

# **SUOMEN TEKSTILAUSEIDEN PIIRTEITÄ: KVANTITATIIVINEN TUTKIMUS**

**Auli HAKULINEN - Fred KARLSSON - Maria VILKUNA**

Department of General Linguistics  
University of Helsinki

PUBLICATIONS  
NO. 6  
2010 (1996, 1980)

ESIPUHE.....	3
ESIPUHE 2010.....	3
1. JOHDANTO .....	4
1.1. Tausta ja tavoitteet.....	4
1.2. Muuttujien valinta .....	5
1.3. Koodituksen luotettavuudesta .....	8
2. MUUTTUJAT JA NIIDEN SOVELTAMINEN TEKSTIEN ANALYYSIIN .....	10
2.1. Yleistä.....	10
2.2. Tekniset muuttujat .....	11
2.3. Virkekonstruktio­muuttujat .....	12
2.4. Lauseen perussyntaksia ja -semantiikkaa koskevat muuttujat .....	17
2.5. Konstituenttien rakenteita koskevat muuttujat.....	34
3. LAUSEOPIN KVANTITATIIVISTA METODIIKKA .....	91
3.1. Aineiston valinnan ja otannan periaatteet .....	91
3.2. Syntaktisen otoksen koon määrittäminen .....	95
3.3. Juoksevien frekvenssien "jäätymispisteet" .....	96
3.4. Alatekstilajien eroista .....	102
4. MUUTTUJIEN FREKVENSSIT JA NIIDEN KORRELAATIOITA .....	108
4.1. Lauseen perussyntaksia ja -semantiikkaa koskevat muuttujat .....	108
4.2. Konstituenttien rakenteita koskevat muuttujat.....	113
4.3. Lause- ja sanamääriä koskevat muuttujat .....	126
4.4. Sanajärjestykseen vaikuttavista tekijöistä .....	135
4.5. Tekstuaalisten tekijöiden yhteispeli .....	145
KIRJALLISUUTTA .....	156
LIITE 1: Kooditussuunnitelma .....	160
LIITE 2: Analysoitujen tekstien luettelo.....	171
LIITE 5: Taulukkoluet­telo .....	176

## ESIPUHE

Tämä kirjanen on Suomen Akatemian rahoittaman "Suomen kielen kvantitatiivista tekstisyntaksia" -nimisen projektin loppuraportti. Keskeisiä osia siitä on julkaistu myös Auli Hakulisen ja Fred Karlssonin artikkelissa "Finnish Syntax in Text: Methodology and Some Results of a Quantitative Study" (*Nordic Journal of Linguistics* 3:2, 1980, 93-129).

Auli Hakulinen on kirjoittanut jaksot 4.1., 4.4.-5. sekä yhdessä Fred Karlssonin kanssa luvun 1 ja yhdessä Maria Vilkun kanssa luvun 2. Fred Karlsson on kirjoittanut luvun 3, jaksot 4.2.-3. sekä yhdessä Auli Hakulisen kanssa luvun 1. Maria Vilkuna on yhdessä Auli Hakulisen kanssa kirjoittanut luvun 2. Projektin suunnittelu ja toteutus on tehty yhteistyönä. Auli Hakulinen on ollut päävastuussa muuttujien (kooditusjärjestelmän) laadinnasta, Maria Vilkuna tekstien analyysistä ja Fred Karlsson tietokoneajoihin liittyvistä järjestelyistä.

Kiitämme Jussi Salmelaa ja erityisesti Viljo Kohosta, jotka ovat monin tavoin auttaneet meitä projektin teoreettisessa ja teknisessä toteutuksessa. Tietokoneajot perustuvat ratkaisevasti Kohosen ja Salmelan kehittämään CHITAB-ohjelmaan. Olemme kiitollisia myös Valtion humanistiselle toimikunnalle, jonka rahoittamana projektin perusosa on voitu toteuttaa, sekä Turun yliopistosäätiölle, jonka myöntämän avustuksen turvin loppuraportti on julkaistu.

Jyväskylässä, Turussa ja Helsingissä huhtikuussa 1980

Tekijät

## ESIPUHE 2010

Tämän kirjan teksti on skannattu tiedostoksi 1990-luvulla ja viimeistelty maaliskuussa 2010. Alkuperäiseen kolmenkymmenen vuoden takaiseen käsikirjoitukseen on tehty vain pieniä teknisiä ja oikoluvun taseisia muutoksia. Pois on jätetty sivun 102 kuvio 1 sekä liitteet 3 (kuva reikäkortista) ja 4 (esimerkki ristiintaulukoinnista). Typografia on lukijaystävällisempi. HKV-korpus kokonaisuudessaan eli tekstit ja niiden analyysit (koodit) on myös saatettu tiedostomuotoon, koodit reikäkorteista lukemalla vuonna 1987 ja tekstit skannaamalla sekä käsin kirjoittamalla 1980- ja 1990-luvuilla. Kooditiedostoon olen korjannut muutamia kymmeniä havaitsemiani kooditusvirheitä; lauseiden lukumäärä on nykyään 10140. Maaliskuussa 2010 ryhdyin rakentamaan vuorovaikutteista kyselyliitettä, jonka avulla korpuksesta voi suorittaa hakuja koodeilla ja niiden yhdistelmillä: HKV-korpus herää siis uudestaan henkiin. Kyselyliitettä tulee yleiseen käyttöön CSC:n tietokoneilla. Kyselyliitännän dokumentaatio ja käyttöohje tulevat valmistuttuaan löytymään osoitteesta [www.ling.helsinki.fi/~fkarlss/hkvdoc.pdf](http://www.ling.helsinki.fi/~fkarlss/hkvdoc.pdf). Alkuperäisiin muuttujiin ja myös kooditiedostoon olen lisännyt muuttujan nro 64 "tekstityyppi", jonka arvot ovat luvussa 3.1 esiteltävät tekstityypit 1-6. Tämä mahdollistaa hakujen kohdistamisen myös tekstityypeittäin.

Helsingissä maaliskuussa 2010

Fred Karlsson

# 1. JOHDANTO

## 1.1. Tausta ja tavoitteet

Tässä raportissa selostetaan vuonna 1975 aloitettua suomen kielen tekstilauseiden kvantitatiivista tutkimusta. Tavoitteena oli alun perin vertailla suomen ja ruotsin (sekä eräiltä osin englannin) keskeisimpiä tekstisyntaktisia ja -semanttisia piirteitä samanlaisissa tyyleissä. Suomi ja ruotsihan poikkeavat toisistaan sanajärjestykseltään paljon, ja halusimme selvittää, miten tämä näkyy lauseista, kun ne ovat tekstin osina. Projektit eriytyivät kuitenkin sittemmin, ja suomen kieltä varten tarkoitettua kooditusjärjestelmää muokattiin alkuperäisestä jonkin verran edelleen. Ruotsin kielen osuudesta vrt. Andersson & Östman (1978), englannin osuudesta Enkvist & von Wright (1978).

Kysymykset, joihin halusimme vastausta, olivat seuraavantapaisia: Millä tavoin nominaalisen lausekkeen sijaintiin lauseessa vaikuttaa sen tarkoitteen tunnettuus, sen tekstuaalinen sidonnaisuus ja sen lauseenjäsenyys? Mikä näistä "voittaa" tai on ratkaisevin? Mitkä seikat vaikuttavat eniten esim. objektin topikaalistamiseen? Onko suomessa päälauseilmiöitä, joita sivulauseista ei lainkaan tavata? Miten konstituentin kompleksisuus ja pituus puolestaan vaikuttavat sen paikkaan lauseessa?

Laatimamme *kooditusjärjestelmä* eli muuttujien joukko (vrt. liitettä 1 ja sen tarkempaa esittelyä luvussa 2) kokoa myös tietoa monista perussyntaksiin kuuluvista lauseiden piirteistä. Näitä ovat esim. subjektin, objektin, predikaatiivin ja adverbiaalien suhteellinen taajuus, lauseiden ja virkkeiden keskimääräinen pituus sanoina ja konstituentteina, sivulauseiden sijoittuminen virkkeissä, jne. Tekstilingvistiikkaa ei tutkimuksemme ole sikäli, että kokonaisten tekstien analyysiin ei ole pyritty. Tarkastelun perusyksikkönä on lause, joskin aineisto on koottu kokonaisista teksteistä eikä satunnaisotannalla poimituista katkelmista.

Kooditusjärjestelmän ensimmäinen versio käsitti 66 syntaktista ja tekstuaalista muuttujaa, jotka oli valittu siten, että ne soveltuisivat sekä ruotsin että suomen kuvaamiseen. Tätä luonnosta testattiin aineistolla, jonka muodostivat 600 lausetta Astrid Lindgrenin teoksesta *Veljeni Leijonamieli* ja Bo Carpelanin teoksesta *Kaari*; kummastakin oli käännökset suomeksi ja englanniksi.

Esitutkimuksesta saatujen kokemusten perusteella muuttujistoa muokattiin useiden välivaiheiden kautta paremmin nimenomaan suomen kielen analyysin vaatimuksia vastaavaksi. Kooditusjärjestelmän lopullista muunnosta on nyt käytetty 123 tekstin analysoimiseen. Aineistossa oli 10149 lausetta (5016 ortografista virkettä, 66851 sanetta). Yksitoista kuukautta kestäneen kooditustyön eli kaikkien tekstien kaikkien lauseiden analyysin käyttöön otettujen 63 muuttujan mukaan teki Maria Vilkuna, joka toimi suomenkielisen projektin tutkimusapulaisena.

Lauseopin tutkimus on yleensä kvalitatiivista, ja kvantitatiivista tutkimusta joka pyrkii vastaamaan sellaisiin kysymyksiin kuten jonkin kategorian tai rakenteen suhteellinen taajuus tai marginaalisuus, on usein pidetty lauseopin kannalta irrelevanttina - se kuuluu tämän näkemyksen mukaan pikemminkin tyylintutkimukseen. Ei liene silti epäilystäkään siitä, että

naiivillakin kielenpuhujalla on ainakin jonkinlainen "tilastollinen kompetenssi" eli intuitiivinen tuntuma siihen, mitkä kategoriat ja rakenteet ovat yleisiä, mitkä eivät.

Lisäksi on kieliopin ilmiöitä, jotka eivät ole puhtaasti kvalitatiivisia (diskreettejä) vaan muodostavat kvantitatiivisia jatkumoit. Hyvä esimerkki on lausetason sanajärjestys suomessa. Siihenhän on samanaikaisesti vaikuttamassa varsin moni tekijä, eikä voida aina varmasti sanoa, mikä seikka "aiheuttaa" kussakin yksittäistapauksessa esiintyvän sanajärjestyksen. Ei voida toisin sanoen laatia yksinkertaista säännöstöä, joka määrittäisi vaihtoehtoisten ilmausten rakenteelliset ehdot. Pelkkään intuition turvautuminen ei myöskään riitä osoittamaan miten sanajärjestys kulloinkin aktuaalistuu. Adekvaattiin kuvaukseen yltää näin ollen vain edustavasta ja luotettavasta otoksesta tehty kvantitatiivinen analyysi. On samalla kuitenkin syytä pitää mielessä, että kvantitatiivisen analyysin pätevyys riippuu ratkaisevasti siitä *teoreettisesta analyysistä*, jolle tilastointi perustuu.

Päätavoittemme oli tutkia mahdollisimman *neutraalin kirjoitetun asiaproosan* lauseiden keskeisiä piirteitä: sekä perussyntaksia että etenkin niitä ominaisuuksia, jotka ovat tekstistä riippuvaisia. Tätä tarkoitusta varten valitsimme aineistoksi tiedottavaa, ei-fiktiivistä kieltä, jota käytetään esim. lehtiartikkeleissa ja pääkirjoituksissa (ks. tarkemmin lukua 3). Yleistäminen yhdestä tai muutamasta tekstilajista kielen yhteisainekseen (common core) ei toki ole ongelmatonta. Saadaksemme jonkinlaisen tuntuman siihen, missä määrin aineistomme on tyyllillisesti heterogeenista, olemme vertailleet muutamia asiaproosan osatyylejä toisiinsa useiden muuttujien frekvenssierojen tilastollisen merkitsevyyden suhteen (ks. jaksoa 3.4.).

## 1.2. Muuttujien valinta

Lause lauseelta etenevä analyysi laadittiin 63 *muuttujan* kooditusjärjestelmäksi (liite 1). Jokainen muuttuja koostuu *alakategorioista*, joita muuttujaa kohden yleensä on 2-15. Kukin lause kooditettiin joukoksi alfanumeerisiä merkkejä (koodeja), jollaisena aineisto syötettiin tietokoneeseen. ATK:n avulla laskettiin sitten mm. muuttujien alakategorioiden taajuudet ja todettujen erojen merkitsevyydet (yhden muuttujan sisällä yhtä hyvin kuin tyylien välilläkin). Myös koodituksen ja lävistyksen tarkistusajot on suoritettu automaattisesti, samaten useat sadat ristiintaulukoinnit (esimerkki liitteessä 4), joiden avulla muuttujien väliset korrelaatiot voidaan paljastaa. Tällainen toimitus on taloudellinen, koska se säästää lävistysaikaa, kortteja ja muistitilaa. Abstraktit symbolit on toisaalta tällä tavoin erotettu varsinaisesta korpuksesta (esimerkki liitteessä 3), eikä "vastauksiksi" koskaan saa itse esimerkkejä suoraan. Tulokset on tulkittava kooditusjärjestelmän avulla ja esimerkit on haettava viitteiden mukaan alkuperäisaineistosta. Käytetyn CHITAB-tietokoneohjelman ovat laatineet ja sitä lähemmin selostaneet Salmela & Kohonen (1977); vrt. myös Kohonen & Salmela (1978:20-7).

Tällaisen tutkimuksen ensimmäinen ja päättehtävä on tutkittavien muuttujien valitseminen. Haluttujen syntaktisten piirteiden valinta, niiden muuntaminen riittävän moneksi täsmälliseksi, keskenään vaihtoehtoiseksi alakategoriaksi, osoittautui työn aikaavievimmäksi vaiheeksi. Mitkä olisivat vaikkapa subjektin konstituenttirakenteen vaihtoehdot? Niitähän voisi olla kymmeniä, jos tehdään hienojakoista erittelyä, tai vain kolme, jos jaetaan karkeasti (subjekti puuttuu; subjekti on yksisanainen; subjekti on pitempi kuin yksisanainen). Lähdimme liikkeelle saatavilla olevasta suomen lauserakenteita koskevasta tiedosta sekä joistakin yleisperiaatteista, ja päädyimme kooditusjärjestelmän lopulliseen versioon lukuisten yritysten, erehdysten ja

kokeilujen kautta. Tämä oli empiiristä työtä sanan varsinaisessa merkityksessä: alustavien hypoteesien asettamista, niiden testaamista pienellä aineistolla, luopumista hypoteeseista osittain tai kokonaan, uusien kokeilemista jne. Vaikka lopulliseen järjestelmään sisältyy eräitä ongelmia, sillä voi suhteellisen hyvin analysoida ja jäsentää aineiston siihen tarkoitukseen, jonka itsellemme alussa asetimme.

Teleman (1974) on esitellyt viisi aluetta, joita Lundin laajoissa Talsyntax- ja Skrivsyntax -projekteissa on pyritty tutkimaan. Nämä ovat (i) *kompleksisuus*, joka on määritelty mm. virkkeen lausemääräksi, lauseen sanamääräksi ja sanan morfeemimääräksi; (ii) *topologia*, jolla tarkoitetaan konstituenttien sijaintia lauseessa - tässä selvitetään peruslausetyyppejä ja aräitä sanajärjestysilmiöitä; (iii) *spesifisyys* eli ilmauksen täsmällisyys; tämä ominaisuus näkyy etenkin nominaalilausekkeen määriteosan rakenteesta; (iv) *modaalisuus*, johon kuuluvat mm. puhujan asenteiden ja arviointien ilmaisimet; sekä (v) *tekstuaaliset sidokset* eli koheesiota ilmentävät mekanismit. Näiden luokkien tarkempaa esittelyä sekä tuloksia löytyy teoksista Jörgensen (1970), Westman (1974) ja Einarsson (1978).

Mainittuja alueita voidaan pitää tekstisyntaksin keskeisimpinä, ja useimmat niistä on meidänkin kooditusjärjestelmässämme otettu huomioon. Sen avulla saa tietoja seuraavilta neljältä alueelta: (i) erirakenteisten lauseiden ja konstituenttien suhteelliset taajuudet juoksevassa tekstissä. Tähän kuuluvat tiedot eri lauseenjäsenten yleisyydestä sekä se, millä tavoin kompleksisuudeltaan erirakenteiset konstituentit jakautuvat lauseessa. Koko lausetta koskevia rakenteellisia tietoja saadaan mm. muuttujista (15: lauserakenne) sekä (18: pintajärjestys). (ii) virkkeen kokonaisrakenne ja lauseen virkeasema. Muuttujat (4) sekä (6-8) antavat tietoja tavallisimmista virkekonstruktiosta, sivulausetyyppien suhteellisesta taajuudesta ja sijoittumisesta virkkeessä. Tällainen tieto on paitsi kiinnostavaa sinänsä, myös taustatietoa, jota vasten erilaiset yksittäislauseen sisällä vallitsevat temaattiset ja muut tekstuaaliset suhteet saavat lisävalaisua.

Kolmatta aluetta voisi luonnehtia funktionaaliseksi lauseperspektiiviksi. Järjestelmässä on joukko muuttujia, joita edellä mainittujen muuttujien kanssa ristiintaulukoimalla saadaan kokonaiskuva niistä tekijöistä, jotka vaikuttavat tekstilauseen sanajärjestykseen. Viidestä nominaalisesta konstruktiosta, nimittäin subjektista, objektista, predikatiivista sekä alku- ja loppuadverbiaalista kerätään paitsi rakenteelliset tiedot, myös tieto siitä, onko niissä mainittu tarkoitettu vai uusi ja minkälainen mahdollinen sidoskeino konstituentissa on. Voimme selvittää, kuinka tekstuaalinen tunnettuus ja/tai sidonnaisuus vaikuttaa konstituentin sijoittumiseen lauseessa ja samalla pitää silmällä sitä, missä määrin eri ratkaisuja säätelee myös lauseen virkeasema. On ajateltava, että sanajärjestykseen on vaikuttamassa kaksi rinnakkaista, osittain ehkä päällekkäistäkin tekijää: tässä mainittu funktionaalinen eli tekstiperspektiivi ja toisaalta kunkin konstituentin rakenteellinen kompleksisuus ja pituus. Tässä työssä emme ole pystyneet näitä kahta tekijää sovittamaan toisiinsa, vaan esitämme tulokset rinnakkaisina.

Lähtiessämme liikkeelle meillä oli se näkemys, että lauseen teemaa ei voi yksiselitteisesti tunnistaa, sen sijaan alkukonstituentin tekstuaalisen sidonnaisuuden ja tunnettuuden voi. Nykyperspektiivistä tuota lähtökohtaa voinee pitää liiallisena varovaisuutena; tosin teeman löytäminen olisi edellyttänyt nimenomaan kokonaisten tekstien analysoimista aivan kuten lauseen tekstifunktionkin määrittäminen (ks. seur.). Tämän oivaltaminen on tapahtunut vasta työn kuluessa, ja niinpä teeman käsite puuttuukin muuttujien luettelosta.

Neljänneksi olemme keränneet tietoja lauseiden modaalis-pragmaattisista ominaisuuksista: onko lauseessa syntaktisesti merkittyä (9) taikka leksikaalis-morfologista modaalisuutta (10), onko se kielteinen vai myönteinen (14) onko siinä pragmaattisia partikkeleita (13) ja vihdoin mikä tekstifunktio (tekstiakti) on kyseessä (11). Viimemainittu on näistä muuttujista vähiten onnistunut, mikä johtuu siitä, että yksittäisen lauseen tekstiakti on määritettävissä oikeastaan vasta kun koko teksti on ensin analysoitu esim. makrostruktuureihin, nämä puolestaan edelleen virkkeisiin. Koska emme ole kuten mainittua analysoineet kokonaistekstejä, on oletettavissa, että myös yksittäisten lauseiden tekstifunktion määrittäminen on jäänyt liian subjektiiviseksi eikä sikäli täysin luotettavaksi.

Relevanttien muuttujien valinnan on perustuttava teoreettisiin ratkaisuihin ja selviin kriteereihin. Yksi perusoletuksistamme oli, että yksittäisten sanojen morfologiset (etenkin johto- tai taivutus)piirteet ovat yksittäisen lauseen sisäisiä ilmiöitä, joilla ei ole juurikaan tekemistä lauseen tekstuaalisten kytkentöjen kanssa. Pidimme tekstisyntaktisesti keskeisinä yksikköinä *lausetta* ja *lauseketta*. Tärkeä on luonnollisesti myös *virke*.

Tiedot taltioitiin teksteistä käyttäen lausekortteja, sillä virke olisi ollut koodituksessa liian karkea, lauseke puolestaan liian hienojakoinen kooditusyksikkö. Lähestymistapamme poikkeaa tässä suhteessa muista vastaavista, vrt. Telemann (1974), Westman (1974), Ikola & Karjalainen (1977) ja Ellegård (1978). Kokonaan ei morfologiaa ole kuitenkaan syytä sivuuttaa. Olemme ottaneet huomioon liitepartikkelit (13), subjektin, objektin ja predikatiivin sijaitavuuden (20,28,34) sekä persoonaverbin kongruenssin (26), jotka liittyvät mm. tunnettuuteen sekä kvantifiointiin.

Esitutkimuksen muuttujat oli suurelta osalta laadittu teoreettisten pohdiskelujen pohjalta. Kokeiluvaiheessa testaus osoitti joiden muuttujien soveltumattomuuden, joko kokonaan tai osittain. Tällaisia olivat esim. verbiryhmän semantiikka, pituus sekä tekstuaaliset ominaisuudet, jotka oli pakko jättää tämän kartoituksen ulkopuolelle. Suppeasti otettuna ne eivät olleet hyödyllisiä, laajennettuna ne olisivat paisuttaneet järjestelmää liiaksi. Joissakin muuttujissa oli tarpeettomia alakategorioita, esim. teoreettisesti mahdollisia transformaatioita kuten lohkeamat, ns. TOUGH-siirto ja nosto, joita ei kirjoitetussa suomessa kuitenkaan tullut vastaan. Toiset muuttujat oli uudelleen arvioitaessa jaettava kahtia tai kolmiakin. Näin kävi tekstuaalisten muuttujien kohdalla. Aluksi arveltiin, että kustakin nominaalisesta jäsenestä riittää, kun merkitään muistiin siinä käytetty sidoskeino. Kävi ilmi, että tällainen muuttuja on liiaksi kuormitettu, ja se on jaettava kahtia siten, että toisessa muuttujassa otetaan huomioon tarkoitteen luonne (tunnettuus) ja toisessa sidosmekanismin morfosyntaktinen laji, jotka ovat toisistaan riippumattomia ilmiöitä. Myöhemmin oli käyvä ilmi, että tämäkään jaottelu ei ollut riittävän hienosyinen (ks. jaksoa 2.7.).

Perinteisen kieliopin nominaaliset päälauseenjäsenet, subjekti ja objekti, eivät ole suomeen sovellettuina ongelmattomia; vrt. Hakulinen & Karlsson (1979: 163-6, 174-81). Eräät marginaaliset subjektit eivät esim. ole monenkaan kriteerin perusteella kovin subjektimaisia, ja objektilla on tietyissä tapauksissa (subjektittomissa lauseissa) puolestaan subjektin piirteitä (nominatiivimuoto, lauseenalkuisuus). Arvelimme, että mm. sanajärjestysmuuttujan (18) avulla paljastuu monia lisätekiäjiä, jotka esim. saavat lauseenloppuisen subjektin muistuttamaan käyttäytymiseltään enemmän objektia kuin lauseenalkuista subjektia niin, että ehkä tulokset osaltaan osoittaisivat lauseenjäsenyyksen nykyisen "valta-aseman" yksipuolisuutta pyrittäessä ymmärtämään lauseen rakennetta.

Yksi alue, jota tämä tutkimus ei pyri eikä voikaan kovin seikkaperäisesti selvittää, on adverbiaalien käyttäytyminen tekstilauseissa. Välttääksemme kooditusjärjestelmämme ylenmääräisen paisumisen jouduimme rajoittamaan tarkastelumme vain kahteen adverbiaaliin lausetta kohden: lauseenalkuiseen ja -loppuiseen. Näiden arveltiin olevan tekstuaalisesti tärkeimpiä. Kokonaiskuva siitä, millä tavoin ja kuinka usein adverbiaalit sijoittuvat esim. verbiryhmän osien väliin taikka verbin ja objektin väliin sekä siitä, millä tavoin ns. liikkuvat adverbiaalit eli kommentti- ja kehysadverbiaalit käyttäytyvät minkintyyppisessä lauseessa, jää tässä yhteydessä täysin käsittelemättä.

Koodituksessa ei rajoitettu pelkkien ilmipiirteiden kirjaamiseen, vaan joukko muuttujia edellytti lauseiden transformaatiohistorian ottamista huomioon. Lähdimme nimittäin siitä, että lauseen perussanajärjestys on SVOA n/1 ja jos jokin tekstilause tästä poikkeaa, siihen on sovellettu topikaalistusta (54), subjektin lykkäystä (56), verbin siirtoa (57) taikka dislokaatioita (58). Nämä muuttujat tavallaan täydentävät pintajärjestysmuuttujaa (18) mahdollistaessaan muiden kuin tavallisimpien perussanajärjestysten tarkan kuvauksen.

Edelleen lähdimme ajatuksesta, että peruslauseessa on kaikki konstituentit läsnä ja jos tekstilauseesta jokin sen merkityksen analyysin kannalta tärkeä konstituentti puuttuu, se on poistettu ellipsisäännöllä (59). Ellipsisit on lisäksi kooditettu myös konstituenttirakennetta kuvaaviin muuttujiin, esim. (19,27).

### *1.3. Koodituksen luotettavuudesta*

Kooditusjärjestelmässä olevien muuttujien määrää voidaan pitää niin suurena, että tekstilauseiden tulkitseminen niiden mukaan käy liiankin hitaaksi. Lisäksi monet muuttujat edellyttävät koodittajalta teoreettista tietoa, sillä aineiston jaottelemisen oikeisiin alakategorioihin ei ole suinkaan mikään mekaanisesti ratkaistavissa oleva asia; erityisen selvästi näin on tekstuaalisten muuttujien kohdalla. Sen ratkaiseminen voi yksittäistapauksissa olla hyvinkin vaikeaa tai jopa mahdotontakin, onko jonkin konstituentin tarkoite tuttu vai ei, ja jos on, minkälaisesta tuttuudesta on kyse. Koodituksen luotettavuus vähenee sitä mukaa kuin järjestelmä mutkistuu ja käytettyjen käsitteiden teoriapitoisuus kasvaa, mutta ellei kooditettava aineisto ole ylettömän laaja (vrt. lukua 3), kooditusvaiheesta ei välttämättä tule kohtuuttoman aikaaviepä ja vaativa.

Tietokoneistetun työn yksi vaikeus on ilmaistu tunnetussa periaatteessa "garbage in - garbage out": konekäsittelyn tulokseksi ei saada kvalitatiivisesti erilaista, vaan ainoastaan eri tavoin järjestettyä ja käsiteltyä tietoa kuin liikkeelle lähdetessä on. Tästä syystä tutkittavien kategorioiden valinta ja seulominen sekä niiden teoriapohjan tarkka kartoitus on koko tiedon keruun ratkaisevin vaihe.

Kooditustyöhön sisältyy joka tapauksessa monia mahdollisia virhelähteitä, ja näitä on pyritty eri tavoin välttämään. Koodituksen luotettavuutta on testattu pääasiassa automaattisesti, mutta hankalimpien muuttujien kohdalla on aluksi tehty useita tarkistuksia siten, että kukin kolmesta työryhmän jäsenestä kävi läpi samat tekstit tietyn muuttujan kannalta ja tuloksia vertailtiin; ristiriitaisten ratkaisujen tapauksessa neuvoteltiin soveltuvimmasta tulkinnasta. Näin päästiin hitaasti riittäväksi arvioituun intersubjektiviisuuteen. On silti korostettava, että aineiston

lopullisen analyysin suoritti tutkimusapulainen yksin, ja tekstilauseiden tulkinnan oikeellisuus on hänen työnsä tarkkuuden ja johdonmukaisuuden varassa.

Tietokoneen avulla tehdyt tarkistusajot olivat seuraavat (ks. tarkemmin Kohonen & Salmela (1978:22): (i) kunkin muuttujan *ala- ja ylärajojen tarkistukset*, jolla pyritään lähinnä karsimaan lävistysvirheet; jos muuttujassa on esim. 7 alakategoriaa, tätä korkeampi alfanumeerinen arvo on ko. muuttujan kohdalla virheellinen; (ii) *haarukointiajot*, joilla alustavasti tarkistetaan, miten koodiarvot jakautuvat kussakin muuttujassa. Ristiintaulukoimalla vierekkäiset muuttujat keskenään ja vertailemalla eri sarakkeista saatujen määrien summia eri tavoin, voidaan paikantaa virheitä. (iii) "*mahdottomat*" *taulukoinnit*, joissa ristiintaulukoidaan sellaisia muuttujia keskenään, joilla ei ole mitään yhteistä ja joista ei pitäisi löytyä yhtään tapausta. Kaikki löydökset ovat tässä tapauksessa virheellisiä. Kun esim. subjektin konstituenttirakenne (19) ja sanajärjestys (18) ristiintaulukoidaan, ei saa löytyä yhtään subjektin esiintymää pelkistä V-sarakkeista, jotka per definitionem ovat subjektittomia. Tällä tavoin on laadittu 41 tarkistustaulukkoa. (iv) *yhdenmukaisuustarkistukset*, joissa verrataan samaa kategoriaa käsittelevien muuttujien keräämien tapausten määrää toisiinsa. Jos lukumäärät poikkeavat toisistaan, kyseessä on tavallisesti kooditusvirhe. Niinpä esim. objektikonstituenttia käsittelee sangen moni muuttuja, ja niissä kaikissa pitäisi olla sama loppusumma. Tällaisia tarkistusajoja on tehty 9, ja niillä on tarkistettu yhteensä 44 muuttujakombinaation kooditukset.

Automaattisen virheiden etsinnän jälkeen korjaus on luonnollisesti tehtävä käsin, mutta on huomautettava, että vastaavanlaajuista tarkistusta ei voisi ajatellakaan käsin tehtäväksi. Tarkistusajot paljastivat melko runsaasti (2-3 %) virheitä ensimmäisten 6000 lauseen koodituksesta, minkä jälkeen virhemäärä laski ratkaisevasti. Jäljelle jääneet kooditus- ja lävistysvirheet ovat ilmeisesti siksi harvalukuisia, että niiden merkitys on mitätön (vähemmän kuin 1 %). Ainoa jäljelle jäävä merkittävä virhelähde on siis perusratkaisujen tasolla: esim. jokin muuttuja on alun perin jaettu alakategorioihin tavalla joka ei ole mahdollisimman havainnollinen, taikka on tehty systemaattisia virheratkaisuja esim. määrätessä, mikä on lause tai ei, mikä on ellipsi tai ei, mikä on tuttu tarkoite tai ei, jne.

Aineisto osoittautui riittävän laajaksi melkein kaikkien muuttujien taajuuksien selvittelyä varten (vrt. lukua 3). Ristiintaulukointeja tehtäessä kävi kuitenkin usein ilmi, että oli alakategorioita, joista oli kertynyt riittämättömästi esiintymiä, minkä johdosta kaikkia tarpeellisia  $\chi^2$ -testejä ei voitu suorittaa ryhmittelemättä alakategorioita osittain uudestaan.

## 2. MUUTTUJAT JA NIIDEN SOVELTAMINEN TEKSTIEN ANALYYSIIN

### 2.1. Yleistä

Tässä luvussa esitellään kooditussuunnitelma kohta kohdalta. Pidämme tällaista menettelyä tärkeänä kahdestakin syystä. Ensinnäkin vastaavanlaisia tekstin käsittelyyn liittyviä tutkimuksia on toistaiseksi tehty vähän, ja sen vuoksi tällainen työ voi koitua muille samantyyppistä työtä suunnitteleville avuksi. Todellista hyötyä konkreettiseen kooditustyöhön ei kuitenkaan voi saada pelkistä tuloksista sinänsä, vaan on nähtävä ikään kuin kädestä pitäen, miten itse kooditustyö tehdään ja miten yksittäiset esimerkkitaapaukset ratkaistaan. Toiseksi voimme kutakin muuttujaa esitellessämme huomauttaa myös sellaisesta, minkä nyt jälkeen päin ymmärrämme toisin.

Eräisiin vaikeuksiin, joiden parissa jouduimme analyysivaiheessa ponnistelemaan, on selviä syitä. Tärkein näistä on se, että kvalitatiivista perustutkimusta oli tuolloin käytettävissämme liian vähän. Tämä koskee sekä puhdasta lauseopillista tutkimusta että etenkin tekstilingvististä, konkreettisia tekstejä käsittelevää tutkimusta. Vaikka korjasimme esitutkimusvaiheessa juuri tekstuaalisia muuttujiamme kaikkein eniten, on tällä alueella vielä paljon ajateltavaa. Yksi tärkeä kooditustyön "sivutuote" onkin Maria Vilkunan tutkimus (1980), joka olennaisesti selvittää sellaisia viittaussuhteisiin liittyviä kysymyksiä, joita ei juuri ollut tekstilingvistiikassa syvemmin pohdittu 1975, jolloin projektimme aloitettiin. Toinen tekstilingvistinen heikkous on se, että emme tuolloin onnistuneet kehittämään mitään sellaista menetelmää, joka olisi analysoinut kokonaiset tekstit, ja sitten tekstin kokonaisuudesta käsin koettanut selvittää yksittäisten virkkeiden tai lauseiden tekstuaalista funktiota. Lähdimme tekstifunktiossa (muuttuja 11) liikkeelle yksittäisestä lauseesta, mikä oli liian atomistinen lähtökohta. Tässä on tulevalle tekstilingvistiikalle laaja tutkimuskenttä.

Se 63 muuttujan lopullinen kooditussuunnitelma, jota tässä esittelemme, on parin vuoden kehittelyn tulos, ja tämän järjestelmän mukaisesti on koko aineistomme (10149 lausetta) kooditettu siten, että kukin lause on saanut oman korttinsa. Esim. Lauseopin arkiston koodituksesta poiketen emme siis ole pitäneet lähtökohtana sanaa, vaan *lausetta*.

Muuttujat on numeroitu juoksevasti samoin kuin liitteenä 1 olevassa systeemissä (1-63), ja kunkin muuttujan nimike ja sisältö on erikseen mainittu esittelyn alussa. Muuttujissa on vaihteleva määrä *alakategorioita*, jotka yhdessä kattavat tarkasteltavan, muuttujaksi kooditetun kategorian alan. Yksinkertaisimmillaan muuttuja voi olla kaksijakoinen, esim. lause on joko kielteinen tai myönteinen. Useimmissa muuttujissa on alakategorioita tätä enemmän. Tätä varten ohjelmassa on mahdollisuuksia siten, että voidaan käyttää juoksevia numeroita 1-9, tämän jälkeen kirjaimia A,B,-H. Paljon alakategorioita on tarvittu etenkin niissä muuttujissa, joissa on kyse lukumäärästä: sanojen määrästä lauseessa tai lausekkeessa. Alakategorioiden olennainen piirre on, että ne ovat toisensa poissulkevia. Niinpä lause on tyypiltään joko päälause tai sivulause, aktiivi- tai passiivimuotoinen, ei koskaan molempia. Eniten työtä

esitutkimusvaiheessa vaadittiin siihen, että alakategoriat olisivat toimivia, ettei niitä olisi liikaa tai liian vähän ja että ne antaisivat vastauksia mielekkäisiin kysymyksiin. Alakategorioiden järjestys yksittäisessä muuttujassa ei välttämättä heijasta mitään näiden kategorioiden keskinäistä hierarkkisuutta. Pyrkimystä kuitenkin on ollut siihen suuntaan, että taajafrekvenssisimmät tulisivat ensiksi ja harvinaiset lopuksi. Tähän ei ole kaikkialla päästy.

Yleiskuvan saamiseksi useiden yksittäisten muuttujien kohdalla täytyy ottaa alakategoria *irrelevantti* (i) mukaan yhdeksi vaihtoehdoksi. Tämä tarvitaan, jotta saadaan muuttujaan oikea kokonaissumma. Esimerkiksi subjektin sija (20) merkitään irrelevantiksi silloin, kun kyseessä on lause- tai infinitiivisubjekti; lauserakenne (15) on irrelevantti silloin kun kyseessä on analysoimaton fragmentti, joka on myös sanajärjestyksen (18) suhteen irrelevantti.

Tämäntyyppisessä kooditustyössä ei *itse aineisto* joudu missään vaiheessa tietokoneeseen, vaan se pysyy paperilla tekstinä. Tietokoneeseen syötetään vain kustakin lauseesta saatu joukko numeroita, jotka on merkitty yhdelle kortille (esimerkki liitteessä 3) ja näinä numeroina aineistoa myös käsitellään edelleen. Kun halutaan tietystä ilmiöstä esimerkkejä, ne on etsittävä tekstin, virkkeen ja lauseen numeron perusteella alkuperäisestä tekstiaineistosta: kone ilmoittaa vain haluttujen tapausten paikantamiseen tarvittavan numerosarjan. - Kun jatkossa käytämme esimerkkejä, numeroimme ne ao. tekstiin ja virkkeeseen viitaten; siis esimerkin jälkeinen jakso 60.31 ilmoittaa kyseessä olevan tekstin 60 virke 31, tai sen osalause.

## 2.2. Tekniset muuttujat

Kustakin lauseesta haluttiin joukko teknistä tietoa, joka auttaa paikantamaan lauseen; teknisten tietojen tarkkuuteen vaikutti lähtökohtamme, joka oli alun perin kontrastiivinen. Käsitelimme aluksi samaa tekstiä suomeksi ja ruotsiksi, ja tätä varten tarvitsimme tarkan tiedon lauseen asemasta virkkeessä (vrt. Andersson & Östman (1978)).

### (1) *tekstin numero*

Alle 1000 tekstin korpusta varten tarvitaan kolmen sarakkeen muuttuja. Meidän aineistossamme oli 123 tekstiä.

### (2) *virkkeen numero tekstissä*

Jos kukin yksittäinen teksti on alle 100 virkkeen mittainen (kuten tässä aineistossa), tekstin virkkeet merkitään juoksevin numeroin kahteen sarakkeeseen.

### (3) *lauseen numero virkkeessä*

Samaan virkkeeseen kuuluvat lauseet numeroidaan esiintymisjärjestyksessä juoksevin numeroin. Koska on aivan poikkeuksellista, että yhdessä virkkeessä olisi enemmän kuin 10 tai H lausetta, lauseen numeroa varten riittää yksi sarake.

### 2.3. Virkekonstruktio­muuttujat

Onko tekstin perusyksikkö virke vai lause? Useimmiten väitetään, että virke olisi perusyksikkö. Kuitenkin virkkeeseen kuuluvien lauseiden välillä vallitsee samanlaisia *sidossuhteita* (esim. pronominin käyttö) kuin virkerajan ylitsekin. *Tematiikkaa* käsiteltäessä ei yleensä kukaan puhu virkkeen teemasta, vaan lauseen. Juuri yksinkertaisen lauseen rajoissa sanajärjestys tuntuu palvelevan informaatorakenteen tarpeita. Toisaalta on ilmeistä, että tätä säätelee osittain lauseen asema virkkeessä (onko kyseessä pää- vai sivulause, virkkeen ensimmäinen vai viimeinen lause). Näistä syistä päädyttiin sellaiseen ratkaisuun, että tekstit kooditetaan lauseittain mutta tietoa saa järjestelmän avulla myös lauseiden muodostamista erityyppisistä virkekonstruktioista sekä siitä, miten nämä vaikuttavat yksittäisten lauseiden sisäisiin temaattisiin ratkaisuihin. Jos analyysin perusyksiköksi olisi valittu virke, seula olisi muodostunut ylen karkeaksi. Tällöin olisi myös ollut mahdotonta selvittää samalla kertaa puhtaasti syntaktisia seikkoja, joista olimme kuitenkin myös kiinnostuneita.

Mikä sitten on lause ja mikä ei ole? Tämä puhutun tekstin keskeinen ongelma on kirjoitettua kieltä käsiteltäessä paljon vähäisempi. Silti joudutaan tekemään monia yksittäisratkaisuja, jotka eivät ole yksiselitteisiä. Lähdemme siitä periaatteesta, että *lauseiden määrä on sama kuin finiittiverbien määrä*. Edelleen tärkeimmät "lauseenvastikkeiksi" sanotut ei-finiittiset rakenteet (ks. muuttujaa 8) käsitellään lauseina. Lisäksi on mainittava seuraavat yksittäisratkaisut:

1. Rinnasteiset nominaalikonstituentit kuuluvat samaan lauseeseen, rinnasteiset finiittiverbit eivät. Näin ollen ilmaukset [*Mikko ja Kalle*] lähtivät retkelle (subjektissa rinnastus) ja *Mikko osti [ruoan ja liput]* (objektissa rinnastus) käsitellään kumpikin yhdeksi lauseeksi. Sen sijaan ilmaus *Kalle söi / ja lähti* on luettu kahdeksi lauseeksi, koska siinä on kaksi finiittiverbiä. Silloinkin siis, kun subjekti on elliptisesti poistettu, olemme pitäneet kaksilauseisuudesta kiinni. Tämä menettely kasvattaa siis jonkin verran yksinkertaisten lauseidemme lukumäärää verrattuna perinteiseen käsitykseen lauseesta, ja verrattuna myös esim. Westmanin (1974) kirjoitettua ruotsia koskevaan tutkimukseen, jossa rinnastettuja finiittiverbejä määritteineen pidetään konjunktiolausekkeina, ei lauseina. Saattaisi kuitenkin olla suositeltavampaa pitää seuraavanlaisia rinnastuksia verbiryhmän rinnastuksena:

Esikoeryhmältä tiedusteltiin, kuinka he [kokivat ja tulkitsivat] filmin. 8.20; -- ei liene liioiteltua sanoa hänen teoriansa [kokoavan yhteen ja jäsentävän] suuren osan neuvostopsykologian sosiaalitutkimusta -- 2.45.

2. Kun finiittiverbi on rinnasteisesta lauseesta elliptisesti poistettu (muuttuja 59), pidetään jäljelle jäänyttä verbitöntä jaksoa yhä lauseena, jos rinnastuksen jälkiosassa on jäljellä vähintään kaksi konstituenttia:

Tämä johtuu akateemisen työttömyyden yleistymisestä / ja toisaalta 0 0 pidentyneestä äitiyslomasta. Se syö leivänmuruja / ja 0 0 nimensä mukaisesti sokeria. 90.21.

Samalla tavalla on katsottava ellipsin hävittäneen verbin *ei-vaan* -rakenteista:

Historiallisesti muuttuvaa ja kehittyvää ei tällöin ole ainoastaan korkeampien henkisten toimintojen sisältö / vaan 0 0 jopa niiden rakenne ja muodot /- näkökanta, joka poikkesi ratkaisevasti perinteisen psykologian staattisista ja ylihistoriallisista katsomuksista. 2.23

Tässä jakso *näkökanta, joka ...* on esimerkki fragmentista, lauseeksi koodattavasta verbittömästä tai muuten syntaktisesti vajaan rakenteesta. Vaihtoehdon sisältävät tai polaaraisuudeltaan vastakohtaiset jaksot on tuntunut luontevalta koodittaa eri lauseiksi, toisin siis kuin konjunktiolla *ja* aikaansaatu rinnastus:

--- soittaako Piirpauke omilla ehdoillaan, / sävelmateriaalin ehdoilla / vai yleisön ehdoilla. 63.26

--- suurin osa siirtomaissa tehdyistä investoinneista oli lähtöisin emämaan valtion kassasta / eikä 0 0 siirtomaassa muodostetusta yksityisestä pääomasta. 7.63

Yksityistapauksissa on kuitenkin käytettävä omaa harkintaa. Tämän osoittaa esim. seuraava tapaus, jossa rinnasteisten objektien luettelossa on välistä sivulauseita; tällöin ilmaus on pilkottava eri lauseiksi:

Junassa olen kohdannut miehen, / joka tappoi kepakolla rantavesissä kutevan 11-kiloisen hauen; / naisen, / joka oli kirjailut kaksisataa kertaa kanavakankaalle sanat "Oi kallis Suomenmaa!" --; / isoäidin, / jolla oli niin kunnollinen poika, / ettei tämä ollut koskaan astunut soittoruokalaan; / miehen, / joka -- 79.16

Tästä katkelmasta kertyy siis yhteensä 9 "lausetta".

3. Ei-finiittiset rakenteet ovat eri määrin lausemaisia (vrt. Hakulinen & Karlsson (1979, luku 14)). Tässä on otettu syntaktiseksi "lausemaisuuksien" kriteeriksi se, voiko rakenteella olla oma (tav. genetiivimuotoinen) subjekti. Tällöin lauseeksi luettavia ovat ns. partisiippirakenne (tosin vain objektina), ja temporaalirakenne, myös aikasekvenssiä merkitsemään käytetty "modaalirakenne", vaikka sillä ei olekaan subjektia:

Karhu kuuluu vihapäissään pystyvän pirstomaan hirsistä rakennettuja aitoja aiheuttaen rajavälikohtauksia, joilla kuitenkin ei ole suurpoliittista merkitystä. 9.54

Toinen poikkeus tästä on se, että kooditamme subjektina olevan I infinitiivin määritteineen omaksi lauseekseen (vrt. muuttujaa 8:E). Sen sijaan omiksi lauseikseen ei kooditeta seuraavanlaisia 3. infinitiivin taikka 1. inf. translatiivin muodostamia inf.lausekkeita, vaan ne muodostavat ns. verbiketjun (vrt. 25:3):

-- voi sielunsa silmin nähdä kaukaisen viinituvan, / jossa vanhat miehet levittävät käsivartensa käydäkseen Zorban tanssiin. 63.28

Siri Derkertille on ollut luonnollista / ilmentää surunsa ja ilonsa tanssimalla ateljeen pöydällä tai hämmentyneen kuninkaan ja avajaisyleisön edessä Liljevalchsissa. 68.19

Paitsi eristämistä ulkomaailmasta pyritään kohde eristämään myös lähiympäristöstään mm. vaihtamalla tiheästi työ- yms. ryhmien jäseniä kiinteiden henkilö-yhteyksien estämiseksi sekä soveltamalla käyttöön jäljempänä esiintyvää joukkodynamiikkaa. 13.15

Tällaiset rakenteet on syytä koodittaa lauseiksi sitä varten, että ne sisältävät paljon informaatiota ja paljon uusia tai mainittuja referenttejä, joita ei saada "kiinni" (vrt. kuitenkin muuttujaa 53), jos rakenne merkitään vain tavalliseksi adverbialiksi niin kuin tässä on tehty. Mikäli tarkoituksena on nimenomaan esimerkiksi uusien ja vanhojen referenttien tekstiin sijoittumisen tutkiminen, sitä parempi mitä pienemmiksi yksiköiksi teksti saadaan jaettu; tällöin taas joudutaan vesittämään lauseen käsitettä. Kompromissiksi voisi harkita ei-finiittisen konstruktion koodittamista omaksi lauseekseen (esim. lausetyypiksi 8:F) siinä tapauksessa, että se sisältää vähintään kaksi omaa NP:tä (tai on itsessään rinnasteinen ja kumpikin osa sisältää omat NP:nsä):

--- / mm. vaihtamalla tiheästi työ- tms. ryhmien jäseniä kiinteiden henkilöyhteyksien estämiseksi / sekä soveltamalla käyttöön jäljempänä esiintyvää joukkodynamiikkaa.

