

TARTU RIIGLIK ÜLIKOOL



Keel ja struktuur

Töid struktuuralse ja
matemaatilise
lingvistikalt

TARTU 1965

TARTU RIIKLIK ÜLIKOOI

EESTI KEELE KATEETER.

Keel ja struktuur

Toid strukturaalse ja
matemaatilise
lingvistikaka alalt

[1.]

TARTU 1965

Toimetuskolleegium:

O. M u t t, H. R ä t s e p (toimetaja), E. U u s p õ l d.

S a a t e k s .

Käesolev kogumik sisaldab Tartu Riiklikus Ülikoolis eesti keele kateedri juures individuaalplaani alusel struktureaalset lingvistikat õppinud või õppivate üliõpilaste uurimusi ning on esimene struktureaalsele lingvistikale pühendatud kogumik meie vabariigis.

Enamik kirjutisi on valminud otseselt õppetöö käigus. Nii on Haldur Õimu ja Mati Ereli artiklid kokkuvõtted nende kursusetöödest, Priit Järve kirjutis kokkuvõtte dots. L. Võhandu juhendamisel tehtud seminaritööst, Jaan Kaplinski pikem artikkel osa diplomitööst.

Kogumiku esimene artikkel pärineb struktureaalse lingvistika alal töötavalt aspirandilt Ellen Uuspõllult ja on valminud väitekirja ettevalmistamise käigus.

LAUSE STRUKTUUR
JA
VAHETUTE MOODUSTAJATE MEETOD

E. U s p ö l d

O. Strukturaalse uurimise meetodid tänapäeva keeleteaduses lähtuvad F. de Saussure'i õpetuse põhiteesist, et keel ei kujuta endast mitte isoleeritud elementide summat, vaid on terviklik süsteem, milles iga element on määratletav vaid oma vahekordade kaudu teiste elementidega. Keele elemendid on iseloomustatavad kahest küljest: esiteks häälikulisest, s. o. kuuldavast või nähtavast, ja teiseks tähenduslikust, sisulisest küljest. Vastavalt sellele räägitakse väljenduse ja sisu plaanist keeles. Et keele väljendusvormid on uurijale objektiivses tegelikkuses vahetult antud, peetakse lingvistika primaarseks uurimisobjektiks väljenduse plaani.

Uute ideede rakendamine grammatikas tähendas kõigepealt loobumist igasugustest tähenduslikest kriteeriumidest grammatiliste kategooriate kirjeldamisel ja nende asendamist ainuüksi vormide struktuuriomadusi arvestava iseloomustusega. Formaalse grammatikakäsitluse metodoloogia loomisel ja praktiliste uurimistulemuste poolest on hinnatavaid saavutusi Ameerika deskriptiivse lingvistika¹ esindajatel. Nüüd

¹ Iseloomustavat deskriptiivse lingvistika kui koolkonna kohta vt. H. Rajandi, Morfeemide liigitus deskriptiivses lingvistikas. "Keel ja Kirjandus" 1963, nr. 6, lk. 357-358.

on juba võimalik rääkida L. Bloomfieldi koolkonna teatavast läbitud arenguetaapist, mis piirneb aastatega 1933 ja 1957, s. o. L. Bloomfieldi teose "Language" ja N. Chomsky teose "Syntactic Structures" ilmumisaastatega. N. Chomsky genereeriva grammatika teooria tähendas keeleteaduses üleminekut seniselt empiiriliselt vaatluselt ja kirjeldamiselt formaalsete teooriate ja neil põhinevate loogilis-matemaatiliste mudelite loomisele. Selle otseseks ettevalmistuseks oli eelnenud kahe ja poole aastakümne jooksul väljatöötatud formaalse analüüsi meetodika. Lõpliku ja mõjuvaima õigustuse sai deskriptiivlingvistiline uurimissuund aga praktikas, sest tormiliselt arenev elektroonika koos küberneetikaga tõstsid päevakorrale hoopis uued keeleteaduse rakendusala, nagu automaatne tõlkimine, dispetšerisüsteemide automatiseerimine, õpetavad automaadid jne. Traditsioonilise grammatika reeglid polnud masinate jaoks kasutatavad. Küll aga võimaldasid koostada masinaile "arusaadavaid" juhiseid andmed, mis saadi teksti analüüsimisel selliste formaalsete meetoditega nagu distributiivne analüüs ja vahetute moodustajate meetod.

Alljärgnevas on püütud anda ülevaade vahetute moodustajate (VM) meetodist kui ühest olulisemast deskriptiivse lingvistika meetodite hulgas. Kuigi VM meetod on ette nähtud universaalseks analüüsimeetodiks keele süsteemi kõigi tasandite tarvis, pakub kõige rohkem huvi selle meetodi rakendamine lause struktuuri uurimisel. Seepärast peatume gi VM analüüsil kui süntaksi uurimisevõttel. Sellega seoses tuleb aga kõigepealt juttu teha struktuuralse süntaksi mõningatest põhimõistetest ja nende vahekorrast traditsioonilise grammatika vastavate mõistetest.

1. Deskriptiivses lingvistikas on kõnevoolu analüüsil lähteüksuseks *l a u s u n g* (utterance). See on kõigis tuntud keeltes ühtmoodi defineeritav: kõnelõik, millele eelneb ja millele järgneb paus. *L a u s e t* iseloomustatakse sellest lähtuvalt kui sõltumatut lingvistilist vormi, mis antud lausungis ei kuulu mõne grammatilise konstruktiooni koosseisus suuremasse lingvistilisse vormi.² Seega võib iga lause esineda iseseisva lausungina. Lause kui lingvistilise vormi sõltumatuse tunnuseks on iseloomulik lõpetatud intonatsioon.³ Näiteks Ta kirjutab↓ on sõltumatu vorm, mis erineb vastavast morfeemijärjendist väljendist Ta kirjutab pojale kirja↓ vaid intonatsioonilt. Teisest küljest koosneb iga lausung vähemalt ühest lausest ja järelikult tuleb lauseteks pidada ka selliseid predikatsiooni mittesisaldavaid lausungeid nagu Jaani! Tere hommikust! Aga miks? jne. Viimased eraldatakse küll nn. vähemiklausetest (minor sentences), s. o. harvem esinevate lausete rühmana tavalistest, nn. enamiklausetest (major sentences). Need Bloomfieldilt pärinevad seisukohad⁴ näivad olevat deskriptivistide lausekäsitlustes tänini au sees.⁵ Tähelepanu äratav aga Z. S. Harris'e uus suund tema värskeimas lause struktuuri käsitlevas töös, kus ta teatavaid lausungeid, näiteks But why? What the! - ilmselt predikatsiooni mittesisaldavaid - iseloomustab kui lause segmente, s. t. kui üksusi, mis on lausest väiksemad.⁶ Lausetele peab ta nii-

2 L. Bloomfield, Language. Chicago, 1933, lk. 170; Ch. C. Fries, The Structure of English. New York and Burlingame, 1952, lk. 21.

3 Ch. F. Hockett, A Course in Modern Linguistics. New York, 1963, lk. 199.

4 L. Bloomfield, op. cit., lk. 170-172.

5 Vt. näiteks Г. Глисон, Введение в дескриптивную лингвистику, М., 1959, lk. 196-197; Ch. F. Hockett, op. cit., lk. 164; H. L. Kufner, The Grammatical Structures of English and German. Chicago, 1962, lk. 3.

6 Z. S. Harris, String Analysis of Sentence Structure. The Hague, 1962, lk. 8.

siis iseloomustavaks mitte ainult teatavaid intonatsioone, vaid ka teatavaid segmentaalseid struktuure.⁷ Lause struktuurist rääkides mõeldaksegi tavaliselt segmentaalseid struktuuritunnuseid, eelkõige predikatsiooni, mis aga mitmesugustel "vähemiklausete" tüüpi juhuslikel lausungeil puudub.

1. 1. Diskreetsetest osadest koosneva lineaarse üksuse puhul (selline üksus on ka lause) moodustavad s t r u k t u u r i m õ i s t e tuuma kaks põhilist vahekordade tüüpi: funktsionaalsed ehk osa/tervik-vahekorrad ja distributiivsed ehk osa/osa-vahekorrad.⁸ Nii iseloomustavad lause kui struktuuri iga elementi 1) lauseterviku ja vastava elemendi vahekord ning 2) selle elemendi vahekorrad lause teiste osadega. Vahekordadelt sarnased elemendid moodustavad elementide klassi. Seega kujutavad endast erinevad lause struktuuri tüübid elementide klasside erinevaid kombinatsioone.⁹ Teiste sõnadega: lause struktuuri seisukohalt ei ole distinktiivsed mitte üksiksõnad ega -morfeemid, vaid sõnaklassid ja morfeemiklassid.¹⁰ Seetõttu on kõik keeles olemasolevad laused tagasiviidavad piiratud arvule sõnaklasside kombineerumise mudelitele (pattern), mida kõneprotsessis rakendatakse üha uute lausete produtseerimiseks, täites mudeli vastavaid positsioone sõnaklasside erinevate liikmetega. Sama sõnaklassi liikmeid ühendab niisiis ühine funktsioon lauses. See ongi sõnaklasside põhierinevuseks traditsioonilistest s õ n a l i i - k i d e s t.

Sõnaliike defineeritakse sõnade morfoloogiliste tunnuste (käändsõnad, pöörd sõnad, muutumatud sõnad) ja edasi nende leksikaalse tähenduse (nimisõnad märgivad olendite,

7 Z. S. Harris, String Analysis of Sentence Structure. The Hague, 1962, lk. 7.

8 A. Juillard, Outline of a General Theory of Structural Relations. s'-Gravenhage, 1961, lk. 24.

9 Z. S. Harris, op. cit., lk. 7.

10 Vrd. Ch. C. Fries, op. cit., lk. 64.

esemete või nähtuste nimetusi) või funktsiooni (asesõnad asendavad olendite, esemete, nähtuste, nende omaduste või arvu nimetusi) põhjal.¹¹ Niisugune erinevaid liigitamise aluseid kasutav klassifikatsioon pole järjekindel. Vormi ja funktsiooni vastuolu teeb võimatuks lõikumatute sõnarühmade saamise, sest sõna võib sageli kuuluda erinevaisele sõnaliikidesse olenevalt oma funktsioonist lauses.¹² Näiteks oma võib olla 1) määrsõna (ta on oma 30 aastat vana), 2) omastav asesõna (oma töös on ta hoolas), 3) omadussõna (omad vitsad peksavad), 4) nimisõna (omade peale koer ei haugu).¹³

Inglise keele sõnade formaalse klassifikatsiooni, mis põhineb ulatusliku materjali (50 tundi lindistatud telefonikõnesid) analüüsil, esitas 1952.a. Michigani ülikooli inglise keele professor Ch. C. Fries.¹⁴ Oma töös rakendab ta Z. S. Harris'e distributiivse analüüsi meetodi põhilist protseduuri - korduvat substitutsiooni. Selle põhimõte on lühidalt järgmine. Olgu vormi A ümbruses C - D võimalik asendada vormiga B, ilma et väljendi grammatiline õigsus häviks (s. t. nii CAD kui ka CBD on antud keeles esinevad väljendid). Sel puhul ütleme, et A ja B kuuluvad samasse substitutsiooniklassi, kusjuures sellesse klassi kuuluvad kõik vormid, mis samadel tingimustel täidavad positsiooni C - D.¹⁵ Ch. C. Fries kasutab substitutsiooniklasside saamiseks spetsiifilisi ümbrusi, nn. kontrollraame. Näiteks sõnad, mida on võimalik paigutada raami (The) - is/was good,

11 Vrd. N. Remmel, E. Riikoja, J. Valgma, Eesti keele grammatika keskkooli IX-XI klassile. Tallinn, 1964, lk. 20-21.

12 Eesti keele sõnaliikidest tõlkimisalgoritmi seisukohalt vt. Ü. Kaasik, A. Korjus, Automaatse tõlkimise keeleteaduslikke küsimusi. "Keel ja Kirjandus" 1960, nr.9, lk. 549-551; Т. Тобиас, О частях речи эстонского языка. Сообщения по машинному переводу, Таллин, 1962, lk. 18-23.

13 Johannes Voldemar Veski keelelisi töid. Tallinn, 1958, lk. 230.

14 Ch. C. Fries, op. cit., lk. 65-141.

15 Z. S. Harris, From Morpheme to Utterance. "Language" 1946, Vol.22, No.3, lk.163. Üksikasjalisemalt morfeemiklasside saamise kohta vt. Z. S. Harris, Structural Linguistics. Chicago, 1960, lk. 243-261.

kuuluvad esimesse klassi.¹⁶ Sellise analüüsiprotseduuri tulemuseks on 4 suurt sõnaklassi¹⁷ (vastavad täistähenduslikele sõnadele) ja 15 vähemat nn. funktsioonisõnade rühma (siia kuuluvad näiteks artiklid, prepositsioonid, konjunktsioonid, abiverbid jne.). Saadud substitutsiooniklasside liikmeile on peale positsioonitunnuste omased ka teatavad morfoloogilised ja derivatsioonilised, s. t. vormitunnused. Kõiki neid tunnuseid kokku nimetab Ch. C. Fries strukturealse (= grammatilise) tähenduse signaalideks. Niisiis kui öeldakse, et lause struktuur on sõnaklasside teatav kombinatsioon, tuleb selle all mõista kogu seda grammatiliste tähenduste kompleksi, mida väljendavad sõnaklasside vormitunnused (s. t. infleksiooni- ja derivatsioonisufiks), nende tunnuste abil väljenduvad seosed (kongruents, rektatsioon) ja iga sõnaklassi asend teiste sõnaklasside suhtes (sõnade järjekord). Kõik need süntaktilise struktuuri karakteristikud moodustavad süsteemi, sest iga tunnus omandab strukturealse tähenduse alles teatava terviku osana.

1. 2. Sõltuvalt keele grammatilise süsteemi tüübist võivad olla lause struktuuri signaalidena ülekaalus kas vormi- või positsioonitunnused, s. t. nende vahetõrk on pöördvõrdeline: mida komplitseeritum on keele morfoloogia, seda vabam on sõnajärg, ja vastupidi.¹⁸ Näiteks inglise keel on morfoloogiliselt vähekomplitseeritud¹⁹, küll aga väga kindla sõnajärgjega keel, eesti keele morfoloogiline vormirikkus seevastu võimaldab sõnajärjes küllalt suuri vabadusi. Vastavalt sellele on ka süntaktilise seose eri vormide, s. t. külgnemise, rektiooni ja kongruentsi vahe-

16 Ch. C. Fries, op. cit., lk. 78.

17 Ch. C. Fries tähistab need klassid lihtsalt järjekorranumbritega 1.-4. Üldkasutatavaks on aga kirjanuduses saanud sõnaklasside tähelised sümbolid N, V, A ja D.

18 Ch. F. Hockett, op. cit., lk. 180-181.

19 Morfoloogilise komplitseerituse näitajana on kasutatud sõna kohta tulevat keskmist morfeemide arvu, mis inglise keeles on veidi alla 2 (vt. Ch. F. Hockett, op. cit., lk. 181).

kord erineva grammatikatüübiga keeltes erinev. Domineerivalt süntaktilise grammatikaga keeltes saab kindla sõnajarje ja sõnade morfeemivaesuse tõttu külgnemine tunduvalt suuremaid esinemisvõimalusi kui nn. morfoloogilise grammatika puhul, kus kongruents on iseloomulikum süntaktilise seose vorm. Seda fakti ei tohiks ignoreerida lause struktuuri mitmesuguste analüüsimetodite kasutamisel ja hindamisel. Peab veidi ettehaaravalt märkima, et kui O.S. Ahmanova ja G. B. Mikaeljan kritiseerivad Z. S. Harrist distributiivse analüüsi rakendamise pärast süntaktilisel tasandil²⁰ ja kui N. A. Sljussareva eelistab süntaksi uurimisel vene keeleteaduses väljatöötatud sõnaühenditeooriat VM analüüsimetodile,²¹ siis etendab selle juures mitte just tähtsusetut osa ühe grammatikatüübi baasil väljatöötatud uurimismetodite hindamine teistsuguse grammatikatüübiga keele seisukohalt. Enakeeles domineerivad struktuuritunnused mõjustavad lingvistilist paratamatult ka uurimisvõtete väljatöötamisel ja valikul.²² Kuid sellest ei tule nii aru saada, et mingi keele struktuuri kirjeldamisel annab efektiivseid tulemusi vaid antud keele grammatikatüübile kohandatud uurimisprotseduur. Et mingi keele teksti analüüsiga avastada selle keele struktuuri võimalikult kõik, ka vähem ilmsed, peidetud seaduspärasused, tuleb kasutada maksimaalselt mitmekülgset analüüsivõtete arsenalit.

20 O.C. Ахманова, Г.Б. Микаэлян, Современные синтаксические теории. Изд. МГУ, 1963, lk. 15.

21 Н.А. Слосарева, Лингвистический анализ по непосредственно-составляющим. "Вопросы языкознания" 1960, № 6, lk. 107.

22 Vrd. A.A. Зализняк, Е.В. Падучева, О связи языка лингвистического описания с родным языком лингвиста. - Программа и тезисы докладов в летней школе по вторичным моделирующим системам. Тарту, 1964, lk. 7 - 8.

2. Lause vahetuteks moodustajateks jaotamise idee pärineb L. Bloomfieldilt. Oma taksonoomilise analüüsi printsiipide kirjeldamisel vihjab ta, et analüüs peab lähema vahetutest moodustajatest,²³ mingeid jaotamise kriteeriume ta aga ei anna.²⁴ K. L. Pike, kes proovis L. Bloomfieldi printsiipide põhjal analüüsida üht Mehhiko indiaanlaste keelt, avaldas 1943.a. artikli,²⁵ milles ta kritiseerib L. Bloomfieldi taksonoomiat selle praktilise rakendatavuse seisukohalt ja visandab omapoolse parandatud analüüsivariandi.

Järgneva viie aasta jooksul arutletakse ajakirja "Language" veergudel korduvalt nii VM meetodi teoreetilisi probleeme kui ka praktilise kasutamise üksikasju. Üksteise järel ilmuvad artiklid Z. S. Harriselt (1946),²⁶ R. S. Wellsilt (1947)²⁷ ja R. S. Pittmanilt (1948)²⁸.

2. 1. Z. S. Harris suunda iseloomustab distributiivse analüüsi järjekindel kasutamine keele kõigil tasanditel. Kui morfoloogilisel tasandil on identifitseeritud individuaalsed morfeemid, järgneb nende klassifitseerimine substitutsiooni alusel, millega saadakse morfosüntaktilise

23 L. Bloomfield, op. cit., lk. 161, 209.

24 L. Bloomfieldi analüüsiprotseduur seisneb analüüsitava kõnelõigu lingvistiliste vormide ja neid ühendavate grammatilise organiseerituse (arrangement) tunnusoonte, nn. takseemide kindlakstegemises. Pearõhk langeb konstruktsioonitüüpide määramisele, mistõttu probleemiks pole mitte konstruktsiooni vahetute moodustajate vaheline piir, vaid nende grammatiline seos.

25 K. L. Pike, Taxemes and Immediate Constituents.

"Language" 1943, Vol. 19, No. 2, lk. 65-82.

26 Z. S. Harris, From Morpheme to Utterance. "Language" 1946, Vol. 22, No. 3, lk. 161-183.

27 R. S. Wells, Immediate Constituents. "Language"

1947, Vol. 23, No. 2, lk. 81-117.

28 R. S. Pittman, Nuclear Structures in Linguistics.

"Language" 1948, Vol. 24, No. 3, lk. 287-292.

tasandi üksused - morfeemiklassid. Z. S. Harris on veendunud, et ranget meetodit lause struktuuri üldistamiseks saab luua ainult üheselt identifitseeritavast morfeemist lähtudes.²⁹ Seepärast ongi tal süntaktilise tasandi analüüsiprotseduuri elementideks ainult morfeemid ja morfeemijärjendid (morpheme sequences). Protseduurina kasutab ta sama üksikmorfeemide substitutsiooni tehnikat, kuid laiendab seda nüüd ka morfeemijärjenditele. Kui morfeemiklassi moodustasid morfeemid, mis olid substitueeritavad ümbrusse C - D, siis järgnevalt uuritakse, kas ümbruses C - D esinevad peale morfeemide A ja B ka näiteks morfeemijärjendid AE ja FGH. Kõik sellised morfeemijärjendid, mis on substitueeritavad ümbrusse C - D väljendi grammatilise õigsuse säilimise tingimusel, moodustavad ühise substitutsiooni-klassi. Viimane on aga juba morfeemijärjendite klass, milles ühemorfeemilised järjendid on erijuhtudeks.³⁰

Toome konkreetse huvide triviale näite. Lause grammatilist õigsust säilitades võib paigutada ümbrusse

Kas - on seal?

mitte ainult sõnad

raamat
auto
professor
Jaan
jne.,

29 Z. S. Harris, From Morpheme to Utterance. "Language" 1946, Vol. 22, No. 3, lk. 161.

30 Sealsamas, lk. 165-166. Vt. ka Z. S. Harris, Structural Linguistics. Chicago, 1960, lk. 262-298.

- vaid ka (1) see raamat
 see uus raamat
 see uus raamat, mille ma eile ostsin
 jne.
- (2) sinine auto
 tehase direktori sinine auto
 tehase direktori sinine auto, mis äsja
 ette sõitis
 jne.

Niisiis on lausungis üksikmorfeemid asendatavad morfeemijärjenditega, mis oma koosseisult võivad mitmeti erineda. Kõigi nende järjendite ühiseks omaduseks on i n k l u s i i v s u s , s. o. omadus olla tagasiviidav üksikmorfeemi staatusele, taanduda nn. resulteeruvaks liikmeks (toodud näidetes on resulteeruvateks liikmeteks (1) raamat ja (2) auto).

Selline on kõige üldisemal kujul Z. S. Harrise protseduuri mõte. Jõuti välja e k s p a n s i o o n i kontseptsioonini: kaks morfeemijärjendit, millest üks on teise ekspansioon, kuuluvad samasse substitutsiooniklassi.³¹ Selles on ühtlasi VM teooria tuum. VM analüüs toob esile lause struktuuri kihid, struktuuri h i e r a r h i l i s e e h i t u s e , jaotades lause mittelõikuvateks osadeks (s. o. ükski moodustaja ei lõiku teisega osaliselt, võib ainult kas ise teises sisalduda või teist endas sisaldada). Vahetuteks moodustajateks jaotamine on üldreeglina dihho-toomiline: analüüsi iga samm annab ühe vahetute moodustajate paari. Näiteks lause Sinu kiri, mis eile saabus, oli meeldivaks üllatuseks vahetute moodustajate paarid on

31 R. S. Wells, op. cit., lk. 186.

1. kihis: { Sinu kiri, mis eile saabus
oli meeldivaks üllatuseks
2. kihis: { Sinu kiri { oli
mis eile saabus { meeldivaks üllatuseks
3. kihis: { Sinu { mis { meeldivaks
kiri { eile saabus { üllatuseks
4. kihis: { eile
saabus

Kokku on selles lauses 14 moodustajat, millest 8 on analüüsi lõppüksused, nn. lõplikud moodustajad³² (ultimate constituents).

Oma artiklis ei piirdu Z. S. Harris morfeemilt lausungile viiva protseduuri teoreetilise küljega, vaid demonstreerib ka selle praktilist rakendamist, esitades inglise ja hidatsa keele substitutsioonivõimaluste süsteemi. Substitutsioonid annab ta morfeemiklasside sümbolites³³ võrdustena. Näiteks AN = N tähendab, et morfeemijärjend AN võib igal pool asendada N-klassi iga morfeemi (good boy 'hea poiss' asendab man 'mees'), rikkumata lausungi grammatilist õigsust. Võrdusest DA = A tuleneb, et järjend DA asendab igasugust A-d, ka AN-is, järelikult kui AN = N, siis samuti DAN = N, näiteks poiss = tark poiss = väga tark poiss ümbruses see on - . Tavaliselt on aga substitutsiooni kordumine piiratud (näit. asendamist AN = N, AAN = N, AAAAN = N jne. ei saa lõpmatuseni jätkata). Sel juhul eraldatakse mitteekvivalentsete sümbolid kasva-

32 Erinevalt Z. S. Harrise teooriast on käesolevas näites analüüsi lõppüksusteks sõnad.

33 Tüvimorfeemiklasside sümbolid vastavad sõnaklasside sümbolitele (N - substantiiv, V - verb, A - adjektiiv, D - adverb), sufiksiklasse märgivad sümbolitele lisatud indekseid (näit. V_v - verbisufiks, N_v - substantiividest verbe tuletavad sufiks, jne.).

vate astmenäitajatega, üldkujul $YX = X^2$, mis näitab, et järjendis YX olev X ei ole uuesti asendatav YX -iga (s. t. ei kehti võrdus $YXX = X$). Nii tähistavad võrduste paremal poolel olevad resulteeruvad sümbolid järjendiklasse, mille inklusiivsus kasvab koos sümboli astmenäitaja kasvuga. Näiteks verbifaas tegi toas tööd on väljendatav võrdustega:

$$\begin{array}{ll} \text{tegi} & - VV_V = V^2 \\ \text{tegi tööd} & - V^2 N^2 = V^3 \\ \text{tegi toas tööd} & - V^3 D = V^4 \end{array}$$

(verbi tüvimorfeemist ja sufiksmorfeemist koosnevas järjendis VV_V ei saa tüvimorfeemi korduvalt asendada VV_V -ga, seepärast $VV_V \neq V$; samal põhjusel $V^2 N^2 \neq V^2$ ja $V^3 D \neq V^3$).

Analüüsi tulemusena leiab Z. S. Harris, et inglise keele lausungid on enamikus asendatavad $N^4 V^4$ -ga.

