Semantikakomponent grammatikateooriates

Lumme Erilt
Eesti Humanitaarinstituut

1. Sissejuhatus

Käesoleva artikli eesmärk on selgitada semantikakomponendi olemust (roll, defineerivad omadused) käibivates formaallingvistilistes keeleteooriates. Enamus kaasaja formaallingvistilisi teooriad määratlevad end eelkõige kui grammatikateooriad, see ilmneb kõige otsesemalt nende teooriate nimedes: transformatiivne generatiivne grammatika (Chomsky 1957) leksikaalfunktsionaalne grammatika (Bresnan 2000), peajuhitav fraasistruktuurigrammatika (Pollard, Sag 1994), kategoriaalne grammatika (Oehrle, Bach, Wheeler 1988), erinevad sõltuvusgrammatikad (Mel'čuk 1988; Hudson, van Langendonck 1991) ja funktsionaalsed grammatikad (Dik 1980; Halliday 1994), konstruktsioonigrammatika (Goldberg 1995) jne.[1] Nendest nimedest võib järeldada, et mainitud keeleteooriate ülesandeks on anda keele jaoks grammatiline kirjeldus, või et võrreldes keele muude aspektidega on grammatilise struktuuri kirjeldamine olulisem või prioriteetsem. Samas on enamiku teooriate ambitsioon avaram kui lihtsalt keele ühe aspekti kirjeldus, ulatudes psühholoogiliselt tõsikindlate keelereeglite postuleerimisest väideteni keelelise kogemuse päritolust ja neuronitevahelise infovahetuse modelleerimiseni.

Kui me üldiselt arvame, et inimese keelekasutuse eemärgiks ei ole produtseerida grammatiliselt korrektseid lauseid, vaid öelda midagi tähendusrikast, siis tekib küsimus tähenduskomponendi perifeersusest neis teooriates ning selle perifeersuse põhjustest. Semantikakomponendi tahaplaanile jäämise seletustena käsitlen järgnevalt ajalis-sotsiaalset konteksti, millest kaasaja keeleteooriad pärinevad (jaotis 2) ja keele formaallingvistilisest kirjeldusest tulenevaid mõjusid (jaotis 3).

---

[1] Viiteid süntaksialastele õppematerjalidele ja internetiallikatele leiab leheküljelt: http://www.ehi.ee/ehi/anglistika/courses/2syn_plan00.html Viiteid strukturalistliku keeleteaduse klassikalistele tekstidele on toodud leheküljel http://www.ehi.ee/ehi/anglistika/courses/20cLx_plan.html

2. Tähenduskomponendi perifeersuse ajaloolis-sotsiaalsed põhjused

Üheks oluliseks teguriks, mis mis tahes teooria eesmärgipüstitust mõjutab, on intellektuaalne keskkond, millesse see teooria sünnib, sest just see määrab suuresti ära fenomenid, mida teooria peaks seletama, probleemid, mida lahendama ning metodoloogilised kriteeriumid, mille suhtes teooria pädevust hinnatakse. Intellektuaalne keskkond hõlmab nii teadusele teataval ajal seatavaid üldisi eesmärgipüstitusi kui ka konkreetse teadusala seniseid tõekspidamisi ning saavutusi.

Formaallingvistiliste teooriate teket 20. sajandi teisel poolel mõjutas üldine arusaam, et keeleõpetamine ja keelekirjeldamine tähendab keele grammatika kirjeldamist, kuna grammatika on see, mis keeleti erineb, miski, mida ei pea lihtsalt üles kirjutama või pähe õppima, vaid mille kohta saab esitada reegleid. Keeleteaduse siseselt kuulub Chomsky-eelse ameerika strukturalismi pärandi hulka antropoloogilise keeleuurimise lähtumine keele välisest vormist, mis seni tundmatute keelte puhul oli vajalik, aga ka möödapääsmatu. Samuti biheivioristliku keeleteaduse (Bloomfield 1933) ideoloogiline keskendumine meeltega tajutavale ning empiiriliselt kontrollitavale keelelisele käitumisele ning teadlik antipsühhologistlik hoiak.

Kuigi Noam Chomsky enda huvi vastandina biheivioristidele oli mentalistlik, on üks põhjus, miks ka Chomsky keele vormiaspektide kirjeldusest alustab, sama, mis Bloomfieldil, nimelt püüd olla teaduslik, ning kirjeldada seda, mida on võimalik olemasolevate vahenditega teaduslikus mõttes täpselt ning usaldusväärselt kirjeldada.