Emme ole tässä selvityksessä kiinnittäneet huomiota attribuuttina oleviin partisiippirakenteisiin, joihin kuitenkin voi sisältyä runsaasti erilaista informaatiota. Haluttaessa ottaa nekin huomioon, ne voidaan luokitella omaksi alakategoriakseen esim. relatiivilauseiden joukkoon. Raja partisiipin ja adjektiivin välillä on epämääräinen, mistä tulee tällöin oma ongelmansa.

Muuttujat (5: sivulauseen matriisilauseen numero), (6: lauseen paikka virkkeessä) ja (7: konstruktioyhtyyppi) antavat yhdessä kuvan koko virkekonstruktion kompleksisuudesta ja käsiteltävän lauseen asemasta siinä. Muuttujassa (6) tarkastellaan lausetta itseään ja merkitään sen sijainti virkkeessä; tämä tehdään erikseen pää- ja sivulauseiden osalta. Muuttujan (7) kohdalla taas katsotaan toisaalta sitä, onko käsiteltävä lause itse kompleksinen (ts. "roikkuuko" siitä muita lauseita) ja toisaalta sitä, millaisessa konstruktiossa se itse on osana; ei merkitse mitään, onko kysymyksessä päälause vai ei.

#### (5) *sivulauseen matriisilauseen numero*

Haluttaessa tietää, mihin virkkeen lauseista jokin lause on upotettu merkitään kunkin upotteen lauseen kohdalle tätä hallitsevan lauseen numero. Etenkin kontrastiivisessa tutkimuksessa tällä tiedolla on oma mielenkiintonsa: onko kummassakin kielessä voitu upottaa samalla tavoin vai onko virkestrategiaa jouduttu käännettäessä muuttamaan.

#### (6) *lauseen paikka virkkeessä*

- 1 päälause yksin tai ensimmäisenä
- 2 päälause rinnastuksen alussa
- 3 päälause keskellä
- 4 päälause lopussa
- 5 sivulause päälauseensa edellä
- 6 sivulause päälauseensa sisällä
- 7 sivulause päälauseensa jäljessä
- 8 sivulause virkkeen lopussa

Tällä muuttujalla saa alakategoriat yhdistämällä perustiedon siitä, miten päälauseet (1-4) ja sivulauseet (5-8) suhtautuvat toisiinsa lukumäärältään. Tämän lisäksi siitä saa lukemattomia taustamuuttujia haluttaessa tarkastella yksityiskohtaisesti sitä, mikä lauseen virkeasema vaikuttaa eniten vaikkapa sen sanajärjestykseen. Voidaan vertailla toisiinsa esim. virkkeenloppuisia pää- ja sivulauseita ja katsoa, onko niiden välillä kenties enemmän samankaltaisuutta kuin virkkeenalkuisen pää- ja sivulauseen välillä (vrt. 4.4.). Muutamia esimerkkejä kooditusratkaisuista:

(6:3, päälause keskellä): Junaihmiset ovat erilaisia kuin lentoihmiset: *He suhtautuvat matkustusprosessiin rauhallisemmin / ja asettuvat taloksi sosiaaliseen, rytmikkäästi keinahtelevaan kulkuneuvon, / -- 79.3*

(6:4, päälause lopussa): Luettelo junissa tapaamistani tulisi liian pitkäksi, / jos kertoisin kaiken, / *joten mainitsen vain muutamia tapauksia. 81.14*

(6:6, sivulause päälauseensa sisällä): Ihmistä, *jolla on hukkaajan luonto, / tärisyttää henkisesti tällainen suurpiirteisyys, / jota Annen poikaystävä osoittaa. 82.7*

(6:7, sivulause päälauseensa jäljessä): -- hampaiden välissä keinui muovikassi, / *jonka kirkkaanpunaisella pinnalla mainostettiin jotakuta vaatetusliikettä / ja josta -- 79.8*

(6:8, sivulause virkkeen lopussa): -- hampaiden välissä keinui muovikassi, / *jonka kirkkaanpunaisella pinnalla mainostettiin jotakuta vaatetusliikettä / ja josta kohosi ilmoille tuoreiden savusilakoiden huumaava tuoksu. 79.8*

(7) konstruktioyyppi

1 yksinkertainen	
2 kompleksinen	
3 rinnasteisen konstruktion yksinkertainen	alkuosa
4	keskiosa
5	loppuosa
6 rinnasteisen konstruktion kompleksinen	alkuosa
7	keskiosa
8	loppuosa
9 irrallinen parenteesi	

Tästä muuttujasta näkyy edellisten tietojen ohella se, onko kukin yksittäinen lause yksinkertainen vai kompleksinen ja missä kohden virkettä tämä yksinkertainen tai mutkikas lause sijaitsee. Hypoteesina voi pitää sitä, että kompleksiset lauseet pyrkisivät virkkeen loppuun ja että mitä syvemmälle kompleksiuushierarkiassa mennään, sitä vähemmän lauseessa on topikaalistusta.

Havainnollistamme muuttujien (5-7) käyttöä muutaman kompleksisen virkkeen avulla:

	(5)	(6)	(7)
1 Vain hyvin kovakuorinen katsoja voinee väittää,	0	päälause yksin	kompleksinen
2 ettei elänyt kipeästi mukana Käthen tuskaa poikiensa kuolemasta	1	sivulause päälauseen jäljessä	rinnasteisen yksinkertainen alkuosa
3 ja lopulta sitä luomisen kamppailua,	1	sivulause päälauseen jäljessä	rinnasteisen kompleksinen loppuosa
4 johon puristui hätä koko maailman nuorisosta ja aseiden murentamasta maailmasta	3	sivulause päälauseen jäljessä	yksinkertainen
1 Hänellä oli,	0	päälause rinnastuksen alussa	rinnasteisen kompleksinen alkuosa
2 mitä katsella,	1	sivulause päälauseen jäljessä	yksinkertainen
3 sillä kaulassani riippui nauhuri,	0	päälause keskellä	rinnasteisen yksinkertainen keskiosa

4 hampaiden välissä keinui muovikassi,	0	päälause keskellä	rinnasteisen kompleksinen loppuosa
5 jonka kirkkaanpunaisella pinnalla mainostettiin jotakuta vaatetusliikettä	4	sivulause päälauseen jäljessä	rinnasteisen yksinkertainen alkuosa
6 ja josta kohosi ilmoille tuoreiden savusilakoiden huumaava tuoksu.	4	sivulause lopussa	rinnasteisen yksinkertainen loppuosa
1 Kieroa hymyilyä ja avointa pilkkaa aiheuttaa se,	0	päälause yksin	kompleksinen
2 että Teutorin itsensä ollessa kuusikymmenvuotias	3	sivulause päälauseensa edessä	yksinkertainen
3 synnyttää hänen 14 vuotta nuorempi vaimonsa vielä kaksoset.	1	sivulause päälauseensa jäljessä	kompleksinen
1 Olen lukenut,	0	päälause yksin	kompleksinen
2 että Anne Pohtamon poikaystävä Jacques oli ostanut Annelle timanttinauhan kaulalla kannettavaksi	1	sivulause päälauseensa jäljessä	rinnasteisen yksinkertainen alkuosa
3 mutta hukannut sen New Yorkissa	1	sivulause päälauseensa jäljessä	rinnasteisen kompleksinen keskiosa
4 ennen kuin ehti antaa lahjan armaalleen,	3	sivulause päälauseensa jäljessä	yksinkertainen
5 ja nyt hän ilmoittaa iloisesti ja huolettomasti lehdille:	1	sivulause päälauseensa jäljessä	rinnasteisen kompleksinen loppuosa
6 Täytynee ostaa uusi kadonneen tilalle.	5	sivulause päälauseen jäljessä	yksinkertainen

## 2.4. Lauseen perussyntaksia ja -semantiikkaa koskevat muuttujat

lausetyyppi (8)  
lausemuoto (9)  
modaalisuus (10)  
tekstifunktio (11)  
geneerisyys (12)  
pragmaattiset elementit (13)  
kielto (14)  
lauserakenne (15)  
pintajärjestys (18)

### (8) lausetyyppi

- 1 päälause
- 2 konjunktiolla alkava päälause
- 3 sitaatti (pää- tai sivulause)
- 4 relatiivilause subjekti-NP:ssä
- 5 relatiivilause objekti-NP:ssä
- 6 relatiivilause muualla
- 7 *että*-lause
- 8 epäsuora kysymyslause
- 9 adverbiaalilause
- A komparatiivilause (pää- tai sivulause)
- B partisiippirakenne
- C temporaalirakenne
- D "väärin" käytetty modaalirakenne
- E subjektina oleva infinitiivirakenne

Lausetyyppi käsitetään suurin piirtein perinteisestä kieliopista tutulla tavalla; kriteerit ovat muodolliset. Muuttujan jakaminen juuri näihin alakategorioihin on paljolti hakuammuntaa, koska on ollut vain vähän taustatietoa erityyppisten lauseiden taajuudesta tai muustakaan käyttäytymisestä asiatekstissä. Jälkeenpäin arvioiden ei ole ehkä ollut useinkaan valaisevaa erottaa kolmea relatiivilauseeryhmää sen perusteella, missä konstituentissa ne sijaitsevat; tuonnempana tehtävät ristiintaulukoinnit osoittavat kuitenkin näiden välisiä eroja. Alakategoriana 7 oleva *että*-lause saattaa puolestaan olla liian karkea. Voisi olla kiintoisaa jakaa se kahtia, subjektina ja objektina olevaan *että*-lauseeseen, joskin edellinen näistä on aika vähäkäyttöinen. Adverbiaalina olevat *että*-lauseet tyyppiä *siihen että*, *siksi että* kuuluvat alakategoriaan 9. Adverbiaalilauseet ovat tällainen mahd. liian karkea kategoria; harkitsemista ansaitsisi *kun*-, *vaikka*- ja *koska*-lauseiden erottaminen toisistaan. Komparatiivilauseista ei ole ollut riittävä perustutkimusta, joten se on erotettu omaksi kategoriakseen niissä tapauksissa, joissa on selvästi ollut kyse lausemaisesta ilmauksesta. Näin on päädytty alisteisuudeltaan kirjajaan kategoriaan:

Yleisö tuntui usein vaivaantuneelta, / *ikäänkuin kritiikki kohdistuisi jokaiseen salissa istujaan* 62.36 (päälauseena)

Muutoin voi sitten sienestää / *niin kuin lystää*, -- 113.45 (itsenäinen adverbiaalilause)

-- kieltäytymisemme avoimesta keskustelusta aiheuttaa otaksunan, että olemme vieläkin vähemmän kykeneviä päättämään omista asioistamme / *kuin todellisuudessa olemme* 75.42 (predikatiivissa)

Jokainen kone, jokainen laite, joka maassamme tehdään, sisältää kätkeytyä liikevaihtoveroa useita prosentteja enemmän / *kuin vastaava ulkomailta ostettu kone tai laite* 78.35 (kvanttorilausekkeessa)

-- suotavien reaktioitten palkitseminen näytti tuottaneen enemmän tuloksia / *kuin ei-suotavien reaktioitten rankaiseminen* 8.65 (subjektina)

Toinen alakategoria, joka sisältää sekä pää- että sivulauseita on 3 eli sitaatti. Suoran esityksen asema ei ole tämän tutkimuksen kannalta ollut kiinnostava. Siksi tekstit on pyritty valitsemaan siten, että sitä olisi mahdollisimman vähän.

Eräissä aikakauslehtiartikkeleissa, jotka ovat periaatteessa haastatteluja, raja haastateltavan puheenvuorojen ja muun tekstin välillä suorastaan häviää. Näin ollen on voitu pitää "sitaatti" samassa muuttujassa kuin muut lausetyypit. Jos suora esitys katsotaan tutkimukselle relevantiksi tai jos käsitellään esimerkiksi romaanitekstiä, se kannattaa erottaa omaksi muuttujakseen (lause esiintyy suorassa esityksessä/ei), sillä sitaateissahan voi esiintyä mikä tahansa kyseisistä lausetyypeistä miten tahansa monimutkaisessa konstruktiossa, mikä aiheuttaisi toisin kooditettuna ainakin teknisiä hankaluuksia. Esimerkkejä:

- *Ei tämä päde* (sitaatti sivulauseena), sanoo johtokunta. - *Mieshän sai loparit* (sitaatti päälauseena). 84.13-14

Kompleksisesta lauseesta, joka kuuluu suoraan esitykseen, merkitään "sitaatiksi" vain hallitseva lause: siitä riippuvat lauseet saavat omat lausetyyppikoodinsa:

"*Puolukkamarkkinoilla on usean viikon katkos toimituksiin* (sitaatti sivulauseena; hallitseva), *silloin vaaditaan hermoja* (sitaatti sivulauseena) kun hintoja kytätään (adverbiaalilause)", selvittää Jansa. 98.11

Sitaatti- ja komparaatiotapauksia lukuun ottamatta lauseet siis jaetaan pää- ja sivulauseisiin. *Sivulauseeksi* katsotaan lause, joka on toisen lauseen konstituenttina: subjektina, objektina tai adverbiaalina, tai sitten jonkin NP-konstituentin attribuuttina. Sivulause merkitään myös hallitsevan lauseen kohdalle konstituentiksi. *Päälause* on lause, joka ei ole toisen lauseen konstituenttina. Tällainen muodollinen jako pää- ja sivulauseisiin ei aina ole semanttisesti tyydyttävä. Perinteellisesti on esitetty, että pääasia ilmaistaan päälauseessa, sivulause puolestaan jotenkin rajoittaa, spesifioi, tuo sivuinformaatiota tms. Tämä pitää paikkansa ainakin sikäli että yleensä alisteinen lause ei muodosta yksin puheaktia. Sentyypisissä virkkeissä kuin:

Kun Suomen Elokuva-Arkisto ryhtyi ajamaan elokuvan aseman vakiinnuttamista Helsingin yliopistossa, sillä oli mielessä ennen muuta elokuvan estetiikan tutkimus. 3.9

päälause ilmaisee varsinaisen assertion, *kun*-lauseen informaatiota on luonnehdittu presupponoiduksi (Hooper & Thompson (1973)). Tällaisessa tapauksessa voi sanoa, että sivu- ja päälause ovat myös semanttisesti sivu- ja päälause. Sama pätee restriktiivisiin relatiivilauseisiin: niiden tehtävänä on spesifioida, identifioida NP jonka määritteinä ne ovat:

On olemassa paljon sellaisia sairauksia potevia, *joiden parantamiseen lääketieteellä ei vielä ole keinoja* -- 11.1

Sen sijaan apposiitivinen relatiivilause voi hyvinkin muodostaa itsenäisen assertion - semanttisesti ja tehtävältään se on siis päälauseen arvoinen:

Työttömän taas tuli ilmoittautua kuntansa työttömyyslautakunnalle, joka todettuaan asianomaisen täyttävän annetut ehdot osoitti hänet työmaalle. 10.9

Erityisen hyvin tämä näkyy tapauksista, joissa päälauseen ja apposiitiivisen relatiivilauseen puhe- tai tekstifunktio ovat erilaiset:

Suomen maa-alue kasvaa eräitten laskelmien mukaan jopa 1000 hehtaaria vuodessa, *mikä kiitollisin mielin todettakoon.* 9.10

Restriktiivisten ja apposiitiivisten relatiivilauseiden erottaminen toisistaan vaatii tosin usein taustatietoa: onko esimerkiksi seuraavassa kysymys ensimmäisestä työttömyyden torjuntatoimenpiteet yhdistävästä laista vai ensimmäisestä työllisyyslaista ylipäätään?

-- säädettiin 1956 ensimmäinen työllisyyslaki, *joka yhtenäisti työttömyyden torjuntaan kohdistuvat toimenpiteet* -- 10.36

Kun restriktiivisten relatiivilauseiden tehtävä on tarkoitteen identifioiminen, apposiitiivisten funktio on puolestaan tekstuaalinen. Se kuuluu "konjunktion" piiriin Enkvistin (1975) esittämässä mielessä: se on toisin sanoen kirjoittajan tekstistrategian mukaista valinnaista lauseiden yhdistämistä samaan yksikköön. Restriktiivisten ja apposiitiivisten relatiivilauseiden koodittaminen erikseen voisi siis jossakin tutkimuksessa olla perusteltua. Mutta restriktiivisyyden käsitettä voidaan soveltaa useimpiin sivulausetyyppeihin. Erityisesti *koska* - ja *vaikka* - lauseet sekä *että* -lauseet voivat eri esiintymäyhteyksissä eri määrissä toimia semanttisesti päälauseina eli muodostaa oman assertionsa (oman tekstuaalisen funktionsa; vrt. Urmson (1963)). Esimerkiksi tapauksessa:

Luulenpa että harvoin teatteria on rakennettu niin taidokkaaksi dialogiksi kuvan ja näyttelijän välille. 61.5

muodollinen päälause suorastaan kalpenee parenteesin kaltaiseksi. Voimme kokeilla samaa keksittyihin lauseisiin:

? Katsoja tuskin tulee ajatelleeksi, *että harvoinpa teatteria on rakennettu näin taidokkaaksi dialogiksi kuvan ja näyttelijän välille.*

Näytelmä ei oikein toiminut, *vaikka teatteria on rakennettu harvoin näin taidokkaaksi dialogiksi kuvan ja näyttelijän välille.*

Näytelmä ei oikein toiminut, *vaikka harvoinpa teatteria on rakennettu näin taidokkaaksi dialogiksi kuvan ja näyttelijän välille.*

Näytelmä oli jotakin aivan uutta, *koska teatteria on harvoin rakennettu dialogiksi kuvan ja näyttelijän välille.*

Näytelmä oli jotakin aivan uutta, *koska harvoinpa teatteria on rakennettu dialogiksi kuvan ja näyttelijän välille.*

Esimerkeistä käy ilmi, että semanttisesti itsenäinen tulkinta seuraa *harvoin* -sanana topikaalistuksesta, jonka aiheena on ilmeinen emfaasi. Onkin luonnollista, että emfaattisesti painottunut lause tarjotaan omaksi tekstiaktikseen. On tapana puhua "päälauseilmiöistä", jollaisia nimenomaan ovat erilaiset topikaalistustransformaatiot, samoin kommenttiadverbiaalit ja erilaiset pragmaattiset elementit (vrt. *harvoinpa*). Eikö nyt tekstilingvistisessä tutkimuksessa olisi syytä merkitä restriktiiviset ja "itsenäiset" relatiivilauseet eri lausetyypeiksi? Jälkimmäisten näkeminen päälauseensa konstituentin määrittäenä on jotakuinkin vastahankaista. Eronteko on kuitenkin liian vaikea kannattaakseen näin laajassa työssä. Sen sijaan, jos nyt jaamme lauseet lausetyyppeihin muodollisin kriteerein, niitä alkavan konjunktion perusteella, voimme erikseen tarkastella topikaalistusilmiöiden esiintymistä eri lausetyypeissä kun tulokset ovat käsissä.

Konjunktiolla alkava päälause (alakategoria 2) voi olla joko virkkeen alussa tai tavallisemmin sen keskellä tai lopussa. Kyseeseen tulevat konjunktiot *ja, sekä, -kä, mutta, vaan, tai, vai, sillä,*

*joten, jolloin*. Muutamat konjunktiot aiheuttavat erityisongelmia. *Kun taas* -lausetta on vaikea pitää päälauseensa adverbiaalina, mutta muotokriteereistä kiinni pitääksemme se on kuitenkin siksi koodattu; sama koskee seuraaventyypistä "pseudo"-*kun* ja *siinä missä* -lausetta:

*Kun Ralf Långbackan Lillaniin ohjaama Shveik kuvaa tavallisen pikkumiehen mahdollisuuksia nousta vastarintaan*, kuvaa Jouko Turkan Turandot eli puuvillaprinsessa fasismiin nousun syitä ja strategiaa. 60.2

*Siinä missä Långbackan Brecht tulvii lämmintä huumoria*, on älykkäästi intiimi ja samalla asioista kiinni pitävä työ, tekee Turka hyttävän suoraa paljasta teatteria. 60.4

(Näihin lauseisiin eivät esim. topikalisaatiot käy: \**näytelmä X ei oikein toiminut, kun taas harvoinpa teatteria on rakennettu niin taidokkaasti kuin näytelmä Y:ssä.*) Nykyisessä asiaproosassa tiheään esiintyvä *jolloin* esiintyy hyvin harvoin alkuperäisessä temporaalisen relatiivipronominin tehtävässään, aitona restriktiivisenä sivulauseena:

-- teorioiden mukaanhan täystyöllisyys on vain tilanne, *jolloin työvoiman kysyntä ja tarjonta ovat yhtä suuret*. 10.2

Joskus siitä on mahdollista tehdä kaksi tulkintaa:

Eerik aloitti etenemisen Baltian suunnalla 1561, *jolloin Tallinna ja Pohjois-Viro alistuivat Ruotsin valtaan*. 1) 'jona aikana', 2) 'ja sen jälkeen' 20.21

Useimmiten *jolloin*-lause kuitenkin on täysin itsenäisen päälauseen arvoinen ja myös kooditettu sellaiseksi:

Toisinaan puheessa on vain substantiiveja ja infinitiivimuodossa olevia verbejä, *jolloin [ja silloin] puhutaan agrammatismista ja sähkötyylistä*. 12.11

Ei-finiittiset rakenteet lienevät muuten itsensä selittäviä, paitsi subjektina oleva infinitiivirakenne (alakategoria E). Kuten edellä on todettu, finiittiverbin ja infinitiivin katsotaan yleensä muodostavan ns. verbiketjun, jonka objektina pidetään infinitiivirakenteen objektia:

Bunuelin filmit *pitäisi nähdä* ainakin kolme kertaa. Ensimmäisen kerran *ehtii seurata* konkreettisia tapahtumia valkokankaalla, -- 62.1

Subjektina olevaan infinitiivirakenteeseen verbiketjuratkaisu käy vain silloin, kun infinitiivi on yksinään subjektina ja koko rakenne on hyvin lyhyt ja kevyt (*on ihanaa elää*). Mutta usein subjektina oleva infinitiivirakenne sisältääkin virkkeen varsinaisen asian. Siksi siinä myös voi olla useita omia konstituentteja, ja se kannattaa koodittaa omaksi lauseekseen siitä huolimatta, että siltä puuttuu oma subjekti. Subjektin puuttumisen voisi merkitä muuttujan (19) kohdalla ellipsisiksi – tai geneeriseksi 0:ksi. Esimerkkejä:

On väärin / *luulla esitystä - tai Kollwitzin taidetta – suorasukaisesti poleemiseksi*. 61.18

Tarinan yksityiskohtia on turha / *paljastaa näin etukäteen*, / kun filmi saadaan melko varmasti Suomeenkin. 62.14

On suurenmoista / *havaita*, / *että yhtye on vielä harjoitellut ja latautunut käynnissä olevaa kiertuettaan varten aivan kuin olisi valloittamassa suuren maailman pääkaupunkeja*, -- 63.17

Siri Derkertille on ollut luonnollista / *ilmentää surunsa ja ilonsa tanssimalla ateljeen pöydällä tai hämmentyneen kuninkaan ja avajaisyleisön edessä Liljevalchissa*. 68.9

Seuraavanlaisessa tapauksessa *että* -lauseen alkuun siirretty genetiivimuotoinen NP *Bergelin ja orkesterin* on tulkittava subjektirakenteen omaksi subjektiksi, joka on topikaalistettu (vrt. 55:D) yli lauserajan.

Sen synnyttämän jälkikuvan vivahteita lisäsi vielä osaltaan se, / että Bergelin ja orkesterin oli ilmeisesti helpompi / seurata Raution rohkean vapaanakin kurinalaista rakentelua / kuin Laurilan oikukasta menoa. 67.24

Tämä esimerkki osoittaa samalla sen, miten vaikea on pysytellä pelkässä pintarakenteen analyysissä. Oikeastaan esimerkkinne koostuu seuraavanlaisista osalauseista:

oli ilmeisesti helpompi / Bergelin ja orkesterin seurata Raution rohkean vapaanakin kurinalaista rakentelua / kuin Bergelin ja orkesterin seurata Laurilan oikukasta menoa.

(9) *lausemuoto*

- 1 väitelause
- 2 *kO*-kysymys
- 3 kysymyssanakysymys
- 4 käskylause
- 5 huudahduslause

Lauseen modaalisuus koodittuu kolmen muuttujan avulla. Muuttuja (9) tallettaa lauseen syntaktisen moduksen eli lausetyypin. Tässä lähtökohta on puhtaasti muodollinen eikä kooditettaessa ole otettu kantaa ns. epäsuoriin puheakteihin eli siihen, onko ehkä kysymystä käytettykin käskynä tms.

(10) *modaalisuus*

- 1 verbimodus (*-isi-*, *-ne-*)
- 2 modaalinen apuverbi verbiketjussa
- 3 modaalinen adverbi
- 4 modaalinen predikaatti

Tähän muuttujaan on pyritty koodittamaan modaalisuus, ts. kirjoittajan asennoituminen lauseen propositionaaliseen sisältöön sikäli kuin se näkyy morfologis-leksikaalisella tasolla. Kun tämän muuttujan tuloksia yhdistellään muuttujista (9, 11-13) saataviin, hahmottunee jokin käsitys yksittäisen tekstin subjektiivisesta värityneisyydestä (vrt. 4.1.). Verrataanpa esimerkiksi täysin "väritöntä", toteavaa lausetta:

Tietoisuutta koskevassa tutkimustyössään Vygotski otti teoreettis-metodologiseksi lähtökohdaksi marxilaisen filosofian ajatukset 1.7 lauseeseen

*Uskaltaisikohan vielä kysyä mahtaako metsästäjän iloon joskus sekoittua murhamiehen riemua?* 73.38

Esimerkkejä alakategorioista:

1. *verbimodus*: optatiivi, konditionaali ja potentiaali. (Näistä tosin ainoastaan potentiaali on aina nimenomaan modaalisessa tehtävässä.)

Piirpauke *saisi* suuren suosion mistä hyvänsä kootun yleisön edessä. 63.27

-- *ei liene* liioiteltua sanoa hänen teoriansa kokoavan yhteen ja jäsentävän suuren osan neuvostopsykologian sosiaalistumistutkimusta sekä näyttävän suuntaa vastaiselle tutkimukselle. 2.45

2. *modaalinen verbi* verbiketjussa. Tällaisiksi on käsitetty pakkoa, välttämättömyyttä, lupaa, mahdollisuutta ilmaisevat verbit, kuten *täytyy, pitää, saa, voi, saattaa* (myös: *on -ttAvA*), *mahtaa* sekä myös *näyttää, tuntuu*. Keskeiset modaaliverbit *voi* ja *täytyy* eivät kuvasta kirjoittajan asennoitumista silloin kun niillä on ns. juurimerkitys (vrt. esim. Matihaldi (1979)), toisaalta taas episteemiset verbit *näyttää* ja *tuntuu* kuvastavat hyvin kirjoittajan asennoitumista sanottavaansa; vrt. seuraavia esimerkkejä:

Kun Hasse Walli heittäytyy Jimi Hendrixin perinteen mukaiseen sähkökitarasooloon makedonialaisen kansanlaulun tulkinnassa, *voi* sielunsa silmin nähdä kaukaisen viinituvan, - 63.28

Tämän pitemmälle ei Helsingin yliopisto nykyisellään *näytä* haluavan mennäkään. 3.4

3. *modaalinen adverb*i: *ehkä, varmasti, tuskinpa, ilmeisesti, luultavasti, kai*

Tarinan yksityiskohtia on turha paljastaa näin etukäteen, kun filmi saadaan *melko varmasti* Suomeenkin. 62.14

*Ehkäpä* vaikuttavinta Liisi Tandefeltin suorituksessa on vanhuuden tulkinta. 61.28

4. *modaalinen predikaatti*: kokonainen pintalause tai predikaatti voi muodostaa modaalisen elementin:

*On mahdotonta* kiivailla siitä, soittaako Piirpauke omilla ehdoillaan, sävelmateriaalin ehdoilla vai yleisön ehdoilla. 63.25

Verbimodus voidaan katsoa irrelevantiksi lauseenvastikkeiden ja muiden infiniittirakenteiden kohdalla, mutta muita modaalisia elementtejä niissä voi esiintyä, siksi muuttuja kokonaisuutena ei ole koskaan irrelevantti. Lause on kooditettu modaalisen adverbien sisältäväksi (modaalisesti värityneeksi) silloinkin, kun modaalinen adverbiaali ei ole varsinaisesti lauseen vaan jonkin sen lausekkeen konstituentti: *ehkäpä* vaikuttavinta, muilta *mahdollisesti* kysymykseen tulevilta seikoilta.

(11) *tekstuaalinen funktio*

- 1 tiedotus, toteamus, kuvaus
- 2 selitys, tarkennus, esimerkki
- 3 päätelmä
- 4 arvio, arvoarvostelma
- 5 käsky, kieltö, ohje
- 6 vetoamus, toivomus
- 7 kysymys
- 8 vastaus, responssi
- 9 metateksti
- i irrelevantti

Kuten edellä on mainittu, tämä muuttuja on osoittautunut vähiten onnistuneeksi. Lähtökohtanamme oli se viimeaikaisen kielitieteen paljon käsittelemä seikka, että lausetyypin muodollinen tunnistaminen ei aina takaakaan sitä, että lausetta käytetään juuri siihen funktioon, josta muodollinen tuntomerkki antaa vihjeen (vrt. esim. Hakulinen & Karlsson (1979, luku 13)). Pragmatiikassa on pyritty kehittämään mahdollisimman aukotonta puheaktien luokitusta, joka lähtisi funktiosta ja puhetilanteesta eikä ainoastaan muodollisista tuntomerkeistä. Tässä on ajautettu ongelmiin jo periaatteellisessa luokittelussa (teorian kritiikkiä ks. esim. Allwood

(1977)).

Kirjoitettuun tekstiin puheaktiteoriaa ei ole ainakaan laajoissa empiirisissä tutkimuksissa sovellettu. Perusongelma on sama kuin puhutuissakin teksteissä: yksittäisen jakson tunnistaminen yksiselitteisesti tiettyä vain yhtä puhe- tai tekstiaktia edustavaksi ei osoittadukaan mahdolliseksi. Vaikuttaa siltä, että on vaikea löytää sellaisia tekstiaktityyppejä, jotka olisivat sikäli primitiivejä, että ne olisivat toisensa poissulkevia ja sopisivat täten samaan muuttuun vaihtoehtoisiksi alakategorioiksi. Tässä emme ole omassa yrityksessämme onnistuneet. Vaikka pyrimme tekemään tekstuaalisen funktion luokittelun nimenomaan neutraalin asiaproosan tutkimisen tarpeita vastaavaksi, perusselvitys on ollut ilmeisen riittämätön, sillä alakategoriat eivät nykyisellään ole toisiaan poissulkevia. Varsinaisia, puhuja-kuulija -suhteen huomioon ottavia puheaktikategorioita ovat alakategoriat 5-8, kun taas tekstin organisointiin liittyviä ovat neljä ensimmäistä sekä mahdollisesti alakategoria 9.

Toinen ongelma aiheutuu siitä, että olemme koodittaneet aineiston lause lauseelta emmekä virke virkkeeltä. Ei ole läheskään aina selvää, että "puheaktia" vastaa kirjoitetussa monologissa juuri lause tai että kullakin lauseella olisi oma illokutiivinen sävynsä. Tästä voimme selvittää osittain siten, että jokainen sivulause, joka ei poikkea tekstifunktioltaan päälauseensa funktiosta, merkitään tämän muuttujan kohdalla irrelevantiksi (i). Käytännössä siis varsin usein tekstuaalista funktiota kantaa päälause mahdollisine alisteisine lauseineen. Tämäkin voi olla vielä liian hienojakoista, sillä on kuviteltavissa, että ainakin rinnasteinen, toisiinsa löyhästi kytketyistä päälauseista koostuva virke edustaa yhtä funktiota, on jopa ajateltavissa, että kokonaisella kappaleella, ns. makrostruktuurilla (ks. van Dijk (1977)) on joskus vain yksi tekstifunktio. Vastaisuudessa jonkin jakson tekstifunktion määrittäminen edellyttääkin, että kukin yksittäinen teksti analysoidaan ensin makrostruktuureihinsa ja sitten tutkitaan, missä määrin tällaiseksi hahmotettavat yksiköt erottuisivat toisistaan juuri illokutiiviselta sävyiltään. Tekstifunktion koodittaminen edellyttää siis nykyistä paljon perusteellisempaa kokonaisten tekstien tutkimusta, jota ilman pelkkä koodittaminen jää liian subjektiiviseksi.

Joitakin esimerkkejä lienee silti kiintoisaa esittää tästäkin yrityksestä:

2. *selitys, tarkennus, esimerkki*. Yhteisotsikoksi tähän sopisi ehkä "tekstuaalisesti alisteinen jakso". Tekstiin tuodaan informaatiota, joka on tarpeen selittämään, tarkentamaan, ehkä korjaamaan edellä mainitun "päälinjan" jotakin kohtaa.

Aluksi luulin, että hän haluaa esittävien taiteilijoiden tapaan saada teoksesta pois liialliset modernisuudet, mutta pian kävi ilmi, että hän työskenteli huippuluokan muusikon tavoin: *musiikillisen sanoman luonteisuus oli hänelle pääasia, ei tyylin modernius tai vanhanaikaisuus*. 64.8

Hänen muuntautumiskykyään, joustavuuttaan ja liikkuvuuttaan on pakko ihaila ja ihmetellä: *vain kaksi päivää aikaisemmin hän oli johtanut Turussa Händelin Messiaan, nyt jo Salzburgissa sille monessa mielessä lähes täysin vastakkaisen teoksen*. 65.25

Onni oli suosinut Amerikan maalla. *Näinhän on tapana sanoa, kun on menestytty*. 71.16-17

3. *päätelmä*: tekstiä kokoava jakso.

Suunnitelmia ja päätöksiäkin jatkokoulutuksen kehittämisestä siis on, mutta koko ongelmaa niillä ei vielä ratkaista. 29.21

Näin on jouduttu kiistaan, joka saattaa jatkua pitkäänkin, elleivät osapuolet suostu harkitsemaan asioista uudestaan. 31.13

Varkaudet ovat lisääntyneet. Tyypillinen myymälävaras on annettujen tietojen mukaan yli 20-vuotias mies. *Kyse ei ole siis pelkästä lasten ja nuorten makeisten ja tupakan näpistelystä.* 38.7

4. *arvio, arvoarvostelma.* Erona tavalliseen tiedottavaan virkkeeseen on se, että tässä kirjoittaja ottaa kantaa esittämäänsä asiaintilaan: sen eettiseen, esteettiseen tms. arvoon tai mahdollisuuteen, todennäköisyyteen. On huomattava, että kyseessä on nimenomaan kirjoittajan arvio; tapaukset joissa kirjoittaja referoi jonkun muun tai vaikkapa oman itsensä suhtautumista, eivät kuulu tähän:

*Aluksi luulin, että hän haluaa esittävien taiteilijoiden tapaan saada esityksestä pois liialliset modernisuudet, -- 4.13*

*Maallikon silmään juttu näyttää hullulta.* 84.5

*Hänen muuntautumiskykyään, joustavuuttaan ja liikkuvuuttaan on pakko ihailta: -- 65.25*

*On ikävä kyllä todettava, että demokratia ei ole meillä pelannut enää ainakaan viimeisen vuosikymmenen aikana.* 76.10

5. *käskey, kielto, ohje.* Tässä kirjoittaja ikäänkuin suoraan pyytää tai vaatii jotakin - lukijalta tai yleisemmin joltakin muulta taholta:

*Mutta kyllä siinä puuhassa pitäisi välillä vetää henkeä ja laskea katse sinne, missä tämän päivän todellinen puute majoilee.* 72.25

*Hallitus sanoo ohjelmajulistuksessaan käyvänsä hyvään yhteistyöhön etujärjestöjen kanssa. Pankoon teollisuus tuon julistuksen nyt koetteelle.* 74.53-54

*Anna pienokaisen harjoitella konttaamista ilman potkukousuja.* 119.7

6. *vetoamus, toivomus (suositus, lupaus).* Tässä on kyse jostakin lievemmästä vaatimisesta kuin edellisessä kohdassa, usein vihjauksenomaisesta, ja nimenomaan vetoavasta: "Eikö niin, että olet samaa mieltä kanssani?"

*Löytyipä ohjelmasta kuitenkin fraasi: "rauhantahtoinen puolueettomuuspolitiikka". Suoraan valtiotierailujen kommunikatioista lainattu. Kuka tuntee sodantahtoisien puolueettomuuspolitiikan?* 71.21-23

*Nykyisiä tavattoman ahtaita liputtamisen rajoja pyritään väljentämään. Sovitaan, että se pikkuisen korvaa sen, että virallisista liputuspäivistä suunnitellaan karsittavaksi pois puolustusvoimain lippujuhlapäivä* 70.41-42

*Ei kai ketään voida pakottaa tilaamaan jotakin paikkakunnan lehteä vain siksi, että jotenkuten saisi selville, minne maksuja maksetaan.* 73.33

7. *kysymys.* Kysymyksen status kirjoitetussa tekstissä on epäselvä. Koskaan sen funktio ei voi olla sama kuin puhutussa kommunikaatiossa (paitsi kirjeissä): informaation pyytäminen. Aito kysymys tekstissä kai pikemminkin fokusoi, tähdentää informaatiota, joka sitä seuraa tai edeltää, tai sitten se on mieluummin luokiteltava vaikkapa vetoomukseksi (ks. ed.).

*Uskaltaisikohan vielä kysyä, mahtaako metsästäjän iloon sekoittua murhamiehen riemua?* 73.38

*Vai onko niin, ettei Schumannin toisessa tämän kummempaa henkeä olekaan ja riipeä meno siis vain peitti sävellyksen puutteet, mene ja tiedä?* 65.43

8. *vastaus kysymykseen, responsi*. Paitsi suoraa vastausta suoraan kysymykseen esim. repliikeissä tähän kuuluvat "vastaukset", joita ei välttämättä olekaan kysymyksellä pyydetty:

Vanha hokema väittää, että musiikki olisi kansainvälinen kieli. *Ei pidä paikkaansa*. 63.1-2

Voiko ihmisellä olla lempi-itikkaa? *Miksei voisi*. 90.1-2

Mutta mihin rahat uppoavat. *Kuin Kankkulan kaivoon, järjestöjen päällekkäisten organisaatioitten ylläpitämiseen* -- 91.4-5

9. *metateksti* eli tekstiin orientoiva jakso (vrt. Enkvist (1975:115)):

*Asiaan liittyy vielä toinenkin näkökohta*. 25.8

*Tämän pahanillisen mielen synnyttämän havainnon jälkeen* pitää ottaa osittain komento takaisin. 70.16

Nyt tietysti ollaan uteliaita että minkälaisen lakun noin vähällä rahalla saa irtoamaan. *Minä kerron*. 70.30-31

*Tässä oli johdanto tämän kertaiselle asialle*. 71.26

(12) *geneerisyys*

1 spesifinen lause

2 geneerinen lause

3 irrelevantti

Geneeristen ja ei-geneeristen lauseiden erottamiseen toisistaan päädyttiin kahdesta syystä. Toisaalta uskottiin olevan mielenkiintoista kytkeä lauseiden geneerisyys niiden erilaiseen tekstifunktioon. Toisaalta havaittiin, että muissa kuin kertovissa toimintateksteissä, jollaisia esitutkimuksen tekstit olivat olleet, geneerisyys on yksi referenssin ja samaviitteisyyden toimintaan vaikuttava tekijä (tästä tarkemmin muuttujan 21 kohdalla). Koska koko lauseen geneerisyyteen vaikuttaa sekä finiittiverbin geneerisyys että subjekti- ja usein objekti-NP:n geneerisyys vs. spesifisyys, voisi vastaisuudessa harkita tämän piirteen tarkkailemista näissä kolmessa kohden.

Koska Maria Vilkun (1980) spesifisyyttä ja geneerisyyttä koskeva seikkaperäinen tutkimus perustuu tähän samaan aineistoon, ei asiaa ole tässä yhteydessä tarpeen käsitellä laajasti. Geneeriseksi tulkitsemista ei suomessa ole tukemassa enempää erityinen aikamuoto (vaikka tyypillisin tempus geneerisessä lauseessa onkin preesens) kuin erityinen artikkelinkäyttökään. Tulkinta jää siis kokonaan kontekstin varaan. Tärkeää on tässä yhteydessä ymmärtää ns. ei-faktuaalisten kontekstien vaikutus lausekkeiden spesifisyyteen. Jos NP on modaalisen elementin vaikutusalassa, sillä ei ole spesifistä referenttiä (esim. Minulla ei ole *autoa*; Haluaisin *uudet sukset*). Toisaalta konteksti voi olla kvantifioitu (esim. Jokaisessa huoneessa on *yksi tentittävä*, tai Joka aamu lensi *kurkiaura* talomme yli), mikä myös saa aikaan epäspesifisen tulkinnan kvantifioidussa kontekstissa olevalle NP:lle. Niinpä on ajateltavissa, että kukin NP voitaisiin luokitella referenssiltään spesifiseksi, epäspesifiseksi taikka geneeriseksi. Finiittiverbissä ei ole tällaista spesifisyyttä, mutta voidaan toki erotella erityyppistä geneerisyyttä. Esim. lauseet *Jussi polttaa* ja *Jussi ajaa kiitolinjaa* ovat lähinnä habituaalisia, kun taas lause *Jussi suuttuu helposti* ilmaisee dispositiota. Väljästi tulkiten nämä ovat geneerisen lajeja. Spesifisyyden ja geneerisyyden rajalla liikutaan silloin, kun subjekti viittaa luokkaan:

*Sokeritoukka* asuu kellareissa ja kylpyhuoneissa.

*Sokeritoukat* asuvat kellareissa ja kylpyhuoneissa.

Monikkotapauksissa on hiuksenhieno ero aidosti "geneerisen" tulkinnan ja "pelkän yleistyksen" välillä, mutta sitä on vaikea tavoittaa. Ehkä kyse on vain siitä, että kun luokkareferenttiä voi ilmaista käyttämällä yksikkömuotoa koko luokkaa edustamassa, edustavuuden tunnekin on tällöin vahvempi. Seuraavassa on muutamia tyypillisiä esimerkkejä lauseista, joissa jokin keskeinen nominaalinen konstituentti on ei-spesifinen ja verbi myös geneeriseksi tulkittava:

Jumala vain *leikittelee ihmisillä* istuttamalla heihin toivon ja haavekuvan paremmasta, mutta tekemällä samalla tyhjiksi kaikki heidän pyrkimyksensä. (habituaali) 56.30

Mehiläisyhteiskunnan sydän on *kuningatar*, lisääntymiskykyinen naaras. 21.9

*Sen* apuna on suuri joukko lisääntymiskyvyttömiä naaraita, työläisiä. 21.10

*Susiperhe tai koskelopoikue* metsästää yhdessä tehokkaammin kuin yksinään. 21.19

*Mehiläinen* saa talvella ravintoa yhteiskunnan varastosta. 21.20

*Haukka* epäröi syösyä tiheään pikkulintuparveen, koska sillä ei ole selvää kohdetta hyökkäykselleen, *naurulokit* karkoittavat joukkovoimallaan pesimäalueelle yrittävät pedot ja *tiivinä muodostelmana uivat pikkukalat* ovat suojassa isommilta kaloilta. 21.24

On kiinnostavaa havaita, miten luova kirjoittaja voi manipuloida koreferenssillä ja liukuilla luokkareferenssin, ei-spesifisen referenssin ja spesifisen referenssin välimailloilla niin kuin seuraavassa:

Voiko ihmisellä olla lempi-itikkaa? Miksei voisi. Minun erityinen suosikkini itikoiden parissa on *sokeritoukka*, *Lepisma saccharina* eli "*sokerinen suomukas*" (laji). Se (laji) asustaa kellarissa ja kylpyhuoneessa, missä on sopivan lämmintä ja kosteaa. Väliin kun kirjoittelen jotakin iltayöstä työlampun valossa se (kvantitatiivisesti spesifinen) tulee työhuoneen lattialle minua tapaamaan. Yhtäkkiä se (epäspesifinen) vain on ilmestynyt jostakin ja 0 loistaa hopean hohtavana sinisellä matolla. *Sokeritoukka* (laji) on oikeastaan vahinkoeläin, ainakin kovin runsaana esiintyessään. Se syö leivänmuruja ja nimensä mukaisesti sokeria. Se syö myös kirjanselkiä. Mutta en raatsi tappaa *sitä* (epäspesifinen). -Se (laji) on elänyt neljäsataa miljoonaa vuotta omaa pimentoelämänsä. Mikäpä minä, tulokas ja nousukas, olisin *sitä* (laji - vai epäspesifinen?) vainoamaan.-- -

### (13) pragmaattiset elementit

1 -hAn

2 verbi + kin

3 liikkuva -kin alussa

4 liikkuva -kin lopussa

5 -pA

6 edellisten yhdistelmiä

7 *kyllä*

8 *sitä*

9 *toki*

Koko lausetta määrittäviä pragmaattis-modaalisia aineksia ovat myös erilaiset liite- ym. partikkelit. Näiden koodittamista pidimme tärkeänä siitä syystä, että halusimme saada talteen mahdollisimman tarkoin sen, kuinka usein eri tekstifunktiot on osoitettu jollakin morfologisella tai leksikaalisella merkillä. Tämäntyyppiset partikkelit ovat neutraalissa kirjoitetussa kielessä melko harvinaisia, mutta niitä sentään esiintyy. Esimerkkejä -kin -partikkelin käytöstä:

## 2. verbi + *kin*

Vaikka näyttelyn nimessä *onkin* sana *sexualite*, kukaan katsoja tuskin odottaa näkevänsä Pigallen tyylistä seksishowta -- 57.14

Täydellinen kaksineuvoisuus *lieneekin* sukupuolten tasaveroisuuden ainoa ehdoton ja lopullinen muoto. 115.8

Sinne on menossa niin suuri seurakunta, *ettei* siellä ilmeisesti paljon muuta *ehditäkään* tehdä kuin kuunnella toinen toisiansa. 83.4

## 3. liikkuva *-kin* alussa

*Hekin* käyvät koulu.