Z. S. Harris vastandab oma protseduuri senisele Bloomfield-Pike'i lausungi analüüsi meetodile. Morfeemist lähtuv substitutsiooniklasside moodustamine on tema arvates seotud vähemate teoreetiliste raskustega kui kriteeriumide leidmine lausungi rangelt järjekindlaks jaotamiseks. Aasta hiljem rõhutab aga R. S. Wells oma artiklis, et mõisted ekspansioon ja substitutsiooniklass moodustavad VM teooria tuuma sõltumatult analüüsi suunast, sest niihästi morfeemist kui ka lausungist alustades on moodustajad ikka samad.³⁴

2. 2. Z. S. Harrise ekspansiooniideed aluseks võttes teeb R. S. Wells katset välja töötada lausungi vahetuteks moodustajateks jaotamise range teooria. Selle sõlmküsimuseks, nagu öeldud, on vahetute moodustajate vahelise p i i r i määramine.

³⁴ R. S. Wells, op. cit., lk. 198.

Juba K. L. Pike oli püüdnud anda juhiseid vahetute moodustajate vahelise piiri määramiseks konstruktsioonides. L. Bloomfieldi järgi jaotatakse nii morfoloogilised kui ka süntaktilised konstruktsioonid vastavalt nende vahetute moodustajate vahekorrale tervikuga 1) eksotsentrilisteks, kus terviku grammatiline karakteristik ei ühti mitte kummagi VM omaga (poiss jookseb, üle metsa), ja 2) endotsentrilisteks, kus see ühtib ühe (hea poiss) või mõlema VM (vend Jaan) grammatilise karakteristikaga. Esimesel juhul on endotsentriline konstruktsioon subordineeriv, s. t. üks tema VM on pea ja teine atribuut. K. L. Pike'i üldpõhimõtteks on paigutada piir eksotsentrilistes konstruktsioonides põhikomponentide vahele (poiss / jookseb, väike poiss / jookseb kiiresti jne.), subordineerivas konstruktsioonis pea ja atribuudi vahele (väike / poiss, jookseb / kiiresti jne.). Edasi püüab K. L. Pike praktilises töös tehtud tähelepanekute najal täpsustada, missugused faktorid nimelt ühel või teisel juhul piiri asukoha otsustavad. Süntaktilistes konstruktsioonides on oluline jaotamine positsioonilise struktuurinõrkuse (positional structural weakness), s. o. struktuuri kõige nõrgema seose põhjal (näiteks tõepoolest, / ma pean, sest lauselaiendi seos ülejäänud lauseosaga on nõrgem kui noomeni- ja verbi-fraasi vaheline seos) ja atribuudi astmete põhjal (väga hea / poiss, mitte väga / hea poiss, sest väga on omakorda hea atribuudiks).³⁵

R. S. Wellsi peamiseks töövõtteks on kõigi võimalike piiride uurimine edasise analüüsi seisukohalt: lausung tuleb jaotada sellisteks osadeks, m i s o n o m a - k o r d a e k s p a n s i o o n i d. Seepärast tuleb jaotada The king of England / opened Parliament 'Inglismaa kuningas avas parlamendi', mitte The king / of England opened Parliament, sest The king of England on king eks-

35 K. L. Pike, op. cit., lk. 74-78.

pansioon ja opened Parliament on opened ekspansioon, of England opened Parliament pole aga millegi ekspansioon.

Teiseks tuleb arvestada nõuet, et VM-d peavad oma distributsioonilt olema teineteisest võimalikult s õ l - t u m a t u d, s. t. sobima maksimaalsesse ümbruste hulka ja kuuluma võimalikult erinevatesse substitutsiooniklassidesse. Selle kriteeriumi põhjal on jaotus the / king of England eelistatavam kui the king / of England, sest king of England sobib samadesse ümbrustesse nagu king, the king aga ei sobi näiteks ümbrusse poor - . Järelikult kuulub king of England suuremasse substitutsiooniklassi kui the king või of England.

Kolmanda kriteeriumina arvestab R. S. Wells s õ n a p i i r i. Iga sõna on moodustaja.

Esineb juhtumeid, kus samal morfeemijärjendil on kahe- sugune tähendus, näiteks vanad mehed ja naised võib tähendada 1) 'naised ja vanad mehed', 2) 'vanad mehed ja vanad naised'. R. S. Wellsi järgi peegeldab seda tähenduslikku erinevust kahe- sugu- ne VM analüüs: 1) vanad mehed / ja / naised, 2) vanad / mehed ja naised,³⁶ mis tuleneb konstruktsioonide erinevusest (esimesel juhul AN + ja + N, teisel A + N ja N). Järelikult peab VM analüüs arvestama konstruktsioone, sest vastasel korral võivad konstruktsioonide homonüümia juhud jääda tähele panemata.

2. 2. 1. Seni on juttu olnud VM-te kui pidevate järjendite piiritlemisest. Kuid keeles on ka selliseid juhtumeid, kus üks VM võib asetseda kiiluna teise sees, tekitades k a t k e s t a t u d (discontinuous) moodustaja. Näiteks ei paikne ühendverbi komponendid lauses sugugi alati kõrvuti (ä r a t a oma sõber ü l e s), verbi ja objekti vahel võib olla adverbiaal (a n n a poisile r a a - m a t), lause keskel võib asetseda kiillause (t a t u l e b, olen selles kindel, v a r s t i t a g a s i) jne. Erinevaid arvamusi on seoses põimlausetega. K. L. Pike kaldub

³⁶ R. S. Wells, op. cit., lk. 193.

positsioonilist struktuurinõrkust silmas pidades eelistama lause When he comes, I will go home 'kui ta tuleb, lähen ma koju' jaotamist pea- ja kõrvallause piiri mööda jaotusele 1) I 'ma , 2) when he comes ... will go home 'kui ta tuleb, lähen ... koju'.³⁷ Ei julge selle üksiku näite põhjal koos N. A. Sljussarevaga järeldada, nagu soovitaks K. L. Pike igasuguse põimlause esimeses kihis eraldada pea- ja kõrvallause.³⁸ See oleks järeldus traditsioonilise grammatika üksustest lähtuva mõtteviisi vaimus. Seepärast on ka kahtlane väita, et H. A. Gleasoni näites The old man who lives there / has gone to his sons house³⁹ 'vana mees, kes seal elas, on läinud oma poja maja juurde' peaksid K. L. Pike'i järgi olema VM-d 1) The old man ... has gone to his sons house 'vana mees on läinud oma poja maja juurde', 2) who lives there 'kes seal elas'.⁴⁰ Kuigi K. L. Pike'il puuduvad vastavad näited, ei ole tõenäoline, et noomenifraasiga tugevasti seotud täiendlauset saaks pida analoogiliseks lauselaienditega nagu really 'tõepoolest', of course 'muidugi', millega K. L. Pike võrdleb oma näite kõrvallauset.⁴¹ Viimase seost verbifraasiga ei pea ta nähtavasti tugevamaks pea- ja kõrvallause vahelisest seosest.

Katkestatud moodustajate küsimust püüab R. S. Wells lahendada lihtsa juhisega: mittepidev järjend on moodustajaks, kui mingis ümbruses esineb moodustajana vastav pidev järjend, kusjuures mõlemad konstruktsioonid on semantiliselt kooskõlas.⁴² See reegel õigustab K. L. Pike'i jaotust, sest katkestatud järjend when he comes ... will go home ei esine pidevana üheski ümbruses. H. A. Gleasoni lause

37 K. L. Pike, op. cit., lk. 77.

38 Vrd. H. A. Слюсарева, op. cit., lk. 101.

39 Г. Глясон, op. cit., lk. 188.

40 Vrd. H. A. Слюсарева, op. cit., lk. 101.

41 K. L. Pike, op. cit., lk. 78.

42 R. S. Wells, op. cit., lk. 200.

analüüsiks jätab reegel mõlemad võimalused. Korrigeerima peab R. S. Wells aga K. L. Pike'i koordineerimise the large books, the important papers, the new pens 'suured raamatud, tähtsad paberid, uued suled' jaotust kaheks katkestatud moodustajaks: 1) the large ... the important... the new, 2) books ... papers ... pens.⁴³ R. S. Wellsi järgi on siin tegemist kolme endotsentrilise konstruktsiooniga, mis ühtlasi on ka VM-d.⁴⁴

2. 2. 2. Nagu juba öeldud, kehtib VM analüüsis dihhotoomilise jaotuse printsiip. Järjekindlalt peab sellest kinni K. L. Pike, rohkem kui kahe VM võimalust mõnivad R. S. Wells, Ch. F. Hockett jt. Mitmikmoodustajate (multiple constituents) määramisel kasutab R. S. Wells reeglit: kui ei leidu mingit põhjendust järjendi ABC analüüsimiseks pigem AB / C kui A / BC või vastupidi, tuleb analüüsida A / B/C.⁴⁵ Järelikult on mitmikmoodustajatega tegemist just koordineerimise puhul. Sidesõnu peab R. S. Wells täieõiguslikeks moodustajateks, analüüsides A / ja / B.⁴⁶ Ch. F. Hockett eraldab nad aga moodustajatest, nimetades neid ja ka teisi abisõnu struktuuraseteks markeerijateks (structural marker), sest nad ei ole iseseisva tähenduse kandjaks, vaid markeerivad ainult sõnadevahelist seost.⁴⁷

2. 3. Z. S. Harris ja R. S. Wellsi VM teooria andis võtme mitte üksnes lause struktuuriüksuste piiritlemiseks, vaid tõi ühtlasi esile lause hierarhilise ehituse, mis tuleneb inklusiivsusevahekorra eest eri kihtide moodustajate vahel. Igasugune struktuuri kirjeldus on eeskätt elementide vahekordade kirjeldus. Täpsemalt: iga elemendi kirjeldamisel tuleb arvestada tema vahekordi struktuuri kõigi ülejäänud elementide ja nende kõikvõimalike kombinatsioonide

⁴³ K. L. Pike, op. cit., lk. 78.

⁴⁴ R. S. Wells, op. cit., lk. 201.

⁴⁵ Sealsamas, lk. 199.

⁴⁶ Sealsamas.

⁴⁷ Ch. F. Hockett, op. cit., lk. 153.

dega. VM printsiip lihtsustab lause struktuuri kirjeldust tunduvalt, sest kõik võimalikud ümbruse faktorid taanduvad lihtsalt antud elemendi kui VM juurde kuuluvaks teiseks VM-ks.⁴⁸ Nii elimineeritakse kirjelduse igal tasandil kõik ebaoluline. Näiteks lauses Sinu kiri, mis eile saabus, oli meeldivaks üllatuseks on konstruktsiooni sinu kiri puhul arvesse tulevaks ümbruseks mis eile saabus, sinu puhul aga ainult kiri.

Kõige sagedam on vahetute moodustajate subordinatsioonivahekord: üks neist on domineeriv, nn. tuum (nucleus), teine, nn. satelliit, alistub tuumale.⁴⁹ R. S. Pittman püüab oma artiklis "Nuclear Structures in Linguistics" lahendada küsimust, missuguste printsiipide põhjal saab vahet teha tuuma ja satelliidi vahel. Peale tuuma sõltumata ja endotsentrilisuse printsiibi peab ta oluliseks statistilisi kriteeriume, mis arvestavad

1) klassi mahtu, kuhu antud VM kuulub (suurema mahuga klassi kuuluv element on tuum, mistõttu näiteks verbi ja adverbi ühendis on tuumaks verb);

2) potentsiaalse esinemisvaldkonna suurust (tuum on see VM, mis võib esineda suurema arvu eri klassidest koosmoodustajatega, näiteks adjektiivi ja substantiivi ühendi tuumaks on substantiiv, sest viimane esineb mitmekesisemates ümbrustes);

3) klassi sagedust (suurema esinemissagedusega klassi kuuluv VM on tuum);

4) individuaalset sagedust (väiksema esinemissagedusega individuaalsed moodustajad on tuumad, sest sagedussõnastikes on kõrgem esinemisprotsent peamiselt satelliidi tüüpi sõnadelt).

Lisakriteeriumidena tulevad arvesse ka prosoodia ja tähendus.

48 R. S. Pittman, op. cit., lk. 287.

49 Bloomfield-Pike'i vastavad terminid on p e a ja a t r i b u u t (vt. lk. 15).

Tuuma ja satelliidi küsimust ei teki koordineerivate konstruktsioonide puhul. Subjekti ja predikaadi suhtes on aga raske öelda, kumb kumma üle domineerib. R.S. Pittman sellele selget vastust ei anna, vaid nimetab neid lihtsalt kollateraalseteks klassideks, mis loetletud kriteeriumide seisukohalt on peaaegu võrdsed, kuid mitte identsed.⁵⁰

Originaalselt on selle küsimuse lahendanud S. K. Šaumjan ja P. A. Soboleva, kes käsitlevad domineerimist liigendatult: 1) lause arendamise ja 2) lause taandamise aspektist. Esimesel juhul domineerib alus, sest lause ülesehitusprotsessis tingib ühildumine öeldise sõltuvuse alusest. Teisel juhul domineerib öeldis, sest lause ei taandu mitte aluseks, vaid öeldiseks, mis on minimaalne predikatsioonijuhtum.⁵¹

2. 4. Senistest käsitlustest hoopis erinevalt läheneb VM analüüsi teooriale S. Chatman.⁵² Ekspansiooniprintsiip tema arvates lause struktuuri ei ammenda, sest see arvestab ainult kvantitatiivseid momente, eristamata seejuures moodustajatevahelisi tähtsamaid piire vähemtähtsatest. Piiride hierarhia kindlakstegemiseks määrab S. Chatman lauses igale morfeemile järgnevad potentsiaalsed ümbrused Ch. C. Friesi sõnaklasside terminites. Mida rohkem sõnaklasse võib mingile morfeemile järgneda, seda tähtsam on struktureaalne piir. Analüüsitud lauses The boy played near the house 'poiss mängib maja lähedal' on näiteks fraasile The boy järgneval klassil 15 varieerumisvõimalust, fraasile The boy played järgneval aga 10, seega tähtsam piir on boy järel. Protседuur tõestab, et lause struktuuri hierarhiat on võimalik määratleda informatsiooniteooria seisukohalt, mõõtes iga järgneva elemendi määramatuse ast-

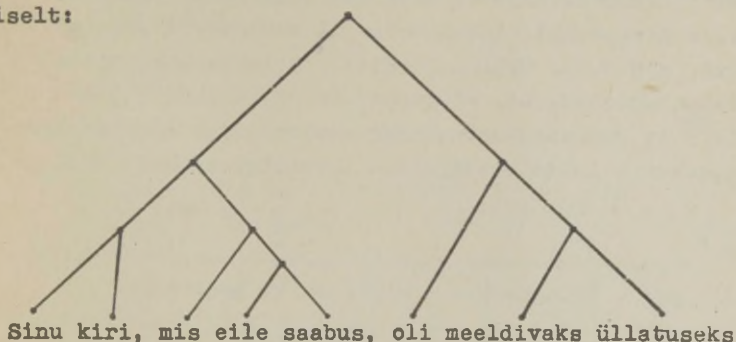
⁵⁰ R. S. Pittman, op. cit., lk. 292.

⁵¹ С.К. Шаумян, П.А. Соболева, Аппликативная порождающая модель и исчисление трансформаций в русском языке. Москва, 1963, lk. 17.

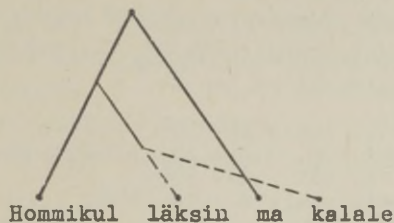
⁵² S. Chatman, Immediate Constituents and Expansion Analysis. "Word" 1955, Vol. 11, No.3, lk. 377-385.

me, mis on väljendatav järgneda võivate sõnaklasside arvuga.

2. 5. Lause VM struktuuri ülevaatlikumaks kujutamiseks kasutatakse mitmesuguseid skeeme. Neist kõige levinum ja ühtlasi otstarbekaim on hargmik (ingl. tree, vn. дерево). Eespool toodud näitelause on hargmiku abil kujutatav järgmiselt:



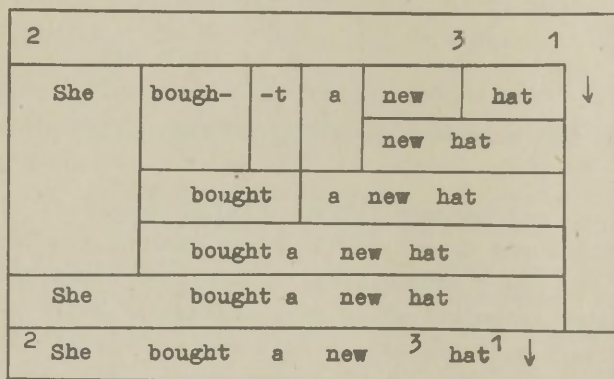
Sõlm, millest hargmik lähtub, kujutab lausetervikut. Iga järgnev sõlmede tasand vastab ühele analüüsikihile (vrd. lk. 13). Lause moodustajaks on iga tema sõnajärjend, millele vastab hargmikul sõlm, kust laskumisel kõiki võimalikke harusid pidi jõutakse ainult selle järjendi sõnadeni, näiteks mis eile saabus on moodustaja, kiri, mis eile saabus aga mitte. Hargmikskeemiga on kujutatav ka katkestatud moodustajaid sisaldava lause struktuur,⁵³ näiteks



53 Vrd. E.B. Падучева, О способах представления синтаксической структуры предложения. "Вопросы языковедения" 1964, № 2, lk. 106.

kus moodustaja läksin kalale on katkestatud moodustajaga ma.

VM eri analüüsikihte ja inklusiivsust kujutab ilmekalt ka kastdiagramm. Selle diagrammitüübi võttis kasutusele K. L. Pike⁵⁴ ja mõnede täiendustega rakendab seda Ch. F. Hockett⁵⁵. Viimane kujutab diagrammil ka lause suprasegmentaalse komponendi (intonatsiooni) eraldamist segmentaalsest, mis R. S. Wellsi järgi⁵⁶ on lause maksimaalselt sõltumatud moodustajad. Järgnevas Ch. F. Hocketti diagrammis⁵⁷ on suprasegmentaalseks komponendiks inglise keele jutustavale lausele tüüpiline intonatsioonikurv 231↓.



Selline kastdiagramm sobib ainult lühikeste, pidevatest moodustajatest koosnevate lausete kujutamiseks. Katkestatud moodustajate puhul peab Ch. F. Hockett abivahenditena kasutama sulgusid või noolt.⁵⁸

⁵⁴ K. L. Pike, op. cit., lk. 70.

⁵⁵ Ch. F. Hockett, op. cit., lk. 151-155 jm.

⁵⁶ R. S. Wells, op. cit., lk. 206.

⁵⁷ Ch. F. Hockett, op. cit., lk. 158.

⁵⁸ Sealsamas, lk. 155.

VM eraldamiseks on kasutatud ka ainult sulgusid, mis asetatakse nii, et iga VM oleks eraldatud ühe sulgudepaariga, näiteks

(Ta) (ostis (uue (kõbara))).

Seegi viis tuleb arvesse ainult pidevate moodustajate puhul. Lause hierarhilise struktuuri näitamiseks pole ta kuigi ülevaatlik.

* *

.

3. On olemas kaks lingvistilise kirjelduse võimalust:

1) kõne⁵⁹ faktide, s. t. süntagmaatiliste nähtuste analüüsist tuletatakse järeldused keele süsteemi kohta;

2) keele süsteemist lähtudes antakse teatav operatsioonide loend kõne süntagmaatika reprodutseerimiseks.

Esimesel juhul on tegemist antud keeles esinevate lausete mingi piiratud hulga analüüsiga, teisel juhul - antud keeles esineda võivate lausete piiramat hulga sünteesi reeglitega. Viimase all mõeldakse tänapäeval genereerivat grammatikat (generative grammar), mis kujutab endast sünteesivat mehhanismi mingi keele kõigi grammatiliselt õigete järjendite tekitamiseks.⁶⁰

Keele kui tegelikkuse komplitseeritud objekti tunnetamise protsessis on genereeriv grammatika kõrgemaks astmeks kui mis tahes meetodiga tehtud analüüs. Kuid süntees pole mõeldav eelneva analüüsita. Seepärast põhineb iga genereeriv mudel mingi analüüsiprotseduuri tulemusena formuleeritud üksustel.

Analüüsiprotseduuri tuleks hinnata selle järgi, kui võrd-saadud tulemused peegeldavad keele süsteemisuhteid.

59 Kõne ja keele kui lingvistiliste terminite kohta vt. H. Rätsep, Keele ja kõne eristamisest. - Nonaginta. Tallinn, 1963, lk. 243-259.

60 Vrd. H. Хомский, Синтаксические структуры. Новое в лингвистике, вып. II. Москва, 1962, lk. 416.

Konkreetsete kriteeriumide puudumise tõttu on siin aga otsuse langetamine raske. Genereeriva mudeli adekvaatsust on võimalik hinnata, kõrvutades mudeli põhjal genereeritud järjendid keeles faktiliselt olemas olevate fraasidega. Parem on see mudel, mis genereerib rohkem keele fraasidele vastavaid järjendeid ja vähem mittevastavaid.⁶¹ Sel teel genereerivale mudelile antud objektiivse hinnangu kaudu saab lõppeks hinnangu ka analüüsimeetod, mis oli mudelile aluseks.

Deskriptiivse lingvistika VM analüüsil põhineb N. Chomsky VM mudel ehk fraasi struktuuri grammatika.⁶² Kontrollides selle adekvaatsust, leiab N. Chomsky, et VM grammatika reeglite abil on otstarbekas genereerida ainult kõige lihtsamat tüüpi lauseid, kuna keerukamate, näiteks koordinatsiooni sisaldavate, samuti passiivsete lausete genereerimine selle grammatikaga on raske.⁶³

VM reeglitega genereeritavad nn. t u u m l a u s e d ja nende ekspansioonid moodustavad grammatika tuuma. Selle tuuma mitmesuguse ümberkujundamise ehk transformeerimise tulemuseks on keele lausete lõpmatu mitmekesisus. Keele tuumlaused kujutavad endast lõplikku hulka elementaarseid lausekonstruktsioone, mis on üksteisest maksimaalselt sõltumatud.⁶⁴ Näiteks R. T. Harmsi eesti keele genereerivas grammatikas on saadud fraasi struktuuri reeglite põhjal 8 tuumlausete tüüpi,⁶⁵ mis on interpreteeritavad järgmiselt:

- 1) transitiivne, ühe objektiga (ma teen tööd);
- 2) transitiivne, kahe objektiga (ma annan sulle raa-matu);

61 Vrd. И.И. Рёван, Модели языка. Москва, 1962, lk. 14.

62 Н. Хомский, op. cit., lk. 430-438.

63 Sealsamas, lk. 439-450.

64 З.С. Харрис, Совместная встречаемость и трансформация в языковой структуре. Новое в лингвистике, вып. II. Москва, 1962, lk. 629.

65 R. T. Harms, Estonian Grammar. 1962, lk. 178-180.

- 3) predikatiivne (ilm on ilus);
- 4) intransitiivne, verbi laiendita (bundid surid);
- 5) intransitiivne, adverbiga (ma lähen sinna);
- 6) subjektita (sajab);
- 7) subjektita, adjektiivpredikatiiviga (on külm);
- 8) possessiivse olema-konstruktsiooniga (mul on raa-
mat).

N. Chomsky transformatoorse grammatika järgi moodustavad mingi keele lausete koguhulga tuumlaused koos oma ekspansioonidega + nende mitmesugused transformid.⁶⁶ Deskriptiivses lingvistikas VM analüüsi rakendamisel tekkinud raskusi (mitmikmoodustajate ja katkestatud moodustajate probleeme) ongi seletatud lihtsalt sellega, et VM meetodi abil püüti analüüsida kõiki lauseid, ka selliseid, mis VM meetodile ei allu, sest nad on teiste lausete transformeerimise tulemus.⁶⁷

Missugust osa peavad grammatikas etendama transformatsioonid, on genereerivate mudelite teooria praegusel arenguastmel alles vaieldav. N. Chomsky on transformatoorne mudel kogu grammatikateooria keskpunktiks. Eri keelte lausestruktuuri selgitamisel on N. Chomsky mudelit juba edukalt rakendatud.⁶⁸ S. K. Šaumjani ja P.A. Soboleva aplikatiivses genereerivas mudelis aga ei ole transformatsioon enam lausete genereerimise vahend.⁶⁹ Ilma transformatsioonireegliteta, ainult fraasi struktuuril põhineva

66 Н. Хомский, op. cit., lk. 452. Vrd. З.С. Хэррис, op. cit., lk. 627.

67 З.С. Хэррис, op. cit., lk. 632.

68 Sellel põhinevad näiteks R. B. Lees, The Grammar of English Nominalizations. Baltimore, 1960; R.T. Harms, op. cit.; M. Bierwisch, Grammatik des deutschen Verbs. Studia Grammatica II. Berlin, 1963; W. Motsch, Syntax des deutschen Adjektivs. Studia Grammatica III. Berlin, 1964.

69 С.К. Шаумян, П.А. Соболева, op. cit., lk. 6.

generatiivse grammatika adekvaatsust põhjendab G.H. Harman.⁷⁰ VM kui keele tuumstruktuure peegeldava mudeli väär-
tus on aga juba praegu vaieldamatu.

⁷⁰ G. H. Harman, Generative Grammars without Transformation Rules. A Defense of Phrase Structure. "Language" 1963, Vol. 39, No. 4, lk. 597-616.

TULEMA, SAAMA JA PIDAMA
TÄHENDUSTE STRUKTURAALNE ANALÜÜS

H. Õ i m

O. S i s s e j u h a t u s. Käesolevas töös esitatakse eesti keele verbide tulema, saama ja pidama tähenduste klassifikatsiooni. Klassifitseerimiseks kasutatakse strukturealse lingvistika aparatuuri, eriti distributiivse ja transformatoorse analüüsi meetodeid. Nende rakendamisel on põhiliselt lähtutud J. D. Апресяни¹ ja V. I. Perebeinose² poolt väljatöötatud metoodikast.

Analüüs toetub eeldusele, et erinevus leksikaalsete tähenduste vahel, kui see on antud keeles relevantne, avaldub ka mingites struktuurierinevustes, mistõttu võime alati, kui oleme kindlaks teinud teatava struktuurilise vastanduse, oletada ka olulise semantilise vastanduse olemasolu. Antud metoodika puhul on nendeks struktuuritunnusteks, mille põhjal otsustatakse tähenduste erinevuse või samasuse üle, süntaktilised tunnused (lähemalt vt. 1.2. ja 1.3.).