Grammatiliste suhete kirjeldamisel oli teaduslikkuse nõue täidetavam, kuna grammatilisi kategooriad on arvuliselt vähem kui sõnu, ning grammatiliste kategooriate kombineerumisvõimalused seetõttu piiratumad. Ka on grammatilised kategooriad oma abstraktsuses ja polüseemsuses suurema esinemissagedusega ja nende regulaarsust oli lihtsam näha.[2]

Chomsky pürgimust võimalikult täpselt modelleerida üht keeleaspekti kui idealiseeritud autonoomset moodulit ning jätta kirjeldamata teised, esimest vaieldamatult mõjutavad osad, on kritiseeri-

---

[2] Selles mõttes ei realiseeri Chomsky ratsionalistlikku metodoloogiat kõigis selle võimalustes.

nud paljud keelevariatiivsust ja keelemuutumist uurivad lingvistid, samuti keele funktsionaalsusest huvitatud pragmaatikateadlased. Chomsky uuemates töödes (Chomsky 2000) on siiski nähtav keele meelelise representatsiooni (I-keele) väljundi otsene sõltuvus sensomotoorse (artikulatoorse ja akustilise) ning mõistelis-kavatsusliku süsteemi poolt seatud loetavustingimustest.

Chomsky ja temast lähtuvalt paljude teiste keeleteadlaste suurendatud huvi grammatikaküsimuste vastu oli põhjustatud veel muustki: Zellig Harrise, Chomsky mentori ja juhendaja, süntaksialase uurimistöö mõjust Chomskyle; keeleteadlaste huviorbiidi nihkumisest sõnatasandilt lausetasandile; modulaarsuspõhimõtete kasutamisest informaatikas ja küberneetikas, mis rõhutas eri valdkondade soovituslikku autonoomsust; neopositivistlikust objektiivsusetaotlusest teadustes; arvutiteaduse arengust, mis neutraliseeris Chomsky mentalistlikku huvi inimkeele vastu ning tingis ekstensionaalsemate keelekirjelduste tekke. Lisaks ajavaimust tulenevatele üldistele ideaalidele hakkasid juba postuleeritud teooriate käekäiku mõjutama ka nende raamistikes töötavate uurijate teooriavälised tõekspidamised.[3]

Soovisin ülalöelduga rõhutada, et tähenduse modelleerimise marginaalne positsioon tekkivais formaallingvistilistes teooriates ei olnud teadlik, programmiline, ja nagu sageli arvatud, lühinägelik tähendusanalüüsi vajalikkuse eiramine, vaid tingitud mitmetest asjaoludest. Suurel määral tuleb seda aga mõista kui veel läbikriipsutamata ülesannet töödenimekirjas, mis alles ootavad tegemist ning mis lükatakse edasi paremaid ja targemaid aegu ootama.[4] Teisest küljest on aga oluline, et kord juba sisse viidud eristus grammatika ja semantikakomponendi vahel määras edasiseks ära vaatekohaeelistuse keeleteooriate poolt antavates keelemudelites ning selle, mil moel lingvistilised teooriad keelelist tähendust kontseptualiseerivad.

---

[3] Generatiivse grammatika esiletõusuga seonduvat sotsiaalset võistlusmomenti on kirjeldanud Randy Harris (Harris 1993). Chomsky poliitiliste vaadete seost tema keeleteooriaga käsitleb Lyons 1970.

[4] Seda isegi juhul, kui neid aegu ei peaks kunagi saabuma (vrd Chomsky (2000) eristus põhimõtteliselt lahendatavate "probleemide" ja inimkognitsiooni piiratusest tingitult alatiseks lahendamatuks jäävate "saladuste" vahel).

3. Tähenduskomponendi perifeersuse formaallingvistilised põhjused

3.1. Reeglipõhisus ja vormipõhisus

Eelnevalt väljatoodud probleemiasetustest tingituna sai formaallingvistiliste keeleteooriate põhiülesandeks anda loomulikele keeltele reegli- ja vormipõhine kirjeldus, mis suudaks abstraheerida ja üldistada keeltes sisalduvaid seaduspärasusi. Reeglipõhist kujutamist võib mõista kui hulga defineerimist predikaatnotatsioonis, st mitte hulga kõigi elementide loeteluna, vaid spetsifikatsioonina selle kohta, missuguste omadustega elemendid teatavasse konkreetsesse hulka kuuluvad: {S | S on eestikeelne lause, mis koosneb nimisõnast ja sellele järgnevast verbist} Predikaatnotatsiooni kasutatakse hulkade määratlemiseks siis, kui elemente neis on üles loetlemiseks liiga palju, või isegi lõpmatu arv. Vormipõhine kujutamine tähendab seda, et keelereeglid on formuleeritavad distributiivse esinemuskonteksti põhjal sõltumatuna semantilisest või pragmaatilisest informatsioonist (süntaksi autonoomsus), ning sõltuvana üldisest süntaktilisest struktuurist, mitte aga üksikutest sõnadest (struktuurisõltuvus, kontekstivabadus).

Siinkohal tuleb täpsustada, mil viisil keelereeglid tegelikult esitatakse. Ilmselt on võimalusi mitmeid, alates kirjeldavatest suunistest, näiteks

> **1.** Kui sa räägid millestki, mida on mitu, siis tuleb seda väljendada mitte ainult selle sõna endaga, vaid ka nende omaduste ja tegevustega, mis seda sõna iseloomustavad.

kuni täpsemate, lisanduvaid kategooriaid kasutavateni:

> **2.** Kui lause aluseks olev nimisõnafraas esineb mitmuses, siis peavad ka temaga ühilduvad omadussõna ja tegusõna olema markeeritud mitmuslikult.