Enempää en siitä ainakaan tällä erää saa irti, joten *lempi aiheellenikin* jää tilaa. 80.15

Ne eivät ole korjattavissa ykskaks, mutta *kehnoissakin hökkelissä* voidaan vielä asua jos pettäviin paikkoihin pannaan jokin tukeva lankku. 87.20

## 4. liikkuva *-kin* lopussa

Vientiä tarvitaan, jotta voidaan tuoda, mutta yhtä ja toista voidaan saada aikaan *kotimaisista raaka-aineistakin*. 36.7

Vain palkkatyöläinen ei voi, jos voisi niin kai pinnaisi *hänkin*. 84.5

Lahjoja joskus tarjotaan *toimittajapahasellekin*, saati silmäntekeville. 70.11

Samaa kuuluvat miettineen *jotkut muutkin*. 89.15

## (14) *kielto*

- 1 myönteinen
- 2 kieltosanakielto
- 3 *tuskin, tokkopa*

Vaikka kiellon taajuutta onkin suomessa laskettu eri korpuksista (ks. esim. Savijärvi (1977)), se otettiin huomioon tässä systeemissä jo senkin vuoksi, että sitä saatetaan tarvita erilaisissa modaalisten ainesten yhteistarkasteluissa. Varsin vähämerkityksiseksi osoittautui alakategoria 3, joten kaksijakoisella muuttujalla olisi hyvin tultu toimeen.

## (15) *lauserakenne*

- 1 aktiivi transitiivi
- 2 aktiivi intransitiivi
- 3 passiivi transitiivi
- 4 passiivi intransitiivi
- 5 omistusrakenne
- 6 tulorakenne (*pojasta tuli mies*)
- 7 sää, tila (*sataa*)
- 8 psyykinen kausatiivi (*minua kaduttaa*)
- 9 idiomi, kangistuma
- i fragmentti (rakenne ei tunnistettavissa)

Peruskiinnostus oli tämän muuttujan kohdalla syntaktinen eikä tekstuaalinen: haluttiin selvittää aktiivi- ja passiivilauseiden keskinäiset taajuussuhteet tekstissä ja katsoa, miten usein kumpikin perustyyppi on käytössä transitiivi- tai intransitiiviverbistä. Näiden lisäksi muuttujaan otettiin eräitä sellaisia rakenteita, jotka eivät alistu aktiivi-passiivi -vaihtelulle ja joille on ominaista joko subjektin tunnusmerkitön lauseenloppuisuus (5-6) tai sen puuttuminen (7-8). Tämä jaottelu on viety pitemmälle teoksessa Hakulinen & Karlsson (1979, luku 6), joka ei vielä ollut valmis tätä muuttujaa suunniteltaessa.

Huomattakoon, että eksistentiaalilauseetta ei ole sijoitettu omaksi alakategoriakseen, vaikka sitä on pidetty jopa toisena perustavana lausetyyppinä "normaalilauseen" ohessa (Hakanen (1972)). Projektin alkuvaiheessa eksistentiaalilause olikin mukana omana kategorianaan. Mutta pian kävi ilmi, että sen yksiselitteinen tunnistaminen on mahdotonta. T. Itkonen (1974-75) ja F. Karlsson (1978) ovatkin todenneet, että eksistentiaalilauseeseen kolme muotokriteeriä (subjektin sija, sanajärjestys ja verbin laatu) ovat kaikki jossakin määrin toisistaan riippumattomia, ja jos vähintään kaksi näistä kriteereistä ei esiinny samanaikaisesti, on vaikea puhua yksiselitteisestä eksistentiaalilauseesta. Tämänkertaisen tekstmateriaalin kohdalla tunnistaminen olisi ilmeisesti ollut vieläkin vaikeampaa, sillä varsin harvinaisia ovat sellaiset tyyppilliset tapaukset kuin *kaulassani riippui nauhuri, hampaiden välissä keinui muovikassi*; tekstimme ovat paljon abstraktimpia, "eksistentiaalilauseet" enemmän seuraavantapaista:

*Talous- ja sosiaalipoliitikkojen piirissä syntyi voimakkaastikin esitettyjä ajatuksia välttämättömyydestä muuttaa maatalouspolitiikan suuntaa. 10.57*

Emme siis koodittaneet eksistentiaalilauseetta omaksi lausetyypikseen. Sen sijaan tyyppilliset eksistentiaalilauseet kerääntyvät verkkoon, jonka silminä ovat sellaiset tuntomerkit kuin sanajärjestystyyppi XVS, jossa X on lokaalisytyppinen valenssiadverbiali ja jonka subjekti on siis lykätty, erityisesti jos subjekti on partitiivissa. Sen sijaan eksistentiaalilauseelle läheistä sukua oleva omistusrakenne (*Hänellä on auto*) on syntaktisesti kangistuneempi ja merkitykseltään selvästi erottuva, joten sitä voidaan käsitellä omana lausetyypinään.

Transitiivi- ja intransitiivilauseiden erottamisessa on noudatettu muotoperustetta: jos lauseessa on objekti, se on transitiivilause. Epäröintiä voisivat aiheuttaa "ymmärretyn" objektin sisältävät lauseet, sellaiset kuin *Jussi syö*. Tämän tyyppistä lausetta ei kooditettaisi transitiiviseksi; transitiivinen olisi vain lause, jonka tekstiyhteydestä selviää että siitä on objekti poistettu:

*Pioneerijärjestö arvostaa 0 / ja kunnioittaa lapsia, -- 89.35*

*Siellä oli kolmisenkymmentä tenavaa, vanhimmat tytöt jo niin aikuisen näköisiä etten 0 ohjaajista erottanut, -- 89.26*

*Siinä sitä on, ojenna kätesi ja ota 0. Ei, äläpä ota 0, muista, ettei kukkarossasi ole tarpeeksi rahaa. 85.25-26*

Myös generisen nolla-objektin lauseet ovat transitiivisia:

*Turkan ohjaus sivaltaa 0. 60.5*

Periaatepäättöstä tarvitaan myös aktiivin ja passiivin erottamiseksi. Mikäli yksipersonaiset modaaliset apuverbit *täytyy, pitää, kannattaa*, jotka voivat saada vain genetiivisubjektin, ovat ilman subjektia, on vaikea sanoa, onko kysymys generisestä nolla-subjektista vai passiivista. Tällaiset tapaukset on aina kooditettu passiiviksi. Sitä voi pitää tuntomerkitömmänä ratkaisuna sikäli, että yksipersonaiset verbit eivät pysty ottamaan passiivin morfeemia enempää kuin mitään muutakaan persoonamorfeemia, ja siksi ne voivat ilmaista "passiivisuuden" vain subjektittomuudella (Hakulinen (1973)). Semanttisesti raja passiivin ja generisen nollan välillä

häviää joskus olemattomiin. Esimerkiksi seuraavista lauseista ensimmäinen "tuntuu" enemmän passiivilta kuin toinen:

*Tämmöinen asia olisi pitänyt alistaa vähintään johtokunnalle tai oikeastaan hallintoneuvostolle. -- 84.17*

*Hotellihuoneen seinällä luki ihan selvästi, että jos tulee maanjäristys, pitää pistää huoneessa oleva ämpäri päähän ja mennä palotikkaille. 81.13*

Lauserakenteet ovat kieliopista tuttuja konstruktioita, joten ne tuskin kaipaavat esimerkkejä. Esitettäköön vain tapauksia, jotka on kooditettu joko idiomiksi taikka fragmenteiksi.

9. *idiomi, kangistuma*. Verrattuna seuraavaan kohtaan erona on se, että idiomiksi on syntaktisesti analysoitavissa. Sen sijaan näitä ei voi luokitella edellisiin lausetyyppeihinkään, koska niiden syntaksissa (esim. sanajärjestyksessä) ei juuri voi soveltaa tavanomaisia variaatiomahdollisuuksia. *Kysymys on rikkihappotehtaan suunnittelusta. Kysymyksessä ovat uudet lukutaidottomat. Näin kuitenkin ei ole asian laita. Todellisuudessa pata kattilaa soimaa. Sarbulla on pullat hyvin uunissa. Näin on marjat.*

i. *irrelevantti, fragmentti*. Tähän kuuluvat tapaukset, jotka ovat syntaktisesti analysoimattomissa, esimerkiksi siten että ei voida nähdä niihin sovelletun säännönmukaisia ellipsejä. *Selvä se. -- entäs sitten. -- miten päin vain. Oli miten oli.*

(18) *pintajärjestys*

1 SVX

2 SXV

3 VSX

4 VXS

5 XSV

6 XVS

7 SV

8 VS

9 VX

A XV

B V

i

Sanajärjestys on kooditettu kahdella tavalla. Muuttujat (15) ja (54) antavat kuvan siitä, paljonko perusjärjestystä on muunneltu tekstin vaatimusten mukaisesti. Tällä muuttujalla (18) tarkastellaan lauseen pääjäsenten keskinäistä pintajärjestystä.

S = subjekti (nominisubjekti, genetiivimuotoisenakin; geneerinen 0; elliptisesti poistettu subjekti; (N+) lause; infinitiivisubjekti)

V = verbin finiittinen osa, siis verbiketjun ollessa kyseessä sen ensi osan sijainti; ei-finiittisissä rakenteissa partisiippi- tai infinitiivimuoto.

X = verbin muu kiinteä aktantti, siis objektin, predikatiivin ja valenssiadverbiaalin yhteisnimeke. Kun lauseessa on sekä objekti että valenssiadverbiaali (pitää **jotakin tärkeänä**, otetaan huomioon objektin asema; kun lauseessa on sekä

predikatiivi että valenssiadverbiaali, predikatiivi (*hänestä* tuli **suosikki**), ja kahdesta valenssiadverbiaalista merkitään "olennaisemmalta" tuntuva (joutua *jostakin johonkin*: tulosija on primaarinen).

Esitutkimuksessa otimme huomioon useampia tekijöitä, kuten verbiketjun eri osien sijainnin sekä parin lauseadverbiaalin sijainnin. Näin saimme muuttujaan peräti 50 alakategoriaa. Tämä määrä oli paitsi kooditettaessa varsin epätaloudellinen ja mutkikas, myös sellainen, että useimpiin alakategorioihin tuli hyvin harvoja tapauksia. Tarkistuksen jälkeen päädyimme nykyisiin 11 alakategoriaan, joihin mahtuvat kaikki keskeiset tapaukset. Lauseadverbiaalien sijoittumista tämä ei tosin palvele riittävästi, sillä koodituksessa otamme toisaalla mukaan vain alku- ja loppuadverbiaalin, joten lauseen sisäisten adverbiaalien laatu ja sijainti esim. verbiketjuun nähden jää tämän kooditusjärjestelmän ulkopuolelle. Tarkka adverbiaalitutkimus vaatisikin oman erillisen muuttujistonsa.

Aidosti subjektittomat lauseet: passiivit, subjektina olevat infinitiivirakenteet sekä lauserakenteet (15:6-8), joista subjekti puuttuu, ynnä tarpeen mukaan idiomit, merkitään rakenteilla VX ja XV, missä X on tärkein kiinteä konstituentti. Fragmentit ovat sanajärjestykseltään irrelevantteja, idiomit sen sijaan eivät. Aidosti komplementittomat lauseet (*Pekka nukkuu*; *Syntyi riita*) merkitään rakenteilla SV ja VS, joista edellinen on oikeastaan tyyppin SVX alatapaus. Pelkkä V tarkoittaa passiivia, johon ei liity kiinteää aktanttia (*Juostaan*).

Silloinkin kun pintasubjekti puuttuu generisen nollan ja 1. ja 2. persoonan verbiin inkorporoitumistapauksissa, rakenne merkitään SVX:ksi. Toisin sanoen katsotaan, että subjekti "piilee" nimenomaan neutraalilla paikallaan tapauksissa tyyppiä *siitä pidän kovasti*, joka olisi XSV ja *Suomessa joutuu aina saunaan* olisi SVX. Ainakin generisen tapauksessa tämä ei ehkä ollut onnistunut kooditus, vaan olisi tällöin pitänyt käyttää tyyppiä XV / VX. Tämä ratkaisu on muistettava tulkittaessa sanajärjestystuloksia: se vinouttaa niitä jonkin verran SV-tyypin suuntaan.

Elliptiset rakenteet merkitään järjestykseltään sellaisiksi kuin ne kontekstin perusteella täydentyvät. Elleivät ne ole tällä tavoin täydennettävissä, ne voidaan merkitä sanajärjestykseltään irrelevantteiksi. Muuttujat (19) ja (59) osoittavat sitten, mitkä osat ellipsi on poistanut. Esimerkkejä:

Voiko ihmisellä olla lempi-itikkaa? *Miksei voisi*. 90.1-2 XVS: *Miksei* [IHMISELLÄ] *voisi* [OLLA LEMPI-ITIKKAA].

Se vikeltää matolla *kuin kalanpoikanen liian matalassa vedessä*. 90.8 SV: -- *kuin kalanpoikanen* [VIKELTÄÄ] *liian matalassa vedessä*.

Toukan näköinen se kyllä on aikuisenakin, *tai aikuisen näköinen toukkana, miten päin vain*. 90.12 XSV: -- *tai aikuisen näköinen* [SE ON] *toukkana* -- *miten päin vain* = fragmentti, sanajärjestys irrelevantti

Koska koodittaminen lauseadverbiaalit sivuuttaen saattaa olla hankalaa, esitetään lopuksi kustakin rakenteesta joitakin tyypillisiä kooditusratkaisuja.

### 1. SVX (neutraali pintajärjestys)

Maallikon silmään *juttu näyttää hullulta*.

-- *heidän pitää saada erikoisalennuksia* osuuskaupasta ostaessaan.

Nyt *O olen* vähän odotellut, *että* –  
*O en ole lainoppinut*, enkä *O pysty sanomaan, vastustiko hän* --  
Kesälomalaisena *O katselin, miten* --  
-- ja tästä *O olen ollut* hänelle kiitollinen.  
Ollaanpa taas tilanteessa, jossa *O toivoisi* työväenluokalle *enemmän voimaa ja valtaa* --  
Vahingollista on se, ettei *tieto* läheskään aina *ole oikeaa*, vaan *O johtaa O* harhaan.

## 2. SXV (harvinainen)

### a) *verbin siirto loppuun* (vrt. muuttuja 57:2):

Yksiviivaiset ratkaisut toimistopöydän äärellä eivät välttämättä johda toimiviin tuloksiin, vaikka *teoriat niin väittäisivätkin*. 48.20

-- kuluttajaväestöllä ei ole varaa hallituksen määräämiä korotuksia suurempiin korotuksiin vain sen takia, että *Valio tai MTK niin päättävät*. 39.19

Hallitus puolestaan tietää vanhastaan, ettei *tosiasiallisten kustannusten unohtaminen niitä* mitätöi --31.15

-- kuten *K. Goldstein asian tulkitsee*. 12.34

Harvoin pääsee kokemaan elämän rytmiä niin juhlallisena kuin hoitamattomassa ikimetsässä - mikäli *O sellaisen* Suomessa nykyisin *löytää*. 9.22

Jonkinlaisena piilokuvana kai tuo vallankumous täytynee olla niilläkin kommunisteilla, *joita* ministereinä *porvarillista yhteiskuntaa ylläpitävät*. 77.10

### b) *subjektin topikaalistus*:

Helsingin seudulla kerätään laatikoilla, *joita hänellä on* 1300 kappaletta.

### c) *O-subjekti*:

Vaan kun vierellä käryttää ei piippu eikä paperossi vaan tehtaan korsteeni niin milläs *O sen* hoitelet?.

Kun *O sitä koettaa käsittää* etsimällä vertausta joltakin vähän alemmalta tasolta --

## 3. VSX

### a) *kO-kysymys* (automaattinen V-siirto alkuun; samoin V+hAn)

-- enkä pysty sanomaan, *vastustiko hän maanlunastuslain vahvistamista* juristin omantunnon vai poliittisten päämäärien takia.

### b) *V-siirto* (inttävä tms.)

Mutta *ei se ole ainoa ratkaisu*.

*Ei se, että ulkomaille ei saada kaupaksi yhtä paljon kuin pari vuotta sitten*, merkitse *sitä, että täällä täytyy ruveta kädet ristissä kärsimään nälkää*.

### c) *inversio*

Tässä astuu risteilyohjus kuvaan mukaan –

Ilman sukupuolisuutta *ei puu eikä ihminen* olisi kansoittanut *kivennäisten synkkää maailmaa*.

#### 4. VXS

a) *eks-sanajärjestys + V-siirto alkuun* (kooditettu: subjektin lykkäys 56:1)

Muutaman kuukauden kuluttua *ilmestyy johtokunnan puheille pari ravintoloitsijaa, jotka --  
Voiko ihmisellä olla lempi-itikka?*

b) *ekstrapositio tai vastaava* (56:3)

Minusta näyttää *valheelliselta, että* kun yksi toheloi itsensä kiinni, siitä pidetään kauhea haloo.

Meille suomalaisille *on erityisen imartelevaa se, että -*

Mutta *ei liene liioiteltua sanoa hänen teoriansa kokoavan yhteen ja jäsentävän -- 2.45  
On nimittäin mahdollista, että --*

#### 5. XSV

a) *komplementin relatiivisuus* (automaattinen)

Eivät he ole henkilökohtaisesti niin riippuvaisia ilman puhtaudesta kuin se, *jonka omakotitalon pihaan piipun varjo ulottuu.*

-- ja puuttunut henkilökohtaisesti sellaisiin virkamiesten ylitsekäymisiin, *joista lehtemme on kertonut.*

b) *komplementin topikaalistus*

Mutta *siipiä se ei saa* koskaan.

*Kaskuista pilasadut* ensisijaisesti *eroavat* siinä, että -- 6.19

-- Englanti ja Ranska jakoivat keskenään Afrikan ja Aasian vielä vapaina olevat alueet. *Tämän ne usein tekivät* siellä jo omistamistaan sillanpääasemista lähtien.

#### 6. XVS

a) *eksistentiaalilauseen järjestys*

*Sillä on litteä, sukkulainen, taaksepäin kapeneva ruumis,--*

Vaan kun *vierellä käryttää ei piippu eikä paperossi vaan tehtaan korsteeni* niin milläs sen hoitelet?

Talouselämän *rakenteissa täytyy olla vikaa* silloin, kun ne estävät ihmisiä tekemästä työtä--

b) *subj./obj. -vaihdos*

*Porvarivaltaa varmistaa eduskunnan porvarienemmistö.*

*Minätietoisuuden nopeaa kehitystä* kuvaa esim. *Dragunovan tutkimus, jossa eri-ikäisille koululaisille annettiin tehtäväksi... 2.31*

c) *subj./predvi -vaihdos*

*Merkillepantavaa on se, että tämä "vilkastumisilmiö" ilmaantuu -- 2.16*

*Ajatellaan, että vertauskohdaksi sopii vaikkapa osuuskauppa,--*

*Hyvä on, että kokouksissa keskustellaan.*

*Näin yksinkertainen ei asia ehkä sittenkään ole. 3.16*

d) *muu subjektin lykkäys*

*Joo, siitä on kirjallinen sopimus, sanovat kaverit.*

*Silverfish, hopeakala, sanovat englantilaiset sokeritoukkaa, ja hopeakalasesta puhuvat saksalaisetkin.*

e) *johtolause*

*Joo, siitä on kirjallinen sopimus, sanovat kaverit.*

7. *SV (neutraali järjestys, ilman kiinteää komplementtia)*

*Johtokunta hämmästy.*

*Kuva väkisinkin muuttuisi, jos --*

*Lukemismateriaalin puuttuessa hankittu lukutaito unohtuu.*

*Kenties siellä kuitenkin jotakin ymmärtämystä syntyy.*

8. *VS*

a) *subjektin lykkäys (olla, löytyä, syntyä jne. ks. 56:1)*

*Eipä ollut montakaan, joka ei ilmoittautunut --*

*Enemmän kannattaisi kohista, jos löytyisi yksikin firma, joka maksaa veronsa --*

*Kun sitä koettaa käsittää etsimällä vertausta --, syntyy esimerkiksi tällainen tarina.*

b) *V-siirto*

*Taisivat ne vieraamme ulkopuolelta vähän hämmästyä.*

*Ei tämä päde, sanoo johtokunta.*

9. *VX*

(Subjektiton lause, jonka pääaktantti on verbin jäljessä. Verbin eteen joko jää toinen konstituentti, joka kooditetaan topikaalistetuksi, tai paikka jää tyhjäksi, mikä käy vain joissakin lausetyypeissä.)

*Ajatellaan, että --*

*Ollaanpa taas tilanteessa, jossa toivoisi --*

- *Kävi niinkuin myymäläsuunnittelijat sen olivat ajatelleet.*
- *hänestä oli tuleva Suomen pitkäaikaisin maaherraehdokas.*
- Eihän siitä yleensä Helsingissä muulloin piitata kuin vaivoin vaalipäivinä.*

A. XV (subjektiton lause, jonka pääahtantti on verbin edellä; se kooditetaan topikaalistetuksi)

- *miten oikeuskansleri Risto Leskistä höykytettiin myös Kansan Uutisissa.*
- kun valtion vuosittaisia urheilumäärärahoja jaetaan.*

*Koska kaikenlainen kuntoilu on laskettava myös sairauksia ennalta ehkäiseväksi toiminnaksi, koko rahoitustoiminta olisi järjestettävä terveydenhoitovaroista.*

*Esitystä oli helppo / perustella.*

B. V

*Puhuttiin avomielisesti.*

*Hyvä on, että kokouksissa keskustellaan.*

## 2.5. Konstituenttien rakenteita koskevat muuttujat

subjektin rakenne (19,20)

objektin rakenne (27,28)

predikatiivin rakenne (33,34)

alku- ja loppuadverbiaalin rakenne (38-40,45-47)

verbin rakenne (25,26)

Käsitlemme seuraavaksi yhdellä kertaa eri konstituenttien rakenteeseen liittyvät muuttujat, koska niissä on osittaista päällekkäisyyttä ja koska niitä koskevat ratkaisut ovat samansuuntaiset.

Kaikki subjektia koskevat muuttujat merkitään 0:lla, jos lause on aidosti subjektiton, toisin sanoen jos kyseessä on jokin muuttujan (18) alakategorioista 3,4,6,7,8, taikka imperatiivilause tahi lause on upotettunatoiseen infinitiivisubjektiksi (tällöin voi tosin inf.rakenteelle merkitä elliptisen subjektin). Muussa tapauksessa subjekti kooditetaan kaikkien otsikossa mainittujen muuttujien mukaan. Silloin kun muuttuja ei sovellu, merkitään alakategoriaksi irrelevantti (i). Subjektin rakenteessa tämä ei tule mainituista syistä kyseeseen: lause on joko subjektiton tai subjektillinen.

(19) *subjektin rakenne*

1 rinnasteinen

2 lause

3 pääsana + lause

4 infinitiivi

5 nominaalistuma

- 6 N + selittävä appositio, *kuten tjs.*
- 7 N + adv:li tai inf.
- 8 partisiippi + N
- 9 attrib. (yksi tai us.) + N
- A N
- B inkorporoitu verbiin
- C geneerinen 0
- D ellipsi

Kaikki lausekkeiden rakennetta koskevat muuttujat (subjektin, objektin, predikatiivin, adverbiaalinen rakenne) ovat sellaisia, joiden alakategoriat eivät ole toisensa poissulkevia. Voihan nimittäin rinnasteiseen subjektiin sisältyä attribuuttejakin tai samassa NP:ssä voi olla sekä etuattribuutteja että rinnastuksia. Jos välttämättä haluaa alakategoriat toisensa poissulkeviksi, joutuu tekemään joko hyvin karkean jaottelun (esim. "pelkkä sana", "lauseke", "lause" ja "0") taikka sitten niin hienojakoisen, että se ei enää ole tulosten kannalta tarkoituksenmukainen. Olemme päätyneet kompromissiin siten, että karkeaa jaottelua on eritelty jossakin määrin edelleen: niinpä 0-subjekteja on kolmea tyyppiä (B, C ja D), lausemaisista subjekteista neljää tyyppiä (2-5) ja lauseke-NP:t on jaettu viiteen tyyppiin (1, 6-9). Alakategorioiden numerointi edellyttää myös niiden keskinäistä hierarkiaa: silloin kun tarkasteltavaan subjekti-NP:hen sopisi kaksi kuvausta, valitaan se, jolla on pienempi numero. Tässä muuttujassa ei ole kiinnitetty huomiota siihen, minkä sanaluokan sana yksisanaisessa subjektissa esiintyy. Ristiintaulukoimalla kategoriat 19:A ja 23:2,3,9 saadaan selville, kuinka usein yksisanaisena subjektina on pronomini. Esimerkkejä:

#### 1. rinnasteinen

Hallituksessa on nyt, *5 kepulaista, 4 demaria, 3 kansandemokraattia, 1 liberaali, 1 ruotsalainen ja 1 puolueeton porvari.* 86.2

Ellei niihin ryhdytä, tulevat *sekä miljöön että saaristoasutus katoamaan.* 104.25

#### 2. lause

Hyvä on, *että kokouksissa keskustellaan.* 92.46

Juuri tästä johtuu, *että tavallisimmat eli magneettiartikkelit on sijoitettu aivan myymälätilan äärimmäiseen päähän.* 101.25

#### 3. pääsana + lause

-- vertauskohdaksi sopii *vaikkapa osuuskauppa, demokraattisesti hallittu yhteisö joka harjoittaa taloudellista toimintaa* 84.7 myös: 6

Elinkustannusindeksi lasketaan sellaisen kulutustavan mukaan, johon *kuuluu paljon sellaista mihin kansaneläkeläisellä ei ole koskaan varaa.* 85.6 myös: 9

*Niiden tavarain hinnat, joita kansaneläkeläiset ostavat, ovat nousseet kaksi kertaa kiivaammin kuin hinnat keskimäärin, --* 85.7 myös: 9

#### 4. infinitiivi

Ei, tarkoitus on *pyydystää asiakas liikkeeseen.* 101.68

On parasta tehdä lista tavaroista, joita tarvitsee --. On viisasta varata rahaa mukaan tietty pienehkö summa. On kohtalokasta mennä valintaliikkeeseen vain katsomaan, mitä siellä on. 101.73-75

#### 5. nominaalistuma määritteineen

Seinien maalailu sen sijaan ei auta. 87.29

Näillä hinnoilla loppui puolukan myynti tähän, -- 98.58

-- että elävän saariston säilyttäminen kansallisena luonnonvarana vaatii kiireellisiä toimenpiteitä. 104.2

Näin elokuvan tutkimuksen ja opetuksen (sikäli kuin lehtoraatista on kysymys, lähinnä jälkimmäisen) aseman vakiinnuttaminen jäi edelleen odottamaan -- 3.7

#### 6. nominaalilauseke ja selittävä attribuutti

Minun erityinen suosikkini itikoiden parissa on sokeritoukka, *Lepisma Saccharina eli "sokerinen suomukas"*.

-- että tulevien sopeutumisten ainoa lähde, geneettinen moninaisuus, köyhtyy ja katoaa. 95.44

-- sanoo Itä-Helsingin kansalaisopiston kannatusyhdistyksen puheenjohtaja, teollisuusneuvos Launo Laakkonen. 96.20

#### 7. nominaalilauseke ja adverbiaali tai infinitiivi

Aiheena oli työväenpuolueiden yhteistyö Helsingin politiikassa -- 86.12

Se vanha juttu kapteenista ja perämiehestä on aika hyvä. 89.1

Kehitysmaiden mahdollisuudet saada tieto omista tapahtumistaan välittymään laajalle on kovin rajattua. 92.37

#### 8. partisiippi (määritteineen) nominaalin attribuuttina

Ensimmäinen osa sai minut ajattelemaan, että koululaitos kaipaa peruskorjausta, nyt näkemäni kolmas nosti ihon kananlihalle. 89.21

Turkkilaisen valtionyhtiön Etibankin kanssa allekirjoitettu sopimus oli noin 80 miljoonaa markkaa. 93.27

Laajin hanke, jossa suomalaiset ovat mukana, on samaan kaupunkiin rakenteilla oleva voimala. 93.32

#### 9. muu attribuutti + nominaali

Lasten ja nuorten pahat teot ovat meidän aikuisten syytä. 89.34

Siellä oli kolmisenkymmentä tenavaa, vanhimmat tytöt jo niin aikuisen näköisiä, etten ohjaajista erottanut -- 89.26

#### A. yksi sana

Ihmiset ovat nyt niin kummissaan Saloran veropetoksesta ja siitä, että poliitikot ottavat lahjoja. Eivätkö he siis sitä ennen tienneet? 89.4

## B. inkorporoitu verbiin (1. ja 2. pers.)

Kesälomalaisena katselin, miten oikeuskansleri Risto Leskistä höykytettiin myös Kansan Uutisissa. *En ole lainoppinut enkä pysty sanomaan,--*

Ostoksia kantaessaan nainen mietiskeli, että jälleen kerran hän oli romahtanut halpaan. 101.20

## C. geneerinen 0

Vladimir Jashtshenko, 18, ei varmaan aavistanut, mitä tapahtuu jos 0 hyppää ensimmäisenä maailmassa 233 senttiä korkeutta. 99.1

Alkoholismi on sairaus, josta 0 voi toipua vain täysraittiuden avulla. 123.45

Uskaltaisikohan 0 vielä kysyä mahtaako metsästäjän iloon joskus sekoittua murhamiehen riemua? 73.38

## D. ellipsi

Hän katsoi laskua ja 0 löi sitten kauhistuneena otsaansa: -- 84.26

Tämän tehtaan on suunnitellut 0 ja 0 0 rakentanut Tampella. 93.31

## (20) subjektin sija ja kvanttori

- 1 nominatiivi
- 2 partitiivi
- 3 genetiivi
- 4 lukusana (+N)
- 5 kvanttori *paljon, vähän, liikaa* (+N)
- 6 *kaikki* (+N)
- 7 *eräs, harva, joku, moni* (+N)
- 8 lohjennut partitiivisubjekti
- i irrelevantti

Subjektin sijanvaihtelun tarkkäileminen on tekstuaalisestikin tärkeää, sillä nominatiivin ja partitiivin vaihtelu kytkeytyy läheisesti kvantitatiiviseen määräisyyteen sekä tiettyihin lausetyyppeihin ja sanajärjestyskaavoihin. Nominatiivin ja partitiivin lisäksi olemme ottaneet kolmanneksi sijavaihtoehdoksi genetiivin, esiintyyhän lauseenvastikkeen subjekti usein genetiivimuotoisena. Mutta pidämme genetiivisubjektina myös ns. yksipersonaisten verbien yhteydessä esiintyvää lauseenalkuista nominaalijäsentä (*Minun täytyy mennä*). Verbin ja infinitiivin muodostama jakso tulkitaan verbiketjuksi (*täytyy mennä*), jonka subjektina ao. genetiivimuotoinen nominaali on (vrt. Hakulinen & Karlsson (1979:264-5)). Kun subjektin sijanvaihtelu on olennaisesti kvantifioinnin ilmaisukeino, on syytä samalla tarkastella myös varsinaisten sanakvanttorien käyttäytymistä. Jälkeen päin ajatellen näyttää siltä, että muuttuja (20) olisi kannattanut jakaa kahtia ja sijoittaa sanakvanttorit omaksi muuttujakseen, taipuvathan alakategorian 7 kvanttorit sekä lukusanat itsekin lauseen kokonaisuudesta riippuen nominatiivissa ja partitiivissa.

### 1. nominatiivi

*Mehiläinen* saa talvella ravintoa yhteiskunnan varastosta. 21.20

## 2. partitiivi

Afrikan savanneilla asustaa *suuria sorkkaeläinlaumoja* -- 21.1

-- ns. kuhnureita tarvitaan hedelmöittämään parittelulennolla uudet kuningattaret, *joita* syntyy erityisravinnolla ruokituista toukista. 21.11

-- ettei *luottamuspulaa* pöytäkirjan johdosta synny.

## 3. genetiivi

-- suojelun alueen japaninmakakit oppivat nopeasti pesemään niille annetut hiekkaiset bataatit nähdessään *toisten* tekevän niin, -- 21.21

*Ravinnon* ollessa vähäistä lähekkäiset yksilöt joutuvat kilpailemaan siitä kiivaammin kuin erillään asuvat. 21.35

*Hallituksen* onkin yhtenäisesti pidettävä huoli siitä, ettei *luottamuspulaa* pöytäkirjan johdosta synny -- 26.17

Miten tuhoisia *vaikutusten* pitää olla ennen kuin myrkyt kielletään? 28.8

## 4. lukusana (+ N)

-- yhteensä *288 henkilöä*. Näistä *192* toimi koeryhmissä ja *96* kontrolliryhmissä. 8.35  
filmissä oli keskeisissä osissa *kuusi henkilöä*, joista -- 8.18

## 5. paljon-tyyppi (epämääräistä paljoutta osoittava kvanttori joka aina vaatii pääsanansa partitiiviin: esim. *paljon, vähän, runsaasti, liikaa, tarpeeksi, koko joukko, hiukkasen; puolet, loput*)

*Yhtä vähän samankaltaisuutta* näyttää olevan kehityksen teknis-operatiivisten jaksujen johtavilla toiminnoilla -- 2.40

-- talonpoikia, joilla oli *vähän kosketusta* kaupunkien kulttuuriin. 5.24

*Paljon hikeä* on virrannut uudisasukkaan pyrkiessä muokkaamaan vastahakoisesta luonnosta anteliaampaa ympäristöä. 9.80

Suunnan muutoksen vaatijoihin liittyi kyllä myös *koko joukko maatalouden asiantuntijoita*. 10.58

## 6. kaikki (+ N)

Primitiivisessä yhteiskunnassa *kaikki musiikki* on "kansanmusiikkia". 5.13

*Läheskään kaikki elokuvat* eivät ensinnäkään läpäise sitä sosiaalista kontrollia, -- 3.18

## 7. kongruoivat kvanttorit (heterogeeninen ryhmä *joku, jokainen / (ei) mikään, (ei) kukaan, joka, jokin, kukin, mikin, kumpikin, eräs, moni, harva, usea*)

On *monia erilaisia jouduttamisen välineitä* ja pamflettikirjallisuus on näistä yksi. 4.6

-- *monet jäljet* vievät luolaan, mutta *yhdetkään* eivät tule ulos. 6.6

Niiden aiheisiin pohjaavat *monet Ovidiuksen metamorfoosit*, -- 6.10

Oleellista on, että *moni selkä* on köyristynyt ennen aikojaan. 9.79

Työttömyyden ilmaantuessa oli *kukin kunta* ensin velvollinen työllistämään oman osuutensa -- 10.8

-- *jokin muusikkoryhmä* erikoistuu palvelemaan yhteiskunnan johtavia ryhmiä. 5.20

Eli *jokainen kehitysaskel* ei ole välttämättä edistysaskel. 9.72

-- vaikka *mitään suoranaista saastutusta* ei tapahtuisikaan. 9.93

## 8. lohjennut partitiivisubjekti

Partitiivisubjektin (ja myös -objektin) rakenteellisiin erikoisuuksiin suomen kielessä kuuluu taipumus lohjeta kahtia, subjektin tapauksessa erityisesti *olla*-verbillisissä lauseissa. Partitiivisubjekti ja -objekti ovat tällöin (epämääräisesti) piilokvantifioituja, ts. ilmaisevat epämääräistä paljoutta. Subjektista lohkeava eli yleensä verbin jälkeen heittäytyvänä osana voi olla

(a) useampia adjektiiviattribuutteja tai yksi joka implikoi useutta:

Oppilaita on *ahkeria ja laiskoja*. Oppilaita on *ahkeriakin*. (= On (olemassa) ahkeria ja laiskoja oppilaita)

(b) kvanttori

Puutteita on, kuten sanottu, *paljon*. 54.83 (= Kuten sanottu, on paljon puutteita)

Tällainen lohkeaminen on yhteydessä lauseen tematiikkaan sikäli, että tällä keinolla ikään kuin tasapainotetaan lausetta neutraaliin SVX-kaavaan sopivaksi. Lohkeamat katsottiin soveliaiksi kirjata juuri tähän muuttuinaan, jossa oli muutenkin kvantifioinnista puhe. Toinen vaihtoehto olisi se, joka on esitetty Hakulisen & Karlssonin Nykysuomen lauseopissa (s. 98-9), nimittäin että kyseessä olisi kiteytynyt kvanttorilauseityyppi, joka kuuluisi muuttujan (15) alakategoriaksi. Todisteena siitä, että kvanttorilauseiksi kutsutut rakenteet ovat etäänntyneet puhtaista lohkeamista ovat tapaukset, joita ei voi palauttaa kaksikonstituenttiseksi lähtömuodoksi:

Tärkeintä on, että *kiellettyjä asioita ei ole niin monta*, että kissaa täytyy jatkuvasti komennella. 122.56

\*-- että ei ole niin monta kiellettyjä asioita --

*Niitä* on *paljon* säilynyt. 17.9

\*On säilynyt paljon niitä

Ohjaaja kertoi että on tilaisuus päästä alennuslipulla teatteriin, kuka lähtee? *Käsiä lensi ylös toistakymmentä* heti ja päätöksen myöhempään lykkääviä oli monta. 89.29 x -- heti lensi ylös toistakymmentä käsiä ja oli monta päätöksen myöhempään lykkääviä.

Lohjenneen kvanttorin lauseet olisi syytä koodittaa omaksi lauseityypikseen siitäkin syystä, että kvanttorilauseke voitaisiin tällöin koodittaa erikseen. Tämä on tärkeätä, koska se voidaan puolestaan edelleen kontrastiivisesti tai muusta syystä topikaalistaa:

Mitä kauemmin hän viihtyy, *sitä enemmän* kilisee pennejä kassaan. 101.65

*Kolme* meitä vain on (ei kaksi).

Kvanttorilauseityyppi on marginaalinen mutta ei mitenkään harvinainen - kvanttorilausekkeen topikaalistusta tosin tuskin esiintyy paljon. Ilmeinen tarve tälle konstruktiolle on siinä, että tällä keinoin definiittinen pronomini voidaan kvantfioida: *niitä on 15 vrt. 15 niitä*.

Olemme kuitenkin tässä kooditusjärjestelmässä vielä pitäneet tapauksia lohkeamien kanssa samanarvoisina ja koodittaneet kvanttorilauseet vain tähän muuttujaan. Tällöin on huomattava, että subjektin sanamäärään joudutaan laskemaan mukaan myös kvanttorilauseke. Huomattakoon, että tähän on kirjattu myös sellaiset tapaukset, joissa kvanttori-ilmaus muodostaa elatiiviattribuutin kanssa kokonaisuuden, joka sitten on lohjennut. Tässä tapauksessa lohkeaminen on erittäin yleistä:

Ovien ulkopuolelle *jää pelkästään tämän vuoden ylioppilaista / yli puolet / ja edellisinä vuosina jatkokoulutukseen sijoittumattomista ylioppilaista / lukumääräisesti paljon enemmän.* 29.2 (subjektina)

-- valtio korvasi *opiston toiminnan varsinaisista menoista* kahden ensimmäisen vuoden aikana *vain 35 prosenttia.* 96.24 (objektina)

Lähtömuotoina on pidetty siis seuraavia: *yli puolet pelkästään tämän vuoden ylioppilaista ja vain 35 % opiston toiminnan varsinaisista menoista.*

### (27) objektin rakenne

- 1 rinnasteinen
- 2 pääsana + lause
- 3 lause
- 4 nominaalistuma
- 5 N + selittävä appositio
- 6 N + adverbiaali t. inf.
- 7 partisiippi + N
- 8 attrib. (yksi t. us.) + N
- 9 N
- A geneerinen 0
- B ellipsi

Objektin rakenne noudattelee samoja linjoja kuin subjektinkin. Erot johtuvat objektin luonteesta. Ensinnäkin objektina ei tämän kooditustavan mukaan voi olla infinitiivi, koska objektina oleva infinitiivi on laskettu verbiketjun osaksi ja näkyy siis verbin rakenteesta (muuttuja 25). Edelleen objekti toisin kuin subjekti ei suomessa voi inkorporoitua verbiin.

#### 1. rinnasteinen

Tyypillisinä esimerkkeinä Elkonin mainitsee *Piaqet'n teorian kognitiivisesta kehityksestä ja Freudin teorian motiivien kehityksestä.* 2.9

#### 2. pääsana + lause

Vygotski ei siten hyväksynyt *monille 1920-luvun tiedemiehille ominaista reduktionistista ajattelua, jonka mukaan fysiologian tai jopa fysiikan avulla voitaisiin kattavasti selittää kaikki psyyken ilmiöt.* 1.6

*Vanhastaan tunnettu tosiasia, että lapsi reagoi emotionaalisesti, esim. hymyllä, ihmiskasvojen näkemiseen kolmannelta elinkuukaudestaan lähtien* on Elkoninin tulkinnan mukaan ymmärrettävä pikemminkin viestinnäksi kuin pelkäksi reaktioksi siihen voimakkaaseen ärsykkeeseen, mitä aikuinen edustaa. 2.16

#### 3. lause, lauseenvastike

Tässä hän totesi mm. että laiminlyödessään tietoisuuden ongelman neuvostopsykologia jätti ovet auki kaikelle aikaisemman psykologian dualismille ja spiritualismille. 1.3

-- Elkonin -- katsoo sukupuolisen kypsyyden vaikuttavan vain välillisesti, lapsen ja ympäristön suhteiden kautta. 2.34

#### 4. *nominaalistuma* (määritteineen)

Vastaavasti teorian mukaan puberteetin alku merkitsee *uuden kauden alkamista* -- 2.44

Kun Suomen Elokuva-Arkisto ryhtyi ajamaan elokuvan aseman vakiinnuttamista Helsingin yliopistossa, -- 3.9

-- tarkastella, *minkäläisten tutkimusnäkökulmien omaksumisen* elokuva ylipäänsä sallii. 3.12

-- jos elokuvan tutkimuksen vakiinnuttamisen asemesta olisi vaadittu *populaarikulttuurin tutkimuksen aloittamista*. 3.14

#### 5. *nominaalilauseke* + *selittävä appositio*

*Apartheidia vastaan syntyneet protestiliikkeet*, mm. *Nobelin rauhanpalkinnon saaneen Albert Luthulin johtama African National Congress (ANC) ja Pan-Africanist Congress (PAC)* kiellettiin 1960 ns. Sharpevillen verilöylyn jälkeen, -- 16.38

#### 6. *N* + *adverbiaali tai infinitiivi*

Neuvostopsykologiassa, jossa tähdennetään *toiminnan merkitystä kehitystekijänä*, on varsin yleisesti hyväksytty Leontjevin esittämä johtavaan toimintaan perustuva jaksottelu. 2.4

*Erään painavimmista teoreettisista puheenvuoroista sosiaalistumisen vaiheiden ongelmaa käsittelevään keskusteluun* on äskettäin esittänyt Elkonin. 2.6

Samalla se luo *valmiuden puberteettikauden toisen jakson johtavaan toimintaan, ammatilliseen oppimiseen*. 2.37

#### 7. *partisiippi* (määritteineen) + *N*

Tietoisuuden semanttisella rakenteella Vygotski tarkoitti *käsitteiden merkitysten kehityksessä syntyviä kognitiivisia struktuureja*. 1.16

Neuvostopsykologiassa -- on varsin yleisesti hyväksytty *Leontjevin esittämä johtavaan toimintaan perustuva jaksottelu*. 2.4

#### 8. *attribuutti* + *N*

Tietoisuutta koskevassa tutkimustyössään Vygotski otti teoreettis-metodologiseksi lähtökohdakseen *marxilaisen filosofian ajatukset*. 1.7

#### 9. *N*

*Tietoisuutta* tuli Vygotskin mukaan tutkia juuri em. psyykkisen toiminnan spesifisten rakenteiden konkreettisen kehityksen kautta. 2.18

#### A. *geneerinen 0*

Turkan ohjaus sivaltaa 0. 60.5.

Tämä lamanäytelmä puhuttelee 0 lahjomattomuudessaan. 60.53

-- minulle tuli muuten kyyneleet silmiin. Eikä 0 edes hävettänyt. 63.30

Pois kotiseudulta ei haluttu lähteä eikä voitukaan aina, oma maa ja talo sitoivat 0. 10.44

## B. ellipsi

-- ei liene liioiteltua sanoa hänen teoriansa *kokoavan yhteen 0 ja 0 jäsentävän suuren osan* neuvostopsykologian sosiaalistumistutkimusta 2.45

Pääosa siirtomaissa muodostetusta pääomasta sijoitettiin joko palveluselinkeinoihin, maatalouden tuotantokustannusten pienentämiseen tähtääviin parannuksiin, siis voittojen vaan ei tuotannon lisäämiseen tai *0 vietiin pois maasta sijoitettavaksi Eurooppaan*. 7.60

Majava sen sijaan hävitettiin sukupuuttoon 1800-luvulla, mutta *0 istutettiin uudelleen 1930-luvulla / ja 0 on viihtynyt siitä lähtien hyvin /*. 9.49

### (28) objektin sija ja kvanttori

1 nominatiivi

2 partitiivi

3 genetiivi

4 -t

5 lukusana (+N)

6 *paljon, vähän, liikaa* (+N)

7 *kaikki* (+N)

8 *eräs, harva, joku, moni* (+N)

9 lohjennut partitiiviobjekti

i irrelevantti

Erona vastaavaan subjektin sijan kooditukseen on alakategoria 4, joka koskee pronominiobjektien akkusatiivimuotoa. Tässä yhteydessä on syytä esittää sellainen käytännön huomautus, että tällaista poikkeamaa harkittaessa on punnittava vastakkain kahta mutkistusta. Jos tätä ei olisi objektimuuttujana merkitty omaksi alakategoriakseen, koodittaminen olisi tekijälleen helpompaa kun olisi kaksi identtistä muuttujaa (kooditusvirheiden mahdollisuudet vähenevät). Tiedon pronominiobjektien määrästä olisi saanut tällöin esiin epäsuorasti eli ristiintaulukoimalla muuttujat (27:9) (yksisanaiset objektit) ja (31:2,9), josta saa pronominesintymät. Koodituksen helpottuessa tiedon saanti mutkistuu.

"Nominatiivilla" tarkoitetaan tässä paitsi yksikön nominatiivia myös monikossa sekä nominatiivina että ""akkusatiivina"" käytettyä muotoa. "Genetiivi" tarkoittaa puolestaan perinteellisesti päätteelliseksi akkusatiiviksi kutsuttua *-n*-päätteistä muotoa.