Selline metoodika võimaldab välja selgitada nn. diferentsiaalsed tähendused, s. t. me võime kindlaks teha ainult tähenduste erinevused või samasused (väärtused Saussure'i mõttes), aga me ei saa midagi teada uuritavate

- 1 D.Д. Апресян, О понятиях и методах структурной лексикологии. Проблемы структурной лингвистики. Москва, 1962, lk. 141-162.
- 2 В.И. Перебийнос, К вопросу об использовании структурных методов в лексикологии. Проблемы структурной лингвистики. Москва, 1962, lk. 163-174.

tähenduste sisulise külje kohta. Saadud klassifikatsiooni sisulist olemust võime selgitada ainult interpretatsiooni korras.

Käesolevas töös analüüsitakse verbe tulema, saama ja pidama ainult iseseisvatena. Ei analüüsita juhtumeid, kus uuritavad verbid on fraseoloogiliste üksuste või ühendverbide koostisosad.

Analüüsi materjaliks olid ilukirjanduslikud tekstid.³ Iga verbi kohta on sedeldatud umbes 350 näitelauseid.

1. A n a l ü ü s i m e e t o d .

1. 1. K o k k u l e p p e d . Analüüsi teostamiseks on vajalikud järgmised kokkulepped.

1. 1.1. Me oskame eristada märgitud fraase mitte-märgitud fraasidest.⁴ Märgituiks loeme käesolevas töös ainult need fraasid, mis on õiged nii grammatiliselt kui ka leksikaalselt. Leksikaalne õigsus on seejuures muidugi täiesti tinglik mõiste, sest see sõltub eelkõige kontekstist, milles vastav fraas esineb. Teatavas kontekstis võib fraas värvitatud rohelised ideed magavad raevukalt olla täiesti "õige". Oma töös loeme sellised fraasid siiski mitteõigeks, s. o. me vaatleme neid harilikus ja mitte mingis erilises kontekstis.

1. 1.2. Me oskame kindlaks teha märgitud fraasi struktuuri. Selleks on vaja:

- 1) teada sõnade grammatilisi klasse antud keeles;
- 2) osata iga sõna puhul määrata grammatiline klass, millesse ta kuulub;
- 3) osata kindlaks teha sõnade süntaktilisi vahekordi.

Teisiti öeldes, me eeldame, et meil on teada antud keele formaalne grammatiline mudel. Eesti keele formaalset grammatilist mudelit pole tegelikult muidugi olemas, ent

3 E. Krusten, Kevadet otsimas. Tallinn, 1960; L. Promet, Roosa kübar. Tallinn, 1962; J. Saaul, Jäine raamat. Tallinn, 1962; Fr. Tuglas, Teosed II. Tallinn, 1957.

4 Fraaside märgitusest seoses tähenduse analüüsiga vt. näit. D. A. Апресян, op. cit., lk. 143-144.

meie töö jaoks piisab ka traditsioonilise grammatika seda laadi andmeist, kui neid vastavalt vajadusele modifitseerida.

Eeldame, et meil on antud sõnade grammatilised klassid N, V, A, D ja peale nende loendina kõik abielemendid, s. o. sõnad, mis ei kuulu loetletud nelja klassi: abimäärsõnad, kaassõnad jms. Seejuures tuleb märkida, et antud jaotus ei ühti mitte igas mõttes traditsioonilise jaotusega sõnaliikideks. N-i alla kuuluvad kõik nimisõnad, aga ka teatud ase- ja arvsõnad (mis esinevad nimisõna funktsioonis); V-ga tähistame verbid ja näiteks ka fraseoloogilised ühendid, mis üheskoos esinevad väljendverbina; A-sse kuuluvad omadussõnad ja ka osa ase- ning arvsõnu (mis esinevad omadussõna funktsioonis); D alla kuuluvad (iseselsvad) määrsõnad ja muud sõnad, mis esinevad määrsõna funktsioonis.

Antud eeldustele toetudes võime asuda sõna tähenduste struktuuralsele analüüsimisele. Analüüs jaguneb kahte etappi vastavalt kasutatavale metoodikale. Esimesel etapil klassifitseeritakse sõna tähendused nende distributiivsete iseärasuste põhjal (sõnaklasside tasemel). Teisel etapil analüüsitakse ühesuguse distributsiooniga tähendusi edasi transformatoorse metoodikaga.

1. 2. D i s t r i b u t i i v n e k l a s s i - f i k a t s i o o n. Mitmetähendusliku sõna erinevaid tähendusi võime sageli ära tunda sõna erineva süntaktilise seostumise põhjal. Kui näiteks saama esineb fraasides ma saan raamatu ja ma saan õpetajaks, siis ümbruse erinevuse tõttu võime ka verbi pidada kummalgi juhul erinevas tähenduses olevaks. Selliste erinevate ümbruste väljaselgitamises ja nende põhjal sõna tähenduste klassifitseerimises ongi antud juhul distributiivse analüüsi põhiidee.

1. 2.1. T a a n d a m i n e. Laused, kus esineb meie poolt uuritav sõna, sisaldavad sageli palju selliseid sõnu ja morfeeme, mis uuritava sõna tähenduse seisukohalt on liigsed. Näiteks on ilmne, et lauses punase peaga poiss

sai eile paksu raamatu pole verbi saama tähenduse suhtes (meie töö seisukohalt) mingil määral oluline see, et saaja oli punase peaga, et raamat oli paks või et saamine toimus eile. Selleks et meil ümbruste klassifitseerimisel ja nendega opereerimisel ei oleks pidevalt vaja selliseid asjasse mittepuutuvaid elemente arvestada, tuleb nad juba enne klassifitseerimisele asumist kõrvaldada. Sel otstarbel tehaksegi lausetega kõigepealt läbi taandamisoperatsioon. See seisneb selles, et endotsentrilised konstruktsioonid asendatakse nende pealiikmetega ja jäetakse arvestamata käesoleva analüüsi seisukohalt ebaolulised grammatilised kategooriad. Viimaste olemasolu ilmneb näiteks lausetest ta saab raamatu, nad saavad raamatu, nad olevat saanud raamatu jne., kus saama tähendus on sama, ehkki ümbrused on erinevad. See, missugused kategooriad on olulised ja missugused ebaolulised, otsustame harilikult intuitsiivselt, ent seda võib kontrollida ka range protseduuri abil.⁵ Käesoleva töö seisukohalt on olulisteks (relevantseteks) kategooriateks nimi- ja omadussõnad (N ja A) nende kääne ning verbidel käändelised vormid.

Seega, kui meil tekstis esineb näiteks lause paks pruuni palituga mees tuli aeglaselt lonkides minu poole, siis pärast ebaoluliste sõnade väljajätmist (paks mees → mees, pruuni palituga → palituga, palituga mees → mees, aeglaselt lonkides → lonkides, tuli lonkides → tuli, tuli minu poole → tuli) saame selle kujul mees tuli.

Selliseid taandatud lauseid võib nimetada orienteeritud fraasideks;⁶ lühiduse pärast nimetame neid edaspidi sageli ka lihtsalt fraasideks.

Pärast ebaoluliste kategooriate taandamist (ja kodeerimist - vt. 1.2.3.) omandab eespool toodud fraas kuju $N_n V^a$.

5 Д.С. Уорс, Трансформационный анализ конструкций с творительным падежом в русском языке. Новое в лингвистике, II. Москва, 1962, lk. 645-646.

6 Vt. Ю.Д. Апресян, op. cit., lk. 145.

1. 2.2. Orienteeritud fraaside korrastamine. Pärast taandamist on osaga saadud fraasidest otstarbekas läbi teha veel rida korrastavaid operatsioone, selleks et saada ühtne alus ümbruste võrdlemiseks. Need operatsioonid on mitte niivõrd nõutavad, kui võrd lihtsalt otstarbekad. Toetudes toodud eeldustele (vt. 1.1.), võiksime identsed ümbrused ära tunda ka selliste ümberkorraldusteta; ent praktiliselt on siiski mugav, kui ümbruste erinevus ja samasus on ka n.ö. silmaga nähtav. Esitame siin vaid üldised põhimõtted, millest lähtume ümberkorralduste teostamisel.

(1) Fikseerime järgmise positsioonilise mudeli: esimesel kohal subjekt, seejärel predikaat ning siis selle laiendid. (Seega: kõne pidas esimees → esimees pidas kõne jne.)

(2) Osaaluse transformeerime täisaluseks (poisse tuli → poisid tulid).

(3) Konstruksioonid $V^1 V^2$ da/ma, kus V^2 on uuritav verb (tulema, saama või pidama), asendame V^2 vastava vormiga (hakkab tulema → tuleb, tahab saada → saab, mõtles pidada → pidas).

1. 2.3. Kodeerimine. Viimase operatsiooni nina enne ümbruste võrdlemist ja selle alusel sõna tähenduste klassifitseerimist on vaja konkreetset sõnad asendada vastavate sõnaklasside sümbolitega. Seejuures märgime iga sümboli juurde vajaliku grammatilise informatsiooni kodeeritud sõna kohta (s. o. olulised grammatilised kategooriad, vt. 1. 2.1.).

Käänded märgime järgmiselt (vastavalt käänete järjekorrale): N_n , N_{om} , N_{os} , N_{sse} , N_s , N_t , N_{le} , N_l , N_{lt} , N_{ni} , N_{ks} , N_{na} , N_{ga} , N_{ta} (ja täpselt niisamuti A puhul). Verbidel: V_{ma} - ma- tegevusnimi; V_{da} - da-tegevusnimi, V_{tud} - tud-kesksõna; V_{nud} - nud-kesksõna; V_ϕ - verbi umbisikuline vorm; V - verbi isikuline vorm. Sageli on vaja ka

määrsõnade puhul opereerida detailsema jaotusega. Töös kasutatakse järgmist tähistust: D_{sia} - määrsõnad, mis määravad suundumiskohta (sia, sinna, üles jne.); D_{siit} - määrsõnad, mis märgivad tegevuse lähtekohta (siit, sealt, ülalt jt.); D_{siin} - määrsõnad, mis märgivad olemiskohta (siin, seal, ülal jt.). Ülejäänud määrsõnad tähistame lihtsalt D-ga.

Kui ühes fraasis esineb mitu samasse klassi kuuluvat sõna, siis tähistame nad kodeerimisel erinevate numbritega vastavate sümbolite juures: N^1 , N^2 , V^1 , V^2 jne. Uuritavate sõnade jaoks kasutame erinevat tähistust: V^a - tulema, V^b - pidama, V^c - saama.

Orienteeritud fraasi, kus sõnad on asendatud vastavate sümbolitega, nimetame distributsioonivormeliks. Nii näiteks saame fraasidest ma sain raamatu, lapsed saavad hariduse distributsioonivormeli $N_n^1 V^a N_{om}^2$.

Seoses sellise sümboolse tähistusega tuleb silmas pidada asjaolu, et sümbolid N , V , A , D konkreetseis distributsioonivormeleis ei esinda mitte kõiki vastava sõnaklassi liikmeid, vaid ainult neid, mille puhul antud distributsioonivormel jääb leksikaalselt märgituks.

1. 2.4. Iga distributsioonivormel defineerib uuritava sõna mingi tähenduse; J. D. Apresjani eeskujul võime selliseid tähendusi nimetada lekseemideks.⁷ Distributiivse analüüsi lõpuks saadud lekseemid ei kujuta endast sageli aga mingit ühtset tähendust; üks distributsioonivormel võib esindada mitut ja vahel üsnagi erinevat tähendust. Selleks et neid lahutada, ongi vajalik edasine transformatoorne analüüs.

1. 3. T r a n s f o r m a t o o r n e k l a s s i -
f i k a t s i o o n. Pärast seda kui fraasid on klassifitseeritud distributsioonivormelitesse, allutame igas dis-

7 Ю.Д. Апресян, op. cit., lk. 145.

tributsioonivormelis kõik fraasid reale transformatsioonidele (teoreetiliselt: kõikvõimalikele transformatsioonidele). Kui ilmneb, et kõik fraasid ei reageeri neile transformatsioonidele ühesuguselt, siis võime selle põhjal järeldada, et meil on tegemist ka vastavalt erinevate tähendustega. Näiteks avastame transformatsiooni

$T : N_n^1 V^c V_{da}^2 \rightarrow V_\phi^2 N_{os}^1$ abil distributsioonivormelis

$N_n^1 V^c V_{da}^2$ kaks erinevat saama tähendust, sest osa fraase reageerib transformatsioonile positiivselt: ta saab peksta $\rightarrow$ teda pekstakse, ta sai sõimata $\rightarrow$ teda sõimati, osa aga negatiivselt: ta saab joosta $\nrightarrow$ teda joostakse, poisid said mõelda $\nrightarrow$ poisse mõteldi. Viimase näite kohta tuleb lisada, et ehkki transform poisse mõteldi on märgitud fraas, on tema sisuline erinevus transformeeritavast fraasist liiga ilmne, selleks et nendevahelist vahet luge-da transformatsiooniks.

Transformatoorse analüüsi tulemusena jagunevad ühe distributsioonivormeli alla kuuluvad morfoloogiliselt samatüübilised fraasid reaks alltüüpideks; iga selline alltüüp defineerib palju konkreetsema ning ühtsema lekseemi kui distributsioonivormel.

Sõna lekseemide loendi, mille saame transformatoorse analüüsiga, loeme lõpploendiks. Tähenduslikud erinevused, millele ei vasta distinktiivset vormelit sõnaklasside tasemel ja millel me ei avasta ka mingeid transformatoorseid iseärasusi, peame oma meetodi seisukohalt lugema ebaoluliseks.

1. 3.1. Natuke lähemalt tuleb selgitada töös kasutatavat transformatsiooni mõistet. Operatsioonid, mida me siin kasutame transformatsiooni nimetuse all, on peamiselt kaht tüüpi: (1) operatsioonid, kus fraasile lisatakse juurde (või jäetakse fraasist välja) mõni sõna (vt. näit. 3.1.), ja (2) transformatsioonid kitsamas mõttes, kus transform sisaldab parajasti neidsamu sõnatüvesid, mida

transformeeritav fraaski (vt. näit. 3.1.2., 3.3.2. jt.). Ühel juhul (vt. 3.1.3.2.) on kasutatud kontrolloperatsioon T: $\rightarrow N^a$ on N^b , mille abil võime selgitada N^a ja N^b vahelise seose laadi. Terminoloogilise ühtsuse pärast nimetame oma töös kõiki neid operatsioone transformatsioonideks, toetudes seejuures eelkõige D. S. Worthi eeskujule.⁸

1. 4. Lõpuks tuleb analüüsi meetodi kohta öelda veel järgmist. Antud meetod võimaldab välja selgitada sõna diferentsiaalsed tähendused, s. o. tähenduste erinevused. Antud meetod ei võimalda aga objektiivselt hinnata saadud tähenduste (lekseemide) sisulist vahekorda. Ühe distributsioonivormeli piires võime küll vastavalt transformatsioonide kasutamise järjekorrale kehtestada teatava hierarhia. Ent see hierarhia on puhtformaalne. Sageli võivad ühes distributsioonivormelis koos esineda tähendused, mis sisuliselt tunduvad olevat mõnede teistesse distributsioonivormelitesse kuuluvatele lekseemidele palju lähedasemad kui sama vormeli teistele lekseemidele. Erinevatesse distributsioonivormelitesse kuuluvate lekseemide vahekorda ei võimalda aga antud meetod välja selgitada ja seeõttu pole välistatud võimalus, et mõned erinevatesse distributsioonivormelitesse kuuluvad tähendused on täiesti "samad".

Üks osa kirjeldatud probleemist, nimelt lekseemide üldisusvahekord (missugune lekseem esindab sõna põhitähendust ja missugune kõrvaltähendusi ning missuguses üldisusvahekorras on need kõrvaltähendused) lahendatakse harilikult mitmesuguste abikriteeriumide sissetoomisega. Perebeinos kasutab nendena järgmist kolme:

1) vormeli esinemissagedus (mida sagedamini esineb vormel, seda üldisem on tema poolt defineeritav tähendus);

2) vormeli struktuuri keerukuse aste (mida lihtsam on struktuur, seda üldisem on tähendus);

3) vormeli poolt defineeritava tähenduse sõnatuletus-

8 Д. С. Ворс, op. cit., lk. 637-683.

lik võime (mida enam tuletisi saab moodustada antud tähenduses esinevast sõnast, seda üldisem on tähendus).⁹

Nende abil osutub harilikult võimalikuks välja selgitada sõna tõenäoline põhitähendus, üksikute kõrvaltähenduste vahekorra üle on aga raske otsustada; peab olema läbi töötatud väga ulatuslik materjal, et andmed vormelite sageduste kohta oleksid objektiivsed.

Käesolevas töös on neid kriteeriume kasutatud ainult sõnade tõenäoliste põhitähenduste väljaselgitamiseks.

2. Distributiivse klassifikatsiooni tulemused. Eespool kirjeldatud meetodi järgi teostatud distributiivse klassifikatsiooni tulemusena rühmitusid analüüsitud laused (mõnede eranditega) järgmistesse distributsioonivormelitesse.

2. 1. tulema.

1) $N_n V^a$. Poiss tuli; teade tuli; tervis tuleb; traktor tuleb; tuleb Rooma; tuli mõte; tuleb lahkumine jt.

2) $N_n^1 V^a N_{le}^2$. Ma tulin mõttele; Setšenov tuli järel-dusele; õu tuleb rookimisele; asi tuleb arutamisele jt.

3) $N_n^1 V^a N_{st}^2$. Neist tulevad kotletid; sellest tuleb näidend; poisist tuleb lendur; see tuleb vaesusest; vaesus tuleb rumalusest jt.

4) $N_n^1 V^a A_n$. Kartul tuleb puhas; surp tuli maitsev; õhtu tulevat põnev jt.

5) $V^a V_{da}^2 (N_1)$. Tuli teha remont; tuleb rajada lennu-väli; (meil) tuli kokku puutuda.

2. 2. pidama.

Verbi pidama võime jagada kaheks juba enne süntaktilise analüüsi juurde asumist - ühelt poolt pidama minevikuvormiga pidas ja teiselt poolt pidama minevikuvormiga pidi (tähistame vastavalt V_1^b ja V_2^b). Esimene neist moodus-

9 В.И. Перебейнос, op. cit., lk. 170.

tab rea erinevaid distributsioonivormeleid, teine aga
käitub distributsioonitasandil ühtselt.

2. 2.1. V_1^b

1) $N_n^1 V_1^b N_{os}^2$. Katus peab vihma; direktor peab kõnet;
rahvas peab laata; ta peab kõrtsi; prussakad pidasid paraa-
di; ma pean viha; radistid peavad sidet jt.

2) $N_n^1 V_1^b N_{os}^2 N_{ks/A_{ks}}^3$. Distributsioonivormeleid
 $N_n^1 V_1^b N_{os}^2 N_{ks}^3$ ja $N_n^1 V_1^b N_{os}^2 A_{ks}$ pole põhjust lugeda erine-
vaiks distributsioonivormeleiks, sest vastandus $N_{ks}^3 - A_{ks}$
pole ilmselt seotud verbi pidama erinevate tähendustega.
Võib arvata, et distributsioonivormeli $N_n^1 V_1^b N_{os}^2 A_{ks}$ puhul
on meil tegemist elliptilise konstruktsiooniga; täielik
kuju on ilmselt $N_n^1 V_1^b N_{os}^2 A_{ks} N_{ks}^2$ (ta pidas riiet puna-
seks - ta pidas riiet punaseks riideks; mõnedel juhtudel
ei tundu see siiski niisama loomulik, vrd. näit. ta pidas
palka väheseks - ta pidas palka väheseks palgaks). Sel
juhul taandub A välja ja me saamegi struktuuri
 $N_n^1 V_1^b N_{os}^2 N_{ks}^3$. Näiteid: ta peab mind enesetapjaks; Stanis-
lavski peab elutõde (kunstipärasuse) aluseks; ta peab
palka väheseks; ta pidas mind ebamoraalseks jt.

3) $N_n^1 V_1^b N_{st}^2$. Pead sa temast? Juhan peab suusatami-
sest jt.

4) $N_n^1 V_1^2 N_{os}^2$ D (D all võib esineda ka D , mitte aga
siin

D või D . Ühtlasi võib D positsioonis esineda ka
siia siit
vestavas funktsioonis olev N või A). Pingviin peab pea
kummargil; nad pidasid tütart näljas; kus peeti Marion;
nad pidagu hobune paigal jt.

2. 2.2. V_2^b . Nagu öeldud, esineb V_2^b vaid ühes
konstruktsioonis: $N_n^1 V_2^b V_{ma}^2$. Ma pean ialutama; mõlemad
pidid ära uppuma; peab armastama distantsti; ta pidi naerma
jt.

2. 3. saama.

1) $N_n^1 V^C N_{om}^2$. Seltsimees sai kirja; inimesed saavad maa; haige sai söki; taevas saab (sinise) värvi jt.

2) $N_n^1 V^C N_{ks}^2 / A_{ks}$. Elu saab võimatuks; ta oli saanud peremeheks; minagi saan moralistik; ta sai (Senni-ga) tuttavaks jt.

3) $N_n^1 V^C V_{da}^2$. Karl sai õppida; ma sain vaeva näha; poiss sai sõimata; Benno sai utsitada jt.

4) $N_n^1 V^C N_{st}^2$. Mudelitest saavad paadid; mis saab rahvast; kommunistidest saavad õpetajad jt.

5) $N_n^1 V^C D_{siia}$. Täpsem oleks esitada see distri-butsioonivormel kujul $N_n^1 V^C N_{sse}^2 / N_{le}^2 / N_{om}^2 + \text{kaassõna} / D_{siia}$.

Siia kuuluvad näit. fraasid ta sai laatsaretti; ma sain laevale; ta sai isa juurde; (kuidas) sa siia said? Me võime nad kõik ühte distributsioonivormelisse koondada sel alusel, et N_{sse}^2 , N_{le}^2 , $N_{om}^2 + \text{kaassõna}$ ja D_{siia} on nendes

ühes ning samas funktsioonis. See ilmneb kasvõi sellestki, et nad võivad kõik koos esineda ühes lauses, näiteks: ta sai siia laevale laatsaretti isa juurde; kui igaüks neist annaks verbile erineva tähenduse, oleks see võimatu.

6) $N_n^1 V^C N_{om}^2 D_{siia}$. Täielikum kuju oleks jällegi $N_n^1 V^C N_{om}^2 N_{sse}^3 / N_{om}^3 + \text{kaassõna} / N_{le}^3 / D_{siia}$. Näiteks: saime müraka lennukisse; saime müraka autole; saime müraka kuuri alla; saime müraka sinna.

7) $N_n^1 V^C N_{om}^2 A_{ks}$. Saime laua puhtaks; ta sai roobi punaseks; mehed said auto tühjaks jt.

8) $N_n^1 V^C V_{ma}^2$. Ta sai minema; kiri saab olema käre; sa saad tööd tegema jt.

Sellesse alltüüpi kuulub suurem osa (umbes 90%) ruumilist liikumist väljendavaist fraasidest. Ühtlasi on see kõige sagedam tulema alltüüp üldse (umbes 30% kõigist fraasidest). Ja et ta kuulub seejuures ka kõige lihtsamasse distributsioonivormelisse ($N_n V^a$), siis võime selle alltüübiga esindatavat tähendust pidada tulema põhitähenduseks.

3. 1.1.1.2. T: $\phi \rightarrow V_{ma}^2$. Näiteks: kiri tuli vanglast; telegramm tuli kodunt; tellimused tulid kõikjalt; jutud tulid mõisasakste kõrvu. Erinevus eelmisest alltüübist on kahtlemata üsnagi tinglik ning seda võib mitte tunnistada erinevuseks leksikaalsete tähenduste vahel.

3. 1.1.2. Alltüübis, kus pole võimalik T: $\phi \rightarrow D$ (/) sealt

D ei õnnestunud meil leida transformatsioon, mis siia osutaksid erinevatele lekseemidele sellega, et oleksid võimalikud ühtedel ja võimatud teistel juhtudel, ehkki intuitsivselt tundub olevat mõningaid erinevusi tulema tähenduste vahel. Seepärast leidsime olevat õigema jätta siin küsimus lahtiseks ja mitte kasutada mingeid muid kriteeriume, mis ei vastaks meie meetodile. Esitame vaid mõned näited; tuleb vihm; tuleb kevad; tuleb sõda; tuleb nohu; tulid lahkkelid; tuleb koosolek; (tal) tuli mõte; (siis) tulid küsimused; tuleb juhus jt.

3. 1.2. D i s t r i b u t s i o o n i v o r m e l $N_n^1 V^a N_{le}^2$. Selles distributsioonivormelis sisaldub kaks selgesti erinevat alltüüpi.

3. 1.2.1. Ühel juhul on võimalik transformatsioon T: $\rightarrow V_{\phi}^2 N_n^1$ (V_{ϕ}^2 märgib N^2 -st moodustatud verbi): õu tuleb rookimisele $\rightarrow$ roogitakse õu; tükk tuleb läbivõtmisele $\rightarrow$ tükk võetakse läbi. Teisel juhul on selline transformatsioon aga võimatu: ta tuli mõttele $\nrightarrow$ ta mõeldakse; ta tuli järeldusele $\nrightarrow$ ta järeldatakse.

3. 1.2.2. Omakorda võime aga teist alltüüpi iseloomustada transformatsiooniga $T: \rightarrow N_n^1 V N_n^2$: ta tuli mõttele $\rightarrow$ ta mõtles; ta tuli järeldusele $\rightarrow$ ta järeldas. Alati pole see transformatsioon küll praktiliselt teostatav, sest pole N^2 -le vastavat verbi, nagu näiteks fraasi puhul ta tuli ideele. Sellegipärast ei saa me neid fraase lugeda eri tüübiks, sest põhimõtteliselt on see transformatsioon teostatav ka siin, tarvitseb vaid moodustada vastav verb (näiteks: ta tuli ideele $\rightarrow$ ta ideestus vms.).

3. 1.3. Distributsioonivormel $N_n^1 V^a N_{st}^2$. Ka selles distributsioonivormelis sisaldub kaks suurt alltüüpi. Esimeses väljendab tulema N^1 ja N^2 põhjuslikku seost ning teises N^1 kujunemist N^2 -ks.