Selleks, et reegleid veelgi täpsemateks muuta, esitatakse need keeleteooriates matemaatikast ja arvutiteadusest laenatud vahenditega: algebrast pärit hulkade ja nendel defineeritud suhete ning operatsioonidega (tuletus- ehk ümberkirjutusreeglid), graafiteoorial põhinevate lausepuudena, tõlkimisega erineva väljendusvõimsusega loogikatest pärinevatesse primitiivmõistetesse, vastavusse seadmisega reeglite põhjal keelt analüüsivatele või sünteesivatele arvutusmasinatele (automaatidele). Niisugune matemaatiliselt täpne esitus

on siis justnagu varasematest uurimustest selgem metakeel, millesse tõlgitakse loomuliku keele reeglid, et neid paremini mõista.

Keelekirjelduse esitamisel tuletusreeglite abil antakse kirjeldus hierarhiliste sisalduvussuhete järk-järgulise strukturaalse täpsustamisena, sisuliselt tähendab see keele protseduraalset või derivatsioonilist kirjeldamist. Sisalduvussuhte väljund tuletatakse ehk genereeritakse sel juhul sisendist reeglite abil, kusjuures sisend ei võrdu väljundiga. Derivatsioonilise kirjelduse tugevaks küljeks on esituse ökonoomsus ja liiasusevabadus ning potentsiaal üldistada seda reeglipära, mis keeles eksisteerib. Derivatsiooniline kirjeldamine vastandub deklaratiivsele või representeerivale kirjeldamisele kujutusreeglite abil, kus sisend võrdub väljundiga ning kirjeldus antakse täielikuna, "nagu keel on"

Tuletusreeglite abil defineeritakse keele formaalne grammatika järgmiselt:

3. Formaalne grammatika G on nelik ⟨T, M, L, R ⟩, kus
T on lõpliku suurusega terminaalide (lõplike moodustajate, sõnade) hulk;
M on lõpliku suurusega mitte-terminaalide (sõnaklassinimede, fraasinimede vms postuleeritud, ent füüsiliselt mitteeksisteerivate kategooriate) hulk;
L on tuletusreeglite lähtesümbol hulgast M;
R on tuletusreeglite lõpliku suurusega hulk kujul $\alpha A \alpha \rightarrow \alpha$, kus A on element hulgast M ja α on element hulkade M ja T summast, nii et T sisaldab ka tühja sõna.

Niisugune formaalne grammatika esitab keele tuletades ühtedest sümbolitest teisi, ning keel defineeritakse grammatika kaudu kui kõigi niisuguste lausevormide hulk, mida see grammatika on suuteline genereerima. Vaatame selgituseks järgmist näidet:

4. Koosnegu meie grammatika G järgmistest osadest: M={S, N, V}; T={kitseke, koerake, magab, joob}; L=S; R={S→N V; N→kitseke, N→koerake, V→magab, V→joob}. Siis see grammatika kirjeldab ära keele K={kitseke magab, koerake magab, kitseke joob, koerake joob}. Et antud grammatika suudaks kirjeldada ka teisi lauseid, tuleb lisada juurde uusi reegleid või muuta olemasolevate reeglite kuju.

Põhimõtteliselt püütakse keele grammatika esitamiseks leida tarviliku, ent piisava väljendusvõimsusega reegleid, reeglid, mis oleksid nii võimsad kui vajalik (st mis suudaksid kirjeldada kogu keelt ning ei "alagenereeriks") ja nii nõrgad kui võimalik (reeglid,

mis ei "ülegenereeriks"; samuti on nõrkade reeglite arvutuslik keerukus ja seega ka ajaline mahukus väiksemad).

3.2. Keele kirjeldamine tuletusreeglite abil

Nagu öeldud, oli reegleid võimalik anda eeskätt eksplitsiitselt reeglipärasele osale keelest, s.o grammatikale, mistõttu enamike keeleteooriate lahendus oligi see, et keele grammatika esitati protseduraalselt, ning muu osa deklaratiivsena valikupiirangute ja süvakäändeinfoga täiendatud leksikonis. Enamike keelte grammatikareeglite esitamiseks leiti olevat kõige sobivamad teatud modifikatsioonidega kontekstivabad reeglid. Kontekstivabad reeglid olid võimsamad kui regulaarseid keeli genereerivad lineaarsed reeglid, ent samas ei nõudnud enamik kirjeldatavaid struktuure kontekstitundliku spetsiifilisusega või Turingi masinate võimsusele vastavaid reegleid.[5] Niisuguste reeglite formuleerimine, mis genereeriksid täpselt õige keele, osutus aga keeruliseks, sest reeglid olid kas liiga spetsiifilised ega hõlmanud kõiki keeles esinevaid väljendeid (eelpoolmainitud alagenereerimine) või olid liiga üldised, ning kuulutasid korrektseteks ka niisugused konstruktsioonid, mis keeles tegelikult kasutusel ei olnud (ülegenereerimine). Nii näiteks genereerib järgmine grammatika:

5. S → NP VP
   NP → N
   VP → V (NP)
   N → kitseke, kitsekesed, heina
   V → magab, magavad, toob

   peale korrektsete lausete *Kitseke magab, Kitsekesed magavad ja Kitseke toob heina ka laused *Kitseke magab kitseke, *Kitseke toob, *Kitseke toob kitseke, *Kitsekesed toob, *Kitsekesed magab, *Heina magavad kitseke,* jne.