#### 1. nominatiivi

Samoin hän hylkäsi *refleksologian kaltaiset mekanistiset käyttäytymisopit*, -- 1.5

#### 2. partitiivi

-- jotka eivät tunnustaneet *tietoisuutta* tieteelle soveliaaksi tutkimuskohteeksi eivätkä edes psykologiaa tieteenä tarpeelliseksi. 1.5

### 3. *genetiivi*

*Tietoisuutta koskevan tutkimuksen merkityksen* hän otti esille jo eräässä varhaiskirjoituksessaan, -- 1.2

### 4. *-t*

### 5. *lukusana (+ N)*

Koehenkilöstä muodostettiin *kahdeksan koe- ja neljä kontrolliryhmää* -- 8.37

### 6. *paljon jne.*

Kaikesta huolimatta ankeriaasta tiedetään vielä *aivan liian vähän*, -- 102.6

Jos tietyllä summalla haluttiin sijoittaa töihin mahdollisimman *paljon väkeä* -- 10.23

Hoitometsät antavat *enemmän selluloosaa* hehtaaria kohti, -- 9.42

Yhteiskunta tarvitsee *erityisen paljon pamfletteja* -- 4.8

### 7. *kaikki*

-- ajattelua, jonka mukaan fysiologian tai jopa fysiikan avulla voitaisiin kattavasti selittää *kaikki psyyken ilmiöt*. 1.6

Järjestelmän mukaan edelleenkin siis kutsuttiin *kaikkia tätä tietä työllistettyjä* työttömiksi, vaikkapa he olisivat kuinka ankarasti tahansa juuri silloin tehneet työtä. 10.10

### 8. *eräs jne.*

*Erään painavimmista teoreettisista puheenvuoroista* sosiaalistumisen vaiheiden ongelmaa käsittelevään keskusteluun on äskettäin esittänyt Elkonin. 2.6

-- käytetään rinnan *useita väyliä ja välineitä joista pamflettikirjallisuus on vain yksi*. 4.30

*Mitään itsenäistä laajempaa, kehittyneempää teollisuutta* siirtomaihin ei yleensä luotu. 7.58

Suomalaisella työvoimalla on rakennettu myös *eräitä tilaustöitä* Neuvostoliiton puolella, -- 10.52

### 9. *lohjennut partitiiviobjekti*

Tähän sopii mutatis mutandis se mitä sanottiin lohjenneesta partitiivisubjektista kohdassa 20:8, vaikka yhtä selvää ei ole, olisiko tarpeen erottaa lohjenneen partitiiviobjektin sisältävät lauseet omaksi lausetyypikseen samalla tavalla kuin subjektin tapauksessa. Esimerkkejä:

Mutta *lehtipuita* siellä tapaa *niukasti*. 9.35

Kuntien koululaitokset tarvitsevat syksyisin *marjoja useita tuhansia kiloja*. 97.18

Tavallisempi lohkeama tuntuu olevan passiivilauseiden objekteissa:

Norjassa *hölmöläissatuja* on tavattu *yllättävän vähän, ainoastaan alle kymmenen tyyppiä*. 6.38

Kuluvan vuoden puoliväliin mennessä *jätepaperia* oli kerätty *noin 8,5 miljoonaa tonnia*. 96.22

### Elatiivirakenne:

-- että tuottajille voidaan tilittää *hallituksen lupaamasta korotuksesta vain puolet, pari penniä*. 31.12

Joskus raja lohjenneen kvantifioidun objektin ja yksinkertaisen objektin ynnä epämääräisen, "yleisen" OSMA:n (Tuomikoski (1978) välillä on hämärä ja varsin vaikea hahmottaa:

Suomalaiset lukevat sanomalehtiä kansainvälisestäikin ottaen paljon. (OSMA) 43.1  
passiivisen sanavarastonsa ansiosta vauva ymmärtää *puhettasi enemmän kuin itse* osaa ilmaista. 119.118 (OSMA)

*Nuoret* eroavat *eniten* (OSMA), vrt. *Nuoria* eroaa *eniten* (lohkeama)

Hän pystyy kertomaan *Korppoosta, sen oloista ja ihmisistä paljon*. 117.60

(33) *predikatiivin rakenne*

- 1 A
- 2 määrite + A, A + määrite
- 3 A + A
- 4 A + lause
- 5 vertailu (*kuin kukkanen*)
- 6 N
- 7 määrite + N, N + määrite
- 8 N + N
- 9 N + lause
- A lause
- B nominaalistuma
- c ellipsi

Predikatiivin rakenne poikkeaa tuntuvasti subjektin ja objektin rakenteesta. Tämä johtuu siitä, että predikatiivina voi toimia joko substantiivi (määritteinen) tai adjektiivi (määritteinen), ja halusimme selvittää, kumpi lauseke on asiaproosassa yleisempi. Predikatiiviin liittyy kaksi tärkeää syntaktista pulmaa, ja niitä on aluksi kosketeltava lyhyesti ennen kuin esitämme esimerkein, miten predikatiivit on kooditettu.

Ensinnäkin on ratkaistava kysymys rajan vetämisestä predikatiivien ja predikatiiviadverbiaalien välille. Yleisin käytäntö lienee se, että predikatiiveina pidetään *olla-* ja *tulla-*verbiin kiinteitä nominatiivi-, partitiivi- ja genetiivikomplementteja, kun taas essiivi- ja translatiivimuotoiset lausekkeet luokitellaan predikatiiviadverbiaaleiksi (näin myös Hakulinen & Karlsson (1979:189)). Olemme tässä poikenneet vallitsevasta käytännöstä, joten kaikki seuraavat tapaukset kooditettaisiin predikatiiveiksi:

Jussi on *kasvissyöjä* / *fanaattinen*

Jussi tuli *kasvissyöjäksi* / *fanaattiseksi*

Jussista tuli *kasvissyöjä* / *fanaattinen*

Jussi on *tarjoilijana*

Jussi rupesi *tarjoilijaksi*

Jussi toimii *tarjoilijana*

Jussi esiintyy *strip-tease tanssijana*

Tästä ratkaisusta on enemmän hyötyä kuin haittaa. Ensinnäkin essiivi- ja translatiivimuotoiset lausekkeet saadaan poimittua erilleen muuttujan (34) avulla. Jos ne sen sijaan laskettaisiin adverbiaaleihin, ne "hukkuisivat" useammin, koska niiden kanssa samassa lauseessa voi olla

muitakin adverbiaaleja - koodituksemmehan saa esiin vain yhden loppuadverbiaalain lausetta kohden.

Sen sijaan sellaisia essiivi- ja translatiivimuotoisia lausekkeita, jotka eivät ole pakollisia verbin komplementteja, ei voi laskea predikatiiveiksi, koska ne voivat esiintyä samassa lauseessa nominatiivipredikatiivin kanssa, esim. *Kriteerinä tämä on huono*. Seuraavassa virkkeessä on kahdessa lauseessa predikatiivi ja sen lisäksi ensimmäisessä myös pred.adverbiaali:

Elokuvasta voidaan puhua *populaarikulttuurina* (pr.adv:li) vain jos *kriteerinä* (pred:vi) on lajille tyypillinen tuotantotapa, mutta *tämä on kriteeri, jonka mielekkyyttä ei kovin pitkälle voi puolustaa* (pred:vi) 3.21

Kriteerinämme rajanvedossa on siis tässä pakollisuus.

Ns. objektipredikatiiveja (*Pidän häntä ystävänäni; Luulin häntä ystäväkseni*) jotka ovat harvinaisia, ei voi tässä järjestelmässä merkitä predikatiiveiksi siitä syystä, että verbillä ei ole kahta kiinteää komplementtia (subjektin lisäksi). Ne on siis kooditettava pakollisiksi predikatiiviadverbiaaleiksi (ks. 39:5).

Toinen paljon pohdintaa aiheuttanut syntaktinen ongelma on se, miten voitaisiin eräissä tapauksissa erottaa toisistaan subjekti ja predikatiivi. Tällainen ongelma tulee esiin, kun sekä subjekti että substantiivipredikatiivi ovat nominatiivissa. Predikatiivin tunnistamisesta käyty keskustelu on osoittanut, että eräissä tapauksissa mikään syntaktinen kriteeri kongruenssi, transformointi monikkoon, parafraasi ei tyhjentävästi erota toisistaan subjektia ja predikatiivia. Näissä tapauksissa on kyse ns. identiteettipredikatiivista, jolla on myös oma tarkoite (esim. *Tuo on se mies*). A. Hakulinen (1976:84) on huomauttanut, että näissä tapauksissa subjektin ja predikatiivin erottaminen näyttäisi olevan toisarvoista verrattuna lauseen informaatorakenteeseen. Jos on pakko puhua subjektista ja predikatiivista, voinee turvautua suomessa neutraaliin lausekaavaan SVX, jos syntaktisia kriteereitä ei ole.

Tekstilauseista havaitsemme kuitenkin, että useimmiten ratkaisua on etsittävä *semantiikasta*. Niinpä intuitiivisesti tekisi mieli väittää, että seuraavissa tapauksissa sanajärjestyksenä on XVS eli predikatiivi onkin teemana:

a) Niistä *pisimmälle kehittynyt yhteisö* on koko maapallon kattava ihmisten monivivahteinen yhteiskunta. 21.14

*Tämän aiheen kansainvälisesti tunnetuin toteutus* lienee Odysseiasta tuttu Polyfemos-motiivi, jossa -- 6.27

*Suomalaisten koululaisten ehkä parhaiten tuntema pilasatulaji* ovat hölmöläissadut, jotka -- 6.35

*Toinen perusmenetelmä* on kohteen fyysinen ja psyykinen heikentäminen. 13.16

*Suurimmat maaeläinten yhteisöt* ovat kulkusirkkojen vaellusparvet, jotka liikkuvat jopa 50 km:n levyisinä rintamina. 21.3

*Räntäsateen ja vesikelin parhaat kulkuneuvot* ovat reilut kumisaappaat. 19.7

b) *Hyvä ulkopohjan materiaali* on solukumi. 19.4

*Synteettisten aineiden heikkous* on niiden hiostavuus. 19.11

Toisaalta *tehokas vaikutuskeino* on jyrkimpien vastustajien selittämätön katoaminen joukosta (siirto toiseen laitokseen). 13.31

Intuition perustana lienee se, että näissä lauseissa on pikemminkin predikaatio kuin identifiointisuhte, ja predikaatiosuhteeseen sisältyy "joukkoon kuulumisen" ajatus. Niinpä

voidaan esittää parafraaseina, että ihmisten yhteiskunta kuuluu yhteisöihin, hölmöläissadut kuuluvat pilasatuihin jne., joskin predikatiivia on rajattu, yksilöity jollain superlatiivintapaisella määritteellä. b-ryhmän tapaukset ovat tätäkin selvemmin luokiteltavia: solukumi kuuluu hyviin materiaaleihin, hiostavuus synteettisten aineiden heikkouksiin; subjekti on semanttiselta alaltaan kapeampi kuin predikatiivi.

Sanajärjestyksen käänteisyydellä on yleensä se tarkoitus, että verbin jälkeen siirtynyt subjekti on joko jossakin mielessä (tekstissä) uusi taikka informaatioltaan painava. Tämä näkyy seuraavista esimerkeistä, jotka ovat tyypillisiä tälle ryhmälle:

Niiden *yhteinen nimittäjä* on se, että ne tapahtuvat -- 2.39

Tyypillinen *esimerkki* tästä oli englantilaisen music hallin esitys. 5.39

Tämän maaseudun kokeman nopean muutoksen *seuraus* oli myös nopea kaupungistumisen kiihtyminen, -- 7.47

Mahdollinen *selitys* tälle tulokselle on se, että -- 8.59

*Perussyö* oli maan vähäinen teollistumisaste. 10.55

-- johtoafasia, jonka olennainen *piirre* on sanojen ja lauseiden toistamisen vaikeutuminen. 12.15

Näissä kaikissa tapauksissa predikatiivi on maksimaalisen yleinen, semanttisesti melko väritön ja voimakkaasti sidoksissa edeltävään tekstiin. Subjekti kuuluu välttämättä predikatiivin alaan, inkluusiosuhde on olemassa, ja subjekti tuo esiin uuden informaation. Ensimmäisen lauseen teeman analysoiminen predikatiiviksi lienee näissä tapauksissa aivan kiistaton ratkaisu.

Havaitsemme siis, että tekstilauseissa predikatiivi vallan hyvin saattaa edeltää subjektia ilman, että lauseen ensimmäiseen nominaaliin välttämättä tulee eksistentiaalisia presuppositioita, kuten yksinkertaisiin kielioppiesimerkkeihin pyrkii tulemaan, vrt. *N.N. on erinomainen lingvisti* vs. *Tuo erinomainen lingvisti on N.N.*

Esimerkkejä predikatiivin rakenteen koodittamisesta:

### 1. yksi adjektiivi (myös partisiippi)

Tätä aihetta käyttävissä saduissa on *mielenkiintoista* se, että -- 6.30

On *helppo* nähdä, että -- 4.3

### 2. määrite + A, A + määrite

*Erityisen tehokkaasti järjestäytyneitä* ovat monien hyönteisten *yhdyskunnat* -- 21.8

*Näin yksinkertainen* ei asia sittenkään ole. 3.16

Eerik oli ulkopoliittisessa ajattelussaan *toisenlainen kuin mahdollisimman varovasti edennyt isänsä*. 20.21

### 3. rinnasteiset adjektiivit (määritteineen)

Heidän suosimansa -- *fabliaux-runous* -- oli kaikessa realistisuudessaan *pilailevaa, arvoja häpäisevää ja ajassaan suorastaan kapinallista*. 6.12

"Suhtautumiseni oli *skeptinen ja hylkäävä*", sanoo tri Hasler. 11.42

#### 4. adjektiivi ja lause

Amerikkalainen farmari selittää suomalaiselle pientilalliselle, että Amerikassa maatilat ovat niin suuria, että niiden kiertämiseen autolla menee kokonainen päivä. 6.56

-- kieltäytymisemme avoimesta keskustelusta aiheuttaa otaksuman, että olemmevieläkin vähemmän kykeneviä päättämään omista asioistamme / kuin todella olemme. 75.42

#### 5. vertailupredikatiivi:

Hän on kuin kukkanen.

#### 6. substantiivi

Elokuvasta voidaan puhua populaarikulttuurina vain jos kriteerinä on lajille tyypillinen tuotantotapa -- 3.21

Don Quijote on romaani, vaikka se onkin kehittynyt eepin runon pohjalta. 18.20

#### 7. määritteellinen substantiivi

Hän lähti marxismin klassikkojen tavoin siitä, että tajunta on todellisuuden heijastusta -- 1.8

Don Quijote ja Sancho Panza ovat jatkuvasti osoituksena Cervantesin humanismista, -- 18.36

#### 8. rinnasteiset substantiivit

Hiekan kivennäiskoostumus on yleensä kvartsia ja/tai maasälpää. 19.9

#### 9. substantiivilauseke ja lause

Tietoisuus ei Vygotskin mukaan ollut vain jokin korkeampi psyykkisten prosessien taso, jonka lävitse kokemukset virtaavat -- 1.15

-- oliko ja onko sitten juuri elokuvan estetiikka se alue, jolla tutkimusta yliopistotasolla eniten kaivataan? 3.10

#### A. lause

Nykyisessä kehitysvaiheessani on käsitykseni, että paranormaalilla parantamisella on osansa parannusmenetelmien sinfoniassa. 11.45

#### B. nominaalistuma

Lusikan ja matematiikan samoin kuin ammatillisten taitojen hallitsemisella on se yhteinen piirre, että ne ovat kaikki yhteiskunnan kulttuurin elementtien omaksumista, yhteiskunnallisesti muokattujen esineellisten toimintojen "haltuunottoa". 2.42

-- afasiassa on kyllä kysymys häiriöstä, joka kohdistuu moniin muihinkin opittuihin asioihin kuin kieleen ja ilmenee esim. abstraktisen asennoitumiskyvyn häviämisenä, -- 12.34

#### C. ellipsi

Historiallisesti muuttuvaa ja kehittyvää ei tällöin ole ainoastaan korkeampien henkisten toimintojen sisältö vaan 0 0 jopa niiden rakenne ja muodot -- 2.23

Jos tarkoituksena ei ollut asutuskolonin luominen vaan 0 0 ainoastaan kolonisoidun yhteiskunnan maanviljelyspotentiaalin saattaminen emämaan palvelukseen, -- 7.22

(34) *predikatiivin sija*

- 1 nominatiivi
- 2 partitiivi
- 3 genetiivi
- 4 essiivi
- 5 translatiivi
- 6 kvanttori (+N)
- i irrelevantti

Edellä jo esitettiin, että tämän muuttujan avulla saadaan predikatiiveiksi luokitelluista tapauksista eroteltua haluttaessa pois essiivi- ja translatiivitapaukset. Partitiivipredikatiivissa on semanttisesti kyse eri lajeista (esim. *omenat ovat makeita* luokittelupredikatiivi, *Orivesi on Hämettä* taas inkluusiopredikatiivi). Näiden välisen taajuussuhteen saa esiin ristiintaulukoimalla tämän muuttujan alakategorian 2 edellisen muuttujan alakategorioiden 1-4 ja 6-9 kanssa. Genetiivipredikatiiviin kuuluu kaksi eri tapausta: sekä omistajan genetiivi (auto on *minun*) että partisiippirakenteen predikatiivin genetiivi (esimerkkejä seur.). Nämäkin voi tarvittaessa erotella käyttämällä muuttujaa (8), joka erottelee lausetyypit. Esimerkkejä muista kuin nom.- ja part.predikatiivista:

3. *genetiivi*

-- uskovat juuri oman menetelmän olevan *sen oikean ja pätevän* jokasessa tapauksessa. 11.8

-- järjestö painottaa -- henkiparannuksen tarkoituksen olevan *täydentävän, mutta ei vaihtoehtoisen*. 11.32

4. *essiivi*

-- yliopisto olisi tarttunut syöttiin helpommin, jos *houkuttimena* olisi elokuvan tapaisen itsenäisen ilmiön sijasta laajempi kokonaisuus -- 3.14

*Tämän esityksen tarkoituksena* ei ole niinkään antaa vastauksia -- 3.12

1500-luvun Italiassa akatemit toimivat *filosofian ja kirjallisuuden etuvartijoina*. 14.3

5. *translatiivi*

Ulkopuolisen tason ylläpitämä virka olisi myöhemmin saattanut rahoittajan väsyttyä tulla *sen omaksi taakaksi*. 3.6

Nationalistipuolueen päästyä valtaan 1949 apartheid muodostui *käytännön politiikaksi*, jonka tavoitteena on Etelä-Afrikan rotujen pitäminen toisistaan niin erillään kuin mahdollista. 16.6

6. *kvanttori-ilmaus*

(tähän on laskettu muualle mahtumattomat numeraali-ilmaukset)

On monia erilaisia jouduttamisen välineitä ja pamflettikirjallisuus on *näistä yksi*. 4.6

Alkuaan Hölmölä on ollut *eräs Ruokolahden pitäjän kylä*. 6.43

Dyyneiden korkeus on *30 cm - 30 m* ja niiden aallonpituus *20 cm - 300 m*.

Sen mukaan näiden menojen osuus bruttokansantulosta oli pari vuotta sitten 8,5 prosenttia, mutta nykyisin perusteiden on vuonna 1985 jo kymmenisen prosenttia. 27.5

i. *irrelevantti*

(jos predikaatiivina on lause tai predikaatiivi on poistettu)

(38) *alkuadverbiaalin rakenne* ja (45) *loppuadverbiaalin rakenne*

- 1 yksi sana
- 2 lauseke
- 3 pre- tai postpositiolauseke
- 4 3. infinitiivi t. nominaalistuma
- 5 lauseenvastike
- 6 pääsana + lause
- 7 lause
- 8 rinnasteinen
- 9 ellipsi

Olemme rajoittaneet tässä koodituksessa tarkasteltavien adverbiaalien määrän kahteen, eli otamme mukaan vain lauseenalkuisen ja lauseenloppuisen adverbiaalin. Adverbiaaleja voi lausetta kohden olla jopa seitsemän; lisäksi niiden paikka voi vaihdella suhteellisen vapaasti. Jo neljän adverbiaalin (kaksi lauseenalkuista ja kaksi lauseenloppuista) ottaminen mukaan, jollaista esitutkimuksessa kokeiltiin, mutkistaa näin laajaa esitystä kohtuuttomasti ja kasvattaa muuttujien määrää 14:llä. Katsoimmekin, että adverbiaalit ansaitsisivat oman, pelkästään niihin rajoittuvan jatkotutkimuksensa. Kaksi adverbiaalia otettiin mukaan siitä syystä, että haluttiin saada yhtenäinen kuva lauseenalkuisista nominaaleista ja toisaalta pitää silmällä sitä, vaikuttaako "raskausperiaate" loppuun sijoittavasti adverbiaalien kohdalla samalla tavoin kuin subjektin ja objektin kohdalla. Nämä kaksi adverbiaalia olivat siis tekstuaalisesti keskeiset. Muuttujasta (52) saadaan kuitenkin lauseen adverbiaalien koko määrä selville.

Alkuadverbiaaliksi kooditetaan lauseen alussa (sivulauseessa konjunktion jäljessä) sijaitseva adverbiaalikonstituentti. Huomattakoon, että muuttuja (54) kirjaa tämän adverbiaalin toiseen kertaan, ja tällöin saadaan jopa selville, onko lauseen alussa useampia kuin yksi adverbiaali (54:6,C).

Tietyissä konstruktioissa alkuadverbiaalin edellä voi kuitenkin olla yksi konstituentti, nimittäin siirretty finiittiverbi (vrt. muuttujaa 57). Siirto tulee kyseeseen kun lause on *joko/tai* -kysymys, verbiin liittyy partikkeli *-hAn* tai se on inttävässä tai myönnyttelevässä mielessä siirtynyt lauseenalkuiseksi. Seuraavista esimerkeistä on alleviivattu alkuadverbiaaliksi kooditettava konstituentti:

Onko *Bergelillä* yhtä luova ote perinteiseen sinfoniseen ajatteluun kuin hindemithiläiseen polyfoniaan? 67.8

Onhan *kansan kaikkivaltiaan seassa* siinnyt hyvin sinisiä ajatuksia. 70.74

Eihän *siitä* yleensä Helsingissä muulloin piitata kuin vaivoin vaalipäivänä. 86.15

*Parlamenttivaalien jälkeen* maaliskuussa 1967 hän muodosti toisen hallituksensa -- 24.14

*Päältä katsoen* tällaiseen määritelmään sisältyy looginen noidankehä. 1.13

*Loppuadverbiaaliksi* kooditetaan lauseen lopussa sijaitseva adverbiaalikonstituentti, jos sellainen on. Vastaavasti kuin alkuadverbiaalin kohdalla on tehty täsmennys: loppuadverbiaaliksi kooditetaan myös sellainen adverbiaalikonstituentti, jonka yli transformaatio on heittänyt verbin (57:2,3). Samoin tapauksissa, joissa lauseen objektina tai subjektina on lause, tämä yleensä heittyy hallitsevan lauseen loppuun, ja tällaisissa tapauksissa muodollisesti subjektia tai objektia edeltävä adverbiaalilauseke käsitellään loppuadverbiaalina, onhan se omassa lauseessaan viimeinen. Esimerkkejä:

Epäilen *tosin*, että monet ovat käsittäneet kysymyksen kahdellakin tavalla väärin. 109.36

Nelisen vuotta sitten Ilkka Sumu kirjoitti *tässä lehdessä*, että asuntotuotannossa on tullut pää vetävän käteen. 112.1

Meren aallot eivät tosin Suomen graniittikallioille *suuriakaan* mahda, -- 9.16

Silti ei ole yhdentekevää, minkälaisen muodon vesilainsäädäntö *seuraaviksi vuosikymmeniksi* saa. 50.5

Tiistaisen Rautavaara-konsertin lisäksi Sarbu pääsee säväyttämään yleisöä koko illan verran *torstai-iltana, jolloin* -- 64.17

Coro nousee kuin koko maailman, sen eriväristen ja rotuisten ihmisten yhteisenä huutona rauhan, ihmisyyden, veljeyden puolesta *uhkaavia, tuhoavia voimia vastaan*. 65.16

Eikä hän näe, että nainen hyppää silti junan toiseen vaunuun *mielellään* yhtä *koomiset kostonaikeet, jotka hänen toinen minänsä (Angela Molina) sitten toteuttaa*. 62.11

*Ellipsitapauksissa* (ks. muuttuja 59) tarvitaan joskus ylimääräistä harkintaa adverbiaaleja tulkittaessa. Läpikäyvänä periaatteena on, että poistotapauksissa sanajärjestykseksi oletetaan se järjestys, johon lause edellisen perusteella luontevasti täydentyy (ks. muuttuja 18). Täten seuraavanlaisissa tapauksissa kooditetaan alkuadverbiaali:

-- vain kaksi päivää aikaisemmin hän oli johtanut Turussa Händelin Messiaan, *nyt jo 0 0 hän johti* Salzburgissa sille monessa mielessä täysin vastakkaisen teoksen. 65.37

Myös laulusolistit edustivat luotettavaa tasoa, *ensimmäisessä 0 0 edusti luotettavaa* tasoa sopraano Gabrielle Fuchs, *jälkimmäisessä 0 0* tenori Claudio Nicolai. 65.55

Mathis der Maler on varmaan ennenkin vaikuttanut yhtä lujalta, taidokkaalta ja oppineelta, *harvoin* sitävastoin 0 0 se on *vaikuttanut* näin energiseltä ja loistokkaalta. 67.6; vrt. \*Mathis der Maler on varmaan ennenkin vaikuttanut yhtä lujalta, taidokkaalta ja oppineelta, 0 0 se on *vaikuttanut* harvoin sitävastoin näin energiseltä ja loistokkaalta.

Seuraavissa tapauksissa taas adverbiaali jää keskipaikkaiseksi:

Saavuttaako hän toisentyypisen tehtävän parissa yhtä hyvän yhteisymmärryksen orkesterin jousiston kanssa kuin 0 0 [HÄN SAAVUTTI] *nyt* sen puu- ja erityisesti vaskipuhallinkuorojen kanssa? 67.9

On todellista löytämisen riemua yhyttää kuusikosta ryhmä muhkeita herkkutatteja, 0 [YHYTTÄÄ] *koivikosta* kopallinen kanttarelleja, 0 [YHYTTÄÄ] *männiköstä* sammalikon posket punehduttava haperokasvusto, -- 113.77

Aina ei adverbiaalinen "sijoittaminen" ole helppoa. Esimerkiksi seuraava tulkinta ei välttämättä ole ainoa oikea:

Vain hyvin kovakuorinen katsoja voinee väittää, / ettei elänyt kipeästi mukana Käthen tuskaa poikiensa kuolemasta / ja 0 0 0 / [ETTEI ELÄNYT KIPEÄSTI MUKANA] *lopulta* sitä luomisen kamppailua, / johon puristui hätä koko maailman nuorisosta ja aseiden murentamasta maailmasta. 61.15

Esimerkkejä adverbiaalisten rakenteesta:

### 1. yksi sana

*Valitettavasti / meillä* tällainen tutkimus on edelleen hyvin vähäistä. 53.51

### 2. lauseke

Jo yksistään tällä Bachillaan Kremer todella on huipun tuntumassa. 65.49

### 3. pre- tai postpositiolauseke

*Määräysten mukaan* tyttöjen pitää osallistua myös marjojen ja sienien säilömiseen yhdessä koulujen keittäjien kanssa. 97.27

Realismiin pyrkivällä maalarilla oli epäilemättä kotipainajaisensa *ilman* surrealismiakin. 9.11

### 4. 3. infinitiivi tai nominaalistuma määritteineen

-- tutkanilmaisin ei pelasta *joutumasta kiinni ylinopeudesta*. 25.9

-- olisin halunnut neuvoa häntä *nousemaan ensi tilassa junaan* -- 79.12

Siri Derkertille on ollut luonnollista ilmentää surunsa ja ilonsa tanssimalla ateljeen pöydällä tai hämmentyneen kuninkaan ja avajaisyleisön edessä Lilevalchsissa. 68.9

### 5. lauseenvastike

Meier-Graafe joutui vastaamaan sanoistaan sodan jälkeen *Dixin maalattua teoksen nimeltä Ampumahauta*. 69.27

*Niitä radioonkin analysoidessaan* hän tuli epäilyksettä siihen tulokseen, että -- 64.4

Tähän muuttujaan voi merkitä myös ne tavalliset lauseenvastikkeet, joita ei lauseiksi kooditetakaan:

-- vanhat miehet levittävät käsivartensa *käydäkseen Zorban tanssiin*. 63.28

-- näyttelijä seisoon tovin liikkumatta *antaen rinnastukselle aikaa tehot*. 61.38

### 6. pääsana + lause

Mutta juuri radioyhtymähän ne kautta Euroopan yleensäkin kantavat huolta *siitä, että oman aikamme säveltäjille riittää työtä ja leipää*. 65.22

Samaan aikaan Wrightin työt olivat suorastaan vastatulessa *Yhdysvalloissa, joka oli löytänyt kermakakkuklassismin*. 66.32

Tähän kuuluvat myös rakenteet adv. + lause: *sillä hetkellä kun, joka kerta kun, aina kun*.

### 7. lause

*Ja kun rikas antaa yltäkyläisyydestään köyhälle muka jalomielisiä* almuja, tekee hän siinäkin itsensä narriksi -- 62.34

-- sillä mitäpä Jacques tällaiselle ikääntyneelle lyllerölle tulisi, *kun annea kasvaa tasavallassa joka oksalla* -- 82.9

-- silti kesänäyttelyt vaikuttavat tuhottoman ahtailta, *vaikka akatemia-rakennuksessa on peräti 16 erikokoista salia*. 55.24

*Jos niiden katsomiseen vain riittää kinttuvoimaa*, ne ovat ylettömän hauskoja. 55.27

Lause-, lauseke- ja sana-adverbiaalin sijoittaminen samoihin muuttujiin ei ole ongelmatonta. Toisaalta ratkaisu on yhtäpitävä muiden nominaalisten konstituenttien kuvauksen kanssa; tällä tavoin saadaan tietoa siitä, kuinka usein adverbiaalilause pyrkii virkkeen alkuun (vrt. 4.1.). Se, että adverbiaalilause (esim. aikaa tai syytä ilmaiseva) on virkkeen alussa, ei kuitenkaan estä sitä, että päälause tästä huolimatta voi alkaa adverbisanalla. Nyt emme saa tätä sanaa enää mahtumaan kooditukseen, koska meillä on mallissamme vain yhden alkuadverbiaalin paikka. Toinen tästä ratkaisusta aiheutuva haitta on se, että muuttuja (40), adverbiaalin merkitys, luontuu parhaiten sana- ja lausekeadverbiaaleille. Adverbiaalilauseiden semanttinen luokitus kaipaasi esim. tarkempaa jakoa ehtoa, syytä ja myönnytyistä ilmaiseviin (*jos-*, *koska-*, *vaikka-*) lauseisiin, jotka nykyisellään jäävät kaikki "materiaalisiksi" adverbiaaleiksi semanttisesti.

Ainoa looginen tapa ratkaista tämä ongelma olisi sallia niissä tapauksissa, joissa alku- tai loppuadverbiaalina on lause, myös päälauseen sisällä olevan sana-adverbiaalin koodittaminen alkuadverbiaaliksi. Tämä ratkaisu on puolestaan teknisesti ongelmallinen. Nykyiselläänhan saamme kyllä tietää, kuinka usein lauseessa on lausemaisena alkuadverbiaalin lisäksi jokin toinen alkuadverbiaali, nimittäin ristiintaulukoimalla muuttujan (38) alakategoriat 5-7 muuttujan (54) alakategorian C kanssa, mutta emme saa tietää tarkemmin, mikä adverbialilaji on tällöin kyseessä.

#### 8. rinnasteinen

Kuten subjektin ja objektinkin kohdalla, rinnasteinen on kaikkia sitä seuraavia alakategorioita "vahvempi".

-- mutta tarvittavat varat hankittaisiin heidän mielestään niin, että työläiset tyytyisivät *pienempään palkkaan ja huonompaan sosiaaliturvaan*. 87.13

Paitsi eristämistä ulkomaailmasta pyritään kohde eristämään myös lähiympäristöstään *mm. vaihtamalla tiheästi työ- tms. ryhmien jäseniä kiinteiden henkilöyhteyksien estämiseksi sekä soveltamalla käyttöön jäljempänä esiintyvää joukkodynamiikkaa*. 13.15

9. *ellipsi*. Poistetun adverbiaalin tunnistaminen ei ole niin yksiselitteinen asia kuin poistetun subjektin tai objektin - paitsi selvän valenssiadverbiaalin tapauksessa:

Voitto voi jäädä pienemmäksi, mutta kansantalouden kannalta se ei ole vahingoksi. Työttömyys sen sijaan on 0. 87.21

Maini olisi halunnut keskikoulun jälkeen *Savonlinnan taidelukioon*, mutta ei päässyt 0. 97.32

Mutta jos sävellyksen yksityiskohdat eivät aina tulekaan *perille*, koko teoksen sanoma kyllä tulee 0 eikä voi olla väkevästi vaikuttamatta. 65.12

Samoin jos poistettu on relatiivipronomini tai kysymyssana:

-- ja laajassa osastossa kerrotaan, useinkin humoristisin sarjakuvin *miten* yksilöt tavoittavat ja 0 0 tunnistavat toisensa ja 0 0 saattavat sukusolut toistensa yhteyteen. 95.29

Tuo rakkaustarina, *johon* liittyy Bunuelille ominaista julmaa huumoria mutta 0 0 myös sääliä niitä kurjia raukkoja kohtaan, joita ihmiset ovat -- 62.13

Seuraavassa esimerkissä adverbiaalin ellipsi-ratkaisua puoltaa se, että rinnastuksen lauseet ovat huomattavasti vähemmän hyväksyttäviä, ellei niihin oleteta alkuadverbiaalia:

*Julkilausumassa* kiistettiin taiteilijan pasifismi, 0 muistutettiin Meier-Graafea unohtuneista periaatteista, 0 epäiltiin hänen tietämystään sodan kauheudesta ja 0 toivoteltiin menestystä Anatomianluennon haavan suutelossa. 69.32

Kooditustyöstä saadun kokemuksen perusteella esitämme seuraavan ehdotuksen muuttujaksi, jolla voi koodittaa adverbiaalinen rakennetta. Alakategorioiden hierarkia vastaa paremmin todellisuutta.

1. lause tai lauseeksi kooditettu lauseenvastike
2. pääsana + lause
3. rinnasteinen (jäsenenä mitä tahansa seuraavista)
4. lauseeksi koodittamaton lauseenvastine: *tehden, tehdäkseen*
5. infinitiivi tai nominaalistuma määritteineen
6. pre- tai postpositiolauseke (laajassa mielessä)
7. lauseke
8. yksi sana
9. ellipsi

(39) *alkuadverbiaalinen tyyppi* ja (46) *loppuadverbiaalinen tyyppi*

- 1 valenssi
- 2 kehys
- 3 kommentti
- 4 konnektiivi
- 5 pakollinen predikatiiviadverbiaali
- 6 irrallinen predikatiiviadverbiaali
- 7 muu
- i irrelevantti

Adverbiaalinen, toisin kuin lauseen muiden nominaalisten jäsenten kuvauksessa, ei riitä pelkän konstituentin sisäisen rakenteen merkitseminen. Tämä johtuu siitä, että adverbiaalit eivät ole läheskään samassa määrin yhtenäinen syntaktinen luokka kuin esim. subjekti tai objekti. Tätä kooditusjärjestelmää laadittaessa perustutkimusta adverbiaalinen luokittelusta oli vähemmän kuin nyt: Hakulinen & Karlsson (1979:202-) ovat esittäneet yhden ehdotuksen adverbiaalinen syntaktiseksi tyyppittelyksi, joka olennaisesti selventäisi tästä saatavaa kuvaa. Sen jaottelun mukaan yllä olevaan luetteloon tulisi seuraava muutos:

- 1 verbin aktantti (=valenssil)
- 2 lausetyypin aktantti (valenssi2, esim. *minulla on nälkä*)
- 3 verbin vapaa määrite
- 4 lauseen vapaa määrite (kehys ym.)

Usein adverbiaaleja on luokiteltu samanaikaisesti sekä syntaktisesti että semanttisesti. Nämä voidaan kuitenkin pitää toisistaan erillään, kuten tässäkin kooditusjärjestelmässä on tehty. Seuraavassa esitetään esimerkkejä kooditusratkaisuista, esimerkkeihin on merkitty sekä syntaktinen tyyppi että semanttinen laji, joka käy ilmi seuraavasta muuttujasta.

(40) *alkuadverbiaalinen merkitys* ja (47) *loppuadverbiaalinen merkitys*

- 1 aika
- 2 paikka
- 3 habitiivi
- 4 tapa

- 5 muu materiaallinen
- 6 määrä "OSMA"
- 7 modaalinen
- 8 irrelevantti

### 1,2. valenssi, paikka

- hän on kuulemma palannut *Meksikoon* -- 62.4
- maailman vanhin rauhanjärjestö International Peace Bureau, jonka päämaja on *Genevessä*, -- 110.5
- Mökeissä, joissa vanhuksat talvisin asuvat yksin, tungeksii nyt koko suku. 111.2

### 1,3. valenssi, habitiivi

- Realismiin pyrkivällä maalarilla* oli epäilemättä kotipainajaisensa ilman surrealismin. 9.11
- Onko *pohjoisen ruskan maineella* vastineensa todellisuudessa? 106.3
- on ihmisiä, *joilla* ei työttömiksi jouduttuaan ole turvanaan työttömyyskassaa. 72.9
- 1,4. *valenssi, tapa*. Verbejä, jotka vaatisivat nimenomaan tavan adverbiaaleja, ovat ainakin *viihtyä* ja *käyttäytyä*.
- 1,8. *valenssiadverbiaali, muu*. Useimmiten valenssiadverbiaalille on mahdoton antaa mitään itsenäistä merkitystehtävää sillä se määräytyy verbistä: paikallissijojen alkuperäismerkitys on valenssiadverbiaaleissa hämärtynyt.
- Nyt olen vähän odotellut, sanooko hän jotakin *Miettusen hallituksen salaisesta pöytäkirjasta* -- 84.4
- Mutta *maan hallitsemiseen* liittyy nykyisin entistä enemmän myös sen talouselämän ohjaaminen. 77.11
- Päinvastoin vaikutti *siltä, että Bergel kykeni todella* -- 67.5
- Mutta juuri radioyhtymäthän ne kautta Euroopan yleensäkin kantavat huolta *siitä, että oman aikamme säveltäjille riittää työtä ja leipää*. 65.22
- Niitä radioonkin analysoidessaan hän tuli epäilyksettä *siihen tulokseen että* -- 64.4

### 2,1. kehys, aika. Tähän kuuluvat *nyt, nykyisin, tällä hetkellä, silloin, meidän vuosisadallamme, 1500-luvulla*.

- Samaan aikaan* Wrightin työt olivat suorastaan vastatulessa *Yhdysvalloissa* -- 66.32
- Vuosi sitten* sitä [yhteistyötä] tehtiin kuukauden päivät Rautavaaran kotona ja *maaliskuussa* Rautavaara kävi verestämässä opiskeluaikojensa muistoja Sarbun nykyisessä asuinkaupungissa New Yorkissa. 64.6
- Tässä vaiheessa Piirpaukkeen maanantaista esiintymistä Kårenilla minulle muuten tuli kyyneleet silmiin. 63.29
- Viime vuonna* se oli kaunistelupyrkimyksistä huolimatta puolisen prosenttia nollan alapuolella. 74.22
- Kuluvan vuoden puoliväliin mennessä* jätöpapereita oli kerätty noin 8,5 miljoonaa tonnia. 96.22
- *keväästä 1969 alkaen* yhdistys on järjestänyt kaikissa Itä-Helsingin osissa säännöllisiä paperinkeräyksiä opiston toiminnan tukemiseksi. 96.21

### 2,2. kehys, paikka.

*Isojoella* ei elätellä suuria toiveita. 111.14

Samaan aikaan Wrightin työt olivat suorastaan vastatulessa *Yhdysvalloissa*,-- 66.32

2,8. *kehys, abstrakti*. Sekä syntaktisesti että semanttisesti varsinaisia konkreettisia kehysadverbeja muistuttava luokka. Aivan kuin jossakin paikassa tai jonakin aikana asioita voi tapahtua myös jossakin sosiaalisessa tilanteessa (konsertissa, perhepiirissä) tai vielä abstraktimmin: tietenhaaran sisällä, ja etenkin kirjassa, filmissä:

*Tämänvuotisilla juhlaviikoilla* ei esiintymislavalta vielä näet olekaan liiemmästi sinkoillut kipunoita -- 64.21

Viihdyin jälleen erinomaisesti niin Wolfgang Schneiderhahnin kuin Mozarteum-orkesterin oman kapellimestarin Leopold Hagerinkin johtamissa konserteissa. 65.51

-- suuri osa alkoholisteista tulee perheistä, *joissa* alkoholia käytetään väärin. 123.31

*Meillä* esimerkiksi on yleistä viikonloppujuopottelu. 123.10

3,7. *kommentti, modaalinen*. Modaaliset kommentit ilmaisevat puhujan/ kirjoittajan suhtautumisen lauseen totuusarvoon, sen varmuuteen, todennäköisyyteen jne. (vrt. 10:3): *ehkä, varmaan(kin), varmasti, kai, kaiketi, todennäköisesti, ilmeisesti, epäilemättä, olettavasti, todennäköisesti, arvattavasti*.

Mielessään hänellä olivat *varmaan* Lisztin pianoetydit -- 64.12

*Todennäköisesti* herra Dix on kaikessa yksinkertaisuudessaan halunnut toimia pasifismin puolesta tunnetun pelottelu-teorian tavoin. 69.29

3,8. *kommentti, muu*. (Raja modaalisten ja muiden kommenttien välillä ei ole selvä.) Muut kommentit taas ilmaisevat puhujan arvoväritteisen asennoitumisen lauseeseen: *totisesti, todella(kin), todellisuudessa, mielestäni, käsitykseni mukaan, itse asiassa, onneksi, toivottavasti, valitettavasti, kumma kyllä, ilmankos, ymmärrettävää kyllä, muka, itse asiassa, kerta kaikkiaan, yleisesti ottaen, kuulemma, tietääkseni*. Mukaan mahtuvat myös sellaiset - usein parenteesinomaisesti esiintyvät - metatekstuaaliset kommentit kuin kahdessa viimeisessä esimerkissä:

Intia tunnusti Bangladeshin itsenäisyyden tuettuaan Bangladeshia ja *itse asiassa* valloittettuaan maan. 24.23

*Ilmankos* MTK on ilmoittanut jyrkästi, ettei tehty tuloratkaisu sitä tyydytä -- 36.15

Suomessa vettä näyttää olevan paljon, mutta *todellisuudessa* vesivarat eivät meilläkään ole ylettömän suuret. 50.4

*Valitettavasti* meillä tällainen -- tutkimus on edelleen hyvin vähäistä. 53.51

Molekyylitasolla elämä siis perustuu hiiltä sisältäviin orgaanisiin yhdisteisiin, joissa lisäksi on runsaasti happea, vetyä ja tyyppiä sekä rikkiä ja fosforia, *yleisimmät mainitaksemme*. 22.6

4,7. *konnektiivit*. Konnektiivit ilmaisevat lauseiden välisten suhteiden laatua mutta ovat semanttisesti sisällökkäämpiä kuin konjunktiot. Seuraavaan luetteloon mahtuvat niistä tavallisimmat: *lisäksi, samoin, toisin sanoen, vastaavasti, täten, siten, sitä paitsi; kuitenkin, tosin, silti, toisaalta, muuten, sen sijaan, päin vastoin, siitä huolimatta: sen takia, näin ollen, siitä syystä, siksi, nimittäin; lopuksi, seuraavasti; kaiken kaikkiaan, joka tapauksessa*.

En näe mitään syytä pyrkiä tässä asiassa presidentin kanssa ristiriitaan. *Päinvastoin* tiukentunut käytäntö on aivan oikein. 107.13

Postimerkkeilyn kannatus oli varsin pientä, *samoin* uskonnollisen kerhotoiminnan. 109.74

Tämä johtuu akateemisen työttömyyden yleistymisestä ja *toisaalta* pidentyneestä äitiyslomasta. 115.7

5. *pakollinen predikatiiviadverbiaali*; tarkempaa semanttista luonnehdintaa tämä tyyppi ei kaipaa.

On väärin luulla esitystä - tai Kollwitzin taidetta - *suorasukaisesti* poleemiseksi. 61.18

Tosiaan hämmentävää: filmin nimenkin voisi kääntää *vaikka Halujemme hämmentäväksi kohteeksi*. 62.6

Miksi Bunuel valitsi Arrabalia muistuttavan kääpiön *yhdeksi junavaunun matkustajaksi, psykiatriksi?* 62.26

Muutkin kuin essiivi- ja translatiivimuotoiset komplementit voivat olla tällä tavoin predikoivia - lähinnä *tuntuu joltakin, maistuu joltakin, vaikuttaa joltakin*:

Yleisö tuntui usein *vaivaantuneelta*, ikäänkuin kritiikki kohdistuisi jokaiseen salissa istujaan. 62.36

6. *irrationaalinen predikatiiviadverbiaali*; myöskään sille ei kannata merkitä erityistä semanttista tyyppiä. Se voi tosin "ilmaista" jossakin tapauksessa aikaa (*lapsena kesätkin olivat pitempiä*), syytä (*lääkärinä minun on tunnettava vastuuni*) jne. Tyypillistä tällaiselle Predikatiiviadverbiaalille on, että se liittyy aina yhteen lauseen nominaalikonstituenttiin (paitsi joskus aikaa ilmaistessaan), mutta ei välttämättä aina subjektiin: *Kari piti makkarasta lapsena, Kari piti makkarasta kuumana*.

-- on ihmisiä, joilla ei työttömiksi jouduttuaan ole *turvanaan* [subjektiin liittyvä] työttömyyskassaa. 72.9

Arthur Honeggerin Jeanne d'Arc roviolla esitettiin keväällä Kansallisoopperan tuottamana Finlandiassa *näytelmällisenä produktiona* [objektiin liittyvä], nyt Salzburgissa Itävallan radio-orkesterin ja kuoron kantaesityksenä. 65.30

-- Christina Ortiz, nuori brasilialainen pippuripakkaus, joka *jo 15-vuotiaana* voitti Rio de Janeiron pianokilpailut ja *18-vuotiaana* Van Cliburn-kilpailut. 64.18

Kuka sen oikeastaan on määrännyt, että kuoron tulee laulaa *omana ryhmänään* ja orkesterin soittaa niinkään *omana ryhmänään?* 65.1

7. "*muut*" *adverbiaalit*

7.1. *aikaa ilmaisevat*. Tähän kategoriaan kuuluvat ne ajan adverbiaalit, jotka eivät ole kehyksiä, sekä ns. aikasuhteiset OSMA:t: duratiiviset (*kuinka kauan*), repetitiiviset (*kuinka monta kertaa*), sekä kvanttorin luonteiset (*aina, usein, joskus, harvoin, vielä, taas*).