3. 1.3.1. $T: \rightarrow N_n^1 V^a$ sellest, et N^2 ... Näiteks: peavalu tuleb päikesest $\rightarrow$ peavalu tuleb sellest, et päike (kõrvetab vms.); (paha) tuju tuleb (halvast) ilmast $\rightarrow$ (paha) tuju tuleb sellest, et ilm (on halb) jne.

3. 1.3.2. $T: \rightarrow N_n^2$ on N_n^1 (või N_n^1 on N_n^2), mis esimesel juhul pole võimalik. Näit. lihast tulevad kotletid $\rightarrow$ kotletid on liha; temast tuleb õpetaja $\rightarrow$ tema on õpetaja; teekonnast tuleb tõsi $\rightarrow$ teekond on tõsi jne. Selgesti nähtav on erinevus N^1 -na ja N^2 -na esinevates sõnades (tegelikult me ju nende sõnade põhjal otsustamegi ühe või teise transformatsiooni võimalikkuse või võimatuse üle). Eriti ilmne on erinevus N^1 -na esinevates sõnades: reeglina ei esine esimeses tüübis siin konkreetset eset tähistavaid sõnu. Nii näiteks määrame fraasi lihast tuli kõhuvalu kindlasti esimesse tüüpi, fraasi lihast tulid kotletid aga teise.

3. 1.3.2.1. Teises alltüübis võime omakorda eraldada kitsama tüübi, kus N^2 tähistab N^1 materjali, sest teatavas mõttes käituvad need fraasid teistest erinevalt. Nii on nende puhul loomulikum transformi kuju N_n^1 on N_n^2 ja mitte vastupidi (kotletid on liha, mitte liha on kotletid). Li-

saks sellele aga lubavad need fraasid teostada veel transformatsiooni $T: \rightarrow N_n^1$ on N_{st}^2 , mis teiste fraaside puhul on võimatu: lihast tulevad kotletid $\rightarrow$ kotletid on lihast; marmorist tuleb Veenus $\rightarrow$ Veenus on marmorist jne.

3. 1.4. Distributsioonivormel $N_n^1 v^a A_n$ edasi ei hargnenud.

3. 1.5. Distributsioonivormel $v^a v_{da}^2 (N_1)$. Ka see distributsioonivormel ei jagune meie metoodika seisukohalt enam alltüüpideks. Intuitiivselt võib küll sageli tunda tulema erinevaid tähendusi, ent need on modaalsed tähendusvarjundid ega allu analüüsile meie meetodi järgi. Modaalsed tähendusvarjundid ei sõltu üheselt ümbrusest (ühe lause, veel vähem aga orienteeritud fraasi ulatuses), meie saame aga ainult selle põhjal midagi otsustada. Modaalsed tähendused kuuluvad ilmselt teise liiki kui need, mida on võimalik meie meetodiga välja selgitada.

3. 2. pidama.

3. 2.1. Distributsioonivormel $N_n^1 v^b N_{os}^2$. Seda tüüpi struktuurid võime kõigepealt jagada kaheks vastavalt transformatsiooni $T: v^b \rightarrow v^b$ kinni teostatavusele. Alltüüp, kus see transformatsioon pole teostatav, sisaldab omakorda mitu erinevalt käituvat alltüüpi.

3. 2.1.1. $T: v^b \rightarrow v^b$ kinni: katus peab vihma $\rightarrow$ katus peab vihma kinni; rahakott ei pea raha $\rightarrow$ rahakott ei pea raha kinni.

3. 2.1.2. Antud alltüübis on edasine klassifikatsioon seotud peamiselt N^2 tähendusliku rühmitamisega. Kõigepealt märkame, et ühes osas fraasidest võib N^2 esineda ka omastavas käändes, s. t. on võimalik transformatsioon $T: N_{os}^2 \rightarrow N_{om}^2$, teises osas aga mitte.

3. 2.1.2.1. $T: N_{os}^2 \rightarrow N_{om}^2$: nad peavad lumepidu $\rightarrow$ nad peavad lumepeo; rahvas pidas koosolekut $\rightarrow$ rahvas pidas koosoleku; peremehed pidasid laata $\rightarrow$ peremehed pidasid

laada. Aga ei saa: ta peab sõprust $\rightarrow$ ta peab sõpruse; ta peab sidet $\rightarrow$ ta peab sideme; ta peab poodi $\rightarrow$ ta peab poe jne.

3. 2.1.2.2. Alltüüp, kus $T: N_{os}^2 \rightarrow N_{om}^2$ pole võimalik, sisaldab omakorda kaht tüüpi struktuure.

3. 2.1.2.2.1. $T: V^b \rightarrow V^b$ endal. See on võimalik neil juhtudel, mil pidama tähendust võiks väljendada sõnadega 'omama', 'majandama': ta peab poodi $\rightarrow$ ta peab endal poodi; härja peab talukohta $\rightarrow$ härja peab endal talukohta.

3. 2.1.2.2.2. $T: V^b \nrightarrow V^b$ endal: ta peab sõprust; radistid peavad sidet jne. N^2 tähistab siin peamiselt abstraktseid mõisteid.

3. 2.2. Ülejäänud distributsioonivormelid edasi ei jagunenud. Seejuures sisaldub distributsioonivormelis $N_n^1 V_2^b V_{ma}^2$ küll rida erinevaid modaalseid tähendusvarjundeid, ent nagu öeldud, ei ole võimalik meie meetodi abil modaalseid tähendusi analüüsida.

Põhitähenduse määramine on V_1^b puhul palju raskem ja kunstlikum kui V^a puhul. Ilmselt võib selleks lugeda distributsioonivormeliga $N_n^1 V_1^b N_{os}^2 N_{ks}^3 / A_{ks}$ esindatud tähendust, sest siia kuulub peaaegu 40% materjalist, ehkki see distributsioonivormel pole struktuurilt eriti lihtne.

3. 3. saama.

3. 3.1. D i s t r i b u t s i o o n i v o r m e l $N_n^1 V^c N_{om}^2$. Siin võib vahet teha kaht tüüpi struktuuride vahel.

3. 3.1.1. $T: N_{om}^2 \rightarrow A_{ks}$ (harilikult moodustuvad need omadussõnad line-sufiksi abil). Näiteks: nimi sai (halva) kõla $\rightarrow$ nimi sai (halva)kõlaliseks; taevas sai (sinise) värvi $\rightarrow$ taevas sai (sinise)värviliseks.

3. 3.1.2. $T: N_{om}^2 \nrightarrow A_{ks}$: sain käpikud; sain raadio-
grammi; saame tormi; poiss saab peapea.

3. 3.2. Distributsioonivormel $N_n^1 V^c V_{da}^2$. Selles distributsioonivormelis leiame kaht tüüpi struktuure, mis on meie meetodiga eristatavad. Neid võib nimetada aktiivseteks ja passiivseteks.

3. 3.2.1. Passiivseteks struktuurideks võime nimetada neid, kus on võimalik transformatsioon $T: \rightarrow V_{\rho}^2 N_{os}^1$:

Sa saad petta $\rightarrow$ sind petetakse; Benno saab utsitada $\rightarrow$ Bennot utsitatakse; ta sai sõimata $\rightarrow$ teda sõimati. Seejuures esineb selles alltüübis konstruktiivne homonüümia, sest kõiki neid fraase võib mõista ka n. ö. aktiivses tähenduses (samasuguses, nagu me mõistame näit. fraasi ta saab joosta.). Ent $T: \rightarrow V_{\rho}^2 N_{os}^1$ eraldab siiski fraaside rühma, mida saab mõista ka passiivses tähenduses, vastandina fraasidele, kus seda ei saa.

3. 3.2.2. $T: \rightarrow V_{\rho}^2 N_{os}^1$: ta saab joosta; ta sai kuulda; ta saab lugeda jne. Selles alltüübis võib märgata jällegi rida modaalseid tähendusvarjundeid.

3. 3.3. Ülejäänud distributsioonivormelid edasi ei hargnenud.

3. 3.4. Põhitähenduse määramine on antud verbi puhul veelgi raskem. Sageduselt on distributsioonivormelid $N_n^1 V^c N_{om}^2$ (vt. 3. 3.1.2) ja $N_n^1 V^c V_{da}^2$ (vt. 3. 3.2.2) peaaegu võrdsed (kumbki umbes 20% materjalist). Kolmanda kriteeriumi põhjal (sõnatuletuslik produktiivsus) võime siiski põhitähenduseks lugeda distributsioonivormeliga $N_n^1 V^c N_{om}^2$ esindatud tähendust.

4. K o k k u v ö t e. Esitasime eelnevas kolme verbi tähendusliku analüüsi. Analüüs toimus distributsiooniliste ja transformatsiooniliste kriteeriumide alusel. Nende abil selgitasime välja iga sõna diferentsiaalsete tähenduste (lekseemide) hulga. Kõige mitmetähenduslikumaks osutus saama, mis distributiivsel tasapinnal jagunes üheksaks distributsioonivormeliks; transformatoorsel tasapinnal ja-

gunesid kaks neist kaheks ja seega sisaldab lõpploend 11 lekseemi.

1. $N_n^1 V^c N_{om}^2 \rightarrow N_n^1 V^c A_{ks}$ nimi sai (halva) kõla
2. $\rightarrow N_n^1 V^c A_{ks}$ isa sai käpikud
3. $N_n^1 V^c V_{da}^2 \rightarrow V_{\emptyset}^2 N_{os}^1$ Benno saab utsitada
4. $\rightarrow V_{\emptyset}^2 N_{os}^1$ ma saan joosta
5. $N_n^1 V^c N_{ks}^2$ ma saan õpetajaks
6. $N_n^1 V^c N_{st}^2$ kommunistidest said õpetajad
7. $N_n^1 V^c D_{siia}$ saime laevale
8. $N_n^1 V^c A_n$ kuhi saab ilus
9. $N_n^1 V^c N_{om}^2 D_{siia}$ saime müraka lennukisse
10. $N_n^1 V^c N_{om}^2 A_{ks}$ mehed said laua puhtaks
11. $N_n^1 V^c V_{ma}^2$ ma saan minema.

tulema moodustas 5 distributsioonivormelit ja neist jagunes 3 edasi reaks alltüüpideks; lõpploendis saime 10 lekseemi:

1. $N_n^1 V^a \rightarrow D_{sealt} D_{siia} V_{ma}^2$ poiss tuleb
2. $\rightarrow D_{sealt} D_{siia}$ kiri tuleb
3. $\rightarrow D_{sealt} D_{siia}$ nohu tuleb; kevad tuleb jne.
4. $N_n^1 V^a N_{st}^2 \rightarrow N_n^1 V^a$ sellest, et $N^2 \dots$ vaesus tuleb
rumalusest
5. $\rightarrow N_n^2$ on N_n^1 poisist tuleb lendur
6. $\rightarrow N_n^1$ on N_{st}^2 lihast tulevad kotletid
7. $N_n^1 V^a N_{le}^2 \rightarrow V_{\emptyset}^2 N_n^1$ õu tuleb rookimisele
8. $\rightarrow N_n^1 V^{N^2}$ Setšenov tuli järeldusele

9. $N_n^1 v^a A_n$ õhtu tuleb põnev
 10. $v^a v_{da}^2 (N_1)$ (mul) tuleb õppida

pidama andis 5 distributsioonivormelit, millest üks jagunes alltüüpideks; lõpploendis saime 8 lekseemi:

1. $N_n^1 v_1^b N_{os}^2 \rightarrow N_n^1 v^b$ kinni N_{os}^2 katus peab vihma
 2. $\rightarrow N_n^1 v^b N_{om}^2$ rahvas peab pidu
 3. $\rightarrow N_n^1 v^b N_{om}^2$ radistid peavad sõprust
 4. $\rightarrow N_n^1 v^b$ endal N_{os}^2 härä peab talukohta
 5. $N_n^1 v_1^b N_{ks}^3 / A_{ks}$ ta peab mind moralistiks
 6. $N_n^1 v_1^b N_{st}^2$ ma pean suusatamisest
 7. $N_n^1 v_1^b N_{os}^2 D$ nad pidasid tütarit näljas
 8. $N_n^1 v_2^b v_{ma}^2$ ma pidin jooksuma

Käesoleva töö ülesannetesse ei kuulunud ühe sõna lekseemide hierarhia väljaselgitamine, sest kasutatud metoodika ei võimalda seda. Mitmesugustele abikriteeriumidele toetudes osutasime iga sõna puhul ainult selle tõenäolisele põhithendusele (vt. 3. 1.1.1.1, 3. 2.2 ja 3. 3.4).

Saadud lekseemid kujutavad endast puhtdiferentsiaal-seid üksusi; nende sisulise tõlgitsuse saame, kui interpreteerime neid semantilisel tasandil.

ELATIIVI FUNKTSIOONIDE TRANSFORMATOORSEST

ANALÜÜSIST

M. E r e l t

O. S i s s e j u h a t u s.

Käesolevas töös tehakse katse analüüsida elatiivi funktsioone eesti keeles transformatsioonimeetodi alusel.

Traditsiooniline lähenemine käändetähenduste (-funktsioonide) klassifitseerimisele on enamasti puhtintuiivne. Piirdutakse ainult selle väitmisega, et ühel juhul on vastaval käändel üks funktsioon, teisel juhul aga teine. Et see tõepoolest nii on, jääb lihtsalt tõestamata või peetakse küllaldaseks uuri ja intuiivset tunnetust.

Transformatoorne analüüs püüab intuiivselt oletatavaid erinevusi formaalselt tõestada. Formaalse tõestuse all mõeldakse siin keelesisest tõestust, s. t. tähenduste klassifitseerimisel ei lähtuta keelevälistest erinevustest, vaid uuritavate elementide süntaktilistest vahetkordadest keeles.

Transformatoorne analüüs baseerub järgmistel teoreetilistel eeldustel:

1) igas keelelises üksuses esinevad dialektilises ühtsuses tema vorm ja tähendus: ei saa olla keelelist vormi ilma tähenduseta ja iga tähendus on keeles väljendatud kindla vormiga;¹

2) igale erinevusele tähenduses peab vastama erine-

1 В.И. Перебейнос, К вопросу об использовании структурных методов в лексикологии. Проблемы структурной лингвистики. Москва, 1962, лк. 163-164.

vus vormis ja vastupidi;

3) selleks et näidata erinevust tähenduses, on vajalik ja piisav näidata erinevust vormis.

Analüüs on teostatud umbes 1500 näitelause põhjal. Laused on võetud ilukirjanduslikest tekstidest.²

1. T r a n s f o r m a t o o r n e a n a l ü ü s.

Transformatoorse analüüsi puhul vaadeldakse iga klassifitseeritavat üksust, s. o. konstruktsiooni, milles esineb uuritav element, kahest aspektist: esiteks määratakse, mida kujutab endast antud konstruktsioon, ja teiseks, millistele spetsiifilistele muutustele antud konstruktsioon võib alluda.³ Neid spetsiifilisi muutusi nimetataksegi transformatsioonideks.⁴

Transformatoorne analüüs osutub võimalikuks vaid siis, kui

1) teatakse uuritava keele kõiki põhilisi sõnade klasse;

2) osatakse kindlaks määrata sõnade süntaktilisi seoseid fraasis;

3) osatakse eristada märgitud fraasi mittemärgitust.⁵

Loeme käesoleva analüüsi puhul kõik nimetatud tingimused täidetuks.

2 L. Kibuvits, Elagu inimene. Tallinn, 1962; L. Promet, Meesteta küla. Tallinn, 1962; R. Sirge, Maa ja rahvas. Tallinn, 1959; J. Smuul, Jäine raamat. Tallinn, 1959.

3 Siin ja edaspidi lähtutakse põhiliselt Worth'i poolt esitatud seisukohtadest ning ühtlasi võetakse üle tema poolt kasutatud terminid (vt. Д.С. Уорс, Трансформационный анализ конструкций с творительным падежом. Новое в лингвистике II. Москва, 1962, lk. 637-683).

4 Termin "transformatsioon" kasutamise kohta vt. Н. Хомский, Синтаксические структуры. Новое в лингвистике II. Москва, 1962, lk. 450-451; Э.С. Харрис, Совместная встречаемость и трансформация в языковой структуре. Sealsamas, lk. 536-541.

5 Käesolevas töös mõistetakse märgitud fraasi all sellist fraasi, mis on õige nii grammatiliselt kui ka leksikaalselt.

1.1. Morfoloogiline klassifikatsioon.

Transformatoorne analüüs koosneb kahest etapist:

1) analüüsitavate konstruktsioonide morfoloogilisest klassifikatsioonist⁶ ja 2) morfoloogilisel teel määratletud tüüpkonstruktsioonide ehk fraasitüüpide alltüüpide transformatoorsest klassifikatsioonist.

1.1.1. Reduktsioon.

Iga uuritavat fraasi võib jagada tuumaks ja ümbruseks.⁷ Käesolevas töös vaadeldakse tuumana nimisõna elatiivis. Näiteks fraasis mees väljus autost on tuumaks autost ja ümbruseks mees väljus.

Enamik ümbrusi sisaldab lingvistilisi üksusi, mis on antud tuuma suhtes liiased. Teiste sõnadega, uuritavas fraasis leidub selliseid elemente, mis pole seotud nimi-sõnaga elatiivis ning millede ärajätmine ei muuda fraasi mittemärgituks. Selleks et morfoloogilise analüüsi käigus vältida ebaolulisi detaile, allutatakse kõik fraasid reduktsioonile. Reduktsiooni tulemusena saadakse nn. lias-test ümbrustest optimaalsed ümbrused, s. t. sellised ümbrused, millest ühtki liiget ära jätta ei saa, sest siis muutuks fraas mittemärgituks.

Töös esineb kaht liiki reduktsioone.

1. Endotsentrilistest konstruktsioonidest kõrvaldatakse kõik sõltuvad elemendid, välja arvatud need, mis edasiseks analüüsiks on olulised.⁸ Näiteks sinise kaabuga

6 Morfoloogilist klassifikatsiooni e. nn. distributiivset analüüsi vaadeldakse tavaliselt lahus transformatoorsest analüüsist. Käesolevas töös on aga morfoloogiline klassifikatsioon transformatoorse analüüsi (laias mõttes) üheks etapiks (selline käsitus on ka Worth'il (vrd. Д.С. Уорс, Трансформационный анализ конструкций с творительным падежом. Новое в лингвистике II. Москва, 1962, lk. 637-683).

7 А.А. Холодович. Опыт теории подклассов слов. "Вопросы языкознания" 1960, № 1, lk. 32-43.

8 Endotsentriliste konstruktsioonide kohta vt. Ch. F. Rockett, A Course in Modern Linguistics. New York, 1963, lk. 183-191.

mees karjatas valjusti suurest hirmust redutseerub fraa-
siks mees karjatas hirmust, milles optimaalne ümbrus on
mees karjatas....

2. Kirjeldusest jäetakse välja kõik grammatilised kategooriad, mis on läbiviidavate transformatsioonide seisukohalt ebaolulised. Selleks allutatakse fraas kõigile võimalikele transformatsioonidele ja seejärel loetakse ebaoluliseks need grammatilised kategooriad, mis jäävad kõikide transformatsioonide puhul muutumatuks. Olgu meil näiteks fraas pilt kukkus kohvrast välja. Näeme, et ajakategooria pole oluline sellise transformatsiooni suhtes, kus elatiivis nimisõna kujundatakse konstruktsiooniks nimisõna genitiivis + seest, näit. pilt kukkus kohvrast välja → pilt kukkus kohvri seest välja; pilt kukub kohvrast välja → pilt kukub kohvri seest välja; pilt on kukkunud kohvrast välja → pilt on kukkunud kohvri seest välja jne.

1. 1. 2. K o d e e r i m i n e.

Põhistruktuuriks⁹ (tuum + optimaalne ümbrus) taandatud fraasi võib esitada sümbolite jadana. Iga sümbol tähistab sõnaklassi, millesse antud sõna kuulub. Sümbolid varustatakse indeksitega, mis märgivad relevantseid grammatilisi kategooriaid ning korduvate sõnaklasside järjekorda. Näiteks: naine väljus toast - $N^1 V N_{el}^2$. Käesolevas töös kasutatakse järgmisi sümboleid: N - nimisõna; N_n , N_g , N_p , N_{in} , N_{el} , N_{tr} , N_{ess} , N_k - nimisõna nominatiivis, genitiivis, partitiivis, inessiivis, elatiivis, translatiivis, essiivis, komitatiivis; A - omadussõna; A_{pos} , A_{komp} - omadussõna positiivis, komparatiivis; V - verb; V_{pers} , V_{imp} - verb isikulises tegumoes, umbisikulises tegumoes; V_{irref} - irrefleksiivne verb või verbiühend; $V_{\{ole-\}}$ - mingi vorm verbist olema; $V_{\{saa-\}}$ - mingi vorm verbist

9 Edasises käsitluses nimetame iga põhistruktuuriks taandatud fraasi lihtsalt struktuuriks.

saama; D - määrsõna; N_{sg}, N_{pl} - nimisõna ainsuses, mitmuses; V_{eitav imperat} - verbi imperatiivi preesensi eitav vorm.

1. 1. 3. Kõik redutseeritud ja kodeeritud fraasid grupeeritakse järgnevalt fraasitüüpideks, kusjuures ühte fraasitüüpi kuuluvad struktuurid on morfoloogiliselt identsed.

1. 1. 3. 1. Muus suhtes morfoloogiliselt identsete struktuuride ühteviimist võib takistada erinev sõnade järjekord, näiteks ta värises hirmust ja hirmustvärises tema. Elatiivi seisukohalt pole ilmselt selline erinevus oluline. Selleks et ühendada erineva sõnade järjekorraga struktuure ühte fraasitüüpi, tuleb sõnade järjekord lihtsalt ühtlustada. Iga fraasitüübi kanooniline kuju näitab seega temale vastavate struktuuride sõnade järjekorda.

1. 1. 3. 2. Ühest morfoloogilist klassifikatsiooni võivad segada ka mitmesugused verbi laiendid, mis eraldi võttes pole seotud elatiiviga, kuid mida pole ka võimalik välja jätta, sest vastasel korral muutuks fraas mittemärgituks.

Analoogia põhjal ühendverbidega on otstarbekas verb ja sellised laiendid ühendada ning kodeerida sellist kompleksi nagu ilma laiendita verbi, näiteks ta läks hirmust lolliks (läks lolliks - V); mehed läksid suitsetamissalongist verandale (läksid verandale - V); poiss jooksis toast õue (jooksis õue - V) jne.

1. 2. T r a n s f o r m a t o o r s e d o p e -
r a t s i o o n i d.

Morfoloogilise klassifikatsiooni alusel ühte fraasitüüpi ühendatud struktuuride tuumad võivad olla väga erineva tähendusega (funktsiooniga), mida märkame juba intuitiivselt. Eelduse kohaselt peab aga igale erinevusele tähenduses vastama erinevus vormis. Järelikult, kui meil on tegemist tõepoolest erinevate tähendustega, siis peab

olema võimalik ühte fraasitüüpi kuuluvaid struktuure erinevalt ümber kujundada. Teiste sõnadega, mõningaid struktuure saab muuta ühtviisi, teisi teistviisi. Transformaatorse analüüsi eesmärgiks ongi kirjeldada formaalseid erinevusi morfoloogiliselt identsete struktuuride vahel. Olgu meil näiteks kaks struktuuri: poiss punetas näost ja mees väljus toast. Morfoloogiliselt kuuluvad mõlemad fraasitüüpi $N^1 V N_{el}^2$. Esimene neist allub transformatsioonile $T: \rightarrow N_n^2 V$, teine aga transformatsioonile $T: \rightarrow N^1 V \{ole-\}$ N_{in}^2 , näiteks poiss punetas näost $\rightarrow$ nägu punetas; mees väljus toast $\rightarrow$ mees oli toas, kuid poiss punetas näost $\nrightarrow$ poiss oli näos; mees väljus toast $\nrightarrow$ tuba väljus. Järelikult on meil tõepoolest tegemist elatiivi kahe erineva funktsiooniga.

1. 2. 1. Eespool me ei vaadelnud lähemalt transformaatorse analüüsi põhimõistet - transformatsiooni. Peab ütleva, et lingvistilises kirjanduses ei leidu seniajani transformatsiooni lõplikku definitsiooni. Kõige üldisemalt me mõistame transformatsiooni all keelelise fraasi sellist ümberkujundamist, mille puhul on täidetud kaks tingimust: 1) transformil on oluliselt sama tähenduslik sisu mis lähtefraasil; 2) transform ei tohi sisaldada selliseid iseseisvate sõnade tüvesid, mis lähtefraasis puuduvad.

Mis puutub käesolevas töös kasutatud transformatsioonidesse, siis on leksikaalse samasuse nõudest üldiselt kinni peetud, välja arvatud transformatsioonide $N^1 V N_{el}^2 \rightarrow N^1 V \{ole-\} N_{in}^2$ ja $N^1 V N_{el}^2 \rightarrow N_n^2 V \{saa-\} N_{tr}^1$ korral. Analoogilised ümberkujundused on aga transformatsioonidena kasutusel ka näiteks Worth'il.¹⁰

Tähendusliku samasuse nõue on siiani üheks transformatsioonimeetodi valulapseks. Oluline on eelkõige see, et ühe ja sama transformatsiooni korral peab transformi ja

10

Д.С. Юпп, оп. cit., lk. 637-683.

lähtefraasi vaheline tähenduslik erinevus olema identne mis tahes fraasi puhul. Kui suur see erinevus üldse olla võib, see on juba teine küsimus. Siin on tegemist ühe sünonüümsuse erijuhuga. Senini on aga sünonüümide määramisel lähtutud informandi keelevalistust. Sünonüümsuse (sealhulgas ka transformi ja lähtefraasi vahelise sünonüümsuse) formaalne tõestus on ilmselt võimalik transformatsiooni-meetodist aste kõrgema meetodi abil.