Alaprodutseerimist sai vältida lihtsalt uute reeglite lisamisega. Üleprodutseerimisega seonduvad probleemid lahendati deklaratiivselt esitatud piirangutena (peajuhitav fraasistruktuurigrammatika, leksikaalfunktsionaalne grammatika, optimaalsusteooria (Prince, Smolensky 1993)), mis filtreerisid välja ebasobivad lahendid reeglite kuju täpsustamise teel, või teist tüüpi lisareeglitena, nn transfor-

---

[5] Fraasistruktuurireeglite võimsusel põhineva nn Chomsky hierarhia kohta võib põhjalikumalt lugeda Mare Koidu internetikonspektidest aadressil http://www.cs.ut.ee/~koit/SissejALplaan.html ja http://www.cs.ut.ee/~koit/Mat2/

matsiooni või nihutusreeglitena (transformatiivne generatiivne grammatika), mis baasstruktuuridest tuletasid teisesed järjendid. Need lisareeglid muutsid kontekstivabad reeglid pehmelt kontekstitundlikeks.

Peajuhitav fraasistruktuurigrammatika ja leksikaalfunktsionaalne grammatika kasutasid reeglite täpsustamiseks strateegiat, mida nimetatakse tunnuste ühitamiseks ehk unifitseerimiseks (seetõttu kutsutakse nimetatud formalisme ka unifikatsioonigrammatikateks). Tunnuste ühitamisel piiratakse reegli kehtivust mitteterminaalilt tema alamklassile, ent alamklassi jaoks ei looda uut klassinime, vaid ta defineeritakse teda määravate omaduste, nn tunnuste kaudu (Wintner 1998). Nendeks tunnusteks on antud mitteterminaalile rakenduvad grammatilised kategooriad (sugu, arv, kääne, ajavorm, võrre jne), mis on alamklassile kohustuslikud mitmikpäriluse (*multiple inheritance*) läbi. Põhjus, miks alamklass esitatakse tunnuste väärtustena ja mitte iseseisva üksusena, on vajadus eksplitsiitselt näidata osade tunnuste kohustuslikku ühildumist mõne kaasesineva mitteterminaali vastavate tunnustega. Nii saab eelnevas näites toodud mittegrammatilise lause **Kitsekesed magab* välistada järgmise, ühitavaid reegleid sisaldava grammatikaga:

6. S	→	NP	VP		
		[num: X]	[num: X]		
NP	→	N			
[num: X]		[num: X]			
VP	→	V	(NP)		
[num: X]		[num: X]	[num: Y]		
N	→	kitseke,	kitsekesed,	heina	
[num: X]		[num: X(sg)]	[num: X(pl)]	[num: X(sg)]	
V	→	magab,	magavad,	toob	
[num: X]		[num: X(sg)]	[num: X(pl)]	[num: X(sg)]	

Põhimõtteliselt on üheks kontekstisõltuvuse suurendamise viisiks ka statistilise informatsiooni lisamine, mis kontekstivabade sisalduvussuhete absoluutsust relativeerib vastavalt eelnevalt teadaolevale tõenäosusele. Statistilised piirangud arvutatakse süntagmaatiliselt kaasnevuskontekstilt (tunnuste sõltuvus: kui esineb juht X, siis on tõenäoline, et esineb ka juht Y) või sõltumatute tunnuste esinemissagedusel tunnuste arvust sõltuval juhuslikul valikul (paradigmaa-

tiline esinemus). Nii võib öelda, et mingi reegli 100%-line esinemistõenäosus tähendab, et sellel konkreetsel reeglil ei ole erandeid ega alternatiive, samas kui väiksem tõenäosus osutab reegliga väljendatava suhte alamklassilisele kuuluvusele mingisse üldisemasse jaotusesse, kus on ka teisi liikmeid. Näiteks võib väita, et kui lauses esineb artikkel, siis järgneb talle alati, 100%-lise tõenäosusega nimisõna (ehkki mitte tingimata vahetult), samas kui nimisõnale võib järgneda jagatud tõenäosustega teine nimisõna, kaassõna või verb. Oma dünaamilise iseloomu tõttu ei ole statistiliste piirangute lisamine leidnud laiemat kasutust mitte keeleteooriates, vaid keeletehnoloogilistes rakendustes.