Keräystoimintaa tehdään pienellä kuorma-autolla *ympäri vuoden*. 96.17

Bunuelin filmit pitäisi nähdä *ainakin kolme kertaa*. 62.1

Ja tällä alueellahan asiat eivät lopulta ole niin suuresti muuttuneet *250 vuodessa, jotka ---* 57.6

-- että täällä on eletty itsenäisenä kansakuntana *60 vuotta*. 70.35

7,2. *paikkaa ilmaisevat* - nämä ovat joko valenssiadverbiaaleja (*asua jossakin*) tai kehysadverbiaaleja (*jossakin tapahtuu jotakin*).

### 7,3. *habitiiviset*

Lasku oli yli sata markkaa. Ihmeteltiin tuota *tarjoilijalle*. 84.25

Taitaisi tehdä tilanteen tajuttavaksi *monelle, joka sen nyt kokonaan unohtaa*, jos nuo 137 000 tosiaan olisivat kaikki koossa -- 85.12

7,4. *tavan adverbialit*. Ne kuuluvat kaikki tähän ryhmään (lukuunottamatta *viihtyä-* ja *käyttäytyä-*verbien pakollisia komplementteja). Niille on olennaista, että ne ilmaisevat predikaation ominaisuutta, ja tyypillisimmillään tavan adverbialin sisältävä lause voidaankin transformoida nominaalistussubjektiseksi predikatiivilauseeksi: *hinnat ovat nousseet kiivaasti = hintojen nousu on ollut kiivasta*. Adjektiivin sisältävinä niitä voitaisiin pitää myös siksi, että niillä on vertailuasteet: *kiivaammin, kiivaimmin*. Kategoriaan on lisäksi otettu mukaan infinitiivi-ilmauksia, joita on viimeisissä esimerkeissä.

Niiden tavarain hinnat, joita kansaneläkeläinen nostaa, ovat nousseet *kaksi kertaa kiivaammin kuin hinnat keskimäärin*, nimittäin ruokatavarain. 85.7

En enää *tarkkaan* muista miten kaikki alkoi -- 86.9

-- ihmiset eivät antaisi yhtään kiloa paperia, jos keräyslaatikoiden kyljessä lukisi *reilusti* keräysfirman nimi. 96.44

Ajatus lähti kunnasta, ja sille lämpeni *nopeasti* myös Kuopion työvoimapiiri. 97.4

Päinvastoin vaikutti siltä, että Bergel kykeni todella lumoutumaan Hindemithin samalla kertaa karuista ja maalauksellisista näyistä *uutta luovalla tavalla*. 67.75

-- puolue -- pyrkii edelleen vallankumoukseen *väkivallankin keinona hyväksyen*. 77.5

7,5. *materiaaliset adverbialit*. Tämä on erittäin sisällökäs kategoria - siihen sisältyy syyn, seurauksen, välineen, tarkoituksen jne. ilmaiseminen.

-- sillä *mitäpä* Jacques tällaiselle ikääntyneelle lyllerölle tulisi, *kun annea kasvaa tasavallassa joka oksalla* -- 82.9

-- vastustiko hän maanlunastuslain vahvistamista *juristin omantunnon vai poliittisten päämäärien takia*. 84.2

Aivan oikein, entinen johtaja on vahvistanut lupauksen alennuksista *nimikirjoituksellaan*. 84.12

*Niillä ruokalapuilla* ei saa ostaa tupakkaa eikä keskiolutta, -- 85.29

*Paperinkeräyksestä saaduilla rahoilla* kannatusyhdistys voi peittää noin 30 prosenttia opiston kustannuksista. 96.18

Loput menot oli peitettävä *omin hankinnoin*, etupäässä *keräyspaperituotoin*. 96.25

Jos *yksistään* tällä Bachillaan Kremer todella on huipun tuntumassa. 65.49

Coro nousee kuin koko maailman, sen eriväristen ja -rotuisten ihmisten yhteisenä huutona rauhan, ihmisyyden, veljeyden puolesta *uhkaavia, tuhoavia voimia vastaan*. 65.16

*Jakamalla tuloa toisin* voidaan pienviljelijäin asemaa helpottaa -- 84.32

Huoltolautakunta antaa ruokalappuja niille, joille ei jää tarpeeksi *hengissä pysymiseen*. 85.29

7,6. OSMA. Tähän kuuluvat nimenomaan ns. "yleiset OSMA:t", koska aikasuhteiset ovat joutuneet ajan adverbialien luokkaan.

Mutta maan hallitsemiseen liittyy nykyisin *entistä enemmän* myös sen talouselämän ohjaaminen. 77.11

Tietääkseni Lindblom todellakin tuolloin sai myös karvaasti kokea, *miten vähän* hallituksen korkeittenkin edustajien sanaan sai luottaa. 76.7

7, 8. "Muut" adverbialit - siis eräänlainen jäännöskategoria, jonka sisältöä on vaikea lähteä luokittelemaan.

*Asiamies Sainion mielestä* vallitseva järjestelmä on täysin luonnollinen, koska -- 96.38

*Majurin mukaan* jätepaperin keräys ei ole yrittäjälle mikään kultakaivos -- 96.45

On vahinko, että tämä sulkeutunut omahyväisyytemme estää ulkomaalaisia näkemästä sitä valtavaa kehitystä, joka meillä kaikesta huolimatta on *niin taloudellisella kuin sosiaalisellakin alalla tapahtunut*. 75.45

Tämä anti tulee syyskuun 1. päivänä tapahtuvan korotuksen jälkeen olemaan 750-829 markkaa *kuntaryhmästä riippuen*. 85.4

*Tiistaisen Rautavaara-konsertin lisäksi* Sarbu pääsee säväyttämään yleisöä koko illan verran torstai-iltana -- 64.17

#### (25) verbin rakenne

1 yksi sana

2 liittomuoto

3 verbiketju

4 verbiketju liittomuodossa

5 *on aika / määrä + inf., katsoa hyväksi + inf.*

6 edell. liittomuodossa

7 fraasiverbi (*saada aikaan* yms.)

8 edell. liittomuodossa

9 verbipari

A edell. liittomuodossa

B finiittiverbin ellipsi

Verbi ei muodosta lauseessa samanlaista konstituenttia kuin nominaaliset lauseen osat. Useiden näkemysten (esim. sijakieliopin) mukaan finiittiverbi muodostaa lauseen navan, josta muut konstituentit riippuvat. Se mitä voisimme sanoa verbikonstituentiksi tai *verbiryhmäksi*, ei ole pintarakenteessa niin kiinteä kokonaisuus kuin vaikkapa NP, vaan sen osat joutuvat sangen usein etäälle toisistaan. Etenkin verbin finiittinen osa liikkuu usein yksin lauseen alkuun: myös adverbit ja partikkelit voivat sijoittua verbiryhmän sisäpuolelle - seikka, jonka selvittämisen jätämme tämän tutkimuksen ulkopuolelle.

Verbiryhmän rakenteen kuvaus perustuu tässä kokonaan pintarakenteeseen; tämän voi tehdä riippumatta siitä, missä tapauksissa lähtömuotona halutaan pitää kahta propositiota, joiden yhteen sulautumana muodostuu pintaan verbiketju tms.

#### 1. yksi verbi

2. *liittomuoto*. Liittomuotoihin kuuluvat kiellon ja tempuksen ilmaisimet (*en tee, olen tehnyt, on tekevä, on tehty*) sekä näitten muodostamat yhdistelmät modusten kanssa (*olisin tehnyt, emme olleet tehneet, lienee tehnyt, ei olisi tehty*) ja eräät verbiliittorakenteet: *on tulossa/menossa, on tekemässä, on tehtävä, on tehtävissä, on tekemäisillään, on tekemättä; lienee tulossa, olisi ollut tekemättä, ei olisi tehtävä* jne. (vrt. Hakulinen & Karlsson (1979:234)). Yhteen liittomuotoon mahtuu siis jo paljon kombinointimahdollisuuksia.

Kuten poliisi eri yhteyksissä *on todennut*, tutkanilmaisin *ei pelasta* joutumasta kiinni ylinopeudesta. 25.9

-- yhtye *on harjoitellut* ja 0 0 *latautunut* käynnissä olevaa kiertuetta varten aivan kuin *olisi valloittamassa* suuren maailman pääkaupunkeja -- 63.17

3. *verbiketju*. Käsitteellä verbiketju (vrt. Hakulinen & Karlsson (1979:230)) tarkoitetaan tapauksia, joissa 1. infinitiivi on verbin objektina (*haluan mennä*) tai 3. infinitiivi verbin valenssiadverbiaalina (*jouduin menemään*) ja joissa (mahdollisesti postuloitu) alkuperäinen lauseraja on heikennyt. Verbiketjuun kuuluu siis finiittiverbi ja yksi tai useampi (saattaa *joutua menemään*) 1. tai 3. infinitiivin muoto. Tämän analyysitavan mukaan näyttää pintarakenteen kannalta siltä, että NP *ruotsia* olisi koko verbiketjun eikä verbin *puhua* objekti. Tähän ratkaisuun ei sisälly mitään loogis-semanttisia oletuksia. Verbiketjuja muodostavat tunnusomaisimmin ns. modaaliset verbit (*saada, voida, kehdata, viitsiä, uskaltaa, ehtiä, osata*) kuten myös eräät muut samasubjektisuutta edellyttävät verbit (*aikoa, yrittää, luvata, muistaa* jne). Eräät mm. välttämättömyyttä yms. ilmaisevat modaaliverbit kuten *täytyy, pitää, sopii* käyttäytyvät edellisistä poiketen sikäli että ne saavat genetiivisubjektin (*meidän täytyy*); näissäkin tapauksissa katsotaan modaaliverbin ja sitä seuraavan infinitiivin muodostavan verbiketjun (*täytyy mennä, kannattaa puhua*).

Mainittakoon vielä yksi erityisratkaisu: partisiippirakenteen subjektikseen saavat verbit (*näyttää, kuulua, tuntua*) on tässä tulkittu verbiketjun muodostaviksi (*näyttää tulevan, tuntuu puhuvan*): näitä tapauksia on erittäin harvassa.

3. infinitiiviä vaativista verbeistä mainittakoon *ruveta, joutua, pystyä, pyrkiä, päästä, suostua, sitoutua, oppia, tottua, tyytyä, jäädä, kieltäytyä, vältyä*, jotka ovat kaikki samasubjektisia. Verbi *tulla* luetaan tähän vain futuurimerkityksisenä (*tulla tuntemaan*); liikeverbinä sen katsotaan saavan paikan adverbialin, joka ei ole valenssinmukainen eikä muodosta sen kanssa verbiketjua (*tulla kalastamasta*).

3. infinitiivin aseman ratkaiseminen - sen päättäminen, kuuluuko se verbiketjuun vai onko adverbialina - ei ole aina helppoa. Ratkaisevaa on, miten kiinteän sanaliiton verbi ja infinitiivi muodostavat. Esimerkiksi seuraavanlaisissa tapauksissa verbien liitto on liian höllä tullakseen merkityksi verbiketjuksi:

-- henkilökuntaa *ei riitä opastamaan* taloon tulevia potilaita toimenpidehuoneisiin. 66.47

-- taipuuko hallitus eläkepolitiikassa jälleen kerran toteuttamaan työeläkejärjestelmän rahat rahojen luo -tavoitteita -- 35.21

Informaation saannin kannalta koodittaminen verbiketjuksi on edullisempaa kuin adverbialiksi, sillä infinitiivilausekkeessa useimmiten oleva nominaalikonstituentti saadaan

tällöin lauseen itsenäiseksi konstituentiksi ja sen tekstuaaliset piirteet (esim. tunnettuus) voidaan koodittaa.

Sellaiset verbit, jotka saavat infiniittimäärityksen mutta eivät edellytä samasubjektisuutta, eivät muodosta infinitiivin kanssa verbiketjua, sillä näissä tapauksissa kummallakin verbillä voi olla oma objektinsa. Erisubjektista infinitiiviä vaativia verbejä ovat *kehottaa, käskeä, neuvoa, opettaa, pakottaa, pelastaa, estää, vaatia, kieltää*:

-- tutkanilmaisin ei *pelasta* *O joutumasta* kiinni ylinopeudesta. 25.9

Valtiontaloudelliset syyt ovat lopulta *pakottaneet* ajatusta voimakkaimmin ajaneen keskustapuolueen *perääntymään* vaatimuksistaan. 44.6

-- ja yrittivät *saada* Dikaiopolisin *muuttamaan* mielensä. 58.3

-- olisin halunnut *neuvoa* häntä *nousemaan* ensi tilassa junaan ja kokemaan junamatkan kummajaiset. 79.12

#### 4. verbiketju liittomuodossa

-- ellei ratkaisua *kyetä pitämään* kohtuuden rajoissa, -- 26.2

Luonnollisena on kuitenkin pidettävä, *etteivät* valtion neuvottelijat *ole voineet tehdä* sitä asiasta poliittista vastuuta kantavien ministerien selän takana. 26.12

Vesakkojen lentomyrkytyksiin *ei tarvitse pyytää* lupaa keneltäkään *eikä* lääni tai kunta *pysty kieltämään* lentoruiskutuksia alueellaan. 28.2

5. *nominaalikonstituentin sisältävä verbiliitto: on aika tehdä*. Tällaiset ovat idiomiluonteisia sanaliittoja, joihin liittyy pakollisesti infinitiivi. Tällöin sanaliitto ja infinitiivi lasketaan yhdeksi konstituentiksi.

Sivustakatsojan *tekisi vain mieli sanoa* tämän alan vaikuttajille: tutkikaa organisaatiokaavojaanne. 54.87

Myös hallituksen *on syytä perehtyä* siihen huolellisesti. 50.21

#### 6. edellinen liittomuodossa:

Kun elintarvikkeiden hinnat ovat tänä vuonna muutenkin nousseet nopeimmin, *ei* vauhtia *ole* enää *varaa lisätä* MTK:n keinoitekoisilla hinnan korotuslaskelmilla. 33.9

-- *olisi* varmaan *ollut paikallaan järjestää* lisää suomalaisin voimin esitetyjä konserteja -- 54.66

7. *fraasiverbi*. On huomattava joukko verbin ja nominaalikonstituentin liittomuotoja, jotka ovat idiomiluonteisia eivätkä kovin luontevasti syntaktisesti analysoitavissa. Seuraavassa luettelo idiomaattisista verbifraaseista, jotka on tässä käsitelty yhtenä verbikonstituenttina:

olla olemassa	pitää paikkansa	saada tarpeekseen
olla kiinni jossakin	pitää hallussaan	tulla esiin
olla mukana	päästä(ä) irti	käydä ilmi
olla vuorossa	ottaa oppia	astua voimaan
olla omiaan	ottaa huomioon	sanoutua irti
olla jäljellä	ottaa vastaan	käyttää hyväksi
olla tarjolla	panna kiinni	lähteä käyntiin
olla lähtöisin	saada aikaan	käydä kauppaa

pitää huolta

saada selville

mennä vikaan

Dikaiopilis on attikalainen maanviljelijä, joka kesken peloponnesolaissodan saa taistoista tarpeekseen, -- 58.02

Idiomin nominaalilauseke voidaan tosin marginaalisesti topikaalistaa, ja tällöin se ei tule erikseen kirjattua. Tämä on kuitenkin harvinaista ja käy ilmi sanajärjestysmallista sekä siitä että kysymyksessä on lauserakenteen kannalta idiomi (15:9).

#### 8. *fraasiverbi liittomuodossa* (tai v-ketjussa tai v-ketjun liittomuodossa)

Näitä tekstejä on kasattu päällekkäin, limittäin ja lomittain siinä määrin, että edes tekstikirja kädessä ei aina ole perillä siitä, missä mennään. 65.10

Frank Lloyd Wrightin elämäntyö on hyvä esimerkki tällaisena kriisin aikana, ja sellaisena se lieneekin vedetty esiin kansainvälisen kiertonäyttelyn aiheeksi. 66.18

#### 9. *verbipari*. Marginaalisena tapauksena tulee joskus esiin rinnastettuja verbejä, joita ei voi käsittää erillisiksi predikaateiksi: *otti ja lähti, etsi ja etsi*.

-- että tämä alunalkaen käsittämätön tuotantoon kohdistuva rangaistusmaksu päinvastoin lupauksista huolimatta pysyy ja pysyy. 52.24

B. *finiittiverbin ellipsi*. Tämä seikka on kooditettu kahteen kertaan. Ellipsejä selostetaan yksityiskohtaisemmin muuttujaa (59) esiteltäessä.

#### 2.6. *Lause- ja sanamääriä koskevat muuttujat*

lauseiden lukumäärä virkkeessä (4)

lauseen sanamäärä (16)

lauseen konstituenttimäärä (17)

subjektin, objektin ja predikaatiivin sanamäärä (24,32,37)

alku- ja loppuadverbiaalin sanamäärä (44,51)

adverbiaalien määrä lauseessa (52)

#### (4) *lauseiden lukumäärä virkkeessä*

Virkkeen lauseiden lukumäärä merkitään vain yhteen kertaan: virkkeen ensimmäisen lauseen kohdalle. Muiden lauseiden kohdalle merkitään lukumääräksi nolla. Teknisten tietojen sarakkeissa numerosarja 116/38/1/4 merkitsee siis, että kyseessä on tekstin 116 virkkeen numero 38 neljästä lauseesta ensimmäinen, numerosarja 116/38/3/0 taas tarkoittaisi saman virkkeen kolmatta lausetta.

#### (16) *lauseen pituus sanoina*

Lauseen pituus lasketaan ensin mekaanisesti sanamäärästä. Lauserajat määritellään täysin automaattisesti konjunktion edeltä kulkeviksi:

Tulin ajatelleeksi tätä taas nyt (5)/ kun näin El Supremon 32. filmin Montparnassella, lähellä sitä hotellia (10)/ jonne Bunuelin on tapana majoittua silloin (6)/ kun hän saapuu Pariisiin (4). 62.3

(17) *lauseen pituus konstituentteina*

Kuvaa lauseen pituudesta ja informaatiomäärästä täydentää tieto konstituenttien määrästä. Tämä on välttämätön tieto myös yritettäessä arvioida, miten suuri määrä lauseen konstituentteista jää koodittamatta. Konstituentteja voi lauseessa olla enemmän kuin sanoja: konstituentin voi muodostaa lausekin:

*Ei tämä päde, sanoo johtokunta.* (3 konstituenttia)

Konstituentteiksi lasketaan verbi tai ns. verbiketju (ks. Hakulinen & Karlsson (1979:228)), subjekttilauseke, objektilauseke, predikatiivilauseke sekä jokainen adverbiaalilauseke - myös modaaliset ja kommenttiadverbiaalit - lukuunottamatta sellaisia, jotka eivät koskaan voi saada omia määritteitä. Irralliset, koko lausetta kommentoivat tai sitä toiseen lauseeseen kytkevät ainekset on siis jätetty tässä lauseen varsinaisen konstituenttirakenteen ulkopuolelle: konjunktiot, muuttujissa (13,14) luetellut partikkelit ja niiden sukuiset (esim. *kyllä, sitä, toki, tuskin*), hyvin kevyet konnektiivit (esim. *niinpä, näet, nimittäin, taas* 'sen sijaan', *muuten* 'by the way' ja intensiteettiadverbiaalit, jotka kuuluvat jonkin itsenäisen konstituentin yhteyteen (esim. *ainakin, myös, vain, edes, melkein, liki, juuri, nimenomaan*) ja ne lasketaan aina siihen konstituenttiin kuuluvaksi, jota ne määrittävät. Erityisesti on vielä huomattava, että sellaiset liikkuvat pikkusanat kuin (presidentti) *itse* ja (presidentti) *puolestaan* jäävät myös koodittamatta konstituentteiksi (ks. 56:6). Esimerkkejä:

*/Bunuelin filmit/ /pitäisi nähdä/ /ainakin kolme kertaa/ 62.1 (3)*

*/Tuomittavana ja salakähmäisenä menettelynä/ /on pidettävä/ /sitä, että vasta näissä neuvotteluissa/ /on paljastettu/ / viime helmikuulta/ /ns. salainen pöytäkirja, josta --/ (6) /josta/ /vaiettiin/ /täysin/ vielä toukokuun hallitusneuvotteluissakin/. (4) /Tämä "mätäkuun paperi"/ ei saanut /aikoinaan/ /edes valtioneuvoston virallisen istunnon juridista vahvistusta/. 33.14-15 (4)*

(24) *subjektin sanamäärä*

1-9

A-H 10-17

i irrelevantti

Kooditettavan lausekkeen "raskautta" kuvastaa kaksi mittaria: edellä esitelty konstituentin rakenne sekä konstituentin pituus sanoina. Vaikka sanamäärä sinänsä lasketaan mekaanisesti (esim. nimi *Anne Pohtamo* kahdeksi sanaksi), erityisratkaisua tarvitaan silloin, kun konstituenttina on joko pelkkä lause tai pääsana + lause. Olemme tässä muuttujassa olleet kiinnostuneita lähinnä *sanasubjekteista*. Niinpä jos konstituentin muodostaa pelkkä lause, sen lauseen sanamäärää ei ole laskettu tässä (vaan muuttujassa (16)), vaan tällöin on sanamäärän kohdalle merkitty i. Jos taas konstituentin muodostaa pääsana (mahdollisine muine määritteineen) + lause, jätämme taas lauseen sanamäärän laskematta ja ilmoitamme sen sanamäärän, mikä jää jäljelle. Esimerkki:

Tietoisuutta koskevan tutkimuksen merkityksen [objekti: 4 sanaa] hän [subjekti: 1 sana] otti esille jo eräässä varhaiskirjoituksessaan [loppuadverbiaali: 2 sanaa], joka [subj: 1 sana] ilmestyi Kornilovin toimittamassa "psykologia ja marxismi" -julkaisussa 1925 [loppuadverbiaali: 1 sana]. 1.2

Samalla tavoin on laskettu objektin (32), predikatiivin (37), alku- ja loppuadverbiaalien (44,51) sanamäärä, joten niitä ei tarvitse erikseen selostaa.

(52) *adverbiaalien määrä lauseessa*

1-6

Tarkempia syntaktisia, semanttisia ja tekstuaalisia tietoja olemme keränneet vain alku- ja loppuadverbiaalista. Muualla lauseessa olevat adverbiaalit jäävät siis kooditusjärjestelmämme ulkopuolelle, mutta lauseen adverbiaalien kokonaismäärä rekisteröidään silti muuttujassa (52).

### 2.7. *Nominaalisten konstituenttien tekstuaalisia ominaisuuksia koskevat muuttujat*

subjekti (21-23)

objekti (29-31)

predikatiivi (36)

alkuadverbiaali (41-43)\*

loppuadverbiaali (48-50)

Tekstin ymmärrettävyyteen liittyy kaksi sen ominaisuutta, joita on sanottu *koherenssiksi* ja *koheesioksi*. Näistä ensin mainittu tarkoittaa tekstin sisällöllistä johdonmukaisuutta, eräänlaista ajatuksen kulun ennustettavuutta (vrt. van Dijk (1977: 93-129)). Tähän liittyy yhtenä tekijänä se, että kertomuksessa, raportissa tms. puhutaan samasta tai samoista tarkoitteista, että ajankohdat ja tapahtumapaikat pysyvät samoina ellei toisin ilmoiteta ja että uusina tarkoitteina tuodaan esiin uusia eikä sellaisia, joista on jo aiemmin ollut puhetta. Tekstin semanttista koherenssia kaikkineen tuskin voi kvantitatiivisesti mitata, mutta sikäli kuin se heijastuu koheesiona, näin voi tehdä. Olemme tässä pitäneet silmällä nimenomaan koheesiota eli sidoksisuutta. Tämä tarkoittaa sitä, että olemme kunkin nominaalisen lauseenjäsenen kohdalla katsoneet, onko se tekstuaalisesti vapaa, onko siinä jokin ilmikeino, josta näkyy, että sama referentti on ollut aiemmin esillä, vai edustaako tarkasteltava NP (tai AdvIP) vain semanttisesti samaa merkityskenttää, jonka jokin aikaisemmin mainittu konstituentti on jo esitelty. Tältä pohjalta ovat syntyneet muuttujat (23,31,36,43,50).

Mainitut muuttujat kartoittavat kuitenkin vain pintailmiötä, sitä, mikä yksittäisestä konstituentista näkyy. Esitutkimuksen kuluessa kävi ilmi, että tätä täydentämään tarvittaisiin puhtaammin semanttinen muuttuja, joka kartoittaisi sitä, *millä* tavalla tekstissä käsiteltävät tarkoitteet ovat "tunnettuja". Pelkkä kahtiajako "tunnettuihin" ja "uusiin" referentteihin ei riittänyt, vaan päädyttiin kuusijakoiseen muuttujaan sen mukaan, millä tavoin käsillä olevan konstituentin tarkoite voitiin katsoa entuudestaan tunnetuksi. Edelleen halusimme nähdä, miten sidoskeinon valintaan vaikuttaa etäisyys korrelaatiin.

Se, että muuttujiksi valittiin toisaalta tunnettuus, toisaalta tekstuaalinen sidoskeino eikä esim. "teemaa", johtui siitä, että teema on määritelty niin monella tavoin, ettei määritelmien pohjalta voi muodostaa operationaalista eli yksiselitteistä tapaa tunnistaa kustakin lauseesta teema. Tämän koodituksen tulokseksi voimme ilmoittaa tietoja siitä, kuinka usein minkintyyppinen lauseenalkuinen nominaalijäsen sisältää tunnettua tietoa jne., jolloin voimme luonnehtia niitä tendenssejä, joita teemaksi tulkittavilla konstituenteilla on. Teeman käsitettä hämärtää ehkä eniten se, että suurin osa keskustelusta on käyty irrallisten, tekstiyhteydestään otettujen tai jopa keksittyjen lauseiden parissa. On kuitenkin ilmeistä, että teema on lauseen ominaisuus ennen kaikkea sikäli kuin lause liittyy johonkin suurempaan kokonaisuuteen niin että *lauseen teema on osa tekstin yleisteemaa, aihetta*. Ilmeistä on edelleen, että toiset lauseet ovat teeman

kuljettelu kannalta keskeisempiä kuin toiset: on ei-temaattista ainesta sisältäviä lauseen alkuja ja kokonaisia lauseita, joiden tehtävä on modaalinen, metatekstuaalinen yms.

Vasta laajempi koodittaminen paljasti tämänkin moneen kertaan muokatun tekstimuuttujiston puutteet. Selvin puute on se, johon jo edellä muuttujan (12) yhteydessä on viitattu, nimittäin, että läheskään aina varsinkaan asiateksteissä ei käsitellä spesifisiä, tiettyjä tarkoituksia, vaan puhutaan yleisemmin, geneerisesti taikka lajista (vrt. myös seur. (21:2) ja Vilkuna (1980)). Samoin on monia konteksteja (ns. intensionaaliset ja kvantifioivat kontekstit), joiden vaikutusalueella olevilla nominaalilausekkeilla ei ole tarkoitusta vaikka ne ilmaisivat aivan konkreettisia asioita, esim. Minulla ei ole *autoa* ei viittaa mihinkään autoon; Joka päivä minulle tulee luukusta *mainoslehtisiä* ei myöskään sisällä spesifistä nominaalilauseketta (vrt. tästäkin Vilkuna mt.). Näin ollen tarvittaisiin oikean kuvan saamiseksi tietyn lausekkeen tekstuaalisesta käyttäytymisestä myös tieto siitä, onko se geneerinen, ei-spesifinen vai spesifinen.

Valitut kolme muuttujaa ovat aivan samat subjektin, objektin ja kahden adverbiaalisen kohdalla, joten ne esitellään vain yhteen kertaan, subjektin muuttujina. Predikatiivi on erityisasemassa sikäli että vain poikkeustapauksessa se viittaa eli sillä on oma tarkoite; niinpä predikatiivin tekstuaaliseen kuvaukseen on otettu vain sidoskeinoja koskeva muuttuja.

(21) *tarkoitteen tunnettuus* (myös 29,41,48)

- 1 mainittu
- 2 "heikosti" mainittu
- 3 asiointi, lause, juttu
- 4 implikoitu (= tuttu merkityskenttä)
- 5 edellytetty (tekstinulkoisen tiedon perusteella)
- 6 uusi
- i irrelevantti

Tunnettuus tarkoittaa sitä, että tarkasteltavan ilmauksen tarkoite, siis se, johon pääsanana viittaa, oletetaan syystä tai toisesta yksiselitteisesti vastaanottajan tunnistettavaksi. Tarkoite, joka ei ole millään keinoin tunnistettavissa, on lukijan/kuulijan kannalta uusi. On tärkeää panna merkille, että konstituetti voi olla tekstuaalisesti sidonnainen, sen sisältämät attribuutit voivat sitoa lausekkeen johonkin edellä esiintyneeseen lausekkeeseen samalla kuin pääsanassa esiin tuotu referentti on uusi. Että nämä kaksi seikkaa toimivat toisistaan erillään, käynee ilmi seuraavista esimerkeistä:

- konsertin päättäneessä *Schumannin toisessa sinfoniassa* hän prässäsi orkesterin soittamaan ylinopein tempoin ja näin hävitti *teoksen (sidos) hengen (uusi)*
- hampaiden välissä keinui *muovikassi, jonka (sidos) kirkkaanpunaisella pinnalla (uusi)* mainostettiin jotakuta vaatetusliikettä--

1. Tarkoite on *mainittu* aikaisemmin tekstissä. Kahden esiintymän välillä vallitsee siis *samaviitteisyys*. Mainitun tarkoitteen uudelleenmaininnan tapauksessa tulevat sidoskeinoista kysymykseen täysi ellipsi, anafora, toisto ja parafrasi (ks. näitä kohdasta 23).

Sienikuntoilijan ei toki kannata tehdä kuten *toisen tuttavan äiti*. *Muori* keräsi tiinukaupalla *sieniä* ja säilöi *ne*. *Helsingissä asustava poika* tuli kotipuoleen käymään. *Äidin* kehaistua *sienisaalistaan poika* tuli kysyneeksi, että missäs välissä *muori* on opetellut eri sienilajit tuntemaan. - Ehä mie oookkaa, *äitimuori* tuumi. Tunkioon menivät *säilötyt sienet*. 113.22-26

*Ihminen* muuttaa maisemaa. Hän luo myös *uutta maisemaa*, jossa on uusia aineksia ja muotoja. Hän lisää osaltaan muuttumisen vauhtia. *Ihminen* muuttaa meidän aikanamme jopa maankamaraakin enemmän kuin tulivuoret, jäätiköt, maanjäristykset, hyökyaallot ja kuivuus. 9.62-65 (Sama viitteisyydeksi käsitämme myös tällaiset tapaukset, joissa samaviitteisiä ovat geneeriset tai lajiin viittaavat nominaalit.)

2. Tarkoite on "*heikosti*" mainittu aikaisemmin tekstissä. Tämä käsite tarkoittaa samaa kuin englantilainen termi *co-sense*. Toisin sanoen tarkasteltava ilmaus *merkitsee* samaa kuin jokin aikaisemmin mainittu toinen ilmaus, mutta ei viittaa samaan tarkoitteeseen: esimerkiksi jossakin lauseessa on puhuttu *koirista* yleensä, sitten siirrytään puhumaan *tietystä vahtikoirasta x*. Kun samaa merkitsevään konstituenttiin viitataan pronomiinilla, puhutaan tällaisessa tapauksessa *merkitysanaforasta*: *Mikko syö mateen muhennoksena, mutta minä pidän siitä* [merkitysanafora] *savustettuna*. Eriviitteinen, "heikko" mainittuus käyttää sidoskeinoja jonkin verran eri tavoin kuin samatarkoitteisuus. Pronominit ovat harvinaisempia, ellipsi ja substituutiot (suom. *sellainen*, engl. *one*) tavallisempia. Esimerkkejä:

Joka tapauksessa lukemassani kerrottiin, että *ihmismassat* puserretaan ulos lentopaviljongin porteista, työnnetään *metallilieriöihin*, jotka korkataan ja jotka sitten liukuvat näkymättömiin ja *uusi lieriö* rullaa menneen 0 tilalle, jotta siihen puserrettaisiin taas uusi massa sisään. 79.6

[Ateenalaisen teatteriseurueen vierailusta kertova artikkeli] --meillä on vuosien varrella käynyt vierailmassa *monia maailmankuuluja teatteriseurueita*, mutta tekee mieli väittää, että *yhdenkään toisen 0* kohdalla ei voi puhua yhtä selvästi omintakeisesta ja jäljittelemättömästä ilmaisusta. 58.13

Heljä Angervo lauloi *pyhän Katariinan 0* ja Irja Auroora *pyhän Margareetan roolin*. 65.36

Schneiderhan soitti [Mozartin] C-duuri konserton KV 216 ja Hagenin johtamassa konsertissa kuultiin [Mozartin] B-duuri konsertto KV 207. 65.54

*Eugen Sarbu* [viulisti] ei puolestaan salaile tyytyväisyyttään siihen, että on saanut tällä tavoin räätäiltyönä tehdyn konserton, joka on hänen ikiomansa ainakin siihen saakka, kunnes *muut viulistit* keksivät sen - ja saavat sen harjoitelluksi. 64.10

Heikko mainittuus ei ole kovin tavallista kahden spesifisen, tarkoitteellisen NP:n välillä. Paljon tavallisempaa se on silloin kun sama merkitys (sense) toistuu tekstissä läheisissä lauseissa siten, että vain jompikumpi esiintymä on tarkoitteellinen taikka kun molemmat ovat *epäspesifisiä* taikka *geneerisiä*. Seuraavasta tässä muutamia esimerkkejä: kahdessa ensimmäisessä on kyse "luokkareferenssistä", eräänlaisesta lajista (tavallisesti yksilöön viittaava ilmaus viittaakin luokkaan, jota yksilö edustaa). Muissa on selvemmin kyse geneerisyydestä, puhutaan "jokaisesta relevantista yksilöstä jokaisessa relevantissa tapauksessa" (käsitteistä lisää ks. Vilkuna mt.).

*Majava* sen sijaan hävitettiin sukupuuttoon 1800-luvulla --

*Ihminen* muuttaa maisemaa. Hän luo myös *uutta maisemaa*, jossa on uusia aineksia ja muotoja. -

*Varkaus, myymälävarkauskin* on rikos, jota kenenkään ei kannata vähätellä tai puolustella. 38.9

Tiedetään, että *jalka* venähtää keskimäärin 10 mm vuodessa. 114.7

Jos nyt siis vaikkapa viimeksi mainitun lauseen jälkeen seuraisi lause, jossa on samanmerkityksinen nominaalilauseke viittaamassa spesifiseen tarkoitteeseen (*Tiedetään, että jalka venähtää keskimäärin 10 mm vuodessa. Mutta minun jalkani kasvoi viime vuonna 15 mm!*), esiintymät eivät tietenkään ole samatarkoitteiset, vaan jälkimmäisen kohdalle merkitään "heikosti mainittu". Toiseksi ilmaus voi siis olla epäspesifinen. Kielitieteellisessä

kirjallisuudessa "spesifisyys/epäspesifisyys" jakoa on sovellettu useanlaisiin ilmiöihin, jotka voisimme tässä jakaa kahteen ryhmään. Kvantifioinnin kannalta epäspesifiseksi voidaan luonnehtia ilmauksia, jotka viittaavat epämääräiseen joukkoon aktuaalisia tai potentiaalisia entiteettejä siitä nimenomaisesta syystä, että verbin aikasuhteet edellyttävät tätä. Jos esimerkiksi sanon *olen usein nähnyt haikaran*, tämän lauseen epäspesifinen tulkinta sanoisi, että minulle on tapahtunut useita "haikarannäkemisiä" - näkemäni haikaran ei ole tarvinnut ollenkaan joka kerralla olla sama. (Lauseella on spesifinenkin tulkinta, se jolloin olisin nähnyt usein jonkin erityisen haikaran.) Esimerkkinä tällaisesta kvantifioinnin epäspesifisyydestä:

Minä en osaa *koskaan* tunnistaa *ihmisiä valokuvista*, osaatko sinä? 68.12

Kussakin maanosassa järjestetään *alueellisia kokouksia* joka viides vuosi. 92.14

Epäspesifiksi taas sanotaan nominaali-ilmauksia, joilla ei omassa lauseyhteydessään ole tarkoitetta lainkaan. Kvantifioinnin spesifisyyden tapauksessa tarkoitteen (epämääräisen tarkoittejoukon) olemassaoloa ei sinänsä aseteta kyseenalaiseksi: jos joka viides vuosi järjestettiin yksi kokous, kokouksia todella pidettiin. Mutta eräissä lauseasemissa koko referenttien olemassaolo jää auki. Vahvin tällainen konteksti on kiellon vaikutusala:

Jussin ei onnistunut ampua *sorsaa*.

Mikään koekäsittely ei siis aiheuttanut tilastollisesti merkittäviä eroja vastuu- ja syyllisyyskäsitysten muutosten suuruudessa. 8.36

-- on afaatikkoja joilla ei ole todettavissa *älyllistä taantumaa* -- 12.36

Ensimmäinen Anjalan liittokirjan kappale palautettiin kuninkaalle ilman *allekirjoituksia*. 15.10 /parafraasi: -- niin että siinä ei ollut allekirjoituksia/

Ensimmäinen tulkinta on se, että *sorsaa*, tilastollisesti merkittäviä eroja, *älyllistä taantumaa*, *allekirjoituksia* mahdollisesti ei ollutkaan. Olennaista on taas, että tekstiä voidaan jatkaa näiden lauseiden jälkeen vaikkapa tarkoitteellisella, siis välttämättä "eriviitteisellä" nominaalilausekkeella:

Mikään koekäsittely ei siis aiheuttanut *tilastollisesti merkittäviä* eroja --. *Niitä* syntyi vasta kun muutuja 1 vaihdettiin --

Lauseasemia, joissa nominaalilauseke voi menettää tällä tavoin referenssipotentiaalinsa, kutsutaan sameiksi (opaque) tai intensionaaliseksi konteksteiksi. Niistä vahvin on kiellon vaikutusala; kaikille niille on yhteistä "toteutumattomuuden" merkityskomponentti. Sellaisia ovat edelleen mahdollisuutta tai (epä)todennäköisyyttä merkitsevien modaaliteettien vaikutusalat:

Hän sanoi olevansa sitä mieltä, että ellei ratkaisua kyetä pitämään kohtuuden rajoissa, saattaa seurauksena olla *sekä taloudellinen että poliittinen ketjureaktio*. 26.21

[Hypoteettinen jatkomahdollisuus: ja *sehän / sellainenhan / ketjureaktiohan* siitä sitten syntyikin.]

Kuva olisi varmaan toinen, jos Helsingissä olisi *saman kokoinen oopperatalo* kuin Glasgowissa, eli noin 1300-paikkainen, mutta koskahan *sekin* rakennetaan? 54.73

Lisäksi on joukko ns. intensionaalaisia verbejä, joihin sisältyy sama "toteutumattomuuden" komponentti: *toivoa, etsiä, hakea, haluta*: Jos sanomme *Pentti etsii uutta autoa*, voi olla että hän etsii spesifistä autoyksilöä, mutta (luonnollisempi) epäspesifinen tulkinta jättää auki mahdollisuuden, että tarvittavaa autoa ei ole olemassakaan.

Harvoin pääsee kokemaan elämän rytmiä niin juhlallisena kuin *hoitamattomassa ikimetsässä* - mikäli *sellaisen* Suomessa nykyisin löytää. 9.22

Seuraavassa muutamia esimerkkejä tapauksista, joissa heikkoa mainittuutta on jouduttu soveltamaan:

Se on sangen pitkälle *Suomen historiaa* [predikatiivi], koska *historia* [geneerinen] muodostuu olennaisesti poliittisluontoisista päätöksistä hallituksessa ja eduskunnassa. 41.2

Mutta toisaalta - *työntekijöinä* [predikatiiviadverbiaali: predikoiva, siis tarkoitteeton] nämä ihmiset ovat aivan ihmeellisiä. *Ahkerampia ja tunnollisempia työntekijöitä* [eks. epäspesifinen] saa hakea, enkä silti usko, että *sellaisia* [eks. epäspesifinen] löytyy. 117.67-68

Täysin tarkoitteettomallakin nominaali-ilmauksella voi olla aidosti samaviitteisiä mainintoja, edellyttäen että ne itse kuuluvat samaan "modaliteettiin" korrelaatin kanssa:

Olisi mahdottoman hyvä hetki aloittaa yhteistyö juuri nyt, löytää porvarien kaavailemille devalvaatioille ja palkkasuluille *vaihtoehto, joka* ei merkitsisi kansan verille piiskaamista, ja ajaa se läpi kaikella kootulla voimalla.

Esimerkki mainitsemistamme hyvin hienovaraisista eronteista olisi vielä ns. "laiskanpronominin" käyttö: *se* on seuraavassa esimerkissä oikeastaan selvästi eriviitteinen kuin korrelaattinsa:

Ikäkautta kahdestakymmenestä kolmeenkymmeneen voitaneen siis pitää *naimisiinmenon etsikkoaikana* [predikoiva]. Joka ei *etsikkoaikaansa* [heikosti mainittu] tunne, ottaa tilastojen mukaan suuremman riskin kuin se joka käyttää *sitä* hyväkseen. 118.3-4

(Esimerkin toinen virke on eräänlainen geneerinen lause, jossa kuitenkin täytyy merkitä *joka ei etsikkoaikaansa tunne* ja *se joka käyttää sitä hyväkseen* - kaksi eriviitteistä geneeristä persoonaa! - eriviitteisiksi.)

Olemme koodittaneet "heikosti mainituiksi" myös tapaukset, joissa kaksi ilmausta osittain kattaa toisensa referenssiltään. Tähän kuuluisivat siis "referentti" - "osa referentistä" -suhteet:

Elokuvasta voidaan puhua populaarikulttuurina vain, jos kriteerinä on lajille tyypillinen tuotantotapa, mutta -- Tässä tarkastelussa puhun ennen muuta *siitä elokuvataiteen osasta, jota* voidaan pitää populaarikulttuurina. 3.21-23

3. Nominaalilauseke viittaa *mainittuun*, sitä laajempaan jaksoon: lauseeseen, tekstikatkelmaan, edellä selostettuun asiaintilaan ylipäänsä. Tässä tapauksessa kyse on mainittuudesta aivan samoin kuin 1. kohdassa, mainittu tarkoite vain on asiaintila, juttu, tapahtuma tai muu "intensionaalinen objekti", jota yleensä ilmaistaan yhtä konstituenttia laajemmalla jaksolla. Esimerkeistä näkyy, että sidoskeinona käytetään kokonaiseen lauseeseen tai muuhun isompaan yksikköön viitatessa anaforista pronominia ja/tai etenkin parafraasi-ilmausta.

Mutta ilmeisesti hänen palava henkensä ja loistava virtuositeettinsa veti myös Gridenkosta esille parhaat puolet, sillä heti kun tämä jäi yksin lavalle, se merkitsi selvää tason laskua. 65.46

Ihmiset ovat nyt niin kummissaan Saloran veropetoksesta ja siitä että *poliitikot ottavat lahjoja*. Eivätkö he siis *sitä* ennen tienneet? 89.4-5

Lucian Berio -- on sijoittanut joka pulpettiin orkesterimuusikon partneriksi yhden kuorolaulajan. Ja katso: jo pelkästään *tällä tempulla* musiikki on saanut aivan uudet ulottuvuudet. 65.2-3

- Ei tätä kärsi koko aikaa katsella on se niin kaunista, sanoi mies. Vaimo yhtyi *havaintoon*. 71.11-12

Wrightin kaikki asuintalot kasvavat elimellisesti ympäröivästä luonnosta oli maisema sitten viljelty tasanko tai jylhä vuorenrinne. Tästä luonnon ja rakennuksen yhteensulautumisesta johtuu määre orgaaninen arkkitehtuuri, joka toistuu Wrightia luonnehdittaessa. 66.23-24

4. Tarkoitetta ei ole tekstissä mainittu, *mutta aikaisempi teksti implikoi sen mukana olon*. Jos esimerkiksi sanotaan näin:

Kuten poliisi eri yhteyksissä on todennut, tutkanilmaisin ei pelasta joutumasta kiinni ylinopeudesta. Laite varoittaa *kuljettajaa* nimittäin vasta silloin, kun mittaus jo tapahtuu. Tällöin kuitenkin *syytetylle* jää tieto siitä hetkestä ja paikasta missä mittaus tehtiin, ja hänen oikeusturvansa paranee. 25.9-11

lukija tuskin joutuu kysymään: "ketä kuljettajaa?" tai "mille syytetylle?" tai "kuinka niin oikeusturvaa?", vaikka ei ole ollut puhetta enempää kuljettajasta kuin oikeudenkäynnistäkään, syytetystä puhumattakaan. Teksti käsittelee autoilijoille tarkoitettua tutkanilmaisinta, ja tutkat taas ovat laitteita joilla poliisi voi paljastaa ja saattaa oikeusvastuuseen, siis syytetyiksi ja oikeusturvaa vaativaan asemaan kuljettajat. Kuljettajan ja syytetyn olemassaoloa implikoi se mitä edeltävässä tekstissä sanotaan. Tämä implikointi on pragmaattista luonteeltaan eli perustuu siihen mitä lukija tietää maailman menosta. On tietysti suhteellista, mitä voidaan pitää tällä tavalla implikatiivisesti tunnettuna. Kiistämätöntä on kuitenkin, että olisi suorastaan Gricen kvantiteettimaksiimin rikkomista jos esittelisi erikseen uusina asiat jotka tekstiyhteys riittävän selvästi jo implikoi - siis tekstin

(a) sijasta kirjoittaa tekstin (b) tapaan:

(a) Alkukesästä 1975 Indira Gandhi tuomittiin Allahabadin paikallisoikeudessa esteelliseksi pääministerin viran hoitoon. *Syytettä* oli ajanut Gandhin poliittinen vastustaja Jayaprakash Narayan. 24 .29-30

(b) Alkukesästä 1975 Indira Gandhi tuomittiin Allahabadin paikallisoikeudessa esteelliseksi pääministerin viran hoitoon. Häntä vastaan oli nostettu *syyte*. *Syytettä* oli ajanut Gandhin poliittinen vastustaja Narayan.