1. 2. 2. Igasuguse transformatoorse analüüsi puhul on klassifitseeritavaks üksuseks mingi konstruktsioon, milles esineb uuritav element. Teatavas käändes olev sõna on aga ainult üheks elemendiks mingis konstruktsioonis. Transformatoorse analüüsi abil saame iga konstruktsiooni jaotada reaks allkonstruktsioonideks. Igale sellisele allkonstruktsioonile vastab teatav käändefunktsioon.¹¹ Samal ajal ei tea me aga midagi kindlat väita, missugustes vahetkordades on erinevate konstruktsioonide alltüübid ning seega neile vastavad käändefunktsioonid.

1. 2. 3. Analüüsi tulemusena saame elatiivi tähendused, millest igaüks on iseloomustatav ainult oma erinevuste kaudu teistest. Igasugune sisuline interpretatsioon on väljaspool antud meetodi raame.

Transformatsioonid võib vaadelda teatud diferentsiaalsete (distinktiivsete) tunnusjoontena ning seega käände iga tähendust kui selliste tunnusjoonte kimpu.

2. Morfoloogilise klassifikatsiooni tulemused.

1) $N^1 V N^2_{el}$: piim voolab klaasist välja; sõnnik lõpeb laudast; mees väljub autost; mees läheb uksest välja; ta läks trepist üles; me tõuseme lauast; tuul puhub aht-rist; ta haarab lauast kinni; tuba haiseb seebirasvast;

¹¹ Ühes alltüübis võivad esineda koos ka sellised struktuurid, milles elatiiv on traditsioonilise käsitluse järgi mitmes erinevas funktsioonis.

ta väriseb hirmust; ta ehmatab äikesest; tuba täitub veest; see selgub kirjast; poiss punetab näost; insener saab sinust; mees kasvab poisist; nad räägivad lenduritest; ma lahkun sinust.

2) $N^1 V N_{el}^2 N^3$: ta maksab kanadest 50 rubla; ta tunneb narri naerust; Rein tõmbab tikust tuld; arst puudutab teda õlast.

3) $N^1 N_{el,pl}^2$: keegi magajaist; osa ekspeditsioonidest; grupp meestest; igäiks neist.

4) $N_{el}^1 N^2$: plekist konservikarp; ungarlasest noormees Juhanist vend.

5) $N_{el} A_{pos}$: näost punane; võidusõnumist õnnelik; ärevusest rahutu; õnnest rikas; noorest nõrk.

6) $N_{el} A_{komp}$: ookeanist tumedam.

7) mis $N^1 V N_{el}^2$: mis me raiskame ajast.

8) $N^1 V_{imp} N_{el}^2$: oaas on piiratud jääliustikest.

3. T r a n s f o r m a t o o r n e k l a s s i -
f i k a t s i o o n .

Järgnevalt esitatakse eraldi iga fraasitüübi struktuuride transformatoorne klassifikatsioon. Fraasitüüpide alltüüpe iseloomustatakse vastavate tabelite abil.

3. 1. F r a a s i t ü ü p $N^1 V N_{el}^2$. Selle fraasitüübi transformatoorne klassifikatsioon on antud tabelis 1.

Näiteid:

1) T: $\rightarrow N^1 V \{ole-\} N_{in}^2$: mees väljub autost $\rightarrow$ mees on autos; piim voolab klaasist välja $\rightarrow$ piim on klaasis; sõnnik lõpeb laudast $\rightarrow$ sõnnik on laudas jne.

2) T: $\rightarrow N^1 V N_g^2$ seest: piim voolab klaasist välja $\rightarrow$ piim voolab klaasi seest välja; pilt kukub kohvrast välja $\rightarrow$ pilt kukub kohvri seest välja jne.

3) T: $\rightarrow N^1 V N_{in}^2$: sõnnik lõpeb laudast $\rightarrow$ sõnnik lõpeb laudas jne.

4) T: $\rightarrow N^1 V$ läbi N_g^2 ($\rightarrow N^1 V N_g^2$ kaudu): mees läheb uksest välja $\rightarrow$ mees läheb läbi ukse välja ($\rightarrow$ mees läheb ukse kaudu välja); naine vaatab aknast õue $\rightarrow$ naine vaatab läbi akna õue ($\rightarrow$ naine vaatab akna kaudu õue) jne.

5) T: $\rightarrow N^1 V N_D^2$ mööda: ta läheb trepist üles $\rightarrow$ ta läheb treppi mööda üles; mees ronib raudredelist katusele $\rightarrow$ mees ronib raudredelit mööda katusele jne.

6) T: $\rightarrow N^1 V N_g^2$ äärest: me tõuseme lauast $\rightarrow$ me tõuseme laua äärest jne.

7) T: $\rightarrow N^1 V N_g^2$ juurest: me tõuseme lauast $\rightarrow$ me tõuseme laua juurest; ma lahkun sinust $\rightarrow$ ma lahkun sinu juurest jne.

8) T: $\rightarrow N^1 V N_g^2$ poolt: tuul puhub ahtrist $\rightarrow$ tuul puhub ahtri poolt jne.

9) T: $\rightarrow N^1 V N_g^2$ küljest: ta haarab lauast kinni $\rightarrow$ ta haarab laua küljest kinni jne.

10) T: $\rightarrow N^2 V$ irref $N_{g/p}^1$: ta väriseb hirmust $\rightarrow$ hirm paneb ta värisema; ta ehmatab äikesest $\rightarrow$ äike ehmatab teda; tuba haiseb seebirasvast $\rightarrow$ seebirasv pani toa haise; tuba täitub veest $\rightarrow$ vesi täitis toa; see selgub kirjast $\rightarrow$ kiri selgitab selle jne.

11) T: $\rightarrow N^1 V N_g^2$ tõttu: ta väriseb hirmust $\rightarrow$ ta väriseb hirmu tõttu; ta ehmatab äikesest $\rightarrow$ ta ehmatab äikese tõttu; tuba haiseb seebirasvast $\rightarrow$ tuba haiseb seebirasva tõttu jne.

12) T: $\rightarrow N^1 V N_g^2$ järele: tuba haiseb seebirasvast

→ tuba haiseb seebirasva järele; mees lehkab viinast → mees lehkab viina järele; aed lõhnab sirelitest → aed lõhnab sirelite järele jne.

13) T: → $N^1 V \{ole-\}$ D: ta väriseb hirmust → ta on hirmul; poiss luksub ärevusest → poiss on ärevil jne.

14) T: → $N^1 V N^2_k$: tuba täitub veest → tuba täitub veega; leiliruum täitub aurust → leiliruum täitub auruga jne.

15) T: → $N^2_n V$: poiss punetab näost → nägu punetab; ta paraneb haavast → haav paraneb; tuba haiseb seebirasvast → seebirasv haiseb jne.

16) T: → $N^2_n V \{saa-\}$ N^1_{tr} : insener saab sinust → sina saad inseneriks; maalrisell tuleb temast → tema saab maalriselliks; mees kasvab poisist → poiss saab meheks; öö tuleb päevast → päev saab ööks jne.

17) T: → $N^1 V \{ole-\}$ N^2_n : insener saab sinust → insener oled sina; maalrisell tuleb temast → maalrisell on tema jne.

18) T: → $N^1 V$ sellest, et ... N^2 ... nad räägivad lenduritest → nad räägivad sellest, et ... lendurid ...; ma raporteerin sellest → ma raporteerin sellest, et ... see ... jne.

3. 2. F r a a s i t ü ü p $N^1 V N^2_{el} N^3$.

Käesoleva fraasitüübi struktuurid alluvad valdavas enamikus samadele transformatsioonidele, millele allusid vastavad $N^1 V N^2_{el}$ tüüpi struktuurid. Näiteks struktuurid me tõuseme lauast ja me palume teda lahkuda lauast alluvad mõlemad transformatsioonidele ... N^2_{el} ... → ... N^2_g äärest ja ... N^2_{el} ... → ... N^2_g juurest..., s. o. me tõuseme lauast → me tõuseme laua äärest → me tõuseme laua juurest ja me palume teda lahkuda lauast → me palume teda lahkuda laua äärest → me palume teda lahkuda laua juurest. Taolisi näiteid võib tuua piiramatul arvul. Pole ilmselt mõtet siin selliseid struktuure eraldi vaadelda, kuigi ranges käsitluses tuleks seda teha.

Ülejäänud N^1 v N_{el}^2 N^3 tüüpi struktuure esitab järgmine tabel.

T a b e l 2.

$T: \rightarrow$	N^1	N^2	N^3	
	V	N_g	<u>eest N^3</u>	
	N^1	V	N_g	<u>järgi N^3</u>
	N^1	V	N_k^2	N^3
	N^1	V	N_g^2	N_p^2
	+			<u>ta maksab kanadest 50 rubla</u>
		+		<u>ta tunneb narri naerust</u>
			+	<u>Rein tõmbab tikust tuld</u>
			+	<u>arst puudutab teda õlast</u>

Näiteid.

1) $T: \rightarrow N^1 V N_g^2$ eest N^3 : ta maksab kanadest 50 rubla
 $\rightarrow$ ta maksab kanade eest viiskümmend rubla; voorimees võtab sõidust raha $\rightarrow$ voorimees võtab sõidu eest raha jne.

2) $T: \rightarrow N^1 V N_g^2$ järgi N^3 : ta tunneb naerust narri $\rightarrow$
ta tunneb naeru järgi narri jne.

3) $T: \rightarrow N^1 V N_k^2$ N^3 : Rein tõmbab tikust tuld $\rightarrow$ Rein tõmbab tikuga tuld jne.

4) $T: \rightarrow N^1 V N_g^2$ N_p^2 : arst puudutab teda õlast $\rightarrow$ arst puudutab tema õlga jne.

3. 3. Fraasitüüp $N^1 N_{el,pl}^2$.

Tabel 3.

$T: \rightarrow$	$N^1 N_{g,pl}^2$	$N^1 N_{n,sg}$	$N^1 N_{p,pl}^2$	$N^1 V\{ole-\} N_{el,pl}^2$	
	<u>hulgast</u>				
+	+				<u>keegi magajaist</u>
+		+			<u>osa ekspeditsioonidest</u>
+		+	+		<u>grupp meestest</u>
+					<u>igäuks neist</u>

Näiteid.

1) $T: \rightarrow N^1 N_{g,pl}^2$ hulgast: igäuks neist $\rightarrow$ igäuks nende hulgast; üks õnnetest $\rightarrow$ üks õnnete hulgast; keegi magajaist $\rightarrow$ keegi magajate hulgast; neljas polaar-töötajaist $\rightarrow$ neljas polaar-töötajate hulgast; auahneim geenistest $\rightarrow$ auahneim geeniste hulgast; parim parima-test $\rightarrow$ parim parimate hulgast; grupp meestest $\rightarrow$ grupp meeste hulgast; osa ekspeditsioonidest $\rightarrow$ osa ekspedit-sioonide hulgast jne.

2) $T: \rightarrow A N_{n,sg}^1$: üks õnnetest $\rightarrow$ üks õnnetu; keegi magajaist $\rightarrow$ keegi magaja; neljas polaar-töötajaist $\rightarrow$ neljas polaar-töötaja; auahneim geenistest $\rightarrow$ auahneim geenius; tubli külastest $\rightarrow$ tubli külaste jne.

3) $T: \rightarrow N^1 N_{p,pl}^2$: grupp meestest $\rightarrow$ grupp mehi; osa ekspeditsioonidest $\rightarrow$ osa ekspeditsioone jne.

4) T: $\rightarrow N^1 V \{ole-\} N_{el,pl}^2$: grupp meestest $\rightarrow$ grupp on meestest; rühm pioneeridest $\rightarrow$ rühm on pioneeridest jne.

3. 3. 1. Lahtiseks jääb küsimus, mida teha selliste $N^1 N_{el}^2$ tüüpi struktuuridega, milles elatiivis olev nimisõna (tegelikult küll nimisõna funktsioonis olev noomen, nagu eespool nägime) võib esineda ka ainsuses. Siin on kaht tüüpi struktuure. Kui me vaatleme struktuure osa ekspeditsioonist ja osa ekspeditsioonidest, siis on ilmne, et elatiivil siin tähenduslikku vahet pole, mõlemad struktuurid alluvad transformatsioonile T: $\rightarrow N^1 N_g^2$ hulgast, näiteks osa ekspeditsioonidest $\rightarrow$ osa ekspeditsioonide hulgast; osa ekspeditsioonist $\rightarrow$ osa ekspeditsiooni hulgast. Arvukategooria pole ilmselt oluline. Võiks öelda, et struktuuris osa ekspeditsioonist on nimisõna elatiivis vormilt ainsuses, sisult mitmuses. Struktuuride veerand romaanist ja veerand romaanidest puhul on olukord teistsugune. Struktuur veerand romaanidest allub transformatsioonile T: $\rightarrow N^1 N_g^2$ hulgast (veerand romaanidest $\rightarrow$ veerand romaanide hulgast), kuid struktuur veerand romaanist ei allu (veerand romaanist $\nrightarrow$ veerand romaani hulgast). Järelikult arvukategooria on siin oluline. Struktuur veerand romaanidest peaks kuuluma fraasitüüpi $N^1 N_{el,pl}^2$, struktuur veerand romaanist aga kuhugi mujale.

3. 4. Fraasitüüp $N_{el}^1 N^2$.

Tabel 4.

$\uparrow$	N_n^2	V	$\{\underline{ole-}\}$	N_{el}^1
	N_n^2	V	$\{\underline{ole-}\}$	N_n^1
	N_n^2	N_n^1		
	+			<u>plekist konservikarp</u>
		+		<u>ungarlasest noormees</u>
		+	+	<u>Juhanist vend</u>

Näiteid.

1) T: $\rightarrow N_n^2 V\{\underline{ole-}\} N_{el}^1$: plekist konservikarp $\rightarrow$ kon-
servikarp on plekist; nahast jalgpallid $\rightarrow$ jalgpallid on
nahast; kalevist ülikond $\rightarrow$ ülikond on kalevist; rist-
palkidest tee $\rightarrow$ tee on ristpalkidest jne.

2) T: $\rightarrow N_n^2 V\{\underline{ole-}\} N_n^1$: lätlasest parvekikas $\rightarrow$ parve-
kikas on lätlane; marssalist Griša $\rightarrow$ Griša on marssal;
ungarlasest noormees $\rightarrow$ noormees on ungarlane; Juhanist
vend $\rightarrow$ vend on Juhan jne.

3) T: $\rightarrow N_n^2 N_n^1$: Juhanist vend $\rightarrow$ vend Juhan; marssa-
list Griša $\rightarrow$ marssal Griša jne.

3. 5. F r a s s i t ü ü p N_{el} A_{pos}.

T a b e l 5.

T: →					
	$A_{pos} N_n$				
	$N_g \underline{tõttu A_{pos}}$				
	$N_g \underline{üle A_{pos}}$				
	$N_g \underline{poolest A_{pos}}$				
	$N_{ess} A_{pos}$				
+					<u>näost punane</u>
	+				<u>võidusõnumist õnnelik</u>
	+				<u>äre vaseest rahutu</u>
			+		<u>õnnest rikas</u>
				+	<u>noorest nõrk</u>

Näiteid.

- 1) T: $\rightarrow A_{\text{pos}} N_n$: näost punane $\rightarrow$ punane nägu; kaelast kange $\rightarrow$ kange kael; kraest kitsas $\rightarrow$ kitsas krae jne.
- 2) T: $\rightarrow N_g$ tõttu A_{pos} : ärevusest rahutu $\rightarrow$ ärevuse tõttu rahutu; võidusõnumist õnnelik $\rightarrow$ võidusõnumi tõttu õnnelik jne.
- 3) T: $\rightarrow N_g$ üle A_{pos} : võidusõnumist õnnelik $\rightarrow$ võidusõnumi üle õnnelik jne.
- 4) T: $\rightarrow N_g$ poolest A_{pos} : õnnest rikas $\rightarrow$ õnne poolest rikas; kuulsusest vaene $\rightarrow$ kuulsuse poolest vaene jne.
- 5) T: $\rightarrow N_{\text{ess}}$ A_{pos} : noorest nõrk $\rightarrow$ noorena nõrk; vanast väeti $\rightarrow$ vanana väeti jne.

3. 6. Fraasitüüp $N_{el} A_{komp}$.

Käesolev tüüp on transformatoorselt ühtne. Kõik siia kuuluvad struktuurid alluvad transformatsioonile
 $T: \rightarrow A_{komp}$ kui N_n , näiteks ookeanist tumedam $\rightarrow$ tumedam
kui ookean; endisest hullem $\rightarrow$ hullem kui endine; teistest
ilusam $\rightarrow$ ilusam kui teised jne.

3. 7. Fraasitüüp $mis N^1 V N_{el}^2$.

Kõikidele antud fraasitüübi struktuuridele on iseloomulik transformatsioon $T: \rightarrow V$ eitav imperat N_p^2 , näiteks
mis me raiskame ajast $\rightarrow$ ärgem raisakem aega; mis te rikute
päevast $\rightarrow$ ärge rikkuge päeva; mis sa tänad minust $\rightarrow$
ära täna mind; mis sa väristad käest $\rightarrow$ ära värista kätt
 jne.

3. 8. Fraasitüüp $N^1 V_{imp} N_{el}^2$.

Kõik käesoleva fraasitüübi struktuurid alluvad transformatsioonile $T: \rightarrow N_n^2 V_{pers} N_p^1$, näiteks oas on piiratud
jääliustikest $\rightarrow$ jääliustikud piiravad oasi; ta on vai-
mustatud looduseilust $\rightarrow$ looduseilu vaimustab teda jne.

4. Kokkuvõtte.

Konstruksioonid, milles tuumana esineb nimisõna elatiivis, jagunevad seitsmeks erinevaks fraasitüübiks. Fraasitüüpide loend pole ilmselt lõplik, sest real juhtudel ei suudetud elatiivi relevantset ümbrust määrata. Siin on ühtlasi käesoleva töö üks suurimaid puudusi. Lisaks fraasitüüpide loendi oletatavale mittetäielikkusele pole täiesti kindel, kas ka esitatud fraasitüübid ise on kõik iseseisvad (vrd. $N^1 V N_{el}^2$ ja $N^1 V N_{el}^2 N^3$). Transformatoorne analüüs on igal juhul edukas mingi fikseeritud konstruksiooni piires.

4. 1. Fraasitüüp $N^1 V N_{el}^2$.

Elatiiv esineb 18 funktsioonis. Iga funktsiooni iseloomustavad kindlad transformatsioonid.

- 1) piim voolab klaasist välja
 $T: \rightarrow N^1 V\{\text{ole-}\} N_{in}^2$; $T: \rightarrow N^1 V N_g^2$ seest.
- 2) sõnnik lõpeb laudast
 $T: \rightarrow N^1 V\{\text{ole-}\} N_{in}^2$; $T: \rightarrow N^1 V N_{in}^2$.
- 3) mees väljub autost
 $T: \rightarrow N^1 V\{\text{ole-}\} N_{in}^2$.
- 4) mees läheb uksest välja
 $T: \rightarrow N^1 V$ läbi N_g^2 .
- 5) ta läheb trepist üles
 $T: \rightarrow N^1 V N_p^2$ mööda.
- 6) me tõuseme lauast
 $T: \rightarrow N^1 V N_g^2$ äärest; $T: \rightarrow N^1 V N_g^2$ juurest.
- 7) tuul puhub ahtrist
 $T: \rightarrow N^1 V N_g^2$ poolt.
- 8) ta haarab lauast kinni
 $T: \rightarrow N^1 V N_g^2$ küljest.
- 9) tuba haiseb seebirasvast
 $T: \rightarrow N^2 V_{irref} N_{g/p}^1$; $T: \rightarrow N^1 V N_g^2$ tõttu;
 $T: \rightarrow N^1 V N_g^2$ järele; $T: \rightarrow N_n^2 V$.
- 10) ta väriseb hirmust
 $T: \rightarrow N^2 V_{irref} N_{g/p}^1$; $T: \rightarrow N^1 V N_g^2$ tõttu;
 $T: \rightarrow N^1 V\{\text{ole-}\} D$.
- 11) ta ehmatab äikesest
 $T: \rightarrow N^2 V_{irref} N_{g/p}^1$; $T: \rightarrow N^1 V N_g^2$ tõttu.
- 12) tuba täitub veest
 $T: \rightarrow N^2 V_{irref} N_{g/p}^1$; $T: \rightarrow N^1 V N_k^2$.
- 13) see selgub kirjast
 $T: \rightarrow N^2 V_{irref} N_{g/p}^1$.
- 14) poiss punetab näost
 $T: \rightarrow N_n^2 V$.

- 15) insener saab sinust
 $T: \rightarrow N_n^2 V\{\text{saa-}\} N_{tr}^1 ; T: \rightarrow N^1 V\{\text{ole-}\} N_n^2.$
- 16) mees kasvab poisist
 $T: \rightarrow N_n^2 V.$
- 17) nad räägivad lenduritest
 $T: \rightarrow N^1 V \text{ sellest, et } \dots N^2 \dots$
- 18) ma lahkun sinust
 $T: \rightarrow N^1 V N_g^2 \text{ juurest.}$

4. 2. F r a a s i t ü ü p $N^1 V N_{el}^2 N^3.$

Elatiiv esineb 4 funktsioonis.

- 1) ta maksab kanadest 50 rubla
 $T: \rightarrow N^1 V N_g^2 \text{ eest } N^3.$
- 2) ta tunneb narri naerust
 $T: \rightarrow N^1 V N_g^2 \text{ järgi } N^3.$
- 3) Rein tõmbab tikust tuld
 $T: \rightarrow N^1 V N_k^2 N^3.$
- 4) arst puudutab teda õlast
 $T: \rightarrow N^1 V N_g^3 N_p^2.$

4. 3. F r a a s i t ü ü p $N^1 N_{el,pl}^2.$

Elatiiv esineb 4 funktsioonis.

- 1) keegi magajaist
 $T: \rightarrow N^1 N_{g,pl}^2 \text{ hulgast; } T: \rightarrow \Delta N_{n,sg}^1.$
- 2) osa ekspeditsioonidest
 $T: \rightarrow N^1 N_{g,pl}^2 \text{ hulgast; } T: \rightarrow N^1 N_{p,pl}^2.$
- 3) grupp meestest
 $T: \rightarrow N^1 N_{g,pl}^2 \text{ hulgast; } T: \rightarrow N^1 N_{p,pl}^2;$
 $T: \rightarrow N^1 V\{\text{ole-}\} N_{el,pl}^2.$
- 4) igaüks neist
 $T: \rightarrow N^1 N_{g,pl}^2 \text{ hulgast.}$

4. 4. F r a a s i t ü ü p $N_{el}^1 N^2$.

Elatiiv esineb 3 funktsioonis.

1) plekist konservikarp

$$T: \rightarrow N_n^2 V\{\underline{ole-}\} N_{el}^1.$$

2) ungarlasest noormees

$$T: \rightarrow N_n^2 V\{\underline{ole-}\} N_n^1.$$

3) Juhanist vend

$$T: \rightarrow N_n^2 V\{\underline{ole-}\} N_n^1; T: \rightarrow N_n^2 N_n^1.$$

4. 5. F r a a s i t ü ü p $N_{el} A_{pos}$.

Elatiiv esineb 5 funktsioonis.

1) näost punane

$$T: \rightarrow A_{pos} N_n.$$

2) võidusõnumist õnnelik

$$T: \rightarrow N_g \underline{tõttu} A_{pos}; T: \rightarrow N_g \underline{üle} A_{pos}.$$

3) ärevusest rahutu

$$T: \rightarrow N_g \underline{tõttu} A_{pos}.$$

4) õnnest rikas

$$T: \rightarrow N_g \underline{poolest} A_{pos}.$$

5) noorest nõrk

$$T: \rightarrow N_{ess} A_{pos}.$$

4. 6. F r a a s i t ü ü p $N_{el} A_{komp}$.

Elatiiv esineb ühes funktsioonis.

ookeanist tumedam

$$T: \rightarrow A_{komp} \underline{kui} N_n.$$

4. 7. F r a a s i t ü ü p mis $N^1 V N_{el}^2$.

Elatiiv esineb ühes funktsioonis.

mis me raiskame ajast

$$T: \rightarrow V \text{ eitav imperat } N_p^2.$$

4. 8. F r a a s i t ü ü p $N^1 V_{imp} N_{el}^2$.

Elatiiv on ühes funktsioonis.

oaas on piiratud jääliustikest

T: $\rightarrow N_n^2 V_{pers} N_p^1$.

PRANTSUSE KEELE SÜNTAKTILISE STRUKTUURI KIRJELDAMISEST

J. K a p l i n s k i

O. See töö püüab osutada ühele võimalikule teele keele süntaktilise struktuuri kirjeldamisel. Ta ei ole mõeldud ammendavana ei meetodi andmises ega materjali esitamises.

Kõigepealt mõned üldisemad märkused.

Süntaksi valdkond ei ole tänini täpselt määratletud, eriti kõikuvad on piirid morfoloogia, semantika ja süntaksi vahel. Muidugi ei ole oluline, et neil või üldse mingitel terminitel oleks jääv tähendus, oluline on kõigepealt ainult võimalikkus paremini kirjeldada seda, mis on, olenevata selle nimetusest, sellest, mis sõnaga vastavat nähtust on kutsutud ja kas ongi kindlat terminit selle nähtuse jaoks.

Kui lähtume tänapäeval mitmesugustes teadustes endastmõistetavast ja keeleteaduses endastmõistetavaks muutuvast kirjeldamise meetoodikast, võime jagada kirjeldamise kahte ossa: segmenteerimine ja struktuuri, s. t. segmentide vahetkordade kirjeldamine.

