lausetele tõetingimusi omistades piirab formaalne semantika meie keelevõime (*language faculty*) väljundit, samuti seab ta piirid sellele, kuidas me oma mõistetest kompositsioonilisuse printsiibi abil need tõetingimused arvutame. Võimalikud maailmad (või nende hulgad) on aga matemaatilised objektid. Nad on abstraktsioonid ning neid ei ole võimalik samastada mõistetega, mis meie peas tegelikult on. Seda ei saa teha, sest võimalikud maailmad on meie jaoks liiga suured, et nad saaks olla õiget tüüpi sisend seesuguse psühholoogiliselt reaalse algoritmi jaoks. Kui semantikalt oodatakse seda, et ta analüüsiks mõisteid, mida me väljendite tähendustena kasutame, siis ei saa me piirduda ainult formaalse semantikaga.

¹³ Arutlust erinevate viiside kohta, kuidas selle probleemiga hakkama saada, vt Bremer, Cohnitz 2004.

Kirjandus

- Barwise, Jon, John Perry 1999. *Situations and Attitudes*. Stanford: CSLI.
- Bremer, Manuel, Daniel Cohnitz 2004. *Information and Information Flow: An Introduction*. Frankfurt: Ontos.
- Cann, R. 1993. *Formal Semantics: An Introduction*. Cambridge: Cambridge University Press.
- Chierra, G., S. McConnell-Ginet 1990. *Meaning and Grammar: An Introduction to Semantics*. MIT Press.
- Cohnitz, Daniel 2005. *Is Compositionality an A Priori Principle? – The Compositionality of Meaning and Content*. Volume I: Foundational Issues. Toim. G. Schurz, M. Werning, E. Machery. Frankfurt: Ontos, 23–58.
- Frege, Gottlob 1971. *Schriften zur Logik und Sprachphilosophie*. Hamburg: Meiner.
- Heim, I., A. Kratzer 1998. *Semantics in Generative Grammar*. Blackwell.
- Janssen, Theo 1997. *Compositionality*. – *Handbook of Logic and Language*. Toim. J. van Benthem, A. ter Meulen. Amsterdam: Elsevier, 417–473.
- Kamp, Hans, Uwe Reyle 1993. *From Discourse to Logic*. Dordrecht: Kluwer.
- Larson, R., G. Segal 1995. *Knowledge of Meaning*. MIT Press.
- Lepore, E. 1983. *What model theoretic semantics cannot do*. – *Synthese*, 167–187.
- Löbner, Sebastian 2003. *Semantik: Eine Einführung*. Berlin: de Gruyter.
- Ludlow, P. 1999. *Semantics, Tense, and Time*. MIT Press.
- Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman.
- Meulen, Alice ter 2005. *Logic and Natural Language*. – *The Blackwell Guide to Philosophical Logic*. Toim. L. Goble. Blackwell, 461–483.
- Meulen, Alice ter, Johann van Benthem (ed.) 1997. *Handbook of Logic and Language*. Amsterdam: Elsevier.
- Pajusalu, Renate 2000. *Indefinite Determiners *mingi* and *üks* in Estonian*. – *Estonian Typological Studies IV* Toim. M. Ereht. Tartu ülikooli õppetooli toimetised 14. Tartu, 87–117.
- Pajusalu, Renate 2001. *Definite and indefinite determiners in Estonian*. – *Pragmatics in 2000. Selected Papers from the 7th International Pragmatics Conference*. Toim. Enikő Nemeth T. Antwerp, 458–469.
- Partee, Barbara 1988. *Possible Worlds in Model-Theoretic Semantics: A Linguistic Perspective*. – *Possible Worlds in Humanities, Arts, and Sciences: Proceedings of Nobel Symposium 65*. Toim. Sture Allen. Berlin / New York: de Gruyter, 93–123.
- Partee, Barbara 1995. *Lexical Semantics and Compositionality*. – *Invitation to Cognitive Science*. Second edition. Toim. D. Osherson. MIT Press, 311–360.

- Partee, Barbara 2005. Reflections of a formal semanticist as of Feb 2005. http://people.umass.edu/partee/docs/BHP_Essay_Feb05.pdf (mai 2006).
- Partee, Barbara, Alice ter Meulen, E. Wall 1990. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.
- Pelletier, F. J. 1994. The Principle of Semantic Compositionality. – *Topoi* 13, 11–24.
- Portner, Paul H. 2005. *What is Meaning? Fundamentals of Formal Semantics*. Blackwell.
- Swart, H. de 1998. *Introduction to Natural Language Semantics*. Stanford: CSLI.
- Tamme, Tõnu, Tanel Tammet, Rein Prank 1997. *Loogika: Mõtlemisest tõestamiseni*. Tartu: Tartu Ülikooli Kirjastus.
- Zimmermann, Thomas Ede, Angelika Kratzer 1990. *Formal Semantics*. <http://web.uni-frankfurt.de/fb10/zimmermann/FS.pdf> (mai 2006).