Toisesta näkökulmasta katsoen voidaan sanoa, että implikatiivisesti tuttu tarkoite esiintyy tutussa, jo esitellyssä "merkitys-" tai "käsitekentässä". Jos teksti käsittelee jotakuinkin rajattua aihetta, syntyy pian koko tekstinkin kattavia implikoidun tunnettuuden ketjuja. Esimerkiksi artikkelista "Valitse huolella koulukengät" voi poimia ainakin seuraavat "kenkien-hankkimiskenttään" kuuluvat asiat esiintymisjärjestyksessään: *kasvunvara, jalka, varpaat, kengän kärki, pakkaskengät, kulutus, ulkopohjan materiaali, sisäpohja, nahka, päällismateriaali, hiostavuus, vuori, nahkasaappaat, varsikengät, sukat, kumisaappaat, irtopohjalliset, antura, kaariosa, kantapäät, kärkiosa, korko, solukumipohjat, voimistelutossut, puukengät, jalkapohjan lihakset, jalkapöytä, vaivaisenluut, vasaravarpaat, hankautumat ja hiertymät*. Tästä käy ilmi, että implikatiivinen tunnettuus tuo aina mukanaan tekstiin *leksikaalisen koheesion* - leksikaalinen koheesio vain on laajempi käsite, koska se kattaa muutkin kielen elementit kuin tarkoitteelliset nominaalilausekkeet. Seuraavassa esimerkkejä implikoidusta tunnettuudesta:

Meillä on alusta asti kiinnitetty erityisen suurta huomiota mahdollisimman korkeaan laatuun [mattokutomossa]. *Käytettävät materiaalit* ovat korkealaatuisia ja *työn taso* samoin. Pelkästään käsin kutomalla kärsii laatu, sillä enää iltapäivällä ei *kutoja* jaksa lyödä niin tasaisen tiivistä, että syntyisi *tasalaatuista jälkeä* -- 170.40-43

[samasta mattokutomosta] Olosuhteet täällä saariston sydämessä ovat erikoiset. *Kuljetukset lauttojen ja lossien kautta mantereelle ja sieltä tänne* vievät aikaa, *työvoiman saannissa* on omat vaikeutensa ja *materiaalihankinnat* vaativat maksuvalmiutta. 117.45-47

Korppoossa on vain kansakoulun ala-aste ja sen jälkeen *lapset käyvät* koulua Paraisilla. Kun sitten lapset ovat viikot Paraisilla ja vain viikonloput kotona muuttavat *vanhemmatkin* usein kokonaan täältä pois -- 117.61-62

Taloustekoneita Oikaraiset eivät lainaile. *Pakastearkku* on molemmilla ja pyykin Selma Oikarainen käy pesemässä *tyttärensä automaattikoneella, joka sekään ei ole kaukana*. 129.71-71

Oman implikoidun tunnettuuden tyyppinsä muodostavat sellaiset aina synsemanttiset eli tekstistä täydennystä vaativat sanat kuin *ero, esimerkki, syy, seuraus, tulos, ansio*, jne.:

Vaikka pidinkin Finlandia-talon esitystä tehokkaampana, ei ero kuitenkaan ollut niin suuri kuin olisi saattanut luulla. *Ansio* oli ennen kaikkea sävellyksen: -- 65.31-32

Tätä hyväksi tiedettyä menetelmää ovat vuoden ajan soveltaneet Einojuhani Rautavaara ja Eugen Sarbu; *sen tulokset* kuullaan tiistai-iltana -- 64.2

5. Tarkoite on *edellytetty eli pragmaattisesti tunnettu*. Ennalta mainitsematon tarkoite voidaan olettaa yksiselitteisesti tunnistettavaksi paitsi tekstissä jo sanotun myös kielenkäyttötilanteen ja vastaanottajan pragmaattisen tiedon perusteella. "Kielenkäyttötilanne" on tällöin ymmärrettävä maksimaalisen laajasti: siihen kuuluu kaikki puhujan ja kuulijan yhteisessä kokemusmaailmassa oleva, siis paitsi esineet ja ihmiset myös yhteinen tai yhteiseksi oletettu tieto maailmasta. Joskus on selvä, ettei puhuja/kirjoittaja voikaan tarkoittaa kuin yhtä tiettyä oliota. On asioita, joita on vain yksi: *aurinko* ja (maassa) *kuu, Jumala, Suomen presidentti*. Samoin on asioita jotka ovat puhetilanteessa yksiselitteisesti tunnistettavissa: perheen kesken sanotaan lukemattomia kertoja *sulje ovi* tai *vie koira ulos*, ja aina on selvä mihin oveen tai mihin koiraan viitataan. Mutta on myös asioita, jotka pystyy identifioimaan se joka tietää - *Armi, Anne ja Jacques, Kekkonen, Oulu, Korpilammin kokous, Christina Onassiksen avioliitto, Moron kidnappaus, Ruotsi-Suomi -maaottelu* ovat ilmauksia, joiden referentit keskivertosuomalainen (tämän kirjoittamisen aikoihin!) mitä todennäköisimmin tuntee. Kirjoituksen luonteesta eli siitä kenelle se on tarkoitettu riippuu, mitä lukijan voi olettaa tuntevan: musiikkiarvostelussa säveltäjät, suuret kapellimestarit ja teokset, poliittisessa uutisessa puolueet ja poliitikot, urheilu-uutisissa huippu-urheilijat ja urheilulajit. Pragmaattisen tunnettuuden kategoriaan tuleekin eniten juuri *erisnimiä*: henkilönnimiä ja maantieteellisiä nimiä, mutta myös tunnettuja tapahtumia.

Pragmaattisen tunnettuuden frekvenssi olisi suurempi, jos aineistona olisi luonnollisissa keskustelutilanteissa puhuttua kieltä: kirjoitetuista teksteistä puuttuu fyysiseen lähiympäristöön viittaamisen mahdollisuus. - Myös "implikoitu tunnettuus" perustuu tietysti suurelta osin pragmaattiseen tietoon maailmasta. Ero on siinä, että implikoidun tunnettuuden tapauksessa "ohjeet" referentin tunnistamiseksi löytyvät edeltävästä tekstistä, pragmaattisen tunnettuuden tapauksessa eivät. Kaikki henkilönnimet eivät ole pragmaattisesti tunnettuja: vähemmän tunnetut henkilöt täytyy erikseen esitellä, jolloin he saavat attribuutteja eteensä:

Suuret mestarit ovat esimerkkejä. *Joku Pentti Nikula ja hänen seiväshypyn maailmanennätyskarusellinsa* häivytti aikoinaan heinäseipäätkin pelloilta, kun pikkupojat kautta maan saivat kipinän seiväshyppyyn. *Lasse Viren* vei pikkupojat tiilimurskalle ja maantielle. *Trial- moottoripyöräilyn viimevuotinen maailmanmestari Yrjö Vesterinen* on saanut aikaan sen, että mini-trialista on kuluvan kauden aikana tullut pikkupoikien varsinainen pop-laji eteläisessä Suomessa -- 108.4

Tekstin sidoksisuuden kannalta on syytä huomata, että pragmaattisesti tunnettu nominaalilauseke voi olla *tekstuaalisesti täysin vapaa*. Usein siihen, ainakin sen attribuutteihin, kuitenkin liittyy tekstuaalista koheesiota, kuten edellisessä esimerkissä. Lisäesimerkkejä:

[artikkelissa ennätyksen tehneestä neuvostoliittolaisesta seiväshyppääjästä]  
*Maailmancupin kohinat* eivät olleet vielä ehtineet Vladimirin korviin. 99.19

[poliittisessa pakinassa] *Sorsan kakkonen* on ottamassa mittaa *taloudellisista vaikeuksista* ja sovittelemassa askelmerkkejään. 74.1

Kun postin pääjohtaja kansanedustaja Pekka Tarjanne saapui viime tiistaina *toiselle työpaikalleen eduskuntaan*, kävivät monet kysymässä mitä hän siellä tekee. Erityisesti *keskustapuolue* odotti innokkaasti hänen sijaistaan, ensimmäistä helsinkiläistä kansanedustajaansa Kyllikki Stenroosia. Ongelma Tarjanteen kansanedustajan paikasta syntyi kun *Tasavallan presidentti* heinäkuun alusta nimitti Pekka Tarjanteen postin pääjohtajaksi. 107.1-3

6. *uusi*. Koska uusia tarkoituksia voi tulla tekstiin muutakin kautta kuin tarkkailemamme neljän lausekkeen mukana, muuttujaan (53) kooditetaan uusien tarkoitteiden kokonaismäärä.

i. Tunnettuus on *irrelevantti*. Kuten jo on todettu, predikatiivi on harvoin tarkoitteellinen, ja siksi ao. konstituentilla ei ole koko tunnettuusmuuttujaa. Muita tunnettuudeltaan irrelevantteja tapauksia:

- lause, tai sellaiseksi kooditettava rakenne (myös *se, että* ---)
- konstituentti, joka on "inkorporoitu verbiin" (subjekti) tai 0 (geneerinen ellipsi)
- eksoforiset *minä, sinä* (ne ovat tunnettuja per definitionem)
- kysymyssanat ja -lausekkeet (*kuka? kuka poika?*)
- refleksiivi- ja resiprookkipronominit: (*toisiaan, itsensä*), koska ne ovat lauseensisäisiä samaviitteisyyselementtejä
- pääsanattomat kvantorit (ellei ole selvää pääsanana ellipsiä): *kaikki, jokainen, joku, ei kukaan, paljon, ei mitään, harvat*; lukusanat ja määrän ilmaukset: *100 %, kilo, markka, metri*
- idiomien pakolliset jäsenet: *Sarbullalla on pullat hyvin uunissa*.

Pahin jäljelle jäävä epäloogisuus lienee se, että predikatiivin tai muun referentittömän ilmauksen relatiivipronomini kooditetaan relatiivipronominien yleiseen tapaan "aidosti samaviitteiseksi" korrelaattinsa kanssa:

Tuskin on olemassa *ihmistä, joka* suhtautuu välinpitämättömästi kissaan.

Kalle on *ihminen, joka* suhtautuu välinpitämättömästi kissaan.

*Kaikki* ottivat varaslähdön. 98.46

Partiopoijat ja heidän järjestönsä eivät tosiasiallisesti kerää *yhtään kilo*a. 96.3 [ei tulkittavissa niin että NP *paperia* olisi poistettu]

Kaikesta huolimatta ankeriaasta tiedetään *aivan liian vähän*, jotta suunnitelmallisiin yleiseurooppalaisiin istutuksiin voitaisiin lähteä. 102.6

-- tässä nyt tarjotaan *jotakin erikoista*. 101.43

15-vuotiaana Vladimir hyppäsi *203*, seuraavana vuonna *212*, vuonna 1976 *222* ja tänä kesänä *233*. 96.57

*Pääpaino* on kuitenkin vaatteilla, joita on pidetty harvoin ja varoen --- [näyttelyssä] 103.9

Jollakin on jonkinlainen vaikutus, suuri merkitys, jokin ominaisuus jne. 13.7

Sen loppuraportissa ei luvata *hyvää eikä paha*a. 94.28

Ns. predikatiivinen objekti (objektipredikatiivi) kooditetaan tunnettuudeltaan irrelevantiksi. Tällä tarkoitamme sellaisen rakenteen objektia kuin.

mikä tekee elokuvasta *suositun*

mikä tekee elokuvasta *populaarikulttuuria*.

Nämä lauseet ovat selvästi paralleellisia lausetyypille:

-- elokuvasta tulee *suosittu*

-- elokuvasta tulee *populaarikulttuuria*

ja objekti on niissä selvästi predikoiva, mikä näkyy siitäkin, että "objektina", voi olla adjektiivi.

Adverbiaalit on merkitty samalla tunnettuusmittarilla kuin muutkin nominaaliset lausekkeet. Usein kuitenkin adverbiaalina esiintyy pelkkä adverbi: *aina, joskus, mielellään, onneksi, pian, tänään*, jne. Näillä ei ole tarkoitetta, ja siksi ne merkitään tunnettuudeltaan sidoksisiksi. Ajan ilmaukset, kuten *1800-luvulla, viime viikolla*, on katsottu tunnettuudeltaan irrelevanteiksi - oikeastaanhan ne ovat pragmaattisesti tunnettuja. Paikan ja valenssiadverbiaalit sen sijaan voidaan vaivatta merkitä tunnettuudeltaan, koska niillä on tarkoite. Pre- ja postpositiolausekkeenomaiset fraasi-ilmaukset, kuten *jonkun puheilla, nimissä, avulla, puolesta, merkeissä, tiimoilta, lisäksi, tavoin, tiellä, selän takana* jne. kooditetaan luonnollisista syistä tunnettuudeltaan lausekkeen "sisältösanan", ei syntaktisen pääsanan eli "p-position" mukaan.

Seuraavassa luettelo adverbiaalityypeistä, jotka kooditetaan tunnettuudeltaan irrelevanteiksi: lauseeksi kooditettavat rakenteet kuten subjektin ja objektin kohdalla sanottiin; kaikki "pääsana + lause"-rakenteet, joiden pääsana ei ole selvästi tarkoitteellinen sisältösana (siis: *ilman että, siitä huolimatta että, aina kun, siitä syystä että* jne.); kysymyssanat, eksoforiset ilmaukset, kvanttorit ja muut (kuten subjektin ja objektin tapauksessa); infinitiivilausekkeet (rakennetyyppi 4); predikatiiviadverbiaalit; OSMA:t ja määrän ilmaukset yleensäkin; kommentit, konnektiivit ja modaaliset adverbiaalit, paitsi mikäli konnektiiviin liittyy sisältösana tai edelläolevan kanssa samaviitteinen elementti; tavan adverbiaalit: *hyvin, kauniisti, ilkeästi, ylinopein tempoin*, sekä adverbit.

Kokemus on osoittanut, että tunnettuuden käsittely on sitä vaikeampaa, mitä abstraktimpi käsiteltävä ilmaus on. Ei tietenkään voi lähteä siitä, että *tarkoitteita* olisivat vain konkreettiset, aikapaikkaiset oliot. Koodittaja tuntee liikkuvansa heikoilla jäillä miettiessään, onko *Vygotskin merkitys neuvostopsykologian kehitykselle* "sama" merkitys kuin *Piagetin merkitys kehityopsykologialle*. (Tässä tapauksessa vastaus olisi kielteinen: konkreettiset määritteet "yksilöivät" abstraktinkin ilmiön, ja mikäli nämä kaksi NP:tä esiintyisivät tekstissä lähekkäin, kyse olisi korkeintaan "heikosta mainittuudesta".) Myös rajanveto spesifisten ja geneeristen abstraktien välillä voi olla hankalaa.

Kooditustyön aikana kokeilimme jonkin aikaa eräänlaista tukimuuttujaa. Koodittaja kirjasi jokaisesta kooditettavasta nominaalijäsenestä sen "abstraktiotason". Tukena muuttujan laadinnalle oli Lyonsin (1977:438-450) jako ensimmäisen, toisen ja kolmannen asteen entiteetteihin, joihin viittaavat ensimmäisen, toisen ja kolmannen asteen nominaalit. Muuttuja voisi näyttää seuraavanlaiselta:

1 inhimillinen tai personifioitu, ts. aloitteelliseen toimintaan kykenevä olio tai sellaisten joukko: *minä, Eugen Sarbu, Helsingin yliopiston opiskelijat*

2 esine tai esinejoukko, paikka: *moottoripyörä, Helsinki, työpaikkani*

3 yhteisö, instituutio: *perhe, SAK, Keskustapuolue, Historiallis-kielitieteellinen osasto*

4 aine (konkreettinen, jaollinen tarkoite): *kahvi, jauhot, puolukka* milloin kysymyksessä ei ole yksityinen marja

5 teko, tapahtuma, prosessi (Lyonsin "toinen aste"): *konsertti, esitys, näyttely, ekskursio, vallankumous, sade, kysymys, tuomio, ympäristön tuhoutuminen*. Suuri osa verbijohdoista nomineja, erityisesti *-minen*-johdokset, kuuluvat tähän ryhmään.

6 abstrakti (Lyonsin "kolmas aste"): *taide, tiede, rakkaus, valtion kassatilanne, surrealismi, asenne, näkemys, todellisuus, tavoite*

i irrelevantti: milloin konstituentin tunnettuus merkitään irrelevantiksi

On huomattava, että status ratkaistaan kulloisestakin tarkoitteesta: esim. *Valintatalo* voi yhtä lailla viitata liikkeeseen (2) kuin instituutioon (3).

Tällaiseen muuttajaan kannattaa turvautua, jos tarkasteltavana on erityisesti tunnettuus, samatarkoitteisuus jne. Mitään objektiivista muuttujaa tällaisesta tuskin saa, pikemminkin sen avulla ovat paikannettavissa epävarmuuden alueet. Muuttujia ristiintaulukoitaessa voidaan esimerkiksi pudottaa pois abstrakteimmat ryhmät, jolloin jäljelle jäävät kiistattomimmat ja siis luotettavimmin kooditetut tapaukset. Itsenäistekin käyttöarvoa abstraktiotason silmällä pitämällä tietysti voi olla, esim. tarkasteltaessa kokonaisten tekstien ominaisuuksia tai tyyliä vertailtaessa.

(22) *tarkoitteen mainittuus* (myös 30,42,49)

- 1 mainittu edellisessä lauseessa
- 2 mainittu edellisessä virkkeessä
- 3 mainittu kauempana samassa yhteydessä
- 4 mainittu eri yhteydessä, kaukana
- i irrelevantti

Kun tunnettuus-muuttujassa on alakategoria 1,4 tai 5, kooditetaan tähän muuttajaan se, miten etäällä edeltävässä tekstissä aikaisempi maininta on. Etäisyys lasketaan aina välittömästi edeltävästä maininnasta, ei ensiesiintymästä. Huomattakoon, että edellinen "esiintymä" ei itse ole aina konstituenttina kirjattu, sillä se voi olla kokonainen lause tai muu pitempi jakso, tai se voi olla lausekkeessa attribuuttina (tuolta tulee *Pekan* auto. *Hän* ajaa lujaa), jopa lauseen keskipaikkaisena adverbialina, jota järjestelmässämme ei kooditeta. "Maininnaksi" lasketaan myös tekstuaalinen ellipsi. Jos samaa NP:tä ei ole ennen mainittu, ts. jos se on implikaatiivisesti tai pragmaattisesti tunnettu taikka kokonaan uusi, sarakkeeseen merkitään 0. Tämä muuttuja on irrelevantti vain silloin kun edellinenkin on.

Toisin kuin kooditussuunnitelman mukaan olemme tehneet, suositamme tähän kohtaan yhtä lisäkatgoriaa. Virkehän voi koostua hyvinkin monesta lauseesta, ja mitä useampia lauseita on, sitä useampia nominaalilausekkeita voi tulla kahden maininnan "väliin". Siksi kannattaa erottaa

"välittömästi edeltävä lause" ja "kauempana samassa virkkeessä oleva lause". Tarkennettu järjestelmä näyttää seuraavalta:

1 edellinen maininta saman virkkeen välittömästi edeltävässä lauseessa:

Junassa olen kohdannut *miehen, joka* tappoi kepakolla rantavesissä kutevan 11-kiloisen hauen; -- 79.15

Anne Pohtamon poikaystävä Jacques oli ostanut *Annelle timanttinauhan* kaulalla kannettavaksi mutta hukannut *sen* New Yorkissa ennen kuin ehti antaa *lahjan armaalleen*-- 82.6

2 edellinen maininta samassa virkkeessä kauempana:

*Anne Pohtamon* poikaystävä Jacques oli ostanut *Annelle* timanttinauhan kaulalla kannettavaksi mutta hukannut *sen* New Yorkissa ennen kuin ehti antaa *lahjan armaalleen*-- 82.6

3 edellinen maininta välittömästi edeltävässä virkkeessä, siis yhden pisteen takana:

Toivottavasti selviät, sanoi *nuori ammattiveli*, joka itse sinkoilee imateitse maanosien välillä kuin tyhjää vain. En voinut vastata *hänelle* -- 79.11

Kussakin maassa järjestetään *alueellinen kokous* joka viides vuosi. Eurooppaan *Unesco-toimikuntien alueellinen kokous* pidettiin viime keväältävena Helsingissä -- 92.15

4 edellinen maininta kauempana kuin edeltävässä virkkeessä, mutta vielä samassa yhteydessä.

5 edellinen maininta erikoisen kaukana tekstissä, eri tekstiyhteydessä. Edellisen ja tämän kohdan ero ei ole itse mainintojen välimatkassa vaan tekstin luonteessa. Tätä alakategoriaa sovelletaan vain mikäli edellinen maininta on tekstin osassa, joka käsittelee jotakin aivan muuta tai jos kahden maininnan välissä on asian käsittelyssä tullut olennainen katkos tai ekskursio. Tavallinen homogeeninen alle 100 virkkeen teksti, jollaisia olemme käyttäneet, on hyvin harvoin niin selvästi jakautunut, että tätä kategoriaa voitaisiin soveltaa.

(23) *tekstuaalinen sidoskeino* (myös 31,36,43,50)

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparatiivi t. muu kieliopillinen
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen sidos
- 8 vapaa
- 9 eksofora
- i irrelevantti

Tekstuaalista sidonnaisuutta syntyy, kun ilmaus on jossakin mielessä edeltävän tekstin perusteella tunnettu; sidoskeinot ovat osaksi tunnettuuden seurausilmiöitä. Mutta pääsanan ilmaiseman tarkoitteen ei tarvitse olla tunnettu jotta siihen voisi liittyä sidoskeino. Määritteenä oleva pronomini tms. voi sitoa koko lausekkeen johonkin edellä mainittuun, vaikka itse pääsana tuokin tekstiin uuden tarkoitteen. Yhdessä nominaalilausekkeessa voi käyttää useampaa kuin

yhtä sidoskeinoa: määritteessä on jokin keino, pääsanassa toinen. Siksi muuttujan alakategoriat on järjestetty hierarkkisesti. Jos sidoksisuutta ilmenee sekä määritteisä että pääsanassa, valitaan "vahvin", siis numeroltaan pienin alakategoria edustamaan kaikkia. Harkitseminen arvoista olisi, pitäisikö sidoskeinoa varten varatakin kaksi muuttujaa yhden sijasta: toinen pääsanalle, toinen määritteelle.

1. *ellipsi*. Nominaalisen konstituentin tekstuaalisista sidoskeinoista selvästi vahvin on ellipsi. Tässä muuttujassa ei ole kysymys koko konstituentin poistosta, sillä se on kirjattu rakennevaihtoehdoksi kuhunkin konstituenttiin esim. (19:D): Hän katsoi laskua ja *O* löi sitten kädellään otsaansa. Tässä muuttujassa kirjataan ns. *pääsan*an ellipsi (ks. Hakulinen (1978) ja muuttuja (59)). Esimerkkejä:

Heljä Angervo lauloi pyhän Katariinan *O* ja Irja Auroora pyhän Margareetan *roolin*.

Toista päätä ja puolta pitää hallussaan Teutorin sisaren mies, toista *O O* punapäät sisaret.

Vaje oli viime vuonna neljä miljardia ja vuotta aiemmin peräti kahdeksan *O*.

Tässä mielessä *Coro* on kiintoisa *sävelteos*, joskaan ei suinkaan ongelmaton *O*. 65.7

Esimerkeistä näkyy, että pääsanan ellipsi ei ole samatarkoitteisten vaan nimenomaan eritarkoitteisten mutta samamerkityksisten ilmausten sidoskeino.

2 *anafora*. Toiseksi vahvin ja kieliopillistunein sidoskeino on anaforisten pronominien käyttö: elementtien, joilla ei ole muuta tai juuri muuta merkitystä kuin tehtävä edustaa jotakin samaviitteistä ilmausta uudelleen mainittaessa. Anaforisia ovat *hän* ja *se*, kirjoituksessa useimmiten myös *tämä* ja *tuo* (niitä on nimitetty myös tekstuaalisesti deiktisiksi - nehen viittaavat ilmaukseen, eivät tarkoitteeseen, vrt. Lyons (1977:667)), adverbiaalina ovat pronominaali-ilmaukset *silloin*, *siten*, *sitten*, *siksi* jne. sekä nimenomainen eriviitteisyyden anafora *sellainen*. Anaforasta on esimerkkejä etenkin tunnettuuden kohdassa (21:1) edellä.

3 *katafora*. Anaforan vastapooli on katafora: pronominaalielementit edustavatkin jotakin tekstissä myöhemmin seuraavaa. Katafora eroaa siksi ratkaisevasti muista sidoskeinoista: siihen ei lainkaan tarvitse liittyä tuttuutta, se vain signaloi, että kyseisestä elementistä kulkee sidos eteenpäin. Pronominien lisäksi tyypillinen kataforasana on *seuraava*, jota käytetään paljon metatekstuaalisissa kommentteissa. Katafora on harvoin tekstuaalisesti tärkeä, useimmiten se on kompleksisten lauseiden sisäinen syntaktinen selvennyskeino; tällöin se liittyy "pääsana + lause" -muotoiseen konstituenttiin ennakoimaan tulevaa lausetta.

Siirtyminen kaudesta toiseen tapahtuu *silloin*, kun lapsen teknisoperatiiviset mahdollisuudet eivät vastaa *niitä* toimintatehtäviä ja motiiveja, joiden pohjalta ne ovat kehittyneet. 2.14

Elokuvan luonnetta populaarikulttuurina voidaan osuvasti kuvata *seuraavalla* päättelyllä: elokuvat ovat *sikäli* populaarikulttuuria kuin niitä katsotaan suurin joukoin; toisaalta niitä katsotaan suurin joukoin *sikäli* kuin niiden tekemisessä on noudatettu tiettyjä sääntöjä. 3.24

4 *toisto*. Lausekkeen tai sen osan toistaminen peräkkäisissä lauseissa pronominiikäytön sijaan tyyliin *Satu meni saunaan. Satu pani laukun naulaan*. katsotaan useimmiten epäkohesiiviseksi. Mutta mitä enemmän asiaa kahden esiintymän väliin tulee lisäämään harhatulkintojen mahdollisuutta tarkoitteen tunnistamisessa, sitä enemmän tarvitaan toistoa. Tästä esimerkkinä seuraava tekstikatkelma:

Kiovan liikuntainstituutissa opiskeleva *Jashtshenko* on ennakkoluuloton ja rohkea hyppääjä. *Hän* ei aliarvioi kykyjään, mutta *O* ei toisaalta yliarvioikaan. Hän on toistaiseksi vaatimaton mutta terveen itsetuntoinen ukrainalaispoika. *Hän* eroaa melkoisesti tyypiltään

*korkeushypyn showmiehestä Stonesista. Stones on myös hyvä hyppääjä. Hän on myös hyvä näyttelijä. Hän rakastaa esiintymistä ja hänen näyttävään tyyliinsä floppi sopii oikein hyvin, arvioi Telegin. Jashtshenko nyökkäsi myöntävästi. 99.1-8*

Seuraavassa esimerkki toiston ja anaforankin käytöstä sidoskeinona eriviitteisten ja erimerkityksisten ilmausten attribuutteina:

Brahmsin viulukonsertto on klassinen esimerkki säveltäjän ja tulkitsijan jo teoksen syntyvaiheessa alkaneesta yhteistyöstä. *Tätä* hyväksi tiedettyä menetelmää ovat vuoden ajan soveltaneet Einojuhani Rautavaara ja Eugen Sarbu; *sen* tulokset kuullaan tiistai-iltana, kun *Rautavaaran* viulukonsertto saa kantaesityksensä *Sarbun* tekemine kadensseineen. 64.1-2

Elokuvasta voidaan puhua populaarikulttuurina vain, jos *kriteerinä* on lajille tyypillinen tuotantotapa, mutta tämä on *kriteeri*, jonka mielekkyyttä ei kovin pitkälle voi puolustella. 3.21

5 *komparatiivinen sidos*. Tekstuaalisesti sitovia voivat olla (lauserajan yli viitatessaan) komparatiivimuodot sekä komparatiivisanat kuten *sama, eri, erilainen, toinen, vastaava, muu* jne. Myös varsinaiset konnektiivit (39:4) merkitään sidoskeinoltaan komparatiivisiksi (esim. *lisäksi, toisaalta, päinvastoin, silti, kuitenkin*) elleivät ne ole anaforisia. Mikäli kyse on tarkoitteellisista ilmauksista, komparatiivinen sidos edellyttää eriviitteisyyttä.

Kansanmusiikin määrittelyä pidettiin tärkeänä, koska ei-kansanmusiikki haluttiin ideologisista syistä jättää tutkimuksen ulkopuolelle. Tämän artikkelin näkökulma on *toinen*, ja siitä syystä en yritäkään aloittaa populaarimusiikin määritelmällä. 5.2-4

Syntyi uudentyyppisiä huvipaikkoja, joissa musiikkia ja muuta ohjelmaa [huom. komparatiivinen sidos keskellä lauseketta] esitettiin maksua vastaan. Tyypillinen esimerkki tästä oli englantilaisen music hallin esitys. *Muita vastaavia* muotoja olivat kabaree, vaudeville, operettiteatterin eurooppalaiset ja amerikkalaiset muodot jne. 4.38-40

Emotionaalinen eläytyminen filmin sanoman suuntaisesti toimivan elokuvan henkilön ongelmiin varmastikin lisää sijaisvahvistamisen vaikutusta. Jos katsoja taas sattuu samaistumaan sellaiseen filmin henkilöön, joka toimii filmin varsinaisen sanoman vastaisesti ja jonka toimintaa johdonmukaisesti rangaistaan, on seurauksia *vaikeampi* ennustaa. 8.6-7

6 *synonyymi, parafraasi, yläkäsite*. Samaviitteisyyden tapauksessa pelkkä toisto ja pronominaalistus eivät aina täytä tyyli vaatimuksia. Toistamisen sijasta voidaan käyttää toistettavan ilmiön synonyymia tai lähes-synonyymia, parafrasaa eli kiertoilmausta, yläkäsitettä jne. Parafraasin käyttö on yleistä etenkin silloin, kun viitataan kokonaiseen lauseeseen tai muuhun suurempaan kielen yksikköön.

Syntyperästään ylpeät ukrainalaiset ottivat *Vladimirin* heti ottopojakseen. Kun *192-senttinen zaporozhelaispoika* hyppäsi Donetskissa, ei stadionilta saanut kuulua hiirenhiiskausta. 99.13-15

*Työttömän* taas tuli ilmoittautua kuntansa työttömyyslautakunnalle, joka todettuaan *asianomaisen* täyttävän annetut ehdot osoitti hänet työmaalle. 10.9 (*Asianomainen* lienee parafrasasin ja pronominin rajatapaus.)

Hänen viulukurssinsa saavuttivat erinomaisen menestyksen sekä viime vuonna Tapiolassa että tänä kesänä *Rovaniemellä*, ja mikä tärkeämpää, myös Rovaniemen musiikkijuhlat pääsivät hänen alkupotkunsu avulla hyvään vauhtiin: -- parhaat konsertit olivat loppuunmyytyjä ja *kaupungista* eivät turistit ja omatkaan kuulijat vähällä loppu. 64.16

7. *leksikaalinen koheesio*. Tämä käsite on peräisin Hallidayn & Hasanin teoksesta (1976) ja sillä tarkoitetaan sitä yleisestä tekstikoheesiosta seuraavaa asiaa, että peräkkäisten lauseiden sanat jossakin mielessä usein kuuluvat merkitykseltään toistensa yhteyteen. Ne voivat muodostaa keskenään sarjoja, vastakohtapareja, osa-kokonaisuus -suhteita, assosiaatioketjuja, kollokaatioita eli totunnaisesti yhdessä esiintyviä sanapareja jne. Se saattaa olla hyvinkin hienovaraista ja jää usein meidän kooditussysteemimme saavuttamattomiin, etenkin kun leksikaalinen koheesio on sidoskeinojen hierarkiassamme heikoin, ts. jää useimmiten muiden "jalkoihin". Systeemiltämme jää tavoittamatta myös *verbeissä* oleva leksikaalinen koheesio, josta ensimmäinen esimerkki antaa kuvan.

Sopii toivoo, että jälleennäkemisestä syntyy (musiikillisessa mielessä) etelämaalaisittain mahdollisimman *tulinen*. Tämänvuotisilla juhlahiikoilla ei näet olekaan liiemmästi *sinkoillut kipunoita*, jotka olisivat *sytyttäneet* kesäsateissa vettyneen yleisön. 64.20-21

[artikkelissa ympäristön muuttumisesta] Nopeimmin muuttuvat *vesistöt, kasvipeite ja suot. Järvien matalat lahdet* kasvavat umpeen, kun *joet ja purot* tuovat niihin liejua ja *pohjalle* kerrostuu hiljalleen vesikasveja. Soiden *rahkasammal*kerros kohoaa muutaman millimetrin vuodessa ja levittäytyy *kankaille*, kunnes suo vuorostaan kuivuu ja väistyy *metsän* tieltä.--

8 *vapaa*. Tekstuaalisesti vapaaksi konstituentti merkitään silloin, kun se ei sisällä mitään tekstuaalista sidoskeinoa.

9 *eksofora*. Eksoforiset eli deiktiset elementit, jotka viittaavat puhetilanteessa johonkin puhujan ja kuulijan yhteisesti tuntemaan oloon, eivät ole varsinaisia tekstuaalisia sidoskeinoja. Kirjoitetuissa teksteissä ne ovat harvinaisia, koska kirjoittaja ja lukija eivät ole deiktisyyden vaatimassa välittömässä ajallisessa ja paikallisessa yhteydessä. Ainoita itsestään selviä eksoforisia elementtejä ovat 1. ja 2. persoonan pronominit; tavalliset deiktiset pronominiainekset kuten *tämä* tai *tuossa* ovat kirjoitetuissa teksteissä yleensä tekstuaalisissa eivätkä eksoforisissa käytöissä (vrt. ed.). Tekstien tavallisimpia eksoforisia elementtejä ovat seuraavantyyppiset paikan ja etenkin ajan adverbialit: *meillä, maassamme; nyt, tänään, viime vuonna, ensi viikolla, eilen, huomenna viikko sitten* jne. Kooditamme eksoforan tekstuaalisten sidoskeinojen kanssa samaan muuttujaan; niiden tunnettuus on itsestään selvä eli siis irrelevantti.

Hän on selittänyt: " -- *Minä* en osaa tunnistaa koskaan ihmisiä valokuvista, osaatko *sinä*?" 68.12

*Nykyisin* kuulee arvovaltaisillakin tahoilla sanottavan, että kommunistimme ovat muuttuneet sodan jälkeen. He ovat *nyt* muka eräänlaisia kilttejä kotitonttuja, jotka vain vanhasta muistista pitävät vallankumousta leikkikalunaan. 77.1-2

*Viime aikoina* sanomalehdet ovat aina pääkirjoituksia myöten käsitelleet telakkateollisuuttamme ja sen huolestuttavalta näyttävää tulevaisuutta. *Tällä kertaa* kysymys ei ole vain *meidän* omatekoisesta pulastamme, vaan vaikeudet ovat lähes yleismaailmalliset. 78.1-2

## 2.8. *Transformaatiomuuttujat*

topikaalistettu alkukonstituentti (54)  
alkukonstituentin topikaalistuksen todennäköinen syy (55)  
subjektin lykkäys (56)  
verbin siirto (57)  
dislokaatio (58)  
ellipsit (59)

Koodituksen lähtökohtana oli se, että (systeemi)lauseen neutraali sanajärjestys on suomessa SV(X), subjektittomissa lauseissa V(X), joissa subjektin edellä sekä verbin jäljessä voi lisäksi vapaasti esiintyä liikkuvia, valenssiin sitomattomia adverbiaaleja. Tästä ratkaisusta seuraa, että jokainen aineistossamme esiintyvä lause, jossa on subjektin edellä objekti tai muu keskeinen nominaalinen jäsen, jossa subjekti on lopussa - jne.- on selitettävä tämän perusjärjestyksen muunnokseksi. Muunnoksia kuvataan tässä transformaatio säännöin. Subjektin edellä olevat nominaaliset konstituentit katsotaan siihen siirretyiksi eli topikaalistetuiksi (muuttujat 54,55). Ellei topikaalistusta tai vastaavaa subjektin lykkäystä (56) ole aiheuttamassa se, että verbi päättyy lauseen alkuun taikka loppuun, katsotaan, että verbi on näihin paikkoihin siirretty (57). Ratkaisu yksinkertaistaa tosiasioita: käytännössä ei jokainen subjektin lykkäys ole samanarvoinen eikä jokainen lauseen alkuun päätnyt konstituentti joudu sinne samasta syystä. Samaa lopputulokseen (esim.verbiloppuisuuteen) voidaan päätyä eri teitä ja silti vaikutus on sama (vrt. Hakulinen (1976)).

(54) *topikaalistettu alkukonstituentti*

- 1 lausekeobjekti
- 2 lauseobjekti
- 3 osa lauseobjektista
- 4 predikatiivi
- 5 osa predikatiivista
- 6 adverbiaali
- 7 osa adverbiaalista
- 8 adverbiaali + objekti
- 9 objekti + adverbiaali
- A adverbiaali + predikatiivi
- B predikatiivi + adverbiaali
- C adverbiaali + adverbiaali
- D NP (tai kaksi) lauseenvastikkeesta yli hallitsevan lauseen
- E kolme konstituenttia

(55) *topikaalistuksen todennäköinen syy*

- 1 kompensatorinen
- 2 sidonta
- 3 emfaasi
- 4 kontrasti
- 5 syntaktinen yhdenmukaisuus
- 6 syytä vaikea mainita (tunnusmerkitön?)
- 7 kompensatorinen + sidonta
- 8 kompensatorinen + emfaasi
- 9 kompensatorinen + kontrasti
- A sidonta + emfaasi
- B sidonta + kontrasti
- C tunnusmerkitön + kompensatorinen
- D tunnusmerkitön + sidonta
- E tunnusmerkitön + kontrasti
- F tunnusmerkitön + emfaasi

G emfaasi + kontrasti  
H en osaa sanoa

Topikaalistusta tarkastellaan paitsi teknisesti, myös siltä kannalta, mikä on tulkitsijan (=koodittajan) mukaan mahdollisesti ollut topikaalistuksen syynä. Muuttuja (55) on siis subjektiivinen: jos siitä halutaan objektiivisempi, on testattava useiden koodittajien ratkaisujen samansuuntaisuus.

Motivointeja on erityyppisiä. Kompensatorisella siirrolla (1) tarkoitetaan sitä, että topikaalistettu konstituentti täyttää subjektin paikan silloin kun subjektia ei ole lainkaan (esim. passiivilauseissa tai kun lause on geneerisenä 0-subjektillinen) tai kun se on lykätty lauseen loppuun.

Verbialkuisuutta vältetään neutraalipainotteisissa väitelauseissa paitsi lausesubjektin tapauksessa (*On valitettavaa, että ...*).

Topikaalistettu konstituentti voi pyrkiä alkuun myös siksi, että se on tekstuaalisesti sidottu. Tämä yhdistyy usein edelliseen motivaatioon: jos on tyhjä paikka, siihen pyrkii tekstuaalisesti sidotuin lauseenjäsen. Joskus sidonta on pakollista: relatiivipronomini on automaattisesti lauseensa alussa. Kaikki pakolliset siirrot (kuten relatiivistus, kysymyssanan siirto, konstituentin *-hAn* siirto) on kooditettu tähän. *Pakolliset siirrot*> pitäisi tulevaisuudessa sijoittaa omaksi alakategoriakseen.

Emfaasilla tarkoitetaan alkuun siirretyn konstituentin erityistä korostamista; kontrasti on tämän erikoistapaus: topikaalistettava konstituentti tarvitsee korostusta, koska se kontrastoidaan johonkin toiseen edellä tai jäljessä sijaitsevaan konstituenttiin.

Usein tuntuu siltä, että ei ole mitään erityistä tekstuaalis-tyylillistä syytä konstituentin alkupaikkaisuudelle, vaan tämä tuntuu olevan neutraalilla tunnusmerkittömällä paikallaan. On kiinnostavaa nähdä, missä määrin intuitiivisesti alkuun kuuluvat lausekkeet ovat niitä, joiden Hakulinen & Karlsson (1979) olettavat kiteytyneen lausetyypin tunnukseksi tälle paikalle. Heidän mukaansa lausetyyppejä on kolmenlaisia: subjektialkuisia, subjektittomia ja obliikvialkuisia).

Irrelevantti topikaalistuksen syyn etsintä on silloin kun kyseessä on virkkeen alkuun sijoittunut adverbialilause: tämän syyt ovat toisentyypiset kuin sanakonstituenttien. Esimerkkejä topikaalistuksista ja niiden syistä:

### 1. lauseobjekti

-(pakollinen) sidonta:

pöytäkirjasta, *jota* Sorsan hallituksen pitäisi maataloushintojen korottamisessa noudattaa.

- kompensatorinen siirto:

Vaikeata on sitä vastoin päätellä, voisiko muunlaista elämää olla olemassa ja mitä maan elämän piirteistä pitäisi olla mukana, että *nimitystä elämä* voitaisiin käyttää. 22.59

- sidonta ja kompensatorinen siirto:

Itse energia sitä vastoin virtaa yhteen suuntaan. Se saapuu auringon säteilyn muodossa ja *sitä* käytetään kemialliseen ja mekaaniseen työhön. 22.36

[Teksti esittelee erään ongelmallisen tilanteen]

*Ratkaisua* on etsittävä yhteisellä neuvonpidolla. 111.37

-kontrasti (ja muutakin tekstuaalista sidontaa):

Sillä on litteä, sukkulainen, taakse päin kapeneva ruumis, jonka etuosassa on kuusi jalkaa niin kuin kunnan hyönteisellä ainakin, ja otsassa heiluvat pitkät uteliaat tuntosarvet hyönteisyyden merkinä. Mutta *siipiä* se ei saa koskaan. 90.14

Melkein kaikki kunnallinen aktiviteetti on sammunut. Siipyy onkin liitetty Kristiinankaupunkiin. *Lopullisen kuoliniskun* seutu sai viime kesänä, kun ruotsalainen tekstiilitehdas lopetti toimintansa. 111.22

2. *lauseobjekti*, tavallisimmin sitaatti. Topikalisoinnin syy aina irrelevantti.

Unescon työn alalla suurimmat ongelmat ovat kehitysmaissa, Margaretha Mickwitz kertoo. 92.22

3. *osa objektista*: kun objektina on pääsana ja lause (etenkin *se että* -rakenne), etenkin puhekielenomaisissa tapauksissa topikaalistetaan usein pääsana. Seuraavassa siirto on kompensatorinen, erityismotivaationa kontrasti.

*Sitäkään* ei tiedetä, missä määrin käytettyinä vesakkomyrkyt rikkovat luonnon tasapainon. 28.10

4. *predikatiivi*

- pakollinen:

*Millaista* oikein on farkkukansa? 109.1

- kompensatorinen siirto:

*Merkillepantavaa* on se, että tämä "vilkastumisilmiö" ilmaantuu -- 2.16

- kompensatorinen ja sidonta:

Maallikon silmään juttu näyttää hullulta. Kun sitä koettaa käsittää etsimällä vertausta joltakin vähän alemmalta tasolta kuin tasavallan huipulta, syntyy esimerkiksi tällainen tarina. Ajatellaan että *vertauskohdaksi* sopii vaikkapa osuuskauppa, demokraattisesti hallittu yhteisö, joka -- 84. 7

-- kun Paavo Berglund kutsuttiin Luzernin musiikkijuhlille sairastuneen Lorin Maazelin sijaiseksi, paikkasi Albrecht Maazelin jättämän aukon Salzburgissa ja johti Wienin filharmonikkojen konsertin. *Tuloksena* oli muutos solistin ohjelmaan. -- 65.39

- kompensatorinen ja emfaasi:

Pienen teamin yhteispelin on täytynyt olla tiivistä sillä *niin saumatonta* on kokonaisuuden viritys. 61.40

- kontrasti ja sidonta:

Shakkiharrastuksella on merkillisiä alueellisia vaihteluita - *suosituinta* se on Helsingin tienoon sekä maaseudun oppikouluissa. 109.73

Meillä se kantaa toukan nimeä, mutta itse asiassa se on aikuinen, täysimuotoinen hyönteinen, ainakin sen jälkeen kun on jättänyt taakseen nuoruutensa toukkanahat. *Toukan näköinen* se kyllä on aikuisenakin, tai *aikuisen näköinen* toukkana, miten päin vain. 90.12

## 5. osa predikatiivista

*Niin saumatonta* oli kokonaisuuden viritys, että teamin yhteispelin on täytynyt olla tiivistä.

## 6. adverbialli (yleisin topikaalistettu konstituentti).

- pakollinen sidonta:

-- ja mökeissä, *joissa* vanhukset talvisin asuvat yksin, tungeksii nyt koko suku. 111.3

- sidonta; konnektiivi:

Ympäri vuoden lenkkeilen 0-3 kertaa viikossa, --. *Lisäksi* pelailen silloin tällöin jalkapalloa ja koripalloa --. 91.20

- kompensatorinen siirto ja sidonta:

Sekä eläinten, kasvien että virusten toimintoja ohjaa koodi, joka on talletettu nukleiinihappoihin. Nämä ovat nauhamaisia suurimolekyylisiä yhdisteitä, joiden neljän erilaisen perusrakenteen, nukleotidin, järjestys muodostaa kolmikirjaimisen sanoman. *Tähän geneettiseen koodiin eli yksilön perintökäsitteisiin on tallennettu kaikki tieto kyseessä olevasta yksilöstä ja sen ominaisuuksista. Tätä koodia toteuttamaan* tarvitaan toinen biomolekyylien ryhmä, proteiinit. 22.11-14

- kompensatoriseksi siirroksi kooditetaan aina eksistentiaalistyyppisen sanajärjestyksen alkuadverbialli; yleensä sekin on tekstuaalisesti sidottu verrattuna lopussa tulevaan subjektiin:

*Eliöiden muodostamassa kokonaisuudessa* vallitsee suurempi järjestyneisyyden aste kuin niissä aineissa, joita eliöt ottavat ympäristöstään. 22.38

Hänellä oli, mitä katsella, sillä *kaulassani* riippui nauhuri, *hampaiden välissä* keuhkoinen muovikassi, *jonka kirkkaanpunaisella pinnalla* [tässä tietysti myös pakollinen sidonta] mainostettiin jotakuta vaatetusliikettä ja *josta* kohosi ilmoille tuoreiden savusilakoiden huumava tuoksu. 79.8

- tunnusmerkittämiä alkuadverbialleja; tunnusmerkittäväksi kooditetaan aina possessiivi- ja *pojasta tuli mies* -lauseen alkuadverbialli (riippumatta sidoksisuudesta):

*Sillä* on litteä, sukkulainen, taaksepäin kapeneva ruumis,-- 90.13

- muita tunnusmerkittämiä, tällaisia ovat usein kehysadverbialit, kommentit, ajankohtaa ilmaisevat predikatiivadverbialit:

*Vuoden 1956 presidentinvaaleissa* Fagerholm oli SDP:n ehdokkaana -- 23.22

*Kaksikymmentäyksivuotiaana* Indira osallistui kongressipuolueen riveissä itsenäisyystaisteluun -- 24.6

- emfaasi: *Vieläkin* näkyy yleisönosastoissa väliin pelästyneitä kirjoituksia, joissa kannetaan huolta peruskoululaisten liiasta biologisesta valistumisesta. -- 95.17

- kontrasti: Vanha Lapin ihminen saattaa ottaa ruskan merkinä lopullisesta kuolemasta --, Väisänen kertoo. *Nuoreen* ruska ei hänen mukaansa vaikuta samalla tavoin. 106.42

## 7. osa adverbialista - tässä tapauksessa motivaationa emfaasi:

*Niin hienosti, suorastaan-kauniisti* se on toteutettu, että ihmistä, joka siitä pahoittaa mielensä on pidettävä vähintäänkin omituisena ja todellisuutta vierastavana. 95.119 [esimerkki tosin on kooditettu 8:ksi, koska adverbialin lisäksi on topikaalistettu myös objekti]

8. *adverbiaali ja objekti*. Kun topikaalistettuja konstituentteja on kaksi, tulevat tietysti kyseeseen topikaalistussyiden kombinaatiot. Tavallisesti kuitenkin jomman kumman syynä on sidonta ja lähes aina kompensatorisen siirron tarve - mikä näkyy siitä, että vain murto-osassa kahden tai useamman topikaalistuksen tapauksista on subjekti omalla paikallaan: ne ovat yleensä passiiveja tai lykätyn subjektin lauseita.