Segmenteerimata materjali struktuuri ei ole võimalik kirjeldada, seega on segmentide olemasolu esimene deskriptiooni tingimus. Siiski ei saa anda sellise nähtustiku analüüsimiseks, nagu seda on keel, mehhaanilist skeemi ranges järgnevuses: segmenteerimine ja seejärel struktuuri kirjeldamine. Segmenteerimisel tuleb paljuski arvestada struktuuri kirjeldamise nõudeid ja tulemusi. Keele struk-

tuuri kirjeldamine on "iteratiivne protseduur", kus enamik meie tulemustest ei jää esialgsel kujul püsima, vaid me parandame neid mitmeid kordi. Nii on lugu keele foneemisüsteemi kirjeldamisel, veel enam aga morfeemikas ja "kõrgematel" tasanditel, keerukamate ühendite puhul. Kõige selle tõttu on keele kirjeldamine, eriti selle algus, näiteks foneemide fikseerimine, vaevalt algoritmistatav. Vähe-malt ei ole see senini osutunud tasuvaks ja seepärast on lähtutud ikkagi mingist intuitsioonile või traditsioonile toetuvast jaotusest, fonoloogias näiteks "häälikutest" ja mõnedest suprasegmentaalsetest nähtustest nagu rõhk ja intonatsioon. Meil on selgem või udusem pilt sellest, kuidas oleks kõige ökonoomsem kõnevoolu osadeks jagada. Meil on ka ettekujutus sellest, milline umbes peaks olema struktuurikirjeldus, mida me tahame saada. Ning meil on mitmed kriteeriumid, mida jälgides saame hiljem muuta esialgseid tulemusi. Kõik need lisategurid mõjustavad meid rakendama üht või teist meetodit, operatsiooni, loogikat. Nad ei tohi aga muuta juba valitud meetodit tema kasutuspiirkonnas. Meetod peab olema võimaluse piires täpne ja peab olema selge, kus millist meetodit on kasutatud. Meetodit ei saa kasutada poolikult, järjekindlusetult, ranguse nõudeid alandades, nagu seda vahel esineb rangusele pretendeerivais töis. Kui oleme valinud range meetodi, tuleb seda rangelt rakendada, ja kui see osutub võimatuks, tuleb rangest meetodist selgelt loobuda.

Materjali esialgse segmenteerimise juurde kuulub materjali esitamine elementide lõpliku arvu abil, nende mingi kombinatsiooni või kombinatsioonidena. Keeleteaduses esitatakse materjal enamasti abstraktsete elementide (foneemid, morfeemid, vabad vormid, minimaalsed lausungid jne.) järjendina. Kui see esimene samm on tehtud, võib asuda elementide vahekordi kirjeldama. Kuna käesolevas töös oluliselt ei tegelda konkreetsetest segmentidest abstraktsete üksuste tuletamisega, siis teeme selle protseduuri kohta

ainult mõne märkuse. Tavaliselt lähtutakse esialgsel segmenteerimisel juba mingist eelkujutlusest tulevaste abstraktsete üksuste kohta. See annabki segmenteerimise aluse. Mida otstarbekamalt me segmenteerime, seda hõlpsam on üleminek abstraktsetele märkidele, näit. kõnevoolu segmentidelt häälikutele ja siis foneemidele. Nähtavasti on võimalik ükskõik kuidas segmentideks jagatud teksti abstraktsete märkidega esitada, seda näitab kasvõi sõnavahedeta või valede sõnavahedega kirjutatud teksti lugemise-dešifreerimise võimalikkus. Keelt oskav inimene sellise teksti dešifreerijana või näiteks sõnavahede asetajana on väga täpne ja korralik masin, kuigi tema töötamise põhimõtted on nähtavasti küllalt keerulised.

Keele struktuuri kirjeldamine peab võimaldama eraldada keeles võimalikku võimatust. Abstraktsete märkidega antud mudelis saame esialgsetest elementidest – siin on nendeks prantsuse keele kirjatähtedest moodustatud "sõnad" – väga palju mitmesuguseid ühendeid. On vaja anda eeskirjad, reeglid, mis näitavad, millised neist ühenditest on võimalikud, keeles loomulikud, tavaliselt kasutatavad. Võib lisada ka andmeid ühendite tõenäoise sageduse kohta, näiteks nimetades, kas mingi ühend on tavaline, harv või ainult erandjuhtudel esinev.

1. Pärast neid märkusi siirdume prantsuse keele lausungite kirjeldamisele.

1. 1. Kõige väiksemateks elementideks meie kirjelduses on sõnad, täpsemalt grafeemiühendid kahe vahe vahel, millest üks või mõlemad võivad olla ka tähistatud kirjavahemärkidega. Lähtume "kirjasubstantsist" seetõttu, et kõrgematel tasanditel (morfoloogia, süntaks) näib olevat ükskõik, millisest substantsist lähtuda, ja seetõttu, et polnud võimalik hankida kõnematerjali.

On enam kui ilmne, et see, mida me siin sõnaks nimetame, ei vasta sõna mõistele prantsuse grammatikas ega üldse enamikus töödes, kus sõna olemust käsitletakse.

See aga ei takista analüüsimist alustamast sellistest grafeemijärjenditest. Hiljem selgitame, millistesse kindlamatesse ühenditesse rühmituvad meie esialgsed sõnad ja nende klassid. Need saadud ühendid on omakorda elementideks mingil järgmisel, "kõrgemal" tasandil, kus nad moodustavad uusi ühendeid. Selliseid tasandeid on teoreetiliselt lõpmata palju, sest alati on võimalik mingile tekstile liita teine tekst ja saadud liittekstile mingi uus tekst. Praktiliselt aga ei näi olevat lausest ja võib-olla mõningatest lauseühenditest kaugemale-kõrgemale analüüsimisel, edasiste tasandite ühendite kirjeldamisel olulist väärtust. Võib-olla ei ole ka siin visandatud analüüsimeetod universaalse meetodina eriti otstarbekas. Tuleb alati kaaluda, kas on mõtet järgida skeemi: algul elementide sedastamine ja siis ühendite kirjeldamine, kas iga järgmise tasandi elemente on kõige sobivam kirjeldada tingimata eelmise tasandi elementide ühenditena.

1. 2. Esitame nüüd sarja sõnaklasse, mille abil kirjeldame hiljem lausungeid.

Tuleb nimetada, et siin toodud sõnaklasside loetelu ei ole range, algoritmilise protseduuri tulemus, vaid rajaneb suuresti nii traditsioonilisel käsitlusel sõnaliikidest kui ka kaalutlustel, mis tekkisid nende sõnaklasside abil lausungite kirjeldamisel. Niisiis on meie töö mõneti "iteratiivne": tema lõpuosa võib mõjustada muutma tema algust.

Juba sõnaklasside esialgsel andmisel ilmneb, et mõned klassid, mis traditsioonilises käsitluses on fikseeritud, sedastuvad hästi ja eristuvad teistest rangemategi meetoditega nende koosseisus leiduvate spetsiifiliste morfeemide põhjal ja ka ümbruste alusel. Sõnaklasside määramisel tuleb esimesed klassid tavaliselt anda kas lihtsa loendiga (kui klassi kuulub vähe sõnu) või morfeemide abil. Edaspidi kasutame olemasolevaid klasse nende ümbruste kirjeldamisel, milles saavad esineda teised klassid, seega järg-

miste klasside määramisel ümbruste kaudu.

Sõnaklasside fikseerimisel ilmneb, et paljudel klassidel ei ole selgelt ja lõplikult antavaid ümbrusi. Väga kõikuvate ja erinevate ümbrustega on sõnad, mida traditsiooniline grammatika nimetab adverbideks. Adverbe saame jagada klassideks analoogiliselt enamiku teiste sõnadega: erinevused üksikute siin antud adverbiklasside vahel on väiksemad kui erinevused näiteks verbi ja noomeni vahel, niisama on adverbiklassid oma koosseisult palju heterogeensemad. Võib-olla lahendaks raskused adverbide klassideks jaotamisel hierarhiline klasside-alamklasside süsteem, mida siin ei ole siiski püütud rakendada. Samal ajal on aga silmanähtavad olulised sarnasused mitmete sõnaklasside vahel, mis siin on eristatud. Kõik see viitab vajadusele sõnaklasside hoolikaks väljatoomiseks ja nende sarnasuste sedastamiseks, näiteks distributsiooni alusel. Sõnaklasside erinevused ja sarnasused ilmnevad hiljem selgesti ka süntaksis, sõnaühendite kirjeldamisel, kus selgub adverbide eriline osa lause koosseisus.

Kuna selle töö eesmärgiks on kõigepealt sõnaühendite lausungite, mitte aga sõnaklasside kirjeldamine, siis anname sõnaklassid ilma ülalnimetatud probleemidel pikemalt peatumata. Arvestame ainult, et sõnaklasside loend ei oleks liiga "kitsas". Ta võib olla liiga "lai", see aga ei ole edasisel analüüsil oluliseks takistuseks. Kui osutub, et sõnaklasse on eristatud liiga palju ja mõned neist käituvad peaaegu niisamuti kui mõned teised, siis võime alati sarnased klassid ühendada. Kui aga oleme töö algul võtnud üheks klassiks kokku erinevate omadustega sõnad, ei paista see hiljem silma ja tulemuseks on, et kirjeldus ei too esile olulisi erinevusi sõnaühendites. Seega on parem kahtlased sõnad eraldada iseseisvasse klassi, kui jätta nad klassi, mille teiste liikmetega nad ümbrustelt täiesti kokku ei lange. Seda juhust oleme üldiselt järginud, välja arvatud adverbid, mille puhul sageli ilmneb, et peaaegu

iga sõna või väike rühm sõnu on oma spetsiifiliste ümbrustega teistest neile sarnastest sõnadest erinev.

Sõnaklassid antud, vaatleme juba ühendeid ja kirjeldame neid nende sõnaklassilise koostise järgi "valemitega". Prantsuse keeles eristame järgmisi sõnaklasse.

1. N. Sellele klassile saab alati vahetult eelneda klass Det, nii et saame lausungi Det N.

2. Det. Esineb ainult N-i ees, kas vahetult või on nende vahel Adv₁ sisaldav ühend. Klassi Det kuuluvad järgmised sõnad (anname nad loendina): le, la, les, mon, mes, ton, ta, tes, son, sa, ses, notre, nos, votre, vos, ce, chaque, tout₂ ... Loend ei pruugi olla täielik, talle võib alati lisada uusi sõnu, mille kuuluvuse sellesse klassi otsustame juba distributsiooni põhjal.

3. Adj₁ ja Adj₂. Neid eristame omavahel ainult vajaduse korral (eristamine on seega fakultatiivne), tavaliselt märgime Adj; nii oleme tegelikult varjatult sisse toonud alamklassi mõiste. Adj₁ saab esineda ümbruses (Det) - N (un bon garçon), mis on tema "diagnostiline" ümbrus; Adj₂ saab esineda ümbruses (Det) N Adv₄ - (le garçon très énergique). Mõlemad adjektiivid saavad esineda ümbruses (Det) N V₉ - (le garçon est bon; le garçon est énergique).

4. Pps. Esineb ainult ümbruses - V. Siia klassi kuuluvad järgmised sõnad: je, tu, il, elle, nous₄, vous₄, ils, elles, on.

5. V. Kuna eeldame, et morfoloogia on tuntud, loeme klassi V kõik sõnad, mis esinevad ühes kuuest võimalikust pöördest (est, va, allons, viens, défend ...). Sageli ilmneb, et paljud klassi V sõnad erinevad ümbrustelt teistest selle klassi sõnadest. Sel juhul, et klassifikatsioon ei oleks liiga kitsas, eraldame need verbid vastavatesse eri rühmadesse (alamklassidesse), kuhu võib kuuluda üks

või rohkem verbe kõigis oma muudetes (nagu est, suis, sois, sommes ...). Neid rühmi märgine V_e (suis jne.), V_a (a, avons jne.) jne.

6. Prep. Sellesse klassi kuuluvad de, à, pour, en, sous, sur, par, sans, avec ... Loendit võib jätkata, kuid ta ei ole piiramatult. Klass Prep vastab sellele, mida traditsioonilises grammatikas nimetatakse prepositsioonideks. Prep on üks kõige heterogeensemajd klasse adverbide kõrval, kuigi on olemas kõikidele prepositsioonidele ühiseid ümbrusi, nagu Pps V - (Det) N; (Det) N - N; Pps V - Inf.

7. Inf. See klass on kergesti tuntav lõppude järgi ja määratud igas traditsioonilises grammatikas. Vajaduse korral võime infinitiive avoir, être jm. tähistada vastavate indeksite abil, näiteks Inf_a , Inf_e jne.

8. Ppres. Preesensi partitsiip on fikseeritud grammatikates ja on määratav lõpu -ant abil.

9. Ppas. Mineviku partitsiip on fikseeritud morfoloogias lõppude järgi (marché, bu, arrivé). Nii Ppres kui ka Ppas juures võime kasutada indekseid mõnede eri verbide vormide märkimiseks: $Ppres_a$ = eu; $Ppas_e$ = été jne.

10. Pop. Personaalpronoomenid direktse objekti funktsioonis on: me₁₀, te₁₀, le₁₀, la₁₀, nous₁₀, vous₁₀, les₁₀, se₁₀, rien₁₀ ... Neile on karakterne ümbrus Det N - V, millele ei saa järgneda Det N või N, kui sellele ei eelne mõnd kirjavahemärki (koma).

11. Pton. Rõhulised personaalpronoomenid on: moi, toi, lui₁₁, nous₁₁, vous₁₁, eux₁₁, elles₁₁, soi₁₁ ...

12. Pind. Indefiniitsed pronoomenid on tout₁₂, rien₁₂, personne₁₂, chacun.

- Kuna grafeemijärjend me võib esindada mitut sõna, mis kuuluvad eri klassidesse, kasutame siin ja edaspidi klasside järjekorranumbreid sõnade juures indeksitena, et vältida homonüümiat.

13. Pdem. Demonstratiivpronoomenid on cela, celui-ci, celui-là, celle-ci, ça, ce₁₃ ...

14. Pdemp. Siia klassi kuuluvad demonstratiivpronoomenid celui ja celle, mis ei saa esineda vabalt.

15. Ppsind. Selles klassis on personaalpronoomenid indirektse objekti funktsioonis: me₁₅, te₁₅, lui₁₅, nous₁₅, vous₁₅, leur₁₅, se₁₅.

16. Adv₁. Klassi eristavaks ümbruseks on Det - Adj₁N. Siia klassi kuuluvad bien₁₆, tout₁₆, assez₁₆, très, moins, aussi₁₆, plus₁₆, si₁₆ ...

17. Adv₂. Selle klassi ümbrustest on olulised (Det) N - Adj₂, samuti Pps (Pob) (Ppsind) V_e - Ppas. Siia kuuluvad sõnad nagu mal, brillamment, sûrement, pas, infiniment.

18. Adv₃. Siia kuuluvad põhiliselt kohaadverbid, positsioonidest on kõige sagedasem lausungi lõpp, kus Adv₃ (välja arvatud en ja y) võib alati esineda. Adv₃ esineb peaaegu alati verbi sisaldavas lausungis. Sõnadest kuuluvad siia näiteks ici, là, loin, là-dessus.

19. Adv₄. Vastab üldiselt ajaadverbile. Ta võib peaaegu alati esineda verbiga lausungi alguses (vrd. Adv₃!). Näiteid: aujourd'hui, hier.

20. Adv₅. Siia kuuluvad sõnad nagu oui, non, mais₂₀ ja ka adieu, holà, allions₂₀ - seega mitmesugused partiklid, mis võivad esineda nii lausungina kui ka lausungi osana, komaga eraldatult lausungi alguses.

21. Adv₆. Siia kuuluvad sõnad cependant, encore₂₁, aussi₂₁, même₂₁, toujours₂₁, beaucoup₂₁, déjà.

22. Sub. Siin on tegemist sõnadega, mis asendavad teatavaid eelnevaid sõnu. Need on qui₂₂, que₂₂.

23. Rel₁. Selle klassi sõnad saavad esineda ainult verbi sisaldava lausungi järel, kusjuures enne punkti

(hüüu-, küsimärki, koolonit, semikoolonit) peab seisama veel üks verb. Siia kuuluvad que₂₃, si₂₃ ...

24. Rel₂. Siia klassi kuuluvatele sõnadele peab järgnema samas lausungis alusega verb. Siia kuuluvad si₂₄ (näiteks s'il pleut aujourd'hui, je resterai chez moi), que₂₄.

25. Rel₃. Siia kuuluvad sõnad, mis asuvad vähemalt kahe samasuguse klassistruktuuriga ühendi vahel nagu et, ou, mais₂₅ (näiteks lausungis gras mais faible), ni, soit ... soit.

26. Tot. Siia kuuluvad sõnad tout₂₆, toute₂₆, tous₂₆, toutes₂₆ ...

27. Ind. Siia kuuluvad sõnad bien₂₇, beaucoup₂₇, peu₂₇, un, une, tant₂₇, plusieurs ...

28. Quant. Siia kuuluvad põhiarvasõnad (peale un, une).

29. Part. Siia kuulub sõna de₂₉ (de partitiivses funktsioonis).

30. voici. Siia kuuluvad sõnad voici ja voilà.

31. Sõnad que₃₁, ce que₃₁, comme₃₁, nagu näiteks lausungis Qu'il est joli!

32. Sõnad comme₃₂, tel₃₂, telle₃₂, que₃₂, nagu neid tarvitatakse võrdlustes: long comme un jour sans pain.

33. Sõna que₃₃, nagu näiteks ühendis qu'on me donne...

34. Q. Klass Q koosneb õieti mitmest küsisõnade klassist, mille ühiseks tunnuseks on seotus küsimärgiga, mis peab asuma küsisõnaga lausungi lõpus. Kui tunneme küsisõnu ja nende ümbrusi, võime nende ümbruste sarnasuse põhjal klassi Q arvata ka relatiivpronoomenid, mis nendega vormilt ühte langevad, esinevad aga ainult kõrvallauses ega nõua küsimärki lausungi lõpus. Seda kõike saab määrata ka formaalselt. Näiteid: que₃₄, comment, combien, quand, pourquoi ...

35. Sõna *même*³⁵, nagu näiteks lausungis *même les vieillards étaient présents*.

Muidugi ei ole näiteks klassidel 31, 33, 35 olulist osa prantsuse lausungite koostises. Kuid ülevaatlikkuse huvides tõime nad sõnaklasside loendis ära. Tuleb ka märkida, et traditsiooniline sõnaliikide käsitus ei klassifitseeri selliseid sõnu tavaliselt kunagi hoolikalt, luges nad adverbide või partiklite hulka, mistõttu need klassid paisuvad veel suuremaks ja kaotavad suure osa oma lingvistilisest sisust.

Me ei peatu pikemalt raskustel, mis tulid ilmsiks sõnade klassifitseerimisel. Võib ainult öelda, et iga keelt on otstarbekam vaadelda mitme kõrvuti eksisteeriva süsteemi ühendina kui üheainsa süsteemina, kuna viimasel juhul see süsteem oleks otstarbetult keeruline ja täis ad hoc reegleid.

Keeles võib näiteks eraldada fraseoloogiliste ja mittefraseoloogiliste ühendite süsteemid. Esimene neist vajab oma kirjeldusviise, sest kui võtta fraseoloogilised ühendid kõrvuti teistega kokku ühte reeglistikku, siis see ainult hägustaks ülevaadet tänapäeva elava prantsuse keele lausungitest ja ühenditest. Käesolevas töös on käsitletud ainult mittefraseoloogilisi ühendeid.

Fraseoloogilist ühendit saab kõige ülevaatlikumalt ja populaarsemalt määratlada kui ühendit, milles elementide vastastikune sõltuvus on tugevam kui teistes sama tüüpi (sama valemiga kirjeldatavates) ühendites. Võib siiski juhtuda, et ühendit ei saagi mingi valemiga adekvaatselt kirjeldada, näiteks kui esinevad vormid ja sõnad, mida tavaliselt ei kasutata. Näiteks ühendis je, soussigné ei ole je meie süsteemi järgi Pps (ümbrus ei võimalda teda selleks lugeda) ega ka Pton. Antud määratlus ei võimalda selgesti eristada fraseoloogilist mittefraseoloogilisest, kuid nende kahe sfääri piir ei olegi tegelikult selge. Võime vaid öelda, et tegelikult on ühendid

kas rohkem või vähem fraseoloogilised ja et nende fraseoloogilisuse aste võib olla väga erinev.

Järgnevalt esitame illustratsiooniks tabeli, kus on näidatud mitmesuguste sõnarühmade ümbrusi, mille järgi me jaotame need sõnarühmad sõnaklassidesse.

Ümbrused	lau- sung	lau- sung	Det N - V _e	Det N - V _e	Det N - Adj	Det N - V _e	Det N - V _e	Det N - V _e	Det N - V _e
Sõna- rühma nr.									
1	+	+	-	+	-	+	+	+	+
2	+	+	-	-	-	-	-	-	-
3	+	+	-	-	-	-	-	-	-
4	+	+	+	+	-	+	+	+	+
5	+	+	+	+	-	+	+	+	+
6	+	+	-	-	-	-	-	-	-
7	-	-	-	-	+	+	+	+	+
8	-	-	-	-	+	+	+	+	+
9	+	+	-	+	+	+	+	+	+
10	-	-	-	-	+	+	+	+	+

Sõnarühmad on:

- 1) cependant, quand même ...
- 2) oui, non, si (si jaatusena eitava küsimuse järel)
- 3) mais
- 4) ici, là, loin ...
- 5) hier, demain, bientôt ...
- 6) même (nagu lausungis c'est même bête)
- 7) aussi (nagu lausungis lui est aussi étudiant)
- 8) encore (nagu lausungis raconte-moi encore quelque chose)
- 9) toujours
- 10) infiniment, brillamment ...

Tabel lubab meil ühendada antud 10 rühma vähemaks arvuks rühmadeks resp. sõnaklassideks. Ühte klassi on võetud rühmad 2 ja 3 (klass Adv₅), rühmad 1, 6, 7, 8, 9 (klass Adv₆), kuigi nende puhul esineb erinevusi distributsioonis, eriti rühmal 1 teiste suhtes. Eri klassideks on eraldatud rühmad 4 ja 5 (vastavalt klassid Adv₃ ja Adv₄). Rühma 10 sõnad kuuluvad klassi Adv₂.

2. Kui oleme fikseerinud sõnaklassid, tuleb nende abil kirjeldada teatav hulk lausungeid. Alustame sellega, et tähistame iga sõna lausungis tema klassisümboliga, nii et tulemuseks on lausungi klassistruktuuri kirjeldus, valem.

Ei ole kuidagiviisi otstarbekas piirduda selliste kirjeldustega. Oleme saanud alles mingi elementide hulga (elemendid on nii sõnaklasside sümbolid kui ka nende ühendid), kus elementide vahel on teatavad suhted. Mõned elemendid võivad ühendis esineda koos, teised isegi saavad esineda ainult koos, mõned elemendid saavad esineda ainult teatava minimaalse pikkusega ühendeis jne. Süntaksiks võibki nimetada sõnaklasside esinemistingimuste kirjeldust, eriti üksteise suhtes.

2. 1. Kõigepealt vaatame, millised toodud sõnaklassidest saavad moodustada lausungeid iseseisvalt, eraldatuna antud ümbrustest. Sõnu, mis on antud ümbrusest eraldatavad,

mis võivad olla iseseisvalt lausungiks, nimetame vabadeks (vabadeks mingis ümbruses). Ja vastupidi, sõnaklass on seotud, kui ta antud ümbrusest ei ole iseseisva lausungina eraldatav, nagu näiteks artiklid ja personaalpronoomeni rõhuta vorm (Pps) kõigis ümbrustes. Paljud sõnad on mõnedes ümbrustes vabad, mõnedes seotud. Näiteks va on vaba lausungis va chez le tailleur, kuid seotud lausungis va-t-en.

Sõnaklassi omadus olla ülekaalukalt kas vaba või seotud iseloomustab seda klassi mõneski olulises suhtes; vaba sõnaklass on lausungis omamoodi keskpunktiks, sageli on ta tuumaks (tuuma kohta allpool). Vabad sõnaklassid vastavad meie intuiitliivsele kujutelmale "autonoomsetest", "täisväärtuslikest", kesketest sõnadest lauses või ühendis. Informatsioon sellest, kas sõnaklass on vaba või seotud, on ühtlasi informatsiooniks selle klassi osa kohta keele süntaktilises struktuuris. See, kas sõnaklass on vaba või seotud, on oluline mitmesuguste "ellipsite" moodustamisel lausungist, näiteks kordamisel, rõhutamisel, dialoogis, üldse lühidalt väljendumisel. Sõnaklasside vabaduse-seotuse vahekorral on tähtsust ka tüpoloogias, nii näiteks iseloomustab prantsuse keelt suur seotusaste vastukaaluks näiteks ladina keelele.

Opositsiooni seotud-vaba kõrval on oluline opositsioon tuum-ümbris, mis on lähedane traditsioonilisele põhisõnalaiendi opositsioonile. Tuumaks ühendis on sõna, mis oma ümbrustelt on ekvivalentne ühendi endaga (see ekvivalent ei ole peaaegu kunagi absoluutne). Tuumast ranges mõttes saamegi rääkida ainult nn. endotsentriliste ühendite puhul. Endotsentriline on ühend, mis ümbruste poolest on ekvivalentne mõne oma elemendiga, näiteks ühend le garçon, sest ta on peaaegu kõigi ümbruste poolest ekvivalentne oma elemendiga garçon. Ühendeid, mis ümbrustelt ei ole ekvivalentsed ühegi oma elemendiga, nimetatakse eksotsentrilisteks. Selline on näiteks ühend à un garçon, mis ei saa asuda samasugustes ümbrustes kui à, un või garçon. Süntaksi käsitlemisel on palju olulisemad endotsentrilised ühendid, kuna

morfoloogias, "sõnasisesel" tasandil on rohkem tegemist eksotsentriliste ühenditega, milles peale muu on morfeemide võimalik arv enamasti piiratud (nii ei saa verbitüvele järgneda mitut pöördelõppu jne.). Ühendi endotsentrilisus eeldab seda, et ühendis oleks mõni element tingimisi vaba, et saaksime võrrelda tema ja ühendi ümbrusi. Paljudel juhtudel ongi ühendi tuumaks tema vaba element ja üldreeglina on eksotsentrilistes ühendites elemendid tugevamini seotud. Peaaegu iga eksotsentrilist ühendit on võimalik muuta endotsentriliseks mingi elemendi lisamisega, mis prantsuse keeles asub ühendi ees (Prep, Rel₁, Rel₂). Nähtavasti on see vastavus keele jaoks oluline, ta seob mitmesugused ühenditüübid tihedamalt üksteisega, lisades sellega keelele "süsteemsust". Niisama on võimalik mitmesugustelt endotsentrilistelt ühenditelt siirduda teistele endotsentrilistele ühenditele ja ka eksotsentrilistele.