Optimaalsusteoorias genereeritakse kõigepealt potentsiaalsete väljundstringide hulk ning seejärel filtreeritakse neist kandidaatidest välja optimaalseim, kasutades kitsenduste hierarhiat. Kitsendused võivad olla nn harmoonilised kitsendused, mis on oma sisult invariantsed keeleuniversaalid ning mis antakse vaikeväärtusena. Vaikimisi esitatud kitsendusi võidakse asendada keelespetsiifiliste variantsete kitsendustega, mis on seatud hierarhiasse konkreetsele keelele omases tähtsuse järjekorras. Optimaalne väljund on kandidaat, mis rikub kõige vähemolulisemaid piiranguid ning mida siis semantiliselt ja foneetiliselt interpreteeritakse.

Chomsky enda programm on aja jooksul keskendunud süntaktilise struktuuri tuletamisel eelkõige nihutusreeglite väljatöötamisele. Põhimõtteliselt on ta püüdnud muuta baasreeglid niivõrd üldisteks (keerulisteks, kõiki võimalusi hõlmavateks), et need sisaldaksid imaginaarseid kinnitumissõlmi võimalikult paljudele (imaginaarsetele) mitteterminaalidele, s.o sõlmi lausehierarhiat kujutavates puudes, kuhu mitteterminaalid ja terminaalid võiksid nihkuda. Struktuuri keerukuse mõttes ei ole Chomsky teooria värskeim arendus, nn minimalistlik programm, kindlasti mitte minimalistlik, küll aga õigustab seda nime süntaktiliste representatsioonitasemete vähenemine (süva- ja pindstruktuuri puudumine) ning reeglite abstraheerumine.

Sellega seoses on huvitav probleem see, et tegelikult tulenevad mitmed tuletusreeglite võimsusega ja Chomsky reeglite ning nende reeglitega kirjeldatavate keelte hierarhiaga seonduvad vahetegemised mitte reeglite tüübist, vaid sellest, missuguse taseme sümbolid valida grammatika terminaalsesse ja missugused mitteterminaalsesse tähestikku. Tavaliselt valitakse terminaalsesse tähestikku kuuluvaks leksikaalsed sõnad, vahel ka sõnadele vastavad sõnaklassid, sest sõnade

ja sõnaklassipositsioonide vastavus lauses on üksühene. Ent põhimõtteliselt poleks välistatud võimalus, kus sõnaklasside ja sõnade vahele asetatakse sõnaliikide alamkategooriad (loenduvad nimisõnad, sihilised verbid vms), mis praegu esitatakse leksikonis iga terminaalse sõna juurde kuuluva idiosünkraatilise informatsioonina.

Teisest küljest, kui sõnade tähendus esitada kompositsionaalsena, mingitest hüpoteetilistest tähenduskomponentidest koosnevana (nii nagu seda on oma komponentanalüüsis soovitanud Katz ja Fodor), siis oleks võimalik terminaalide piiri liigutada veelgi edasi. Seeläbi saaks protseduraalsena esitada suuremat osa keelest ning derivatsiooniliselt esitatavate reeglite hulka laiendada praegustele leksikaalsemantilistele reeglitele. Viimane nimetatutest on võimalus, mida püüdsid ilmselt arendada generatiivsed semantikud (Lakoff, McCawley, Ross), kes üritasid välja tuua hierarhilisi predikaat-argument suhteid tähenduskomponentide vahel.

3.3. Keele kirjeldamine kujutusreeglite abil

Tuletusreeglite kasutamisele vastandsuunaline esitusviis on nn leksikalistlik lähenemine, s.o niisugune esitus, mis suurendab deklaratiivselt esitatava informatsiooni hulka protseduraalse esituse arvelt. Sisuliselt tähendab see seda, et piiranguid ei esitata tuletusreeglites või spetsiifilisemate mitteterminaalsete kategooriate postuleerimisega, vaid terminaalide (sõnade) idiosünkraatiliste omadustena. Või ehk veel teistmoodi öelduna: lausestruktuuri ehitamist ei alustata üldisemalt lause- ja fraasitasandilt (lausepuu juurest) üha spetsiifilisemaks muutuvate piirangutena suunal mitteterminaalidelt terminaalidele (üldisemalt üksikule), vaid sõnadest (lausepuu lehtedest), iga sõna jaoks antud kaasnevuspiirangutega suunal terminaalidelt mitteterminaalidele (üksikult / konkreetselt üldisele). Siinkohal on oluline märkida, et mainitud vastandus ei ole samane ülalt-alla (*top-down*) ja alt-üles (*bottom-up*) vastandusega. Esimese vastanduse puhul on tegemist lingvistilise teadmise esitamise üldise andmestruktuuriga, teisel juhul selle teadmuse kasutamise algoritmiga mingil konkreetsel analüüsil.