- sidonta ja kompensatorinen siirto:

*Toistamiseen / Fagerholm* nimitettiin hallituksen päämieheksi 3.3.1956 --. 23.23

*Kaksi vuotta myöhemmin/maan neljältoista tärkeintä* pankkia kansallistettiin valtion mahtikäskyllä. 24.17

- tunnusmerkitön ja kompensatorinen:

Kaksikymmentäyksivuotiaana Indira osallistui kongressipuolueen riveissä itsenäisyystaisteluun, ja *1942 / hänet* vangittiin kolmeksi kuukaudeksi. 24.6

- kompensatorinen ja kontrasti:

Kansainvälinen kiertonäyttely vieraili ennen Helsinkiä Pariisin Pompidoun kulttuurikeskuksessa. *Suomeen / sen* on tuonut Rakennustaiteen museo -- 66.8

## 9. *objekti ja adverbiaali*

Systä, *joita / Suomessa* ei yleisesti tunnettu, Neuvostoliitto suhtautui hallitukseen kielteisesti, -- 23.27 [pakollinen sidonta ja kompensatorinen]

- kontrasti (jonkin muun tekijän kanssa yhdessä):

Suunnitelmia ja päätöksiäkin jatkokoulutuksen kehittämisestä siis on, mutta *koko ongelmaa/niillä* ei vielä ratkaista. 29.21

Einojuhani Rautavaara -- sooloteoksensa *Varietuden* epälukuisia esityksiä. Niitä *radioonkin* analysoidessaan hän tuli epäilyksettä siihen tulokseen, että -- 64.4

### A. *adverbiaali ja predikatiivi*

Jos pikkuinen herää yöllä eikä uni tahdo tulla takaisin, tarjoa hänelle lämmintä maitoa tai vettä. Se rauhoittaa hieman. *Muuten / ainoa keino* on kärsivällisyys, -- 119.44

*Ensisijaisesti / miehen* ehkäisytablettien kehittelyn vaikeutena on miehen miljoonamääräinen jatkuva siittiöiden tuotto, -- 121.73

### B. *predikatiivi ja adverbiaali*

*Huolestuttavinta / viimeaikaisessa poliittisessä kehityksessä* onkin se, että puolueiden johtajat eristäytyvät äänestäjistään -- 75.30

### C. *adverbiaali ja adverbiaali*

- *1800-luvun puolivälissä / Englannilla* oli esim. Afrikassa siirtomaita ainoastaan nykyisessä Etelä-Afrikassa sekä Ranskalla Algeriassa ja Senegalissa. 7.3

*Päässä / hänellä* oli iso morsiusseppäle -- 103.30

*Kuitenkin / kapitalistisessa yhteiskunnassa* on joukkioittain ihmisiä, joilta puuttuu asunto, -- 37.11

Nyt odotellaan eläkejärjestelmän suurta uudistusta. *Parhaassakin tapauksessa / siihen* kuluu monta vuotta. 85.21

D. *nominaalilauseke siirretty upotetusta lauseesta matriisilauseen jäseneksi*. Tämä koskee partisiippirakenteita ja subjektina olevia infinitiivirakenteita ja on ainoa tapaus, jossa myös subjekti voidaan katsoa topikaalistetuksi tämän järjestelmän mukaan.

*Maapallon* uskotaan olleen aikanaan liian kuuma -- 22.48 (=Uskotaan, että maapallo oli aikanaan liian kuuma).

Elämän syntyä pintavesissä, joihin näitä yhdisteitä runsaasti liukeni, pidetään mahdollisena. *Tämän* on arvioitu tapahtuneen yli kolme miljardia vuotta sitten. 22.53 (= On arvioitu, että tämä tapahtui yli kolme miljardia vuotta sitten).

*Tähän* lienee kuitenkin välttämätöntä mennä -- 40.16 (Lienee kuitenkin välttämätöntä, että 0 menee tähän).

Joskus lause voi sisältää jopa upotetusta lauseesta matriisilauseeseen topikaalistetun konstituentin ja toisen topikaalistuksen tai kaksi upotetusta lauseesta matriisilauseeseen topikaalistettua konstituenttia. Nämä ovat erilliset ilmiöt, mutta tapauksia on niin vähän, että ne mahtuvat samaan muuttujaan upotteen konstituentin yksinkertaisen topikaalistuksen kanssa.

*Vettä* on koko maapallolla todettu olevan tarpeeseen nähden vähän. 50.3 (=On todettu, että koko maapallolla on tarpeeseen nähden vähän vettä).

E. *kolme konstituenttia topikaalistettu*. Tämä on niin harvinaista, että ei kannata ryhtyä tekemään uusia alakategorioita sen mukaan, mitä lauseenjäseniä näihin kuuluu ja missä järjestyksessä. Tavallisinta on, että ne ovat adverbiaaleja, mutta myös objekti näyttää mahtuvan joukkoon. Topikaalistuksen yhtenä syynä on ainakin useimmissa tapauksissa kompensatio, ts. tällaisissa tapauksissa subjekti on äärimmäisen harvoin paikoillaan.

Laitte varoittaa kuljettajaa nimittäin vasta silloin, kun mittaus jo tapahtuu. *Tällöin kuitenkin syytetylle* jää tieto siitä hetkestä ja paikasta, missä mittaus tehtiin, -- 25.11

Presidentti Kekkonen ei ainakaan vielä tässä haastattelunsa ensimmäisessä osassa -- anna siihen omaa vastaustaan. *Sen sijaan eilen Kansan*

*Uutisissa* mm. ministeri Arvo Aalto esitti seuraavan vaatimuksen -- 87.7

*Syksyllä 1944 Fagerholm jälleen* nimitettiin sosiaaliministeriksi Castrenin lyhytikäiseen hallitukseen. *Seuraavana vuonna Fagerholm ensi kerran* valittiin eduskunnan puhemieheksi, jota tehtävää hän sitten hoiti -- 23.17-18

-- toimiessaan vuodesta 1930 Suomen Liiketoimintajien Liiton puheenjohtajana. *Tästä tehtävästä hänet 1942* nimitettiin Oy Alko Ab:n johtajaksi; Alkon pääjohtaja hän oli 1952-69. 23.7

(56) *subjektin lykkäys*

1 intransitiivisubjekti

2 transitiivisubjekti

3 lausesubjekti

4 infinitiivisubjekti

5 subjektin osa

6 kvanttorin siirto

Koska kaikki lauseet ovat lähtömuodoltaan subjektialkuisia, analysoimme ne tapaukset transformoiduiksi, joissa subjekti on verbin jäljessä. Tapaukset ovat epäyhtenäisiä: tähän joutuvat sekä ns. eksistentiaalilauseet, joissa subjekti usein sisältää uuden tarkoituksen, että sellaiset, joissa subjekti kompleksisuutensa tai pituutensa vuoksi päättyy loppuun (ekstrapositioon).

Sanajärjestys VSX, jossa subjekti on lykätty vain verbin jälkeen, voi saada kaksi sellaistakin tulkintaa, joissa subjektiä ei ole lykätty. Ensinnäkin tämä voi aiheutua verbin siirrosta alkuun (vrt. muuttujaa 57): *Eivätkö* he siis sitä ennen tienneet? Mutta *ei* se ole ainoa ratkaisu. Toiseksi kyseessä voi olla ns. *inversio*: subjektin ja verbin paikat ovat vaihtuneet adverbiaalin t. adverbiaalilauseen jälkeen:

Kun Suomen Elokuvasäätiö elokuva-arkiston aloitteesta vuonna 1972 tarjoutui rahoittamaan elokuvan lehtoraatin, *ei yliopisto* ottanut tarjousta vastaan. 3.5

Huomattavan lyhyessä ajassa *on* siis "*tuhattjärvien maa*" saanut liittää *ratkaisua vaativan vesipulman* monien muiden pulmiensa jonoon. 9.94

Käsityksemme mukaan inversio on aivan toisenluontoinen asia kuin tapaukset, joissa subjekti on lykätty tai verbi siirretty alkuun. Jotkut kirjoittajat ehkä suosivat sitä, jotkut eivät: sen tehtävä on korkeintaan korostaa edellä olevaa adverbiaalia. Sen sijaan sekä verbin siirrolla että subjektin lykkäyksellä on omat temaattiset tehtävänsä. Inversiotapauksia ei siis ole kooditettu subjektin lykkäyksiksi eikä verbin siirroiksi. Ne käyvät ilmi vain sanajärjestystapauksista. Tässä kohdin pintajärjestysmallilla (18) on siis oma itsenäinen funktionsa: muissa tapauksissahan pintajärjestysmallin sisältämä informaatio tulee toisen kerran kirjatuksi topikaalistusmuuttujiin. Esimerkkejä lykätystä subjektista:

1. *intransitiivilauseen subjekti lykätty*. Tähän kuuluvat omistusrakenne, eksistentiaalilause ja predikatiivilause.

Sillä on litteä, sukkulainen, taakse päin kapeneva ruumis, -- 90.13

Hallituksessa on nyt *viisi kepulaista, neljä demaria*, -- 86.2

Eipä ollut montakaan, joka ei ilmoittautunut -- 89.31

Merkille pantavaa on *se, että tämä "vilkastumisilmiö" ilmaantuu* -- 2.16

2. *transitiivilauseen subjekti lykätty* - subjekti ja objekti ovat siis vaihtaneet paikkaa

Sekä eläinten, kasvien että virusten toimintoja ohjaa *koodi, joka on talletettu nukleinihappoihin*. 22.11

Porvarivaltaa varmistaa *eduskunnan porvarienemmistö*. 86.4

3. *lausesubjekti lykätty*: ekstrapositio. Olemme tulkinneet tämän sillä tavalla ahtaasti, että kysymys on vain sellaisen subjektin siirrosta, joka on pelkkä lause; *se että*-lause esimerkiksi olisi tässä tavallisen subjektin veroinen, jos se siirtyy lauseen loppuun kokonaisena.

Ja sääntönä on, että eläkeikää lähestyttäessä suhdanneherkkyys lisääntyy ja ansiotulot vähenevät. 35.8

4. *infinitiivisubjekti lykätty*. Subjektina olevan infinitiivin (lausetyyppi 8:E) lykkäminen loppuun on jotakuinkin pakollista. Infinitiivisubjektin (ja usein myös lausesubjektin)

lykkäyksestä seuraa usein järjestys VXS: näissä tapauksissa verbialkuisuutta sallitaan enemmän. Toisaalta itse infinitiivisubjektista voi siirtyä jokin konstituentti matriisilauseen alkuun (vrt. 54:D).

On väärin luulla esitystä - tai Kollwitzin taidetta - suorasukaisesti poleemiseksi. 61.18

Marivaux'n siroon ja henkevään maailmaan on mukava paeta arjen ja laman puristuksesta. 57.3

5. *subjektin osa lykätty*. Jos *se että-*, *se joka-* tms. pääsana ja lause -tyyppisen subjektin lauseosa on lykätty hallitsevan lauseen loppuun, se merkitään tähän kohtaan.

Jo se herätti hämmästyä, että naispääosaa esittää kaksi toisiaan *muistuttavaa, mutta sisäiseltä tyypiltään erilaista naista* -- 62.37

"Se on lääkäri, *joka tietää näkymättömästä*,--" sanoi aikansa suuri lääkäri Paracelsus. 11.52

Ne ajat ovat olleet ja menneet, jolloin sienet manattiin samaan kastiin käärmeiden, paholaisten ja muiden pimeyden voimien kanssa. 113:78

6. *kvanttorin siirto* ym.. Koska olimme kiinnostuneita siitä, miten yleistä materiaalissamme on ns. kvanttorin siirto, ts. se, että subjettiin (tai myös, etenkin passiivilauseen, objektiin) kuuluva kvanttorisana eroaa pääsanastaan ja "ajelehtii" lauseen keskelle, olemme ottaneet tämän tiedon subjektin lykkäys-muuttujan alakategoriaksi - lykätyn subjektin kvanttori tuskin pystyy tällä tavoin irtoamaan, ainakaan niin usein, että päällekkäisyys häittäisi. Kooditettaessa kävi ilmi, että on muitakin samaan tyyliin ajelehtivia pikkusanoja, joita ei voi pitää oman lauseensa konstituenttina - pääasiassa *itse* ja *sekin / nekin* "myös" - joten päätimme merkitä nekin tähän kohtaan. On huomattava, että kvanttorin ajelehtiminen on eri asia kuin "lohjennut partitiivisubjekti" tai "-objekti" (20:8 ja 28:9), vaikkakin ne varmasti ovat sukulaisilmiöitä. Ajelehtivia kvanttoreita ovat totalitatiivit (*kaikki, molemmat*) ja ne liittyvät nominatiivimuotoiseen subjettiin (tai objektiin). Tämän lohkeamisen ei voi katsoa johtaneen oman lausetyypin syntymiseen niinkuin kvantifioidun partitiivin tapauksessa, koskapa ajelehtiva kvanttori ei voi olla lauseen ainoa muu jäsen: *me olimme kaikki* vrt. *meitä oli paljon*. Esimerkkejä tähän alakategoriaan merkityistä tapauksista:

Taitaisi tehdä tilanteen tajuttavaksi monelle -- jos nuo 137 000 tosiaan olisivat *kaikki* koossa, -- 85.12

Puolukankeruuseen nyt ryhtyisivät nuorista työttömistä Anita Kvick, 16, ja Maini Väyrynen, 17, ovat *kumpikin* syrjäkylän työttömiä Pielaveden Laukkalasta. 97.25

Työläiset ovat *itse* kauan sitten oppineet käyttämään kiristystä eli lakkoa, -- 88.24

Ehkä Cervantesin kiihkeyteen kätkeytyi salattu pettymys, jonka hän oli kokenut lukiessaan *itse* ritariromaaneja: -- 18.8

*Suezin kanavan aukaiseminen* vaikutti *sekin* tietenkin joltakin osalta kuljetuksia vähentävästi. 78.5

(57) *verbin siirto*

- 1 alkuun
- 2 loppuun
- 3 molemmat
- 4 verbiketjun infinitiivi alkuun

Verbin tunnusmerkitön paikka on lauseen keskellä. Usein lauseessa on enemmän kuin yksi verbi eli kokonainen verbiryhmä, jonka jäsenet hajaantuvat pitkin lausetta. Jos finiittiverbi tai muu elementti on lauseen alussa (ja tietyissä rakenteissa sen lopussa), sanajärjestyksessä on jotakin tunnusmerkistä, joskaan ei aina samalla tavoin. Olisi harkittava, kannattaisiko kieltoverbin erottaminen omaksi kategoriakseen, koska se liikkuu muita herkemmin (mm. kiinnittyy usein konjunktioihin: *eikä, ellei, vaikkei*).

Verbialkuisuuden funktiot ovat jokseenkin selväpiirteiset, kun taas verbioppisuus on päälauseissa eri asia kuin sivulauseissa, joissa se heijastelee alkuperäistä SOV-sanajärjestystä. Kuten edellä todettiin, verbialkuisuus ja -oppisuus voivat johtua useista syistä; vaihtoehtoinen analyysi lähtisikin siitä, että ne ovat tavoiterakenteita, joihin voi päätyä usealla tavoin (vrt. Hakulinen 1976:116). Tämä aiheuttaisi kuitenkin koodittamiselle teknisiä ongelmia.

Verbin loppuun heittymisestä on se seuraus, että ensimmäinen konstituentti erityisesti korostuu. Tämä vaikutus on samanlainen kuin jos ensimmäinen konstituentti olisi kontrastiivisesti topikaalistettu. Jos lähdetään neutraalista muodosta *kotimaa näytti kauniilta*, komplementtia voi kontrastiivisesti korostaa topikaalistamalla sen *KAUNIILTA kotimaa näytti*, mutta subjektia ei voi topikaalistaa, se voidaan puolestaan kontrastoida vain heittämällä verbi loppuun: *KOTIMAA kauniilta näytti* (esimerkissä vielä *se* -sana antaisi lisäväriä). Kummassakin tapauksessa lopputulos on verbioppuinen. Vastaavasti jos lähdetään neutraalista muodosta *minulla on koira* (johon itseensä on jo sovellettu possessiivilauseeseen siirto-operaatioita), voi tietysti topikaalistaa subjektin: *KOIRA minulla on* (eikä kissa), mutta jos kontrastoitavana olisi adverbiaali, joka jo sijaitsee alussa, ainoa keino on verbin heitto loppuun: *MINULLA koira on* (eikä Kallella); viimeinen lause on jo huomattavan tunnusmerkkisesti painottunut.

Joskus verbioppuisen lauseen läpikäymä transformaatio ei ole helppo jäljittää. Yleensä asia selviää kun pidetään silmällä konstituenttien järjestystä. Aktiivilauseissa subjektin ja toisten lauseenjäsenten neutraali keskinäinen järjestys on SX/SA; jos järjestys verbioppuisessa lauseessa on tämä, verbi on siirretty.

Omaleimainen, kiinnostava ja ansiokaskin Urpelan romaani on, *vaikka perinnäiset yhtenäisyysvaatimukset vain vaivoin siihen sopivat*. -- 56.47

Jos taas järjestyksenä on XS/AS, adverbiaali tai komplementti on alkuun topikaalistettu: *Omaleimainen, kiinnostava ja ansiokaskin Urpelan romaani on*, - Jos verbi on useampiosainen, siis verbiketju tai liittomuoto, ja infiniittimuoto on lopussa, tulkitsemme asian niin, että infiniittimuoto on siirretty. Eihän ole mitään "topikaalistusta", joka siirtäisi esim. objektin verbin jälkeisestä neutraaliasemastaan verbiosien väliin, vaan kaikki topikaalistukset siirtävät vain alkuun. Seuraavissa tapauksissa siis merkitään verbin siirto loppuun:

Suunnitelmia ja päätöksiäkin jatkokoulutukseen kehittämisestä siis *on, mutta koko ongelmaa niillä ei vielä ratkaista*. 29.21

Muutama juhlatilaisuuksiin liittyvä ilmaisesityskään eri paikoissa ei vielä tilannetta *paranna*. 61.44

Mutta kahden verbielementin välissä voi missä tahansa neutraalissa lauseessa olla adverbiaali, ja jos transitiivilauseesta on topikaalistettu objekti tai muu komplementti, on mahdoton sanoa, onko verbi "siirretty" yli tämän adverbiaalain vai onko adverbiaali kahden verbin välillä "ennestään". Tämäkään ratkaisu ei ole kovin tärkeä käytännössä.

--parhaat konsertit ovat loppuunmyytyjä ja kaupungista eivät turistit ja omatkaan kuulijat vähällä *lopu*. 64.26

Lähtömuoto voi olla joko:

turistit ja omatkaan kuulijat eivät vähällä loppu kaupungista [+ ADVERBIAALIN TOPIKAALISTUS + INVERSIO]

tai:

Turistit ja omatkaan kuulijat eivät loppu vähällä kaupungista / kaupungista vähällä.

1. *verbin* siirto alkuun. Tässä ei katsota verbiksi sellaista *ei*-sanaa, joka on sulautunut konjunktioon (*eikä, ellei, ettei, vaikkei*).

Voidaan kysyä, *oliko ja onko* sitten juuri elokuvan estetiikka se alue -- 3.10

Tämä pitää paikkansa etenkin emämaan valtiontalouden kannalta, *palasivathan* sinne tehdyt investoinnit usein suoraan voittoina emämaan yksityiselle sektorille. 7.63

*Ei* liikennemoraali ole kiinni jostain vempaimesta, jota muutamat harvat käyttävät. 25.15

2. *verbin* siirto loppuun.

Mitä työväenliikkeeseen ja sen taisteluun yhteiskunnan muuttamiseksi *tulee*, se on luonteeltaan mitä isänmaallisinta. 34.14

Biologinen determinismi ja seksuaalisuuden ylikorostus, joka freudilaisuuteen *sisältyi*, torjui pitkän päälle marxilaiset. 69.83

3. *verbin* siirto sekä alkuun että loppuun (tässä on tulkittu myös sulautuneet *ei*:t alkuun siirretyiksi verbeiksi).

-- *ei* onnea niin vain *osteta*. 62.31

Esitys jakaa väistämättä yleisön. Ja *ellei* se sitä *tekisi*, *ei* se Brechtin Turandot *olisikaan*. 60.8

4. *infinitiivin* siirto *verbiketjusta* alkuun - teoreettinen mahdollisuus:

yrittää hän kyllä on luvannut.

(58) *dislokaatio*

1 oikealle

2 vasemmalle

Tämä muuttuja on kirjoitettujen tekstien tutkimuksessa marginaalinen, sillä se koodittaa lähinnä puhutussa kielessä suosittuja rakenteita, jolloin temaattista ainesta sijoitetaan kahteen kohtaan lausetta: joko teemaksi pronomini ja kokonainen lauseke jäljemmäksi (*tämä* on niin raskasta *tämä kirjoittaminen*) tai ensin leksikaalinen aines perusmuodossa ja sitten taivutettu pronomini lauseen "oikeassa" kohdassa (*tämä sähkökoneella kirjoittaminen* - minä en ole päässyt *sen makuun* vieläkään). Asiateksteissä tällaisten ilmiöiden runsaus viitanee tekstin puhekielisyteen. Meidän aineistostamme niitä ei löytynyt.

(59) *ellipsit*

1 aukko: finiittiverbin poisto

- 2 aukko: V + objekti
- 3 aukko: V + adverbiaali
- 4 aukko: V + muuta
- 5 "VP":n ellipsi
- 6 *niin ... -kin*

Voimakkaasti lauseita sitova keino on ellipsi: se että edeltävästä kontekstista selviävää ainesta ei panna lainkaan ilmi. Ellipsi sitoo yleensä viereisten lauseiden, korkeintaan lähivirkkeiden aineksia eikä ole tekstuaalisesti yhtä kauaskantoista kuin monet muut sidoskeinot. Tässä ei käsitellä "pragmaattista ellipsiä" eli tapauksia, joissa lause täydentyy tilanteen eikä tekstin perusteella (vrt. Hakulinen (1978)), ei myöskään näiden kaltaisia fraasiutumia, jotka kooditetaan fragmenteiksi (15:i):

Sanoitko että hänen kesäkotinsa nimi on Leporello, hekotin. *Hyvä nimi.* 80.25-26

Puistossa vastapäätä J.W. Snellmania -- oli tyhjä penkki. *No mitäs:/ siihen istumaan/ ja kenkä jalasta.* 83.3-4

Emme myöskään koodita *lauseensisäistä* ellipsiä sen enempää kuin lauseensisäistä pronominaalistustakaan: seuraavista nolista vain jälkimmäinen on tekstuaalinen: ulottuu yli lauserajan.

Tämä kirkko on saaren toinen 0, ensimmäinen 0 tehtiin jo 100 vuotta sitten.

Tästä muuttujasta puuttuu kaksi tekstuaalisesti tärkeää ellipsin lajia: pääsanan ellipsi, joka on kooditettu sidoskeinoihin (muuttujat 23,31,36,43,50) sekä nominaalisen konstituentin ellipsi (esim. subjektin poisto 19:d). Subjektin ellipsi on selväpiirteisin; sen sijaan objektin ym. ellipsi on usein ns. aukkoamisen osa eli poisto tapahtuu yhdessä verbielementin kanssa. Jos verbi kuitenkin on säilynyt, kirjataan vain yhden konstituentin ellipsi. Seuraavat objektin ellipsitapaukset viittaavat siihen, että näin käy etenkin silloin, kun konstituentti edeltää verbiä (on siis subjektin paikalla).

Kohdetta valvotetaan, 0 pidetään nälässä, 0 rasitetaan fyysisesti ja henkisesti -- 13.17

Majava sen sijaan hävitettiin sukupuuttoon 1800-luvulla, mutta 0 istutettiin uudelleen 1930-luvulla ja 0 on viihtynyt siitä lähtien hyvin. 9.49

Kun kukin finiittiverbi kuuluu omaan lauseeseensa, seuraavissa lauseissa ellipsi poistaa itsenäisiä objekteja ja adverbiaaleja (vrt. Hakulinen (1978:242)):

Esikoeryhmältä tiedusteltiin, kuinka he kokivat (FILMIN) ja (KUINKA) (HE) tulkitsivat filmin. 8.29

-- ja laajassa osastossa kerrotaan, useinkin humoristisin sarjakuvien miten yksilöt tavoittavat [TOISENSA] ja [MITEN] [YKSILÖT] tunnistavat toisensa ja [MITEN] [YKSILÖT] saattavat sukusolut toistensa yhteyteen. 95.29

Tähän ellipsimuuttujaan on siis kooditettu ennen kaikkea verbiryhmään liittyvä ellipsi. Tärkeä se on etenkin silloin, kun poistetaan sellaisia konstituentteja, joita muualla kuvauksessa ei näy (eli muita kuin lauseen alku- ja loppuadverbiaaleja). Harkitsemisen arvoista olisi kuitenkin, voisiko tämän muuttujan sulauttaa verbin rakennetta koskevaan muuttujaan (25).

1. *alakategoriat 1-4*. Aukkoaminen (engl. gapping) on transformaatiokieliopissa yleisesti tunnustettu sääntö, joka soveltuu kahteen rinnasteiseen lauseeseen ja poistaa jälkimmäisestä ainakin finiittiverbin. Sen lisäksi se voi poistaa myös verbiryhmän tai liittomuodon infiniittijäsenet, jos niitä on, ja/tai objektin, predikaatiivin tai adverbiaalinen tai niiden yhdistelmiä. Olennaista aukkoamiselle on, että jotakin poistuu lauseen keskeltä: oikealle - ja myös

mielellään vasemmalle - täytyy jäädä jotakin, ja ellei jää, käytetään täytesanaa *kyllä*. Niinpä tyypillisellä aukkotapauksella on kontrastoiva merkitys:

-- tutkijat on sijoitettu erillisiin tutkimuslaitoksiin ja *opetushenkilökunta* [ON SIJOITETTU] *pelkästään korkea-asteen ammattikoulutusta antaviin korkeakouluihin*. 53.31

Aukkoaminen yhdistyy erittäin usein subjektin ellipsiin, joskus myös pääsanana poistoon:

Vaje oli viime vuonna neljä miljardia ja [VAJE] [OLI] *vuotta* aiemmin peräti kahdeksan [MILJARDIA]. 49.9

Kuten esimerkistä näkyy, jos pelissä on myös subjektin ellipsi, käy helposti niin, että aukon vasemmalle puolelle ei jää mitään. Oikealle sen sijaan jää aina.

Ellipsi-muuttuja sisältää aukkoamis-transformaation neljänä alakategoriana sen mukaan, onko poistettu pelkkä verbi vai verbi yhdessä jonkin muun konstituentin kanssa. Toinen huomionarvoinen rakenteellinen mahdollisuus olisi se, että voidaan poistaa useampijäsenisen verbikonstituentin tapauksessa joko pelkästään finiittimuoto tai koko verbiryhmä. Esittelemme aukkoamistapaukset tämän huomioonottaen. Huomattakoon, että verbin ellipsiin liittyy usein myös subjektin poistoa.

1.a. *jäljellä* (vähintään yksi) *infiniittimuoto*, poistettu vähintään finiittiverbi.

Suuri yhteisö on maastossa helpommin havaittavissa kuin yksilö, joten pedot pystyvät löytämään sen nopeasti ja [PEDOT] [PYSTYV[T] *sieppaamaan saaliikseen yhteisön laitaosien yksilöitä*. 21.36

Markkinointimahdollisuudet olisi kaikesti pitänyt selvittää ja kustantajan järjestämisestä [OLISI PITÄNYT] *varmistua kaikkein ensimmäiseksi*,-- 54.27

-- mm. oppimistulosten arvioinnissa käytettäviä menetelmiä on parannettu ja [ON] *luotu "opetuspaketiksi" kutsuttuja oppikokonaisuuksia* -- 53.16

1.b. *koko verbikonstituentti poistettu*: joko yksi verbi tai verbiryhmä

-- työväenluokalle on tärkeämpää/ taistella poliittisesta vallasta / *kuin* [TAISTELLA] *sellaisten pennien ja markkojen puolesta, jotka menettävät arvonsa* -- 37.16

Lukioon siirtyvien nuorten osuus on tähän mennessä ollut suhteellisen suuri: Uudellamaalla [LUKIOON SIIRTYVIEN NUORTEN OSUUS] [ON OLLUT] *noin kuusikymmentä* [PROSENTTIA OPPIKOULULAISISTA] ja *muualla maassa* [LUKIOON SIIRTYVIEN NUORTEN OSUUS] [ON OLLUT] *neljäkymmentä prosenttia oppikoululaisista*.

-- tutkijat on sijoitettu erillisiin tutkimuslaitoksiin ja *opetushenkilökunta*[ON SIJOITETTU] *pelkästään korkea-asteen ammattikoulutusta antaviin korkeakouluihin*. 53.31

Huom. erityisesti polaarisuudeltaan vastakkaiset lauseet:

MTK aikoo kai nyt näyttää kerta kaikkiaan, että hallituksen on kuunneltava sitä *eikä sen* [OLE KUUNNELTAVA] *hallitusta*. 39.11

Elämä ei kuitenkaan ole tämän lain edellyttämä suljettu järjestelmä *vaan* [ELÄMÄ] [ON] *avoin* [JÄRJESTELMÄ]. 22.42

2.a. finiittiverbi ja objekti poistettu

Energiansa se saa Auringon säteilyenergiasta, jota vihreät kasvit kykenevät sitomaan ja [JOTA] [VIHREÄT KASVIT] [KYKENEVÄT] *muuttamaan edelleen muille eliöille käyttökelpoiseen kemiallisen energian muotoon* -- 22.32

## 2.b. koko verbi ja objekti poistettu

-- hänet on nähty toisaalta mielipuolisena tyrannina, toisaalta [HÄNET] [ON NÄHTY] *loistavan lahjakkaana renessanssiruhkinaana*, -- 20.4

Poliittikona häntä voidaan sanoa osittain jopa nerokkaaksi, mutta toisaalta [HÄNTÄ] [VOIDAAN SANOA] *myös todellisuudelle vieraaksi teoreetikoksi*. 20.8

-- enää ei eläkkeen perusteena olevaa palkkaa otettaisi neljän viimeisen vuoden kahdesta parhaasta, vaan [SE] [OTETTAISIIN] *kahdesta keskimmäisestä*, -- 27.10

## 3.a/b. verbi ja predikatiivi poistettu

Vain harvat eläimet voivat elää kokonaan toisista saman lajin yksilöistä riippumatta: tällaisia ovat esim. suolistolaiset, jotka eivät aivan välttämättä tarvitse lajikumppania munia hedelmöittämään, *tai* [TÄLLAISIA][OVAT] *silmikoitumalla lisääntyvät valtameren polttiais- ja sienieläimet*. 21.1

Korkeakoululaitoksen perusongelma ei ole jokin työajan mittausjärjestelmän tai opetuksen apuvälineiden puute, vaan [KORKEAKOULULAITOKSEN PERUSONGELMA] [ON] *tutkintojen ja opetuksen tavoitteettomuus ja suunnittelemattomuus*. 53.18

## 4.a/b. verbi ja adverbiaali/adverbiaaleja poistettu

Työllä ja tuotannolla voitaisiin luoda juuri niitä hyödykkeitä, joita ihmiset tarvitsevat, [TYÖLLÄ JA TUOTANNOLLA] [VOITAIISIIN] *lisätä kokoyhteiskunnan varallisuutta ja* [TYÖLLÄ JA TUOTANNOLLA] [VOITAIISIIN] *kohottaa kaikkien kansalaisten elintaso*. 37.12

Kansleri Klaus Waris kysyi äskettäisessä esitelmässään, onko henkilöautojen veropoliittisella hinnoittelulla lisätty Ruotsiin muuttoa ja [ONKO] [SILLÄ] *vähennetty asumisviihtyisyyttä Suomessa*. 46.17

On muistettava, että yksityismaanomistuksen piiriin eivät tulleet ainoastaan eurooppalaiset vaan [YKSITYISMAANOMISTUKSEN PIIRIIN] [TULI] *myös osa paikallisesta väestöstä*.

Pohjimmitaan Don Quijotessa asetetaan kyseenalaiseksi koko kaunokirjallisuus, [DON QUIJOTESSA] ei [ASETETA KYSEENALAISEKSI] *pelkääntään ritariromaaneja*. 18.9

5. VP:n *ellipsi* ei poista ainesta lauseen keskeltä kuten aukkoaminen, vaan lohkaisee lauseesta pois sen oikeanpuoleisen osan aina finiittiverbiin saakka, joka jää jäljelle. Aukkoaminen käy vain rinnasteisissa lauseissa, VP:n ellipsi soveltuu sekä rinnasteisten että alisteisten välillä. Lähisukua VP:n ellipsille on substituutti *niin ... -kin*, joka sen vuoksi on otettu tähän yhteyteen. Sen käyttäytyminen on jonkin verran epäselvää, VP:n poistosta siinä kuitenkin ei ole kyse. VP:n poisto hävittää vain infinitiiviainesta:

Sehän on käsite, jolle itsekukin saa antaa sen sisällön, *minkä tahtoo* [ANTAA] 75.19

- *jos voisi* [PINNATA] *niin kai pinnaisi hänkin*. 89.8

VP:n olettaminen muualla kuin em. infinitiivitapauksissa on suomen kielessä epävarmaa (Hakulinen & Karlsson (1979:225)), ja siksi käytämme käsitteen ympärillä lainausmerkkejä. Seuraavassa esimerkkejä ellipsisistä, joka on poistanut verbin loppuosan lisäksi siitä oikealle olevia nominaalisia aineksia riippumatta siitä, ovatko nämä valenssinmukaisia kiinteitä määritteitä vai eivät. Se että subjektin ellipsi esiintyy suhteellisen usein samanaikaisesti tämän ellipsin kanssa, antaa vihjeen siihen suuntaan, että parasta ellipsianalyysia ei suomen kielen suhteen vielä ole esitetty:

Koeryhmät näkivät elokuvan, *kontrolliryhmät eivät* [NÄHNEET ELOKUVAA]. 8.38

Voidaan tietysti sanoa, että metallimies on saanut palkkaa oppiaikanaan, *korkeakouluissa opiskeleva ei* [OLE SAANUT PALKKAA OPPIAIKANAAN]. 76.47

Poika tuli kysyneeksi, että missäs välissä muori on opetellut eri sienilajit tuntemaan. *Ehä mie ookkaa* [OPETELLUT ERI SIENILAJEJA TUNTEMAAN], äitimuori tuumi. 113.26

Olemme koodittaneet VP:n ellipseiksi myös tapaukset, joissa ei olekaan pyyhitty vain verbiketjun infiniittistä osaa vaan koko verbiketju (ns. "ränni", engl. sluicing).

Peruskoulujärjestelmään siirtyminen muuttaa tässä suhteessa tilannetta, *mutta vielä ei tiedetä kuinka paljon* [PERUSKOULUJÄRJESTELMÄÄN SIIRTYMINEN MUUTTAA TILANNETTA]. 29.15

Tämän olen lukenut, *vaikkakaan en muista mistä* [OLEN SEN LUKENUT]. 79.5

Kun verbi on yksisanainen, VP:n ellipsiksi tulee tulkittua tapaus, jossa komplementit on pyyhitty; siihen voi tosin sisältyä verbiainesta joko partisiippiirakenteesta tai verbikiteytymästä:

Ulkomaan elävä huokasi ja kysyi: -Otatko vielä viinin? - *Otan* [VIELÄ VIININ], sanoin. 80.34

-- kukaan katsoja tuskin odottaa näkevänsä Pigallen tyylistä seksishowta, ja *jos odottaa* [NÄKEVÄNSÄ PIGALLEN TYYLISEN SEKSISHOWN] niin pettyy. 95.5

VP:n ellipsistä voi jäädä jäljelle pelkkä kieltoverbi:

Koeryhmät näkivät elokuvan, kontrolliryhmät *eivät* [NÄHNEET ELOKUVAA]. 8.38

Tämä eroaa selvästi aukkotapauksista, joissa kieltoverbi ei jää typistetyn lauseen loppuun, vaan aukon vasemmalle reunalle:

-- suurin osa siirtomaissa tehdyistä investoinneista oli lähtöisin emämaan valtion kassasta *eikä* 0 [OLLUT LÄHTÖISIN] *siirtomaassa muodostetusta yksityisestä pääomasta*. 7.63

## 3. LAUSEOPIN KVANTITATIIVISTA METODIIKKA

### 3.1. Aineiston valinnan ja otannan periaatteet

Empiirisen kielentutkimuksen kaksi metodista perusongelmaa ovat tutkittavan *kielenaineuksen valinta* eli otoksen määrittäminen tutkittavasta perusjoukosta sekä *otoksen koon määrittäminen*. Edellisessä tapauksessa on kyse validiteetista, siitä, miten hyvin otos edustaa perusjoukkoa. Jos tutkittavana on esim. yksi tyyli, vaikkapa sanomalehtikieli, käytännössä ei voi tutkia kaikkia ilmestyneitä lehtiä, vaan on rajoitettava joihinkin eli poimittava otos. Otoksen täytyy edustaa perusjoukkoa (koko tyyliä) mahdollisimman tarkoin, jotta tulokset olisivat kelpoisia perusjoukkoon yleistettäviksi. Otoksen koon määrittämisessä taas on kyse tutkimuksen reliabiliteetista. Optimaalinen otoskoko on sopivan suuri, ei liian pieni, jolloin tulokset eivät ole luotettavia, eikä liian suuri, jolloin on tehty turhaa työtä, toisin sanoen kerätty liian paljon materiaalia.

Ratkaisevimmin aineiston valintaan vaikuttavat tutkimuksen tavoitteet ja niiden yhteydessä tehtävät teoreettiset valinnat. Vasta kun tutkimuksen tavoitteet on täsmennetty riittävän selvästi, voidaan ottaa kantaa aineiston valintaan. Olemme tehneet selkoa perustavoitteistamme kahdessa ensimmäisessä luvussa: haluamme selvittää muutamien mahdollisimman neutraalien tyyli- ja tekstisyntaktisten ilmiöiden yleisyysasteita, näiden frekvenssien relevanssia suomen kieliopille sekä otannan metodiikkaa. Tässä luvussa puututaan tarkemmin lauseopillisen tutkimuksen validiteetti- ja reliabiliteettiongelmiin.

Tarkastelkaamme aluksi niitä periaatteita, joiden pohjalta sellaiset suuret korpuset kuin Brownin (Brown University Corpus), Göteborgin ja Oulun on muodostettu; ks. Kucera & Francis (1967), Allen (1970) ja Saukkonen (1976). Nämä aineistot on koottu lähinnä frekvenssisanakirjojen laatimista varten: niiden päätavoite on siis leksikaalinen. Tässä jaksossa käsiteltäviä seikkoja ovat valaisevalla tavalla selvittäneet Kohonen & Salmela (1978), joiden näkökohdat ovat paljolti vaikuttaneet esitykseemme.

Brownin ja etenkin Göteborgin ja Oulun projektit tähtäävät siis ensisijaisesti sanaston kvantitatiivisen koostumuksen selvittelyyn. Otoksen koko vaihtelee 400000 ja 1000000 sanan välillä. Tällaisessa työssä otantamenetelmien täytyy olla tilastollisesti päteviä valittujen tekstilajien sisällä, ja tärkein metodinen ongelma on tekstilajien alaotosten valitseminen mahdollisimman kattavasti. Intuitiolla ei leksikkoparadigman suuruuden takia voi olla otannassa suurtakaan osaa - leksemejä on kymmeniä ellei satoja tuhansia.

Syntaktisia tarkoituksia varten valittujen korpusten päätavoite on yleensä selvemmin kvalitatiivinen. Esim. laajan Survey of English Usage -projektin tärkein tulos on ollut nykyenglannin kieliopin, etenkin syntaksin kuvaus, mm. Quirk & Greenbaum & Leech & Svartvik (1972). Tämän aineiston puhekielinen osa on parhaillaan muokattavana tietokone- ja teksti-aineistoksi Lundissa, ks. Svartvik (1980), jolloin mm. suoritetaan kieliopillinen ja prosodinen kooditus. Tämä tutkimus muistuttaa Lundin ruotsin kielen Talsyntax-projektia, jonka yhteydessä mm. kehitettiin meilläkin huomiota osakseen saanut makrosyntagma-analyysi (Loman & Jörgensen (1971); vrt. myös Saari (1975)).

Tätä työtä on myöhemmin jatkanut Teleman (1974), joka laati hyvinkin seikkaperäisen kooditusjärjestelmän puhutun ja kirjoitetun ruotsin kielen morfosyntaksin selvittämistä varten. Muuttajat olivat yleensä pintasyntaktisia ja siten valittuja, että ne mahdollistavat sosiolingvistikin relevantin päättelyn (vrt. Anward (1976), Einarsson (1978), Saari (1975)). Lundin Skrivsyntax-korpukseseen kuului neljä tyylilajia: tiedottavaa sanomalehtiproosaa, tietokirjallisuutta, väittelyjä sekä esitteitä. Saneita oli tekstilajeittain noin 22000, yhteensä noin 87000, vrt. Westman (1974), Platzack (1974) sekä Hultman & Westman (1977). Tässä yhteydessä on myös mainittava Alankomaiden Nijmegenin yliopistossa suoritettava Computer Pilot Project, joka tähtää syntaktisesti edustavan korpuksen rakentamiseen. Korpuksen sanemäärä on noin 130000, ks. Aarts (1978). Äskettäin on Ellegård (1978) käyttänyt Brownin aineiston syntaktisesti uudelleen kooditettua osaa (128000 sanetta) tutkimukseen, joka tavoitteiltaan muistuttaa Tal- ja Skrivsyntaxia. Ellegård esittää mm. kvantitatiivisia perustietoja sellaisista muuttujista kuin yhdyslauseen osalauseiden lukumäärä, yksinkertaisten lauseiden konstituentti- ja sanamäärä, sivulauseen tyyppi ja sanaluokat.

On myös joitakin suomen kieltä koskevia tämän tyyppin tutkimushankkeita. Kotimaisten kielten tutkimuskeskuksen Lauseopin arkisto käsittää lopullisessa muodossaan noin 800000 saneen kokoisen korpuksen, jonka pohjalta tähdätään suomen murteiden deskriptiivisen syntaksin laatimiseen; ks. Ikola (1975) ja Ikola & Karjalainen (1977). Lauseopin arkistossa on myös kirjakielistä ainesta, joka saattaa tarjota vertailumahdollisuuksia meidän tutkimuksemme aineistoon ja tuloksiin niissä tapauksissa, missä muuttajat ovat vertailukelpoisia. Lauseopin arkiston kooditusjärjestelmän muuttajat ovat lähinnä perinteisen kieliopin morfologisia ja lauseenjäsenkategorioita.

Tampereen yliopiston suomen kielen laitoksessa Paavo Siro on pannut alulle suomen kielen tietokonemuotoisen korpuksen. Aineisto on kokonaisuudessaan aikakaus- ja sanomalehtikieltä. Tämä korpus pyrkii lähinnä tarjoamaan aitoja lauseopin esimerkkejä nopealla ja helposti palautettavalla tavalla. Tampereen kooditusjärjestelmä noudattelee myös pääpiirteissään perinteistä sanaluokka- ja lauseenjäsenjärjestelmää, ks. Yli-Vakkuri (1972).

Yhteenvedona voidaan todeta, että laajempia kvantitatiivisia lauseopillisia tutkimuksia varten kerätyt korpuksat yleensä näyttävät olevan noin 100000 saneen kokoisia, joka tapauksessa huomattavasti pienempiä kuin kvantitatiivisia sanastontutkimuksia varten tarvittavat. Kuten tuonnempana osoitamme (3.2.), monien syntaktisten muuttujien tekstifrekvenssien selvittelyä varten riittää paljon pienempikin (homogeeninen) aineisto kuin 100000 sanetta validiteetin ja reliabiliteetin ollessa turvattuina.

On useita seikkoja, jotka vaikuttavat syntaktisia tarkoituksia varten kootun otoksen koon määrittämiseen; ks. lähemmin Salmela & Kohonen (1977) ja erityisesti Kohonen & Salmela (1978). Yleisesti voidaan sanoa, että syntaktinen otos voi olla huomattavasti pienempi kuin sanaston tutkimiseen tarkoitettu otos. Syntaktisten muuttujien alakategoriat, niiden paradigmatt, ovat vähäjäsenisiä umpiluokkia leksikon kvantitatiiviseen koostumukseen verrattuna. Esim. lauseen modaalisuustyyppijä on viisi ja NP:n erilaisia rakennetyyppejä yksitoista. Leksikon puolella näitä vastaavat sellaiset elementit kuin konjunktiot ja prepositiit, jotka nekin muodostavat umpiluokkia.