Märgime veel, et tarvitame mõistet "tuum" ka eksotsentriliste ühendite puhul. Eksotsentrilistes ühendites nimetame tuumaks esimest endotsentrilist ühendit, mille me saame antud ühendi eest selle elemendi (resp. ühendi) lahutamisel, mis teeb antud ühendi eksotsentriliseks. Nii on ühendi à un garçon tuumaks selles tähenduses un garçon. Kuna iga ühend on kas endo- või eksotsentriline, ei saa mõistet "tuum" siin valesti mõista.

Järgnevalt esitame tabelina ülevaate sõnaklassidest selle järgi, kas nad võivad olla ühendis tuumad ja vabad elemendid. Tuuma vastand on ümbris, vaba elemendi vastand seotud element. Kui sõna võib olla tuumaks ja vabaks elemendiks, märgime tabelisse +, muidu - . Kui sõna võib olla nii üks kui teine, kui me ei saa teha sellel tasandil üldotsustust pluss ega miinus, kirjutame $\pm$.

Klass	Vaba- seo- tud	Tuum- umb- ris	Klass	Vaba- seo- tud	Tuum- umb- ris
N	+	+	Pob	-	-
Det	-	-	Pton	+	+
Adj	+	+	Pind	+	+
Pps	-	-	Pdem	+	+
V	+	+	Pdemp	-	+
Prep	-	-	Ppsind	-	-
Inf	+	+	Adv ₁	-	-
Ppres	+	+	Adv ₂	+	+
Ppas	+	+	Adv ₃	+	+

Klass	Vaba- seo- tud	Tuum- umb- ris
Adv ₄	+	+
Adv ₅	+	+
Adv ₆	+	+
Sub	-	+
Rel ₁	-	-
Rel ₂	+	-
Rel ₃	-	-
Tot	-	-
Ind	-	+
Quant	+	+
voici	+	+

Ülejäänud viie sõnaklassi suhtes ei ole neil hinnangutel enam mingit olulist sisu ja sellepärast me neid siin ei esita.

Tabeli põhjal võime kiiresti määrata, millised klassid võivad olla ainult vabad, millised ainult seotud, millised vahepealsed. Tuleb aga märkida, et kuna tegelik

olukord ei ole nii jälgalt piiriteldav, siis annab meie tabel sellisel kujul natuke vähe relevantseid andmeid. Seda saab muidugi parandada, kasutades mingit rohkemavalentset loogikat, kasvõi iga pluss-minus otsustust kommenteerides, kas antud klass on eranditult, väga sageli, sageli, keskmiselt, harva või väga harva vaba jne.

Nähtavasti on teatav lingvistiline sisu ka elementide "suhtelise vabaduse" määramisel ühendis, s. o. sellel, missugune element ühendis on vabam, missugune vähem vaba. Näiteks ühendis Pps V on V alati vabam ja võib teatud tingimustel (telegramm, kordamine, mõnes pöördvormis rahva kõnes) esineda iseseisvalt, kuigi ta normaalselt (välja arvatud imperatiivis) on seotud Pps-iga. Sama kehtib üldjoontes N-i suhtes ühendis Det N. Mis puutub klassi Adj, siis Adj₂ on peaaegu alati vaba, Adj₁ väga sageli ka seotud. On tuntud tõik, et ühendite Adj N seas on suhteliselt rohkem fraseoloogilisi ühendeid kui ühendite N Adj seas.

2. 2. Järgnev samm meie kirjelduses on juba ühendite kirjeldamine. Rühmitame ühendid selle järgi, missugused sõnaklassid on nende tuumadeks. Kõigepealt käsitleme nn. minimaalseid lausungeid, s. t. lausungeid, mis ükskõik millise elemendi eemaldamisel neist ei jää enam lausungiteks. Vaatleme, missuguste teiste sõnaklassidega on seotud mingi sõnaklass; see tähendabki, et kirjeldame minimaalset lausungit, milles see klass on tuumaks. Minimaalset lausungit võime teisiti nimetada ka lausungiks, mille kõik elemendid on seotud. Kuna muidugi seotus-vabadus ei ole alati üheselt ja selgelt piiriteldud (nagu eespool ühendi Pps V kohta näitasime), on ka minimaalne lausung mõneti laialivalguv üksus. Lausungi minimaalsus oleneb tihti mitte tema sõnaklassilisest struktuurist, vaid konkreetsest koostisest ja kasutamistingimustest, seega juba keelevälistest teguritest. Ometi on minimaalne lausung keeles reaalselt oluline ja tema käsitlemine ka sel-

lisel väheeksaktsel kujul selgitab meile mõndagi keele struktuurist. Minimaalsed lausungid on need "süntaktilised molekulid", millest moodustatakse keerulisemad ühendid, kuigi see moodustamine ei sünni ainuüksi "molekulaarsel tasandil": minimaalsete lausungite struktuur ja terviklikkus ei jää püsima. Kui tahame anda süntaksimudelit, kus minimaalsetest lausungitest saadakse kõik teised, keerukamad lausungid, tuleb arvestada seda, et lausungit moodustades minimaalsed lausungid muutuvad (il travaille - travaille) ja lausungisse lisandub elemente, mis ei ole minimaalsed lausungid ega vahel ka nende koostisosad (näiteks sidesõnad, meil enamasti märgitud Rel₁ ja Rel₂).

Nüüd esitame illustratsiooniks minimaalsed lausungid, mille tuumaks on N.

Klass N võib olla seotud klassidega Det, Ind, Part, Tot, Quant, N. Teatavalt juhtudel võib ta olla ka vaba, näiteks pealkirjades, siltidel, telegrammis. N on harilikult vaba, kui on tegemist pärisnimega. Selles suhtes jaguneb klass N kahte alamklassi, millest üks (üldnimed) võib alati esineda ümbruses Det -, teine (pärisnimed) võib peaaegu alati esineda vabalt. Erandiks on nimed nagu Le Havre, La Rochefoucauld, mida saab esimese alamklassi sõnadest eraldada teiste tunnuste järgi (Det asemel ei saa esineda Ind jne.).

N-i jaoks ainsuses võime anda sümbolvalemi

(Det $\approx$ Ind) $\rightarrow$ N.

N mitmuses saab olla seotud rohkemate sõnaklassidega. Sümbolvalem on siin oma kohmakuse tõttu väheillustatiivne, seepärast loetleme kõik N-ile seotult eelneda võivad klassid ja ühendid. Need on Det, Part, Ind Part, Quant, Det Quant.

Morfofoneemikasse kuuluvad juba reeglid selle kohta, et Part esineb iseseisvalt N-i ees kujul des, Ind järel aga tihti kujul de või nullkujul (un étudiant; des étudiants; un de nous; beaucoup d'étudiants).

Klassi N seotuse näiteid leiame uuemast keelest: mon frère il est dépanneur radio.

Siin jätame käsitlemata teised sõnaklassid, millest kõige keerulisemaid minimaalseid lausungeid saame klassi- ga V. Vähe ühendeid moodustab enamik teisi sõnaklasse, erandiks on klassile V mõneti lähedased Inf, Ppas, Ppres.

Esitades minimaalsed lausungid valemite või lihtsate sümbolitega, võime nende abil kirjeldada keerulisemaid ühendeid. Prantsuse keeles osutub aga segavaks asjaoluks see, et keerulisemaid ühendeid (resp. lausungeid) ei saa väga sageli kirjeldada lihtsalt minimaalsete lausungite järjendina vähemalt morfonoloogilisel tasandil. Kui tahame anda mudeli, kus minimaalsetest lausungitest moodustatakse kõik kõrgemate tasandite ühendid, tuleb arvestada kõige- pealt seda, et minimaalsed lausungid liitudes muutuvad. Nii annaksid mon frère ja il travaille kirjakeeles ühendi mon frère travaille, kuigi ka mon frère il travaille on juba võrdlemisi sage. Teine tülikas asjaolu on see, et lihtsa järjestuse asemel esineb ühendite moodustamisel inkorporatsiooni, üks minimaalne lausung (või mõni muu keeleüksus) kiildub teisesse, lahutab teise osadeks, mille arv prantsuse keeles (ja nähtavasti ka teistes keeltes) ei saa olla üle kahe. Nii saab lausungit il a beaucoup travaillé vaadelda minimaalsete lausungite il a travaillé ja beaucoup ühendina. Sellest tõigast ja tõigast, et kõiki lausungeid ei saa kaugeltki vaadelda koosnevana ainult minimaalsetest lausungitest (tuleb arvestada kasvõi sidesõnu), järeldub, et keele lausungeid ei saa kirjeldada minimaalsete lausungite lihtsate järjenditena ja ainult nendena. Ühend pour le garçon koosneb näiteks minimaalsest lausungist le garçon ja sõnast pour, mis ei saa kunagi esineda minimaalse lausungina. Seega ei vasta prantsuse keele minimaalsed lausungid näiteks ladina keele sõnadele, mis on vabad ja millest saame moodustada kõik ladina keele lausungid ilma seotud elementide abita (kui jätame kõrvale

paar morfeemi, nimelt -que, -ne ja -ve). Prantsuse keele struktuur on liiga erinev ladina keele või saksa keele omast, et temas võiks leida mingi selge analoogi nende keelte "sõnale".

Kui me kasutame minimaalseid lausungeid edaspidi alusena keerulisemate ühendite kirjeldamisel, ei saa N- ja V-tuumalisi minimaalseid lausungeid märkida ühe sümboliga, vaid tuleb säilitada nende sisemine struktuur. Kuna N-tuumalised minimaalsed lausungid on Det N, Ind N, Q N, Part N, N (⊗ N), mis on struktuurilt sarnased, võime N-tuumalist minimaalset lausungit märkida koondatult ΔN ; $\Delta = (((\text{Det} \approx \text{Ind}) \approx Q) \approx \text{Part}) \approx \otimes$).

V-tuumalisi ühendeid, kus verb on seotud sõnadega klassidest Pps, Pob, Ppsind, Ppas, Adj, võime kokku võtta valemities $\sqcap V, V \sqcap, \sqcap V_a \text{ Ppas}, V_a \sqcap \text{ Ppas}, \sqcap_1 V_a \sqcap_2 \text{ Ppas}, \sqcap V_e N, \sqcap V_e \text{ Adj}, V. \sqcap$ tähistab kõiki pronominaalseid elemente, ta võib seista kombinatsioonide Pps, Pps Pob, Pps Ppsind, Pps Ppsind Pob asemel. Mõnedes ühendites asub osa sellisest ühendist V ees, osa V järel (me l'a-t-il dit). Seda tähistame $\sqcap$ jagamisega osaks $\sqcap_1$, mis tähistab V ees asuvaid pronominaalseid elemente, ja $\sqcap_2$, mis tähistab V järel asuvaid pronominaalseid elemente. $\sqcap, \sqcap_1$ ega $\sqcap_2$ ei jagune enam osadeks, seega on meie märkimisviis sobiv just keerukamate ühendite puhul. Vajaduse korral võime $\sqcap$ asemel kasutada alti Pps, Ppsind, Pob või mõnda nende kombinatsiooni olenevalt ümbrustest, mis vahel teatavat kombinatsiooni ei võimalda. Nii ei saa $V_e \text{ Adj}$ ees seista Pob, küll aga Pps, Ppsind, Pps Pind (cela m'est égal).

Sarnaselt N- ja V-tuumalistele ühenditega, kuigi väiksemal määral, ei ole "kompaktsed" ka Inf-, Ppres- ja Ppas-tuumalised ühendid, millel kõigil on tugevaid analoogiaid V-tuumalistega. Võime kirjutada vastavalt $\sqcap \text{ Inf},$

$\sqcap \text{ Ppres}$ (Ppas on peaaegu alati vaba), kusjuures Inf-i ees ei saa esineda Pps, küll aga Pob või Ppsind või Pob

Ppsind, nagu ka Ppres ees. Tegelikult on V-, Inf-, Ppres- ja Ppas-tuumaliste ühendite vahel võimalik leida vastavusi, transformatsioone, kusjuures igale Inf-, Ppres- ja Ppas-tuumalisele ühendile vastab vähemalt üks V-tuumaline ühend (le lui dire; le lui disant - (il) le lui dit). Kuna Pps esimeses kahes ühendis ei saa esineda, saame alati ükspaljuse vastavuse (ülaltoodud vastavuse kolmandaks liikmeks võiks niisama hästi olla (je) le lui dis).

Kõik minimaalsed lausungid, mille tuumaks on mingi teine klass peale N, V, Inf, Ppas ja Ppres, ei moodusta üldjuhul minimaalseid lausungeid ühenduses teiste sõnaklassidega peale Prep (erand on Adv₁, Adj ja Pob voici). Kõik minimaalsed lausungid, mille esimeseks elemendiks on Prep, on eksotsentrilised ühendid (à lui; à rien; à cela). Erandiks on Adv₃ (ici; d'ici; par ici), kus on raske rääkida eksotsentrilisusest või endotsentrilisusest, kuna näiteks ici ja d'ici on ümbrustelt tunduvalt sarnasemad kui näiteks lui ja de lui. Siin kohtusime veel kord tööga, et adverbid ei oma selgelt piiriteldavaid tüüpilisi ümbrusi sellisel määral nagu teised sõnaklassid, mistõttu ei saa ka leida ümbruste poolest selget ja üldist vahet Adv₃ ja Prep Adv₃ vahel. Nii osutuvad prepositsioonid minimaalsete lausungite tasandil peamiseks elemendi väärtuse, funktsiooni (omaduse esineda teatavais ümbrustes, distributsiooni) muutjateks, elemendi või ühendi üleviijateks teistesse klassidesse.

2. 3. Et kirjeldada keerulisemaid ühendeid minimaalsetest lausungitest lähtudes, on otstarbekas kasutada kolme operatsiooni: substitutsiooni, aplikatsiooni ja inkorporatsiooni. Tuleb ka meeles pidada, et iga sõnaklassi asemel võib seista ^(H), välja arvatud klassid, mis koos moodustavad minimaalse lausungi, millest me antud ühendi oleme tuletanud. Kerge on ka näha, et sama lausung on moodustatav mitmel eri viisil, olenevalt sellest, milliseid operatsioone me tema moodustamisel kasutame.

Kui vajame rangemat kirjeldust, võime inkorporatsiooni asendada aplikatsiooni ja substitutsiooniga, kuna meie kasutada on klass ④. Aitab ka ainult substitutsioonist, kuid selline skeem ähmastaks liialt pilti nendest operatsioonidest, mida keele tarvitaja ise lausungite moodustamisel poolteadlikult tarvitab.

Illustratsiooniks vaatleme N-tuumalisi lausungeid.

Ühendist $\triangle N$ saame inkorporatsiooni abil

$\triangle \text{Adj } N$ (du bon vin; un grand garçon).

Adj asemele on omakorda substitueeritavad mitmesugused ühendid, millele ta on tuumaks, näit. Adv₁ Adj. Adj-klassi sõnu võib üksteise kõrval N-tuumalises ühendis esineda teoreetiliselt piiramata arvul, nii võime anda üldisema valemi

$\triangle \text{Adj}_{(1)} \text{Adj}_{(2)} \dots \text{Adj}_{(n)} N$.

Adj₂ saame lähtevalemile ($\triangle N$) aplitseerida:

$\triangle N \text{Adj}$, üldisemalt jällegi $\triangle N \text{Adj}_{(1)} \text{Adj}_{(2)} \dots \text{Adj}_{(m)}$

Kuna üks ei välista teist, võime kasutada kaht operatsiooni koos ja saame üldisema valemi:

$\triangle \text{Adj}_{1(1)} \text{Adj}_{1(2)} \dots \text{Adj}_{1(n)} N \text{Adj}_{2(1)} \text{Adj}_{2(2)} \dots \text{Adj}_{2(m)}$ (un grand garçon roux). Adjektiivide vahele võib enamasti asetada sõna klassist Rel₃ (sagedamini et), täpsemaid reegleid selle kohta me siin andma ei hakka.

Ühendeid saame samuti Prep $\triangle N$ või $\triangle N$, ka Prep Pton, Prep Inf jt. Prep-ga algavate ühendite aplitseerimise teel.

$\triangle N \triangle N$ (appareil photo; dépanneur radio).

$\triangle N \text{Prep } \triangle N$ (un maison de repos; un livre).

$\triangle N \text{Prep Pton}$ (ce livre à moi; un problème pour toi).

$\triangle N \text{Prep Pdem}$ (un vie sans cela; la solution de celle-ci).

$\triangle N \text{Prep Inf}$ (la machine à écrire; permission d'imprimer).

$\triangle N \text{Prep Adv}_3$ (l'avion en loin).

Neisse ühendeisse on omakorda võimalik inkorporeerida adjektiive. Saame valemid nagu

$\triangle$ Adj N Prep ... (la belle dame sans merci; un bon garçon sans le sou; un beau livre pour enfants).

Adj₂ aplikatsioon on haruldane (la boîte aux lettres bleue; la machine à écrire allemande). Valemid on

$\triangle$ N Prep N Adj₂.

$\triangle$ N Prep Inf Adj₂.

Sellised ühendid ei sisalda tavaliselt enam klassi Adj₁.

Adj₂ kohale on substitueeritavad mitmesugused ühendid, mille tuumaks on Ppas, Ppres, V.

$\triangle$ N Ppas (un garçon travaillant).

Ppas-tuumaline ühend on substitueeritud (või apliteeritud) lausungis la demoiselle s'admirant au miroir.

$\triangle$ N Sub V ja $\triangle$ N Sub V Adj jne. (le garçon qui travaille; les oiseaux qui volent très très haut).

On loomulik, et neis ühendeis võime $\triangle$ N asemel kasutada näiteks ühendit $\triangle$ Adj N Adj või $\triangle$ Adj N Prep N ja teisi (un joli garçon roux qui travaille dans le jardin; les grands oiseaux de proie qui volent très très haut).

Selliste ühendite tasandil on võimatu anda ammendavat loendit võimalikest ühenditest, sest nende tüüpe on lõpmata palju.

Kuna N-tuumalised ühendid (arvestame siin ainult endotsentrilisi ühendeid) on ümbrustelt põhiliselt ekvi-valentsed, märgime vajaduse korral N-tuumalisi ühendeid sümboliga *N*. Lepime kokku, et märgime ka teisi teatavatuumalisi ühendeid tuumaklassi märgiga kursiivis: *V* - V-tuumaline ühend, *Inf* - Inf-tuumaline ühend jne.

Kõikide mitmuslike ühendite *N* ees saab seista Quant, Part, Quant Part, Det Quant, Q Quant (mitte alati), Tot, Tot Det Quant. Ainsuses ei saa *N* ees esineda Quant, ka Tot kasutamine on piiratud. Sel viisil saadud ühendid on järelegi N-tuumalised ja me võime neid märkida sümboliga

$\mathcal{N}$. Muidugi on neil ühenditel ka mõned erinevused distributsioonis. Nimelt ei ole näiteks $\mathcal{N}$ ja Tot $\mathcal{N}$ ekvivalentsed ümbruste Tot - ja Tot Det Quant - suhtes.

Kõrvuti saavad teoreetiliselt piiramata arvul esineda ainult ühendid $\mathcal{N}$, Adj, V, Inf, Pres, Ppres, Adv₂, voici, vähemal määral Pton, teised adverbid ja arvsõnad. See viitab nende klasside suuremale vabadusele, kuna kõik teised klassid on seotud kindlamate keeleliste tingimustega ja saavad esineda ainult seoses vabamate klassidega. Ka tegelikult vaba klass Adv₅ ei ole siin samaväärne klassidega V või $\mathcal{N}$, sest mingi tekst võib koosneda ainult ühenditest V ja $\mathcal{N}$ (tavalises tekstis on ülekaalus ühendid V), ei saa aga koosneda ainult ühenditest Adv₅, mis peavad üksteisest eraldatud olema mingite teiste ühenditega.

Ühendite $\mathcal{N}$ järjestikune aplikatsioon annab jällegi ühendi $\mathcal{N}$ (le général de Gaulle; Arthur Rimbaud, le plus génial de tous les poètes), kui üksikute $\mathcal{N}$ -ide vahel ei ole punkte või punktiga "võrdvõimsaid" kirjavahemärke.

Kuna siin ei ole võimalik üksikasjalisemalt kirjeldada teisi ühendeid, millest kõige tähtsamad ja kõige keerukama struktuuriga on V-tuumalised, toome ära mõned tähtsamad tulemused ja kokkuvõtted.

V-tuumalistes ühendites saavad esineda kõik sõnaklassid, mis üheski teises ühenditüübis ei ole võimalik. Võime rääkida mingi klassi otsesest ja kaudsest sõltuvusest V-st. Otseselt on V-st sõltuv klass, mis saab esineda tuuma V otsese ümbrusena. Otseselt võivad V-st sõltuda $\mathcal{N}$, Adj, Pps, Inf, Ppres, Ppb, Ppsind, Pdem, Pind, Pton, adverbid (peale Adv₅), Sub. Otseselt ei saa V-ga seotud olla Det, Prep, Adv₅ Tot, Ind, Part ja mõned teised. Kuid kõik need klassid (peale Adv₅) võivad olla seotud N-iga, mis seob nad kaudselt V-ga. Selline kaudne seotus (kui elemendid ühendis on vabad, on parem rääkida sõltuvusest) võib konkreetsetel juhtudel olla mitmekordne:

tegemist on V ümbrise ümbrise, ümbrise ümbrise ümbrisega jne. Adv_5 (oui, non) sõltuvus V -st on teist laadi. Adv_5 esineb ülekaalukalt iseseisva lausungina, tema esinemus aga sõltub $\checkmark$ esinemusest. Seega on tegemist juba $\checkmark$ ja Adv_5 secsega, mis kuulub ilmselt kõrgemale tasandile kui eespool toodud seosed $\checkmark$ üksikute elementide vahel.

2. 4. Kui me vaatleme kõigi sõnaklasside või kõigi endotsentriliste ühendite hulki nende elementide vaheliste tuum-ümbris-vahekordade seisukohalt, võime tulemused esitada osalise järjestusena, kus $a \leq b$ tähendab, et b on a tuum või vastavalt, et a on b ümbris. Järjestus on osaline, sest kõik elemendid ei ole omavahel võrreldavad (a ei pruugi olla b tuum ega ümbris). Mõned klassid (Adv_5 , que₃₃ jt.) ei ole võrreldavad ühegi teise klassiga; võime neid ignoreerida ja rääkida järjestusest ülejäänud elementide hulga suhtes. Kõigi elementide puhul kehtib $a \leq v$ (v - kas V või $\checkmark$), ühegi elemendi puhul ei kehti vastupidist seost $v \leq a$. Seega võib elementi v nimetada antud hulga maksimaalseks elemendiks.

Nähtavasti vastab keelereaslsusele paremini veidi erinev matemaatiline mudel, kus ei ole nõutav transitiivsus, s. o. kui on $a \leq b \leq c$, siis ei pea olema $a \leq c$. Sellise mudeli leiame häiteks graafiteooriast. Siin ei ole transitiivsus obligatoorne, kuigi see on lubatav. Elemendi (siin graafi tipu) v ja teiste elementide (resp. tippude) suhtes kehtib, et kui ei ole $a \leq v$, siis on kindlasti $a \leq b \leq \dots \leq k \leq v$, mis lingvistiliselt tähendab vastavalt kas otsest või kaudset sõltuvust V -st.

Ülalvisandatud mudelitel oleks tegelikku väärtust siis, kui osutuks, et graafiteooria teoreemil on rohkem kui triviaalset lingvistilist sisu. Tavaline on küll olukord, et matemaatilise teooria tulemused on liiga üldised, et keele kohta midagi olulist ütelda, keele kirjeldused aga on liiga konkreetsed, et neid tasuks matemaatiliselt sõnastada.

2. 5. Nagu oleme juba vihjanud, on suure osa adverbide ülesanne süntaksis erinev teiste sõnaklasside omast selles. Erandiks on Adv₁ ja Adv₂, millest Adv₁ esineb põhiliselt Adj ees, andes ühendi Adv₁ Adj. Adv₂ võib samuti esineda Adj ees, aga ka seotult V-ga (il le dit poliment; il travaille assidûment). Ka Adv₃ ja Adv₄ on selges sõltuvuses V-st, kuid nende positsioon on väga kõikuv. Seega saame küll otsustada, kas antud ühendis saab esineda Adv₃ või Adv₄, me ei saa aga üheselt fikseerida nende positsiooni, mis teiste sõnaklasside puhul on võimalik. Veel selgemini kehtib see Adv₅ ja Adv₆ kohta. Tihti saab anda ainuüksi mõned positsioonid, kus Adv₅ ja Adv₆ esineda ei saa, mitte vastupidi. Mõni positsioonides on neil sõnadel sageli erinev tähendusvarjund, stiilinüanss, mida senised käsitlused ei ole suutnud selgelt esitada. Seetõttu peab neid nüansse esialgu lihtsalt ignoreerima. Võib küll väita, et teatavas positsioonis on näit. Adv₅ kõige tavalisem, sagedasem, sobivam jms., see aga ei välista teisi positsioone.

Siit tuleneb, et ei ole otstarbekas fikseerida Adv₃, Adv₄, Adv₅ ja Adv₆ ümbrusi nii, nagu tegime teiste sõnaklassidega. Ka muudavad adverbid keeruliseks näiteks V-tuumaliste ühendite kirjeldamise, kus meil tuleb anda nii palju erinevaid valemteid, kui palju on mingil adverbil võimalikke positsioone antud ühendis. See muudaks kirjelduse ebaproportsionaalseks. Ka varjutaks selline märkimine seda lingvistiliselt olulist tõika, et adverbide ja nende ümbruste vahel pole üksühest vastavust. Selle kõige parandamiseks peaksime kasutama kõrgemat kirjeldustasandit. Kuna me eespool lähtusime elemendi positsioonist, seega tunnistades sõnajärje relevantsust, võime nüüd siirduda kirjeldusviisile, mis ei kirjelda enam elementide järjekorda, vaid ainult nende sõltuvusi (a on b ümbris jne.). Neid sõltuvusi võime märkida nagu eespool: $a \leq b$, või kasutada sümbolilist loogikat, kirjutades näiteks Adv₅ → V (Adv₅ ei saa esineda ilma V-ta). Adverbide positsiooni

täpne fikseerimine kuulub juba morfofoneemikasse (või "süntaktofoneemikasse").