Leksikalistliku representatsiooni suunas on arenenud peajuhitav fraasistruktuurigrammatika, kus tuletusreeglite arv on vähenenud tänu X-primm tüüpi abstraktsemale reegliesitusele (fraasistruktuuri koostist väljendatakse vaid fraasi tuuma ning tuumalaiendite suhtena) ja algselt tuletusreeglite piirangutena mõeldud argumenditüübi

ja selle võimalike väärtuste kujul deklaratiivselt esitatud kitsendused on kasvanud keelekirjelduse keskseks osaks. Tunnustüüpide väärtustena esitatakse selles grammatikas nii grammatilist ühildumist määratlev informatsioon kui ka sõnade ja sõnaklasside grammatilised või semantilisemad valiku- ja valentsipiirangud. Derivatsioonilise, hierarhilistel sisalduvusstruktuuridel põhineva esituse asendumine representatiivse esitusega on tekitanud peajuhitava grammatika keeleontoloogias samaväärsuse fraasitasandi ja sõnatasandi vahel, nii et eristus mitteterminaalide (*preterminals*) ja terminaalide vahel kaob (Pollard 1997). Nii sõnu kui ka fraase käsitletakse keelemärkidena (*sign*).

Leksikalistlik on ka Hudsoni sõltuvussuhetel põhinev sõnagrammatika, kus väidetavalt ei lähe tuletusreegleid üldse vaja, vaid kogu keeleline informatsioon esitatakse üksnes leksikaalse andmebaasina. Sõnad esitatakse selles andmebaasis koos ülemklassiks oleva sõnaklassiga, millelt IS-A-tüüpi sisalduvussuhete tõttu päritakse üldisemad kollokatiivsed omadused. Sõnaklassid esitatakse samuti sellessamas nimistus, nii et näiteks kui nimistus oleva kirje *tegusõna* kohta on antud, et temast sõltub (HAS-A) subjekt, mis on (IS-A) nimisõna, ning kui nimistus on antud kirje *armastab*, mis on (IS-A) tegusõna, mis nõuab (HAS-A) objekti, milleks on (IS-A) nimisõna, ning kui selles nimistus on veel kirjed *Jaan* ja *Mari*, mis on (IS-A) nimisõnad, siis on niisugune informatsioon piisav lause *Jaan armastab Mari* genereerimiseks.

Leksikalistlik lähenemine oma sõnapõhiste (st mitte sõnaklassipõhiste) struktuuridega ei tarvitseks üldse eristada grammatikat ja leksikoni, sest kogu grammatika on võimalik deklareerida leksikaalsete üksuste idiodünkraatiliste omadustena. Siiski peab Hudson (Hudson 2000) vahetegemist süntaksi ja semantika vahel oluliseks, kuna sõnade ja nende keeleväliste osutuste vahel ei eksisteeri üksühest vastavust. Nii on võimalikud juhud, kui mitu sõna vastab ühele objektile (kolmesõnaline väljend *minu sõber Colin* vastab ühele inimesele, *vedru välja viskama* vastab suremisele jms), samuti see kui üks sõnatüüp vastab mitmele mõistele (polüseemia), mitu sõna ühele mõistele (sünonüümia) jne.

3.4. Süntagmaatiliste ja paradigmaatiliste suhete väljendamine grammatikates

Eraldi huvitav probleem keele reeglipärasel kirjeldamisel on süntagmaatiliste ja paradigmaatiliste suhete väljendamine. Süntagmaatilised suhted keeleväljendite vahel on distributiivsed sõltuvus- ja täiendussuhted lause- või tekstikontekstis. Paradigmaatilised ehk valikusuhted on suhted alternatiivide vahel, mis võiksid esineda samas süntagmaatilises kontekstis.

Allerton (Allerton 1979: 82–83) väidab, et paradigmaatiliste ja süntagmaatiliste suhete väljendamiseks tuleb kasutada erikujulisi fraasistruktuurireegleid. Nii väljendab reegel kujul NP → Det N süntagmaatilisi koosnevussuhteid, sest see reegel väidab, et nimisõnafraas koosneb määritlejast ja nimisõnast selles lineaarses järjekorras. Koosnevusreeglid kirjutavad lahti mingi lauseosa sisestruktuuri.

Seevastu reegel N→Nsg, Npl väljendab paradigmaatilisi alamklassifikatsioonilisi suhteid, sest ta väidab, et nimisõna võib olla kas ainsuslik nimisõna või mitmuslik nimisõna. Alamklassireeglid kirjutavad lahti alternatiivide valiku, mille hulgast tuleb valida üks. Lausepuus annavad alamklassireeglid sõlmi, kus emasõlmest hargneb vaid üks tütar. Niisuguseks alamklassireegliks on ka reeglid, mis mitteterminaaliga seovad terminaali: N→ kitseke.

Teisest küljest võib öelda, et fraasistruktuurireeglid väljendavad samaaegselt nii süntagmaatilisi eelnevussuhteid kui ka paradigmaatilisi sisalduvus-/koosnevussuhteid, sest näiteks reegel NP→Det N väljendab nii seda, et määritleja eelneb nimisõnale, kui ka seda, et nimisõnafraasi üks võimalik alamkonfiguratsioon on Det N, samas kui on ka teisi võimalusi (NP → N, NP → N PP, NP → Adv Adj N, NP → Det N RelPron S, jne). Kui mitteterminaal kirjutub lahti rohkemaks kui üheks koostisosaks, siis valitseb nende osade vahel sõltuvussuhe, nii et mitteterminaali tuumaks on element, mis on küll temast sõltuvate elementide skoobis, ent mis ainsana selles kontekstis sobib emasõlme alamklassiks.

3.5. Grammatikareeglid ja semantikareeglid

Nagu ilmneb, keskenduvad ülalkäsitletud võimalused reeglite esitamisel eelkõige grammatikareeglite, s.o sõnaklasse puudutavate seoste väljatoomisele. Semantilist informatsiooni esitatakse vaid kaudselt leksikaalsemantiliste hüponüümiasuhetena (ilmselt algselt tulenevana

Carnapi tähenduspostulaatidest) või süvakäänetel ja temaatilistel rollidel põhinevate valikupiirangutena (lähtuvalt Fillmore'ist). Semantika-alase uurimistöö olulisemaid ülesandeid tulevikus peaks olema leksikaalsemantilise, sõnaspetsiifilise poole kokkuviimine rollitasandi abstraktsioonidega (vrd Lappin 1996: 8). Üks viis seda teha on teatavate vahestruktuuridena, nagu leksikaalfunktsionaalne grammatika, mis lisaks sisalduvusstruktuurile (*c(onstituent)-structure*) genereerib ka keeleväljendite lausesiseseid funktsioone eritleva funktsionaalse struktuuri (*f(unctional)-structure*).

Samas on selge, et sõnade lausesiseste funktsioonide kirjeldus ei anna veel lause kogutähendust, ning isegi kui lauseliikmete leksikaalsemantiline kirjeldus on olemas, ei piisa sellest lausetähenduse täpseks määratlemiseks. Niisugust, täpselt defineeritud lausetähenduse arvutamise mehhanismi enamik praeguseid teooriaid ei oma. Chomsky programmide loogiline vorm (LF) ei ole lause semantiline representatsioon, vaid süntaktiline struktuur peale nihkeoperatsioonide teostumist, mida semantiliselt interpreteeritakse, st millele antakse vastavate semantiliste reeglitega tähendus teatava metakeele (tavaliselt esimest järku predikaatloogika) suhtes.

On siiski olemas formaallingvistiline teooria, mis lausetähendusega väga tõsiselt tegeleb, see on Montague semantikal (vt Dowty, Wall, Peters 1981) ja matemaatilisest ning filosoofilisest loogikast (Frege, Tarski, Church, Russell, Kripke) saadud järeldustel põhinev kategoriaalne grammatika (Oehrle, Bach, Wheeler 1988). Kategoriaalses grammatikas genereeritakse kohe paralleelselt nii lause süntaktiline kui ka semantiline struktuur. Mõlemad struktuurid defineeritakse algebraliselt ning esitatakse funktsioon-argument tüüpidena. Semantiline struktuur esitatakse tüüpidena, kus nimisõnadele vastava tüübi osutuseks on entiteedid mingis mudelis, verbitüüpide osutuseks funktsioonid, mille argumendiks on entiteedid ja väärtuseks tõeväärtused jne. Kategoriaalse grammatika kirjeldus on mudeliteoreetiline, sest osutused ja tõeväärtused sõltuvad väärtustest konkreetses mudelis või võimalikus maailmas.

Kategoriaalse grammatika poolt semantilisel esitusel kasutatav formalism on oma keerukuselt võrreldav Chomsky süntaksikomponendi keerukusega, osaliselt on see tingitud mõlema teooria eesmärgipüstitusest esitada reeglid, mis annaksid väljundiks rangelt ühese esituse. Samas näib õigus olevat Barbara Partee'l (1996), kes väidab, et teooriad, mille grammatikakomponent on lihtne, peavad

seda kompenseerima komplitseeritud semantikaga ning lihtsa semantikakomponendiga teooriad keeruka grammatikamooduliga.

Vaatamata erinevustele kategoriaalse ja generatiivse grammatika keeleontoloogias, jagavad mõlemad paradigmad üht ühist puudust. Nimelt ignoreerivad mõlemad formalismid lausetähenduse komponentaalse ülesehituse kõrval tähendussuhteid lauses osalevate ja mitteosalevate keeleväljendite vahel, jättes ühe osa redundantsetest korrapäradest keeles mitte ainult abstraheerimata, vaid ka välja toomata.

Süntaktilise representatsiooni ja semantilise interpretatsiooni tegeliku ühendamisega on kõige sisulisemalt tegelenudki sageli mitte keeleteoreetikud, vaid keele arvutuslingvistilistest rakendustest (vt Rosner, Johnson 1992; Smith 1991: 341 jj) või kognitiivteaduslikest seostest (Fodor 1983) huvitatud uurijad.

4. Kokkuvõte

Käesoleva artikli eesmärgiks oli selgitada semantikakomponendi staatust kaasaja formaallingvistilistes keeleteooriates. Lähtusin seisukohast, et enamus lingvistilisi teooriaid määratlevad end grammatikateooriatena ning peavad keelelise tähenduse modelleerimist vähemoluliseks. Järgnenud arutelus püüdsin tabada nimetatud vaatekohaeelistuse põhjuseid. Olulisemate seletustena semantikakomponendi perifeersusele käsitlesin sotsiaalset konteksti, milles keeleteooriad algselt välja töötati ning lingvistiliste formalismide defineerimiseks kasutatud metateooriate võimalusi.