Täheldame veel, et kuigi Adv₃ -, Adv₄- ja Adv₅-tuumalisi ühendeid peaaegu ei esine, korvab seda puudust ühendite rohkus, mille esimeseks elemendiks on Prep ja teiseks *N*, *Adj*, *Ppres* või ka Adv₃, Adv₄, Pton. Sellised ühendid on tihti ümbrustelt peaaegu ekvivalentsed mõne adverbiga (il l'expliquait énergiquement - il l'expliquait avec énergie, - mõlemas figureerib ühine tüvi, mis võimaldaks vaadelda sellist seost üksühese vastavusena (resp. transformatsioonina)).

Adv₃-le vastavad sageli ühendid Prep *N* (mitte kunagi Prep *Inf*) (il est à la maison; je l'ai rencontré hier dans la rue).

Adv₄-le vastavad ühendid Prep *N*, *Ppres*, en *Ppres* (je serai de retour dans une heure; je l'ai rencontré en achetant des journaux).

On võimalik anda reegleid selliste ühendite positsioonide kohta, kuid selleks tuleb arvestada vaevalt formaliseeritavaid subjektiivseid faktoreid. Meie võime ainult öelda, et sellised adverbiaalsed ühendid sõltuvad ühendist *V* ning ei saa esineda ilma selleta. Milline aga on adverbiaalse ühendi täpne positsioon, selle määrab suures osas ühendi enda ja ühendi *V* konkreetne iseloom.

2. 6. Endotsentrilised V-tuumalised ühendid ei saa kunagi olla ümbriseks mingile teisele ühendile või sõnale. Küll aga võivad seda olla eksotsentrilised V-tuumalised ühendid, mis on saadud endotsentrilistest: Rel₁ *V*, Rel₂ *V* Sub *V*. Rel₁ *V* ja Rel₂ *V* on ümbristeks V-tuumalises ühendis (je ne sais pas s'il arrive aujourd'hui; s'il pleût aujourd'hui, je resterai chez moi). Ühendid Rel₂ *V* on vabamas sõltuvuses ühendist, mille ümbriseks nad on, nende positsioone ei saa tavaliselt üheselt määrata ja sellega sarnanevad nad adverbidele. Traditsioonilises grammatikas peetakse ühendit Rel₁ *V* kõrvallauseks, ühendit

Rel₂ V aga üldiselt üheks rindlauset moodustavatest osalausetest (kui Rel₂ on näiteks esindatud sõnadega parce que, car jt.). See viitab samuti teist tüüpi ühendite suuremale vabadusele, nende kuulumisele kõrgemale tasandile kui on üendid Rel₁ V. Ühend Sub V kuulub ümbrisena N, P_{ton} ja P_{demp} juurde (le garçon qui croyait au ciel; toi qui croyais au ciel; celui qui croyait au ciel). Traditsiooniliselt loetakse ka neid ühendeid kõrvallause-teks.

Meil ei ole põhjust järgida traditsiooni ja rääkida pealausetest, kõrvallausetest, põim- ja rindlausetest. Sama saab järjekindlamalt kirjeldada mitmesuguste ühenditüüpide ja seoste iseloomu abil nendes.

Osa sõnu ja ühendeid sõltub alati otseselt mõnest teisest sama taseme ühendist, sõna sõnast, teatavatuumaline ühend teisest ühendist. Mõned sõnad ja üendid sõltuvad aga kõrgema tasandi ühenditest, nagu nägime Adv₅ puhul. Viimaks võime rääkida ühenditest, mis selgelt ei sõltu ühestki teisest ühendist. Nii on üksikud üendid V (vähe-
mal määral N, P_{ton} jt.) tekstis üksteisest sõltumatud, üks ühend V ei ole teisele ümbriseks ja ühe esinemusest ei saa me otsustada midagi teise esinemuse kohta, kui me ei arvesta sisulisi kriteeriume. Loomulikult võime ühe sellise sõltumatu ühendi põhjal teatava tõenäosusega oletada, milliseid teisi ühendeid või sõnu võiks leida antud ühendi naabruses. See aga ei ole enam seoses keele struktuuriga, vähemalt on suur vahe mõlemas lähenemisviisis. Võime öelda, et strukturealne analüüs lõpeb siis, kui oleme saanud üksused, mille disjunktsioonina on kirjeldatav keele igasugune tekst. "Grammatiliselt õigete" ühendite vaheline seos on disjunktsioon. Siiski ei saa ühtegi mõtestatud teksti ega ka kõigi mõtestatud tekstide hulka vaadelda mingite grammatiliselt õigete ühendite disjunktsioonina: tekstis, kus esineb lause See, et Maa on ümmargune, on õige ei saa (peaaegu kunagi) esineda lause

See, et Maa on ümmargune, ei ole õige. Siit näeme, et ühendite vahel tekstis on enamasti tugevamad seosed kui need, mida saab väljendada disjunktsiooniga. Kõige tugevamad on seosed, kui iga ühendi esinemusest oleneb otsestelt iga teise ühendi esinemus. Selliseid seoseid tuleks kirjeldada konjunktsiooni abil. Tekst, kus kehtivad ainult sellised seosed, oleks vastavate ühendite konjunktsioon. Enamik tekste ei ole ei täielikult disjunktiivsed ega konjunktiivsed. Peatumata pikemalt probleemidel, mis sellisel lähenemisel sünnivad, võime märkida ainult seda, et senini on ühendite sõltuvuse uurimisega tekstides tegelnud peaaesjalikult loogika. Tekstides, millega tegeleb loogika, on ühenditeks tõeväärtust omavad laused ja nende esinemust konkreetset määravateks tingimusteks ainult tõeväärtus ja struktuur. On ilmne, et näiteks luuletuse analüüsimisel selline lähenemisviis midagi ei selgitaks: luuletuses on tegemist ühenditega, mille tunnuseks tõeväärtus vaevast arvesse tuleb ja mille esinemistingimused on nähtavasti keerukamad.

2. 7. Mõned märkused sõnade järjekorrast. Sõnade järjekord on prantsuse keeles üldiselt kinnistatud, kui jätame arvestamata adverbid, mille positsioonide vähesest määratusest oli eespool juttu. Sõnajärje kinnistatusega käib sageli kaasas see, et sõnade järjekord ei tulene sõnade endi iseloomust. Me ei saa ilma teatava lisainformatsioonita taastada lausungit tema koosseisu kuuluvatest sõnadest, sest eksisteerib üksipaljune vastavus sõnade hulkade ja lausungite vahel, mis nendest hulkadest saab moodustada. Nii sõnad André, parle, à, Pierre annavad kaks lausungit: André parle à Pierre ja Pierre parle à André. Arvestamata jällegi adverbe, leiame, et vastavus muutub endotsentriliste ühendite tasandil üksüheseks, kui me eristame ühendite klassides sama ühendi esinemuse $V (U)$ ees ja V järel, kirjutades vastavalt $\mathcal{N}_1$ ja $\mathcal{N}_2$ ($\mathcal{N}_1$ vastab üldiselt subjektile, $\mathcal{N}_2$ objektile või predika-

tiivile), P_{ton_1} ja P_{ton_2} , Inf_1 ja Inf_2 , P_{dem_1} ja P_{dem_2} jne. Kui meil on antud ühendite klassikuuluvus, saame nendest üheselt moodustada lausungi (kui see antud ühendite hulga puhul üldse võimalik on). Sama võime läbi viia endotsentrilistes ühendites enestes, kus vajalikul viisil indeksitega varustatud sõnadest saab moodustada parajasti ühe ühendi. Kui me sel kombel oleme vabanenud vajadusest valemi üleskirjutuses järgida sõnajärge, sest uued sõnaklassid märgivad seda ise, võime juba vabamalt kasutada sümbolilist tähistusviisi. Ühendi $\Delta \text{Adj}_1 N \text{Adj}_2$ võime kirjutada kujul $(\Delta \rightarrow N) \& (\text{Adj} \rightarrow N) \& (\text{Adj}_2 \rightarrow N)$, kus implikatsioonide teine liige N osutab ühendi tuumale. Selliseid valemeid võime vajaduse korral teisendada, lihtsustada, sest saame alati siirduda implikatsioonide konjunktsioonile, nagu on ülaltoodud valem. See hõlbustaks struktuurivalemite automaatset analüüsi. Kui tahame näiteks kirjeldada graafina V -tuumalisi ühendeid, tuleb meil klass $\mathcal{N}$ jagada kaheks klassiks $\mathcal{N}_1$ ja $\mathcal{N}_2$. Vastavalt tuleb talitada ka kõigi teiste klassidega, mis saavad esineda nii V -st eespool kui tagapool. Ainult nii saame graafi, mis kirjeldab põhiliselt võimalikke ühendeid. Kui me ei erista $\mathcal{N}_1 \mathcal{N}_2$ -st, saame muuhulgas valemi $Pps V Prep \Delta N V Prep \Delta N V \dots$ (näiteks il parle à mon père parle à mon père parle...).

Ilmneb, et me võime kõiki lausungeid kirjeldada konjunktsiooni abil, mille liikmeteks on implikatsiooni, disjunktsiooni või mõne muu loogilise tehtega seotud paarid kõigist lausungi sõnadest (üldisemalt: ühendi elementidest). Sellisest kirjeldusest ilmneksid selgelt ühendi elementide vaheliste seoste tüübid ja tugevus. Üldisemaid seoseid keele struktuuris saame kirjeldada tabeli abil, mille read ja veerud vastavad mingi loendi elementidele ja nende lõikekohal asub loogilise tehte märk, mis näitab vastava kahe elemendi seose tüüpi. Juhul kui ühe elemendi esinemine ei sõltu teise esinemisest, on meie mudelis tegemist

samaselt tõese valemiga. Ei ole mõeldamatu kasutada tabelis loogiliste tehete asemel korrelatsioonikordajaid, mis muidugi eeldab elementide seoste statistilist uurimist. Lihtsas loogilises mudelis väljendavad relevantseid seoseid konjunktsioonid, implikatsioonid, mitteekvivalents ja disjunktsioon.

3. Sellega lõpetame ülevaate ühest meetodist prantsuse keele süntaksi kirjeldamisel. Töös on püütud võimalikult lühidalt esile tuua olulisemaid momente, osutada tekkivale probleemidele ja arendusvõimalustele, mida on väga rohkesti. Nende probleemide lahendamine ja kirjeldusvõimaluste ammendamine mingites mõistlikes piirides peaksid andma ülevaate prantsuse keele süntaktilistest konstruktsioonidest. Meie meetod või meetodid peaksid olema vajalike täienduste ja parandustega rakendatavad iga keele puhul, nii et neil võiks olla teatav tähendus tüpoloogia jaoks. Ruumpuudusel on mitmed küsimused töös täiesti nimetamata jäänud. Kuid olles seotud käsitelduga, kerkiwad nad ühel või teisel kujul paratamatult esile siin viisandatu põhjalikumal läbitöötamisel.

3. 1. Märkime lõpuks, et töö kirjutamisel kasutasime põhiliselt kolme raamatut:

Ш. Балли, *Общая лингвистика и вопросы французского языка*. М., 1955 (originaal: Ch. Bally, *Linguistique générale et linguistique française*. Berne, 1950).

Л. Ельмслев, *Прологомены к теории языка. Новое в лингвистике*, вып.1-М.,1960, lk. 264-389 (originaal: L. Hjelmslev, *Omkring Sprogteoriens Grundlaeggelse*. København, 1943; inglise k.: *Prolegomena to a Theory of Language*. Baltimore, 1953).

Z. S. Harris, *Structural Linguistics*. Chicago, 1960.

EESTI KEELE KIRJAPILDI ENTROOPIAST

JA LIIASUSEST

P. J ä r v e

O. Käesolevas töös interpreteeritakse informatsiooni-teooria mõisteid "entroopia" ja "liiasus" keeleteaduse seisukohalt ning tuuakse esialgseid andmeid eesti keele kirjapildi liiasuse kohta.

1. Entroopia on üks informatsiooniteooria põhimõisteid. Tema sisu on lühidalt järgmine.

1. 1. Olgu meil mingi katse α , mis võib lõppeda n erineval viisil tõenäosustega vastavalt $p_1, p_2, \dots, p_n$ (kus $(p_1 + p_2 + \dots + p_n = 1)$). Olles kindlaks teinud, millise tulemuse andis katse, saame saadud informatsiooni hulka avaldada arvuga

$$H(\alpha) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_n \log_2 p_n.$$

Suurust $H(\alpha)$ nimetatakse katse α entroopiaks ja ta mõõdab selle katse määramatust. Saab näidata, et $H(\alpha) = 0$ siis ja ainult siis, kui üks tõenäosustest $p_1, p_2, \dots, p_n$ võrdub ühega ja ülejäänud nulliga, s. t. kui katse tulemus on meile ette teada, siis määramatust pole. Teistel juhtudel $H(\alpha) > 0$. Kõige sagedamini mõõdetakse määramatuse (ja informatsiooni) suurust bittides (binary digit). 1 bit on sellise katse määramatus, millel on kaks võrdse tõenäosusega tulemust, näiteks metallraha langeb pärast ülesviskamist võrdse tõenäosusega ühele oma külgedest.

1. 2. Esinegu meil kõrvuti katsega α veel katse β .

On võimalik, et katse α ja katse β tulemused pole kuidagi seotud. Siis öeldakse, et α ja β on sõltumatud ja et β tulemus ei anna mingit informatsiooni α kohta.

Sageli aga muudab katse β tulemuse teadmine katse α ühe või teise tulemuse tõenäosust ja selle kaudu ka entroopiat $H(\alpha)$. Nii avaldab mõne tekstis esineva tähe teadmine mõju temale järgnevate tähtede esinemise tõenäosustele. Selliste α ja β kohta öeldakse, et nad on sõltuvad. Oletame, et teostasime palju kordi katset β ning saime suurusele $H(\alpha)$ iga kord erineva väärtuse. Leiame $H(\alpha)$ keskvärtuse, tähistame selle $H_\beta(\alpha)$ ja nimetame ta katse α tinglikuks entroopiaks tingimusel β .

1. 3. Sõltumatute α ja β puhul on ilmne, et $H_\beta(\alpha) = H(\alpha)$. Kui α ja β on sõltuvad, siis $H_\beta(\alpha) < H(\alpha)$, s. t. katse β eelnev toimumine vähendab katse α entroopiat.

2. Olgu meil rida omavahel sõltuvaid katseid $\alpha_0, \alpha_1, \alpha_2, \dots$ (näiteks mingi arv katseid määrata tekstis üksteisele järgnevaid tähti vms.). Saadud reaga kõrvutame kahaneva rea vastavatest tinglikest entroopiaist: $H(\alpha_0) \geq H_{\alpha_1}(\alpha_0) \geq H_{\alpha_2 \alpha_1}(\alpha_0) \geq H_{\alpha_3 \alpha_2 \alpha_1}(\alpha_0) \dots$ Kui katsete rida on lõpmatu, siis kahanevate positiivsete arvude $H_0 = H(\alpha_0)$, $H_1 = H_{\alpha_1}(\alpha_0)$, $H_2 = H_{\alpha_2 \alpha_1}(\alpha_0)$, $\dots$ $H_n = H_{\alpha_n \alpha_{n-1} \dots \alpha_1}(\alpha_0) \dots$ piiramatu rida läheneb tingimata mingile piirväärtusele $H_\infty = H$. Keelelise teksti puhul üldiselt $H \neq 0$.¹

3. Ehitame nüüd mõned eesti keele fraaside mudelid, järk-järgult üha rohkem eesti keeles valitsevaid statisti-

¹ Matemaatilise aparaaadi kohta vt. lähemalt И.М. Яглом, Р.Л. Добрушин, А.М. Яглом, Теория информации и лингвистика. Вопросы языковедения 1960, № 1.

lisi seaduspärasusi arvestades. Kasutades eesti tähestikku + nulltähte (sõnavähe), kirjutame algul tähti täiesti juhuslikult (F_0), teisel juhul arvestame üksikute tähtede esinemise tõenäosust (F_1)², siis tähepaaride esinemise tõenäosust (F_2), siis kolmikute tõenäosust (F_3). Võib jõuda järgmiste "fraasideni".

F_0 : ELKUÖRDVIMPKAEKS LUMTIUMNEANTRD IFADOTOS

F_1 : SÖDACR UAAELMSED ORAFEISSNLE NGKAMAM

F_2 : SISERDAMEGA MIBUL TA LID NEIIJUD KA ONNDAGINE

F_3 : TAS PEANUD LESTEMATE KÖNNIIGLA KALTSED ASEG

3. 1. Toodud "fraasidest" nähtub, et iga järgmine on eestikeelsele mõtestatud fraasile lähemal kui temale eelnenud "fraas". Samuti võime öelda, et näiteks fraasi F_3 määramatuse aste on madalam kui F_0 oma. Viimase saame arvutada valemist

$$H_0 = H(\alpha_0) = -\frac{1}{n} \log \frac{1}{n} - \frac{1}{n} \log \frac{1}{n} \dots - \frac{1}{n} \log \frac{1}{n} = -\log \frac{1}{n} = \log n,$$

kus n on tähestiku tähtede arv. Meil seega $H_0 = \log_2 34 = 5,292$ bitti. Analooogiliselt saame leida ka teiste fraaside entroopiad (F_1 puhul näiteks $H_1 = 4,1$ bitti. Võrdluseks: H_1 on saksa keeles 4,1; prantsuse keeles 3,96; hispaania keeles 3,98; inglise keeles 4,03; vene keeles 4,35 bitti). On võimalik välja arvutada mõtestatud teksti ühes tähes sisalduv informatsioon ehk piirentroopia H_∞ , mida loetakse keele tähtsaimaks statistiliseks karakteristikumiks. Arvutustööde suure mahu tõttu on H_∞ täpne väärtus eesti keele kohta jäänud käesolevas töös leidmata.

3. 2. Praktikaks loetakse keele statistilise struktuuri iseloomustamisel mugavamaks suhet $\frac{H_\infty}{H_0}$, kusjuures

2 Osaliselt on tähtede esinemissagedused eesti keeles ära toodud töös A. Saareste, Die Estnische Sprache. Tartu, 1932. Puuduvad andmed leidis autor ligikaudse täpsusega katseliselt.

reaalsete keelte puhul $0 < \frac{H_{\infty}}{H_0} < 1$.

3. 3. C. E. Shannon vaatles vahet $R = 1 - \frac{H_{\infty}}{H_0}$,

mida ta nimetas liiasuseks. Kuigi rahuldavalt täpseid liiasuse ega ka entroopia väärtusi pole saadud ühegi konkreetse keele kohta, on üldist tunnustust leidnud oletus, et kõikide Euroopa keelte liiasus R ületab märgatavalt 50%.

3. 4. Mõistest keele liiasus ei tule aru saada nii, et $R = 50\%$ puhul võib poole mõnest raamatust ära visata, ilma et selle raamatu loetavus selle all kannataks. Liiasuse mõte seisab selles, et teatud erilisel viisil saab $R\%$ tähtedest tekstist ära jätta või asendada, nii et ülejäänud tähtedele $(100-R)\%$ ja keele statistilistele seaduspärasustele toetudes on võimalik taastada veatult kogu tekst. Veel võib öelda, et kui küllaltki ulatuslikus tekstis puudub üle $R\%$ tähtedest, siis seda teksti ei saa mitte mingil juhul taastada.

4. Eesti keele kirjapildi liiasuse ligikaudseks määramiseks kasutas autor C. E. Shannon'i meetodit³, mis seisab lühidalt alljärgnevas.

4. 1. Kasutame fakti, et iga eesti keelt vabalt valdav inimese nn. keelevaistul on suur hulk keele statistiliste seaduspärasuste kohta käivaid andmeid, mida see inimene kasutab automaatselt oma keelelises käitumises. Võtame katkendi eestikeelsest, kuid katseisikule tundmatust tekstist ja laseme tal mõistatada katkendi esimest tähte. Kui katseisik oletas kohe õigesti, teatame talle sellest ja laseme tal mõistatada järgmist tähte. Kui oletus esimese tähe puhul oli vale, siis teatame talle õige tähe

3 Vt. K. Шеннон, Работы по теории информации и кибернетике. Москва, 1963, lk. 673-674.

ja läheme edasi järgmise tähe mõistatamisele. Sõnavahet loeme täheks. Toome siinkohal taolise eksperimendi tüüpilise tulemuse. 1-ga märgitud ridades on algtekst, 2-ga märgitud ridades on õigesti arvatud tähe kohal kriips, valesti oletatud tähe kohal on algteksti täht välja kirjutatud.

1) SELLE PASUNA ÄHVARDAVAT HÄÄLT KUULDES TULEB

INIMSOOLE

2) S-L-E-PASU---ÄHV-----AT-----UU-DES-TU---INI-S-----

1) ÜLETAMATU HIRM PEALE INIMESED PÕGENEVAD OMA VENDADE

2) ÜL-T -----HIR-----I-----ÖG-----OM-----

4. 2. 106 tähest on 72 tähte ehk 68% oletatud õigesti. Vead esinesid kõige sagedamini seal, kus katseisik seisis suure valiku ees (kus määramatus oli suur), nagu sõna ja silbi algused, käändelõpud jne. Read (2) sisaldevad katseisiku seisukohalt niisama palju informatsiooni kui read (1). Katseisikuga matemaatiliselt identne teisik taastaks ridade (2) põhjal eksinatult algteksti. See annabki aluse interpreteerida 68% kui konkreetse katkendi liiasust konkreetse katseisiku seisukohalt.

5. Toodud ilukirjandusliku katkendi keskmiseks liiasuseks viie katseisiku puhul sai autor 62%. Samadele katseisikutele (neist üks oli ülikooli juba lõpetanud, ülejäänud õppisid veel ülikoolis) esitati mõistatamiseks 211-täheline tekst ajalehes ilmunud informatsioonilisest teadaandest. Esimese katkendiga võrreldes oli siin liiasus 13-20% suurem. Keskmiseks kujunes 77,6%. Kolmele katseiskule esitati 2 erinevat luulekatkendit. Esimene neist oli järgmine katkend loodusluuletusest.

Ju metsas mättad paiguti
on pohlavartest haljad
ja künka küljed laiguti
ju lumekattest paljad.

Keskmiseks liiasuseks tuli 68% maksimaalse kõrvalekaldumisega keskmisest 1,7%.

Teine oli järgmine katkend ühest vabavärsilisest luuletusest.

Su hambad on maikella kõlisev kaar täis vallatult
valendavaid õisi

ja naer on peoleo laul tiritamme kasvataval
toomingal metsajões.

Siin oli resultaat veelgi stabiilsem, andes keskmiseks liiasuseks 55,7% maksimaalse hälbe 0,8%.

6. Eespool nägime, et entroopia all mõistetakse keeleteaduses statistilist parameetrit, mis mõõdab teksti ühe tähe kohta tulevat informatsiooni. Selguse mõttes liisame, et tegemist pole sisulise (semantilise) informatsiooniga, mille mõõtmiseks tänapäev rahuldavat meetodit veel ei paku, vaid keele puhtstatistilise struktuuri parameetriga.

6. 1. Liiasus on kitsenduste mõõt, mida keele omapära esitab tema statistilisele struktuurile. Tänu keele liiasusele on võimalik leida trükivigu, lugeda vigast ja ebaselget teksti, mõista halvasti kuuldud lauset jne. Mudelis F_0 aga pole ühegi vea parandamine võimalik, sest seal $R = 0\%$.

6. 2. Käesolevas töös toodud andmeid eesti keele kirjpildi liiasuse kohta tuleb vaadelda kui edasist täpsustamist vajavat materjali. Samal ajal aga näitavad saadud tulemused, et mõiste "eesti keele kirjpildi liiasus" muutub C. E. Shannon'i meetodit kasutades ebasoovitavalt laiaks keeles eksisteerivate eri stiilide tagajärjel. Saadud keskmiste liiasuste suurim vahe on üle 20%. Sama probleem peaks kerkima ka teiste keelte puhul.

6. 2. 1. Täiesti väär oleks seada liiasus kriteeriumiks luule hindamisel - põhimõttel, et mida väiksem liiasus, seda rohkem informatsiooni ja järelikult seda parem

luuletus. Siis oleks maailma parim luuletus mudel F_0 , kus liiasust üldse pole.

7. Kaasajal, mil keeleteaduse saavutused leiavad üha rohkem rakendamist tehnikas, osutuvad andmed mistahes keele entroopia ja liiasuse kohta hädavajalikeks näiteks sidekanalite projekteerimisel, ökonoomsete koodide koostamisel ja mujal. Eesti keele kohta tuleb kõigepealt elektronarvutil koostada sagedustabelid üksikute tähtede, tähepaaride ja -kolmikute esinemise kohta ning võimalikult täpselt välja arvutada entroopia ja liiasuse väärtused eesti keele kirjapildi jaoks.

7. 1. Käesolevas töös mindi mööda eestikeelse kõne entroopiast ja liiasusest, mis on omaette lai uurimisobjekt.

S I S U K O R D

	Lk.
1. E. U u s p õ l d, Lause struktuur ja vahetute moodustajate meetod	3
2. H. Õ i m, <u>tulema</u> , <u>saama</u> ja <u>pidama</u> tähenduste strukturealne analüüs .	27
3. M. E r e l t, Elatiivi funktsioonide transformatoorsest analüüsist	46
4. J. K a p l i n s k i, Prantsuse keele süntaktilise struktuuri kirjeldamisest	67
5. P. J ä r v e, Eesti keele kirja- pildi entroopiast ja liiasusest	97

Тартуский государственный университет
ЭССР, г. Тарту, ул. Юликооли, 18
ЯЗЫК И СТРУКТУРА
Труды по структуральной и математической
лингвистике
На эстонском языке

Vastutav toimetaja H. Rätsep
Korrektor E. Uuspõld

=====

TRÜ rotaprint 1965. Trükipoognaid 6,5. Tingtrüki-
poognaid 6. Arvestuspoognaid 5,1. Trükiarv 300.
Paljundamisele antud 3.VI 1965. MB 04773.

Tell. nr. 203.

Hind 35 kop.