Keeleteooriate formuleerimise ajalis-sotsiaalne kontekst määrab ära küsimused, millele teooriad vastust peavad otsima, probleemid ning suhted, *mida* seletama. See tähendab ka, et teooriad ei räägi asjadest, mille järgi ei küsita. Teooriate väljendamiseks kasutada olevad formalismid määravad ära selle, *kuidas* esitatud küsimustele vastus leitakse. Tulemuseks saadav teooria sõltub nii püstitatud eesmärkidest kui ka nende realiseerimiseks kasutatud vahenditest. Valmis teooria saab omakorda metateooriaks konkreetsete keelte keeleteoreetilistele kirjeldustele. Kuna olemasolev teooria määrab oma võimsuse ja võimalustega ära keelekirjelduste üldkuju, siis ei arenda konkreetsed keelekirjeldused ise tegelikult keeleteooriat edasi.

Grammatika- ja tähenduskomponendi eristamise puhul keeleontoloogiates osutub huvitavaks see piirjoon, millest ühele poole kuuluvat peetakse grammatikaks ja teisele poolele jäävat seman-

tikaks. Kui grammatika hulka kuuluvaks pidada seda osa keelest, mis on reeglipärane selles mõttes, et ta koondab elemente, millel on eritletavaid ühisjooni, ja semantika hulka kuuluvaks idiosünkraatiliste omadustega osa, siis võib arvata, et keeleteaduse arenguga piir grammatika ja semantika vahel nihkub: üha suuremaks muutub grammatika ja üha väiksemaks jääb semantika uurimisobjekt. Kuid on selge, et see ei saa olla ainus kehtiv seisukoht.

Kirjandus

Allerton, D.J. 1979. Essentials of Grammatical Theory: A Consensus view of Syntax and Morphology. London, Boston, Henley: Routledge & Kegan Paul.

Bloomfield, Leonard 1933. Language. Ney York: Holt, Rinehart and Winston.

Bresnan, Joan 2000. Lexical-Functional Syntax. Oxford: Blackwell.

Chomsky, Noam 1957 Syntactic Structures. The Hague: Mouton.

Chomsky, Noam 2000. New Horizons in the Study of Language and Mind. Cambridge: Cambridge UP

Dik, Simon C. 1980. Studies in Functional Grammar. London: Academic Press.

Dowty, David R.; Wall, Robert E.; Peters, Stanley 1981. Introduction to Montague Semantics. Dordrecht: Reidel.

Fodor, Jerry A. 1983. The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, MA.: MIT Press.

Goldberg, Adele E. 1995. Constructions: A Construction Grammar Approach to Argument Structure. Chicago, London: The University of Chicago Press.

Halliday, M. A. K. 1994. An Introduction to Functional Grammar. 2. edition. London: Arnold.

Harris, Randy Allen 1993. The Linguistics Wars. Oxford: Oxford UP

Hudson, Richard 2000. Word Grammar. Birmingham: ESSLLI.

Hudson, Richard A.; van Langendonck, Willy 1991. Word grammar. – Linguistic Theory and Grammatical Description. Toim Flip G. Droste, John E. Joseph. Amsterdam, Philadelphia: John Benjamins. 307–336.

Lappin, Shalom 1996. Introduction. – The Handbook of Contemporary Semantic Theory. Toim Shalom Lappin. Oxford: Blackwell. 1–8.

Lyons, John 1970. Chomsky. Fontana: Collins.

Mel'čuk, Igor, A. 1988. Dependency Syntax: Theory and Practice. New York: State University of New York Press.

Oehrle, Richard T.; Bach, Emmon; Wheeler, Deirdre (toim) 1988. Categorical Grammars and Natural Language Structures. (Studies in Linguistics and Philosophy 32.) Dordrecht: Reidel.

Partee, Barbara H. 1996. The development of formal semantics in linguistic theory. – The Handbook of Contemporary Semantic Theory. Toim Shalom Lappin. Oxford: Blackwell. 11–37.

Pollard, Carl 1997. Lectures on the foundations of HPSG. Unpublished manuscript. Ohio State University.

Pollard, Carl J.; Sag, Ivan A. 1994. Head-driven Phrase Structure Grammar. Chicago, London: University of Chicago Press and CSLI Publications.

Prince, Alan; Smolensky, Paul 1993. Optimality Theory: Constraint Interaction in Generative Grammar. New Brunswick, N.J.: Rutgers University; Boulder: University of Colorado.

Rosner, Michael; Johnson, Roderick (toim) 1992. Computational Linguistics and Formal Semantics. Cambridge: Cambridge UP.

Smith, George W 1991. Computers and Human Language. Oxford: Oxford UP.

Wintner, Shuly 1998. Unification-based Linguistic Formalisms. Saarbrücken: ESSLLI.