Syntaktisilla muuttujilla on siis suhteellisen pieni määrä alakategorioita, meidän järjestelmässämme yleensä vähemmän kuin kymmenen. Kuten Salmela & Kohonen (1977) huomauttavat, ratkaisevimmin otoksen kokoon vaikuttavat tekijät ovat *alakategorioiden*

*lukumäärä* sekä *kussakin alakategoriassa olevien esiintymien lukumäärä* (frekvenssi). Perustavassa muuttujassa, jossa on vain muutama alakategoria, kaikkien alakategorioiden esiintymiä kertyy hyvinkin pienestä otoksesta riittävästi. Viisikategoriaisien muuttujan arvot voidaan siten määrittää tilastollisesti luotettavasti pienemmästä otoksesta kuin kaksikymmentäkategoriaisien. Nyrkkisääntönä pätee, että validisti poimittu otos tuottaa reliabelin frekvenssijakauman, jos jokaisesta alakategoriasta on ainakin viisi esiintymää. Jos otoksessa ei ole viittä esiintymää jostakin alakategoriasta, tilastollisen luotettavuuden arviointi edellyttää tällaisen alakategorian yhdistämistä johonkin muuhun alakategoriaan, jos tämä vain on teoreettisesti mahdollista (niin että kyseiset alakategoriat muodostavat sopivan luonnollisen luokan). Yhdistetty alakategoria on tietysti karkeampi kuin alkuperäiset kaksi, mutta saattaa silti olla hyödyllinen. Useampiakin alakategorioita voi olla syytä yhdistää. Esim. pintajärjestysmuuttujan (18) kymmenen alakategoriaa olemme eräiden hypoteesien testaamiseksi yhdistäneet kolmeksi ryhmäksi, SV (alakat. 1,2,5,7), VS (alakat. 3,4,6,8) sekä V (alakat. 9,A). Tällaiset yhden muuttujan sisäiset uudelleen järjestelyt on tietenkin helppo suorittaa tietokoneella.

Suuria otoksia tarvitaan nimenomaan runsaskategoriaisista muuttujista varten. Suurempia otoksia tarvitaan aina myös, kun alakategorioiden jakauma on vino joidenkin ollessa kovin harvinaisia. Esim. lauserakennemuuttujan (15) kohdalla täytyy kasvattaa otosta huomattavasti, jos haluaa luotettavan kuvan myös omistusrakenteen ja säälauseiden frekvenssistä eikä tyydy vaikkapa pelkkiin aktiivisiin ja passiivisiin transitiivi- ja intransitiivilauseisiin.

Niin kauan kuin alakategorioita on "kohtuullisesti", karkeasti sanottuna vähemmän kuin kymmenen, niiden frekvenssijakauma ei ole kovin vino (niin ettei mikään ole kovin harvinaisen), ja otos on homogeeninen (eikä koostu toisistaan jyrkästi poikkeavista tekstilajeista), *muutaman sadan* lauseen otos on itse asiassa riittävä useimpien keskeisimpien muuttujien frekvenssijakauman luotettavaa selvittämistä varten. Esittämiemme teoreettisten pohdiskelujen valossa ja ottaen huomioon tutkittavien tekstien homogeenisuuden ja valitsemiemme muuttujien luonteen, on ilmeistä, että voimme tyytyä suhteellisen pieneen otokseen. Kuten todettua analyysimme perustuu otokseen, jossa on noin 10000 (10149) yksinkertaista lausetta.

Virkkeiden eli graafisten (yhdyks)lauseiden lukumäärä on 5016; "virke" on jono, jonka kummallakin puolella on iso välimerkki tai väli; "lauseen" määritelmästämmä ks. muuttujan (8) esittelyä jaksossa 2.4. Korpuksemme sanemäärä on 66851, siis selvästi pienempi kuin edellä mainittujen ulkomaisten projektien, esim. Skrivsyntaxin, Nijmegenin ja Ellegårdin käyttämän Brownin osakorpuksen. Kuten kohta käy ilmi, aineistomme koko on silti useimpia tarkoituksiamme varten tarpeeksi suuri. On myös helppo suurentaa aineistoa niiden muuttujien osalta, jotka nyt tulevat epäluotettavasti katetuiksi.

Itse asiassa 10000 lauseen korpus on suurempi kuin tarvitaan pelkkien juoksevien tekstifrekvenssien arvioimiseksi. Mutta näinkin laaja korpus on tarpeen heti, kun ristiintaulukointien avulla alkaa selvittää muuttujien välisiä riippuvuuksia. Minimaalinen otos kutistuu liian pieneksi, kun yhden muuttujan kaikki alakategoriat hajaantuvat pitkin toisen muuttujan kaikkia alakategorioita.

Tutkimme yhtä tekstilajia, *neutraalia asiaproosaa*, jonka sisällä olemme erottaneet seuraavat kuusi toisistaan hieman poikkeavaa alatekstilajia. Aineiston tarkka koostumus on eriteltyinä liitteessä 2.

	lauseet	virkkeet
1. tietokirjat	1270	587
2. ensyklopediat	1052	529
3. pääkirjoitukset (kotimaan aiheista)	953	513
4. kulttuuriartikkelit (sanomalehdistä, aiheina musiikki, näyttämötaide ja arkkitehtuuri)	1556	737
5. pakinat (ei parodioita eikä satiireja)	1690	792
6. tiedottavat artikkelit (sanoma- ja viikkolehdistä, ei politiikkaa koskevia artikkeleita)	3682	1854
yht.	10149	5016

Nämä alatekstilajit ovat melkein kaikki sanomalehtiproosan muunnoksia; ks. Westman (1974) ja Strand (1977), jotka luonnehtivat tarkemmin tämän kielen lajin keskeisimpiä piirteitä.

On syytä korostaa, että tekstit valittiin intuitioon perustuvan seulonnan jälkeen emmekä siis soveltaneet satunnaisotantaa. Kolmen henkilön raati (Auli Hakulinen, Fred Karlsson, Maria Viikuna) kävi läpi suuren määrän potentiaalisesti kyseeseen tulevia tekstejä, ja mukaan kelpuutettiin ne, joita yksimielisesti pidettiin neutraaliin asiaproosaan kuuluvina. Pyrimme minimoimaan dialogin osuutta, ja kaikki tekstit, joissa voitiin todeta ilmeisiä tyyllisiä pyrkimyksiä, jätettiin aineiston ulkopuolelle.

Tätä intuitiopohjaista menettelyä voi arvostella puhtaasti tilastollisesta näkökulmasta. Olemme kuitenkin sitä mieltä, että pelkkä satunnaisotanta ei tilastollisesta oikeellisuudestaan huolimatta välttämättä takaa myös *lingvistisesti merkitsevää* lopputulosta. Näin ei ole varsinkaan silloin, kun otos on sekakoosteinen ja sisältää aineksia monenlaisista tyyleistä. Vaikka aineistomme kokoamistapa ei siis ole tilastollisesti validi, menettelytapamme on silti puolustettavissa, koska lingvistinen merkitsevyys ei ole välttämättä samaa kuin tilastollinen merkitsevyys ja koska toisaalta kolmen arvioijan käyttö on jonkinmoinen tae intersubjektiivisuudesta. Vrt. myös jaksoon 3.4., jossa eritellään tarkemmin alatekstilajien tyyllisiä eroja.

Tutkimuksen tavoitteiden kannalta on ensisijaisen tärkeää, että otos koostuu kokonaisista teksteistä tai ainakin itsenäisistä pitkäköistä tekstinkappaleista, jotka on otettu kirjan tai pitemmän artikkelin alusta. Ovathan monet muuttujistamme sellaisia, että niiden luonne ja yhteisvaikutus on todettavissa vasta kokonaisissa teksteissä eikä yksinäislauseiden sisällä. Irralliset, satunnaisotannalla valitut, tekstuurittomat (vrt. Halliday & Hasan (1976)) tekstijaksot ovat meidän tarkoituksiamme varten käyttökelvottomia.

Otantamenetelmämme merkitsee, että saavutetut tulokset pätevät vain valittuun tekstilajiin eikä niitä tietenkään voi suoraan yleistää muihin. On silti pidettävä mielessä, että neutraalin asiaproosan voi olettaa muodostavan kirjoitetun kielen *ydinaineksen* (common core). Eräät Oulun projektin äsken saavuttamat tulokset tukevat tätä väitettä. Niinpä Saukkonen (1977) on sanojen frekvenssijakauman osalta todennut, että lähinnä kaikkien tutkittujen tekstilajien (yhteensä kuutisenkymmentä) keskiarvoa ovat sanomalehtiartikkelit, etenkin asia-artikkelit, reportaasit, katsaukset ja pakinamuotoiset kirjoitukset (vrt. (1-6) edellä!) sekä lisäksi matkakertomukset. Virkkunen (1977) sai samankaltaisia tuloksia tutkiessaan eräiden sijamuotojen frekvenssijakaumaa (essiivin, translatiivin, abessiivin, komitatiivin ja instruktiivin). Hänen mukaansa seuraavat tekstilajit ovat lähinnä keskiarvojakaumaa: asia-artikkelit, pakinat, mainokset, matkakertomukset ja eräät muut. Tällaiset tulokset osoittavat, että

tekstilajimme muodostavat otoksen, jota sinänsä voi pitää edustavana näytteenä yhdestä kirjoitetun suomen kielen keskeislohkosta.

Tutkimuksen tässä vaiheessa emme ole erityisen kiinnostuneita tekstilajien välisistä syntaktisista ja muista eroista (vrt. kuitenkin 3.4.). Pää tarkoituksenamme on selvittää neutraalin asiaproosan keskeiset tekstisyntaktiset ja -semanttiset ominaisuudet. Olemme kuitenkin selvillä tekstilajien välisten vertailujen tarpeellisuudesta. Tuskin tarvitsee mainita, että kvantitatiiviset tiedot pääsevät varsinaisesti oikeuksiinsa vasta tällaisten kontrastiivisten vertailujen myötä. Useampien tekstilajien selvittely olisi ilmeinen mahdollisuus jatkotutkimukselle. Jo otannan reliabiliteettiongelmiensa valaiseminen edellyttää kuitenkin ensimmäisen askelen olevan yhden tekstilajin mahdollisimman huolellinen selvittely.

### 3.2. Syntaktisen otoksen koon määrittäminen

Kuten todettua 10149 lauseen otostamme ei valittu satunnaisesti, vaan analysoitavat tekstit hyväksyttiin korpukseen vasta, kun kolmen arvioijan raati oli yksimielisesti pitänyt niitä neutraalin asiaproosan intuitiivisesti edustavina näytteinä.

Voidaksemme päätellä, mikä reliaabelin homogeenisestä perusjoukosta (=10149 lausetamme) poimitun otoksen *minimikoko* voi olla, kun halutaan luotettava arvio vain lauseopillisten muuttujien alakategorioiden frekvensseistä (eikä siis muuttujien välisistä korrelaatioista), generoimme tietokoneen avulla perusjoukostamme satunnaisesti viisi erisuuruista alakorpusta, joissa oli 100, 300, 500, 1000 ja 2000 lausetta. Koska alakorpukset on generoitu samasta rajallisesta perusjoukosta, ne ovat osittain päällekkäisiä.

Tämän jälkeen suoritettiin joukko tavanomaisia  $\chi^2$ -testejä, joiden tarkoituksena oli 22 muuttujan osalta testata kooltaan vierekkäisten alakorpusten frekvenssierojen merkitsevyyttä. Taulukkoon 1 on koottu näiden testien kiintoisimmat tulokset. Kun puhutaan tilastollisesta merkitsevyydestä on oltava selvillä siitä, että merkitsevyyden taso (riskitaso eli todennäköisyys, jolla estimaatti on oikea) ei ole mitenkään ennalta tiedossa, vaan se on *valittava*. Juuri riskitason valinnanvaraisuuden takia tilastollinen merkitsevyys ja lingvistinen merkitsevyys eivät välttämättä lankea yksiin: mikä on tilastollisesti merkitsevää ei aina ole lingvistisesti merkitsevää, ja mikä on lingvistisesti merkitsevää ei välttämättä noudata ankaria tilastollisia merkitsevyysvaatimuksia. Olemme valinneet 1 %:n riskitason kuten on tehnyt myös Westman (1974) laajassa ruotsin kirjoitettua kieltä koskevassa tutkimuksessaan. Taulukon 1 luvuista 1 %:n yläpuolella olevat osoittavat siis erojen ei-merkitsevyyttä, 1 %:n ja sen alapuolella olevat luvut taas sitä, että todetut erot ovat merkitseviä (tällä riskitasolla). Mutta riskitason voi tietysti valita toisinkin. Esim. Gellerstam (1970) on asettanut sen niinkin ankaraksi kuin 0.1 prosentiksi, Kohonen (1976) taas on tyytynyt 5 prosenttiin.

Muuttujien alakategoriat on osittain ollut pakko ryhmitellä uudelleen  $\chi^2$ -testin sovellusvaatimusten takia. Jotta testiä ylipäättänsä voitaisiin soveltaa, jokaisesta alakategoriasta on oltava ainakin 5 esiintymää. Esim. muuttujassa (4:lauseiden lukumäärä virkkeessä) olemme muodostaneet uuden alakategorian, johon kuuluvat kaikki virkkeet, joissa on enemmän kuin 4 lausetta. Muuttujassa (16: lauseen sanamäärä) 1-11 sanan pituiset lauseet on pidetty erillisinä kategorioina, mutta ne joissa on 12-50 sanaa on yhdistetty yhdeksi alakategoriaksi. Muuttujassa (20: subjektin sija ja kvanttori) alkuperäiset alakategoriat 5,6,8 on

yhdistetty, jne. Alkuperäisten ja uudelleen järjestettyjen alakategorioiden lukumäärä esitetään taulukossa 1 muuttujan nimen jäljessä, esim. 14/7 muuttujassa (8) jne.

TAULUKKO 1. *Alakorpusten merkitsevyydestit*

muuttujat (alakategorioiden lukumäärä sulkeissa)	verratut alakorpukset			
	100/300	300/500	500/1000	1000/2000
(4) lauseiden lukumäärä virkkeessä (11/4)	68%	72%	91%	92%
(8) lausetyyppi (14/7)	81%	7%	89%	99%
(14) kielto (3/2)	91%	41%	83%	90%
(16) lauseen sanamäärä (33/12)	68%	85%	99%	92%
(18) pintajärjestys (12/7)	49%	69%	94%	99%
(19) subjektin rakenne (13/7)	80%	67%	98%	92%
(20) subjektin sija ja kvanttori (9/7)	47%	85%	94%	94%
(33) predikatiivin rakenne (12/6)	8%	65%	95%	88%
(38) alkuadverbiaalin rakenne (9/5)	9%	78%	88%	97%
(46) loppuadverbiaalin tyyppi (8/6)	6%	11%	86%	93%
(52) adverbiaalien määrä lauseessa (6/4)	1%	3%	49%	74%

Taulukkoon 1 on otettu mukaan kaikki ne muuttujat, joissa alakorpusten erot ovat suurimmat. Osoittautuu, että ei ole merkitseviä eroja edes 100 ja 300 lauseen otosten välillä, ja yleisesti ottaen otosten frekvenssiominaisuudet muodostuvat samanlaisemmiksi otoskoon kasvaessa; tämä näkyy ei-merkitsevyysslukujen kasvuna vasemmalta oikealle esim. muuttujissa (4,18,20,38,46). Tämä tendenssi ei silti ole poikkeukseton, minkä osoittaa muuttuja (8), jossa otokset 300/500 poikkeavat enemmän toisistaan kuin otokset 100/300, ja muuttuja (14), jossa otosten 1000/2000 erot ovat hieman suuremmat kuin otosten 100/300. Keskeistä on kuitenkin, että nämäkään erot eivät ole tilastollisesti merkitseviä.

Siitäkin huolimatta että tuloksemme pätevät vain ehdollisesti (aineistomme on homogeeninen ja vähäesiintymäiset alakategoriat on ollut pakko yhdistellä), voi esittää seuraavan päätelmän: muutamana sadan yksinäislauseen otos on riittävä paljastamaan keskeisten syntaktisten piirteiden (alakategorioiden) frekvenssijakauman; vrt. myös Kohonen & Salmela (1978), Kohonen (1976, 1978). Yksi tärkeä jatkopäätelmä tästä on se, että suhteellisen homogeenisten tekstilajien (tyylien) vertailu voi perustua pieniinkin aineistoihin ilman että tulosten reliabiliteetti kärsii. Silti korpuksen koon määrittämiseen liittyvien metodisten ongelmien pohdiskelu on tärkeä osa mitä tahansa empiiristä syntaksin tutkimusta.

### 3.3. *Juoksevien frekvenssien "jäätymispisteet"*

Käsitteellä "jäätymispiste" tarkoitetaan tässä Kohosen & Salmelan (1978) liittyen sitä otoksen kokoa, jonka yläpuolella merkitseviä muutoksia muuttujien alakategorioiden juoksevilla frekvensseillä ei voi enää todeta. Edellisessä jaksossa käsitellyt merkitsevyydestit osoittavat,

että muutaman sadan lauseen korpus riittää useimpien alakategorioiden frekvenssin luotettavaksi estimoimiseksi, jolloin siis jäätymispiste on saavutettu. Tässä jaksossa osoitamme muutamain esimerkein tarkemmin, miten jäätymispiste saavutetaan hyvinkin nopeasti, jonka jälkeen frekvenssi pysyy suurin piirtein samana siitä riippumatta, miten laajaa aineistoa käytetään. Esiteltävät muuttujat ovat (14: lauseiden lukumäärä virkkeessä), (8: lausetyyppi), (16:lauseen sanamäärä), (18: pintajärjestys) ja (19: subjektin rakenne).

Taulukoissa 2-6 olevat korpuksset (A-D) eivät ole satunnaisesti generoituja kuten edellisessä JAKSossa käsitellyt alakorpuksset, vaan koostuvat peräkkäisistä, yhteen tai muutamaaan alatekstilajiin kuuluvista lauseista: (A) = noin 500 lausetta asiaproosaa (tieto- sekä tietosanakirjoista), (B) = noin 1000 lausetta samanlaista asiaproosaa ((A) on (B):n osajoukko), (C) = noin 2000 lausetta tiedottavaa ja pakinaproosaa, ja (D) = koko aineisto eli 10149 lausetta. (Tämän jakson esitys on sikäli impressionistista, että emme ole suorittaneet merkitsevyystestejä, jotka vasta sitovasti osoittaisivat, ovatko vertailtujen aineistojen erot merkitseviä vai eivät.)

TAULUKKO 2. *Lauseiden lukumäärä virkkeessä (4)*

lauseet	(A)		(B)		(C)		(D)	
		%		%		%		%
1	99	39	197	39	420	41	1929	38
2	82	32	184	36	360	35	1754	35
3	47	19	79	16	164	16	851	17
4	17	7	32	6	43	4	330	7
5	6	2	13	3	23	2	103	2
6	2	1	2	0	6	1	31	1
7			1	0	1	0	10	0
8							3	0
9							3	0
10							1	0
11							1	0
yht.	253		508		1017		5016	

Näkee heti, ettei alakorpusten ja koko aineiston frekvenssien välillä ole mainittavia eroja. 500 lauseen alakorpuksesta (A) saadaan melkein yhtä tarkka kuva virkkeiden lauserakenteesta kuin koko aineiston luvuista. Vajaa 40 % virkkeistä on siis yksilauseisia, noin 35 % kaksilauseisia, noin 17 % kolmilauseisia. Viisi- ja kuusilauseisia virkkeitä on enää kovin vähän, 2 % ja 1 %. Poikkeuksellisen pitkiä 8-11 lauseen mittaisia virkkeitä ei pienimmässä korpuksessa ole yhtään, koko aineistossa kourallinen, silti ei yhteensä prosenttiakaan.

TAULUKKO 3. *Lausetyyppi* (8)

	(A)		(B)		(C)		(D)	
		%		%		%		%
1. päälause	270	53	561	55	1023	52	5336	53
2. konj. + päälause	58	11	128	13	216	11	1099	11
3. sitaatti	2	0	2	0	177	9	356	4
4. rel.lause subj. NP:ssä	24	5	45	4	55	3	336	3
5. rel.lause obj. NP:ssä	18	4	34	3	37	2	244	2
6. rel.lause muussa NP:ssä	28	5	73	7	73	4	448	4
7. <i>että</i> -lause	39	8	55	5	140	7	891	9
8. epäsuora kys.lause	1	0	3	0	24	1	192	2
9. adverbiaalilause	32	6	45	4	140	7	676	7
A. komparatiivilause	8	2	10	1	19	1	117	1
B-E. ei-finiitt.rak.	32	6	56	6	62	3	450	4
yht.	512		1012		1966		10145	

Päälauseiden osuus kaikista lauseista on suurin piirtein sama kaikissa korpuksissa, noin 65 %, samaten *että*-lauseiden noin 8 %:n osuus. Muiden lausetyyppien frekvenssit vaihtelevat siinä määrin, että pienimpiin korpuksiin ei voi luottaa, esim. relatiivilauseiden kokonaisosuus on (A):ssa 14 % mutta (D):ssä enää 9 %. Samaten sitaattilauseiden esiintymisessä on huomattavia eroja (tähän kuuluu sekä pää- että sivulauseita). Adverbiaalilauseiden osuus on melko kiinteästi noin 7 %, ne ovat siis vähän harvinaisempia kuin *että*-lauseet. Pääpiirteittäin voidaan kuitenkin jälleen todeta, että (A) on yllättävän tarkka (D):n kuva.

TAULUKKO 4. Lauseen sanamäärä (16)

sanat	(A)		(B)		(C)		(D)	
		%		%		%		%
1	7	1	9	1	25	1	138	1
2	24	5	33	3	141	7	590	6
3	45	9	80	8	199	10	954	9
4	53	10	107	11	258	13	1297	13
5	62	12	124	12	303	15	1381	14
6	69	13	136	13	303	15	1396	14
7	50	10	112	11	195	10	1061	10
8	56	11	116	11	176	9	962	9
9	31	6	64	6	116	6	665	7
10	35	7	67	7	84	4	518	5
11	19	4	41	4	60	3	385	4
12	18	4	32	3	29	1	217	2
13	10	2	18	2	22	1	183	2
14	6	1	15	1	20	1	126	1
15	4	1	13	1	9	0	17	1
16-20	19	4	33	3	22	1	165	2
21-25	3	1	9	1	2	0	29	0
26-50	1	0	3	0	2	0	12	0
yht.	512		1012		1966		10149	

Kaikki otokset osoittavat samaa: lauseen keskipituus on 5-6 sanaa, ja frekvenssikäyrä laskee tästä tasaisesti kumpaankin suuntaan. Pisimpien lauseiden frekvenssin estimoiminen on tietysti vaikeinta, esim. 10 sanan pituisten lauseiden kohdalla on selvältä tuntuva ero korpusten (A) 7 % ja (D) 5 % välillä, samaten 16-20 sanan pituisten lauseiden kohdalla, 4 % ja 2 %.

Virkkeen lausemäärän (muuttuja 4) ja lauseen sanamäärän (16) laskeminen ovat monien luettavuus- ja ymmärrettävyyksien kaikkein perustavimpia muuttujia, vrt. esim. Platzack (1974), ja niitä käytetään paljon myös tyylintutkimuksessa. Näiden testien laatimisen kannalta on jälleen syytä panna merkille, että kovin suuria aineistoja ei ole tarvis käyttää.

Taulukon 4 suhteen on pidettävä mielessä, että "lauseen" käsitteemme poikkeaa hieman perinteisestä. Koska tulkitsemme myös tietyt tekstuaalisten ellipsien kautta redukoituneet jaksot yksinkertaisiksi lauseiksi (esim. Kalle syö ja *0 juo*), saamamme lauseiden sanamäärät ovat hieman pienempiä kuin eräät muut analyysitavat antaisivat tulokseksi. Esim. Westman (1974) ei tulkitse tällaisia redukoituneita konjunkteja itsenäisiksi lauseiksi, vaan pitää koko virkettä lauseena, johon jälkimmäinen konjunktin kuuluu lausekkeena ("konjunkttilausekkeena"). Hänen analyysinsä mukaan siis äskeinen esimerkki on yksi ainut nelisanainen lause, meidän analyysissämme taas on kyse kahdesta lauseesta, joista toisessa on kaksi ja toisessa yksi sana.

TAULUKKO 5a. *Pintajärjestys* (18)

	(A)		(B)		(C)		(D)	
		%		%		%		%
1. SVX	241	47	548	54	902	46	4928	49
2. SXV	3	1	3	0	15	1	60	1
3. VSX	9	2	9	1	35	2	177	2
4. VXS	13	3	15	1	30	2	233	2
5. XSV	14	3	24	2	77	4	348	3
6. XVS	57	11	117	12	251	13	1139	11
7. SV	42	8	81	8	154	8	746	7
8. VS	12	2	18	2	38	2	218	2
9. VX	65	13	97	10	200	10	970	10
A. XV	40	8	70	7	168	9	861	8
B. V					15	1	36	0
irrelevantti	16	3	30	3	81	4	433	4
yht.	512		1012		1966		10149	

Taulukko 5a on hieman epähavainnollinen. Päätendenssit ilmenevät paremmin, kun äskeiset alakategoriat ryhmitellään uudelleen kolmeksi päätyypiksi, nimittäin subjekti ennen verbiä, subjekti verbin jäljessä ja subjektiton lause.

TAULUKKO 5b. *Pintajärjestys pääryhmittäin*

	(A)		(B)		(C)		(D)	
		%		%		%		%
S - V	300	59	656	65	1148	58	6082	60
V - S	91	18	159	16	354	18	1767	17
V	105	21	167	17	383	19	1867	18
irrelevantti	16	3	30	3	81	4	433	4
yht.	512		1012		1966		10149	

Jo (A):sta voidaan jälleen nähdä kaikki pintasanajärjestyksen peruspiirteet: SVX-tyypin hallitsevuus, noin 50 %:ssa kaikista lauseista on tämä järjestys, sekä XVS-lauseiden suuri osuus, noin 11 %. Myös subjektittomien VX-lauseiden huomattava 10 %:n osuus pistää silmään (tämän alakategorian kohdalla (A) eroaa selvästi muista alakorpuksista). - Näiden sanajärjestykseikkojen teoreettisesta relevanssista ks.jaksoa 4.4.

TAULUKKO 6. *Subjektin konstituenttirakenne (19)*

	(A)		(B)		(C)		(D)	
		%		%		%		%
1. rinnasteinen	22	6	42	5	57	4	303	4
2. pääsana + lause	20	5	38	5	61	4	363	5
3. lause	3	1	3	0	16	1	99	1
4. infinitiivi	7	2	9	1	15	1	110	1
5. nominaalistus	16	4	28	3	24	2	208	3
6. N + appositio	10	3	18	2	21	1	91	1
7. N + advl./inf.	14	4	31	4	49	3	290	4
8. partisiippi + N	20	5	43	5	45	3	330	4
9. muu attr. + N, N + muu attr.	86	22	198	24	353	23	1857	23
A. N	150	38	329	40	667	44	3215	40
B. inkorporoitu	16	4	28	3	86	6	403	5
C. geneerinen 0	4	1	5	1	27	2	160	2
D. subjekti poistettu	30	8	56	7	107	7	593	7
yht.	398		828		1528		8022	

Määritteettömät substantiivit, 40 %, sekä yksinkertaisen attribuuttirakenteen omaavat substantiivit, 23 %, ovat kaikkien alakorpusten suurimmat ryhmät, eikä kovin mainittavia alakorpusten välisiä eroja ole. Perustava frekvenssijakauma käy selvästi ilmi jo 500 lauseen otoksesta (A).

Kun tarkoin tutkii taulukoita 2-6, havaitsee että alakorpus (C) poikkeaa yleensä eniten muista. Tämä johtuu siitä, että huomattavan osan (C):n lauseista muodostavat pakinatekstit. A priorikin on selvää, että pakinoiden voidaan odottaa poikkeavan kaikkein neutraaleimmasta asiaproosasta, ja siksi on luultavaa, että (C):n poikkeamistakin ainakin osa on tyylin merkitseviä erottimia.

Tämän jakson numeeriset tulokset osoittavat jo vailla sitovaa tilastollista merkitsevyyksikäsittelyäkin sen, että hyvinkin pienet otokset ovat riittäviä tutkittaessa paitsi melko "mekaanisten" muuttujien kuten virkkeiden lausemäärän ja lauseiden tai konstituenttien sanamäärän myös lingvistisesti monimutkaisempien ja teoreettisesti kiinnostavampien muuttujien alakategorioiden frekvenssijakaumaa, esim. lausetyyppien (8), sanajärjestyksen (18) ja konstituenttirakenteen (19). Korostamme silti vielä kerran lähestymistapamme rajoituksia: vaikka äskeisten taulukoiden pienemmät alakorpukset eivät yleensä poikkeaa "ratkaisevasti" koko aineiston (D) luvuista, emme toisaalta ole suorittaneet merkitsevyydestejä, jotka ainakin (C)-korpuksen osalta toisivat esiin joitakin merkitseviä eroja (D):hen verrattuna; otoksemme ovat intuitiivisesti valittuja ja suhteellisen homogeenisia; harvinaisimpien alakategorioiden frekvenssien luotettava estimointi edellyttää aina suurempien otosten käyttöä, kuten edellyttävät myös useamman muuttujan ristiintaulukoinnit (vrt. lukua 4).

### 3.4. Alatekstilajien eroista

Taulukoita 2-6 tarkasteltaessa huomattiin, että alakorpus (C) eroaa usein muista. Tämä on ymmärrettävää, kun tietää sen sisältävän huomattavan määrän pakinatekstejä, joiden tyyli ei oletettavasti ole kaikkein neutraaleinta. Tässä jaksossa on tarkoituksena selvittää aineistomme homogeenisuutta sekä sitä, missä määrin alatekstilajit poikkeavat toisistaan ja mitkä muuttujat näyttävät olevan tyyli-indikaattoreita.

Testasimme 22 muuttujan juoksevien frekvenssien tekstilajikohtaisten erojen merkitsevyyttä. Taulukossa 7 seuraavalla sivulla esitetään näiden merkitsevyydestien tulokset: frekvenssilukuja emme tässä esittele. Testatut muuttujat esiintyvät pystytasossa, vaakatasossa on sarakkeittain kulloinkin vertaillut kaksi alatekstilajia. Jokaista alatekstilajia on vertailtu kaikkiin muihin. Numerointi on sama kuin liitteessä 2 esitetty: 1) tietokirjat, 2) ensyklopediat, 3) pääkirjoitukset, 4) kulttuuriartikkelit, 5) pakinat ja 6) tiedottavat artikkelit. Taulukon luvut ovat prosenttilukuja, jotka ilmoittavat riskitason, jolla todettu frekvenssiero on tilastollisesti merkitsevä.

$\chi^2$  -testin sovellettavuuden takia muuttujien alakategoriat on osittain ollut pakko ryhmitellä uudelleen (vrt. lukua 4). Mainitsimme muuttujan numeron sekä sulkeissa yhtäläisyysmerkin vasemmalla puolella uudelleen ryhmitetty alakategoria ja sen oikealla puolella alkuperäiset alakategoriat, jotka ilmenevät liitteestä 1: muuttuja 8 (D=D,E), 9 (2=2-5), 14 (2=2,3), 16 (12=12-50), 17 (6=6-8), 18 (1=1,2,5,7; 2=3,4,6,8; 3=9-B), 19 (1=A; 2=8,9; 3=6,7; 4=2-5; 5=1; 6=B-D), 20 (4=4-8), 23 (1=1,2,5; 8=8,9), 24 (7=7-H), 25 (5=5-A), 27 (1=9; 2=7, 8; 3=5,6; 4=2-4; 5=1; 6=A, B), 28 (5=5,9), 31 (1=1,2,5; 8=8,9), 32 (7=7-H), 33 (3=3-5; 8=8-C), 34 (1=1,6; 2=2,3; 4=4,5), 37 (5=5-H), 38 (4=4,5,8,9; 6=6,7), 45 (4,5,8,9; 6=6,7), 52 (4=4-6), 54 (5=3,5,7,H).

(KS. TAULUKKOA 7 SIVULLA 178.)

Taulukosta 7 näkee heti, että alatekstilajien välisiä eroja on melkoisesti eikä siis aineistomme ole täysin homogeenista. Jos ratkaisevaksi merkitsevyyden tasoksi valitaan edelleen 1 %, voidaan ensin todeta, monenko muuttujan kohdalla eri alatekstilajien väliset vertailut tuottavat merkitsevän eron.

Tämä ilmenee taulukosta 8, jonka luvut on saatu laskemalla yhteen taulukon 7 sarakkeiden 1 %:n tai sen alapuolella (= 0) olevat tapaukset.

## TAULUKKO 8. Alatekstilajien merkitsevien erojen lukumäärä

testatut tekstilajit	erojen lukumäärä
1/2 tieto/ensykl	9
1/3 tieto/pääkirj	7
1/4 tieto/kultt	9
1/5 tieto/pakina	17
1/6 tieto/tiedott	14
2/3 ensykl/pääkirj	11
2/4 ensykl/kultt	13
2/5 ensykl/pakina	18
2/6 ensykl/tiedott	16
3/4 pääkirj/kultt	8
3/5 pääkirj/pakina	15
3/6 pääkirj/tiedott	12
4/5 kultt/pakina	14
4/6 kultt/tiedott	12
5/6 pakina/tiedott	18

Erojen jatkumonluonne on ilmeinen, mutta silti voi esittää muutamia yleistyksiä. Eniten toistensa kaltaisia ovat tietokirjat ja pääkirjoitukset (merkitseviä eroja vain 7 muuttujan kohdalla), pääkirjoitukset ja kulttuuriartikkelit (8 eroa), tietokirjat ja ensyklopediat (9 eroa) sekä tietokirjat ja kulttuuriartikkelit (9 eroa). Ensyklopedioiden ja pääkirjoitusten ero on hieman suurempi (11 eroa eli 50 % testatuista muuttujista). Hyvin selvästi näkyy pakinoiden omaleimaisuus. Suurimmat erot syntyvät vertailtaessa pakintekstejä muihin: pakinat ja ensyklopediat (18 eroa), pakinat ja tiedottavat artikkelit (18 eroa), pakinat ja tietokirjat (17 eroa), pakinat ja pääkirjoitukset (15 eroa) sekä pakinat ja kulttuuriartikkelit (14 eroa).

Tiedottavat artikkelit, jotka on poimittu sekä sanoma- että viikkolehdistä (ja jotka eivät käsittele poliittisia vaan "kevyempiä" aiheita), sijaitsevat pakinoiden ja tietokirjojen, ensyklopedioiden, pääkirjoitusten sekä kulttuuriartikkeleiden välissä. Niiden eroja on tietokirjoihin nähden 14, ensyklopedioihin 16, kulttuuriartikkeleihin 12 ja pakinoihin 18. Tiedottavat artikkelit (siten määriteltynä kuin edellä on tehty) kuuluvat siis selvästi aineiston neutraaleimpaan ytimeen, ja väitteemme koko korpuksemme edustavuudesta saa lisäpontta. Huomattakoon, että alatekstilajeista juuri tiedottavia artikkeleita on eniten eli runsas kolmannes (3628 lausetta).

Sentapainen yksinkertainen menetelmä kuin äsken on esitetty tarjoaa mahdollisuuden tekstilajien (tyylien) kvantitatiiviseen vertailuun melko objektiivisella tavalla. Mutta tällaiset vertailut jäävät kuitenkin ulkokohtaisiksi, ellei pystytä esittämään tulkintaa tai selitystä sille, miksi juuri tiettyjen piirteiden ja rakenteiden yleisyysuhteet vaihtelevat tyylistä toiseen. Seuraava taulukko on myös rakennettu taulukon 7 pohjalle. Taulukko 9 osoittaa, miten suuressa määrin testatut 22 muuttujaa erottelevat alatekstilajeja toisistaan. On tehty kaiken kaikkiaan 15 kahden alatekstilajin välistä vertailua (vrt. taulukkoihin 7,8), ja luvut ilmoittavat, monessako vertailussa kyseinen muuttuja merkitsevästi erottelee verratut kaksi alatekstilajia toisistaan 1 %:n riskitasolla.

TAULUKKO 9. *Muuttujien tyylillinen erottelukyky*

MUUTTUJA	MERKITSEVIEN EROTTELUJEN LUKUMÄÄRÄ
8 lausetyyppi	15
9 lausemuoto	10
14 kielto	8
16 lauseen pituus sanoina	10
17 lauseen pituus konstituentteina	10
18 pintajärjestys	11
19 subjektin rakenne	12
20 subjektin sija ja kvanttori	7
23 subjektin tekstuaalinen sidonnaisuus	13
24 subjektin sanamäärä	8
25 verbin rakenne	10
27 objektin rakenne	13
28 objektin sija ja kvanttori	12
31 objektin tekst. sidonnaisuus	14
32 objektin sanamäärä	11
33 predikatiivin rakenne	1
34 predikatiivin sija	3
37 predikatiivin sanamäärä	2
38 alkuadverbiaalin rakenne	7
45 loppuadverbiaalin rakenne	10
52 adverbiaalien määrä lauseessa	1
54 topikaalistettu alkulauseke	5

Taulukosta 9 voidaan tehdä muutamia päätelmiä. Testatuista muuttujista lausetyyppi (8) näyttää olevan tehokkain tyyli-indikaattori, koska se erottelee kaikki alatekstilajit merkitsevästi toisistaan. Jotta näiden erojen luonne kävisi selvästi ilmi, esitämme lausetyyppien frekvenssit kaikissa alatekstilajeissa (taulukko 10).

TAULUKKO 10. *Lausetyypit eri alatekstilajeissa (8)*

ALAKAT.	TIETO	ENSYKL	PÄÄKIRJ	KULTTUURI	PAKINA	TIEDOTT
	%	%	%	%	%	%
päälause	643 51	603 57	535 56	857 55	779 46	1919 53
konj. alk. päälause	133 10	135 13	85 9	151 10	195 12	400 11
sitaatti	3 0	- -	2 0	39 3	106 6	206 6
rel.lause subj:ssa	54 4	40 4	20 2	65 4	60 4	97 3
rel.lause obj:ssa	54 4	29 3	18 2	41 3	39 2	63 2
rel.l. muussa NP:ssä	70 6	95 9	43 4	69 4	68 4	103 3
että-lause	117 9	37 4	118 12	127 8	202 12	290 8
epäsuora kysymys	17 1	8 1	20 2	34 2	52 3	61 2
adverbiaalilause	85 7	29 3	58 6	94 6	124 7	286 8
komp.lause	13 1	9 1	10 1	19 1	18 1	48 1
partisiippirakenne	29 2	8 1	17 2	17 1	16 1	50 1
temporaalirakenne	29 2	46 4	15 2	17 1	17 1	48 1
väär. käyt. mod.rak.	5 0	7 1	1 0	4 0	- -	3 0
subj. inf.rakenne	17 1	7 1	14 1	23 1	14 1	46 1

Mainitsemme vain muutamia kaikkein silmään pistävimpiä eroja. Pakinoissa on selvästi vähiten pelkkiä päälauseita ja eniten *että*-lauseita sekä sitaatteja. Pääkirjoitusten kieli on rakenteellisesti yksinkertaisempaa kuin muiden alatekstilajien, siinä on suhteellisen vähän relatiivilauseita ja lauseenvastikkeita, mutta esim. 3 kertaa enemmän *että*-lauseita kuin ensyklopedioissa. Lauseenvastikkeita esiintyy vähiten kulttuuri- ja tiedottavissa artikkeleissa sekä pakinoissa, enemmän muodollisemmassa proosassa kuten tieto- ja ensyklopedisessä tekstissä. Tämänkaltaisten erojen lingvistinenkin relevanssi tuntuu kiistattomalta, ja viime kädessä erot olisi voitava selittää viestintätilanteesta sekä tekstin funktiosta ja muodollisuusasteesta käsin, vrt. esim. Anward (1976).

Yllättävältä tuntuu kenties se, että tekstuaaliset sidoskeinotkin (sekä subjektin että objektin) ovat hyvin tehokkaita tyyli-indikaattoreita. Näistäkin eroista ainakin osa selittyy tekstien erilaisista funktioista käsin, esim. miksi ensyklopedioissa on subjektin kohdalla enemmän toistoja (36 %) kuin pääkirjoituksissa (26 %) ja pakinoissa (18 %) ja miksi pakinoissa on enemmän eksoforisia ilmauksia (3 %) kuin muissa alatekstilajeissa (yleensä 1 % tai vähemmän).

Luonnollista on, että subjektin ja objektin konstituenttirakenne erottelee alatekstilajit toisistaan. Mutkikas lauserakenne (muuttuja 8) ja mutkikas konstituenttirakenne (mm. muuttujat 19,27), esim. jälkiasemainen infinitiiviattribuutti taikka useampijäseninen etuasemainen attribuuttiketju, kulkevat tavallisesti kädekkäin. Viestinnällisesti voidaan jälleen ymmärtää esim. miksi usein subjektiivisesti arvioivien kulttuuriartikkeleiden subjekti-NP:ssä on enemmän etuattribuutteja (27 %) kuin muissa alatekstilajeissa (yleensä 20-23 %), tai miksi tieto- ja ensyklopedisessä tekstissä on huomattavasti enemmän partisiippiattribuutteja (6 %) kuin pakinoissa (1 %). Muodolliset alatekstilajit sisältävät aina enemmän raskaita rakenteellisesti kompleksisia lausekkeita.

Hieman yllättävältä tuntune se, että subjektin ja etenkin objektin sijanmerkintäsuhteet muuttujat 20,28) erottelevat merkittävästi niinkin usein kuin 7 ja 12 kertaa (edelleen taulukko 9). A priori odottaisi, että näin vähäjäseniset paradigmat joissa on vain muutamia hyvin yleisiä sijoja eivät olisi tehokkaita tyyli-indikaattoreita. Tässä yhteydessä esittelemme tilanteen vain objektin sijanmerkinnän osalta.

#### TAULUKKO 11. Objektin sijanmerkintä eri alatekstilajeissa (28)

	TIETO	ENSYKL	PÄÄKIRJ	KULTTUURI	PAKINA	TIEDOTT
	%	%	%	%	%	%
nominatiivi	98 17	72 16	72 16	110 16	97 13	312 19
partitiivi	280 48	209 47	234 52	310 45	306 40	672 41
genetiivi	95 16	109 25	66 15	127 18	128 17	193 12
ym.						

Tässä on odottamattoman suuria suuria tekstilajikohtaisia eroja. Ensyklopedisen tekstin genetiiviobjekti (-n) on huomattavan paljon yleisempi kuin muiden alatekstilajien (25 % vs. 16-18 %). Toinen merkille pantava ero on pakinoiden ja tiedottavien tekstien muita alhaisempi partitiiviobjektien osuus (40-41 % vs. 45-52 %). Kuten taulukosta 7 on helppo nähdä muuttujan (28) kohdalta, juuri ensyklopediat ja pakinat ovat toisena tekstilajina melkein jokaisessa merkittävässä vertailuerossa. Näiden sijanmerkintäerojen tilastollisen merkittävyyden

*lingvistinen selittäminen* ei silti ole yksioikoista; viimeksi mainittu ero ei johdu esim. siitä, että pakina- ja tiedottavassa tekstissä olisi vähemmän kieltolauseita ja siksi myös vähemmän part.objekteja.

Taulukko 9 osoittaa myös, että testatuista muuttujista vähiten erottelevia tyyli-indikaattoreita ovat predikatiiviin liittyvät muuttujat (sekä konstituenttirakenne että sijanmerkintä ja sanamäärä, (33,34,37)) ja yhtä yllättävästi adverbialien lukumäärä lauseessa (52). Esimerkkinä alatekstilajien tilastollisesti tasaisesta jakaumasta esitämme adverbialien määrän ja predikatiivin konstituenttirakenteen.

TAULUKKO 12. *Adverbialien määrä lauseessa eri alatekstilajeissa (52)*

	TIETO		ENSYKL		PÄÄKIRJ		KULTTUURI		PAKINA		TIEDOTT	
	%		%		%		%		%		%	
1 advl.	551	56	429	56	424	59	669	59	711	61	1538	59
2	318	32	253	33	224	31	359	32	364	31	793	31
3	91	9	71	9	57	8	85	8	89	8	223	9
4	20	2	10	1	10	1	14	1	7	1	38	1
5	5	1	-	-	-	-	-	-	-	-	6	0
6	1	0	-	-	-	-	1	0	3	0	1	0

Prosenttiluvut ovat todellakin melkein identtiset. Ainut 1 %:n tasolla merkitsevä ero koskee tieto- ja pakinatekstejä (taulukko 7, muuttuja 52). Tämä on hyvä osoitus tilastollisen ja lingvistisen merkitsevyyden eritasoisuudesta. On vaikea keksiä lingvistiksi perusteltua syytä, joka selittäisi, miksi juuri tietotekstit ja pakinat eroaisivat toisistaan toisella tapaa kuin esim. ensyklopediat ja pakinat - ovathan ensyklopedioiden prosenttiluvut melkein tarkalleen samat kuin tietotekstien.

TAULUKKO 13. *Predikatiivin sija ja kvanttori eri alatekstilajeissa (34)*

	TIETO		ENSYKL		PÄÄK.		KULT.		PAKINA		TIEDOTT.	
	%		%		%		%		%		%	
nominatiivi	73	35	78	41	74	41	159	54	122	50	329	51
partitiivi	86	41	65	34	71	39	92	31	86	35	192	30
genetiivi	3	1	-	-	1	1	1	0	-	-	1	0
essiivi	22	10	27	14	13	7	26	9	9	4	47	7
translatiivi	15	7	12	6	8	4	11	4	18	7	38	6
KVANT + N	7	3	8	4	12	7	-	-	6	2	35	5
irrelevantti	4	2	1	1	1	1	4	1	1	1	8	1

Tässä muuttujassa voidaan havaita melkoisiaakin alatekstilajien välisiä eroja, mutta taulukon 7 osoittaman mukaan tilastollisesti merkitseviä ovat vain tietotekstien erot kulttuuri-, pakina- ja tiedottaviin artikkeleihin. Jälkimmäisissä on huomattavasti enemmän nominatiivipredikatiiveja ja jonkin verran vähemmän partitiivipredikatiiveja. Ensyklopediat ja pääkirjoitukset ovat tässä suhteessa tietotekstien kaltaisia. Tämän eron lingvistinen selittäminen ei ole helppoa - miksi tietoteksteissä on vähemmän nominatiivipredikatiiveja vaikka niitä odottaisi olevan enemmän,

koska tietotekstile ominaisten määritelmissä käytettävien identiteettipredikatiivien sija on juuri nominatiivi? Ymmärrettävämpää on sen sijaan se, että tietotekstissä on enemmän essiivinmuotoisia predikatiiviadverbiaaleja, koska nämä ovat lauseenvastikkeenomaisia ja siksi vähemmän omiaan kevyempiin tyyliin.

## 4. MUUTTUJIEN FREKVENSSIT JA NIIDEN KORRELAATIOITA

### 4.1. Lauseen perussyntaksia ja -semantiikkaa koskevat muuttujat

Seuraavassa tarkastellaan eräitä lauseen tekstuaalista käyttäytymistä kuvaavia tuloksia, joita on saatu edellä esitellyistä muuttujista joko suoraan taikka tietoja edelleen yhdistelemällä taikka ristiintaulukoimalla. Esitys ei ole tyhjentävä, sillä tässä vaiheessa ei teknisistä syistä ole saatu käytettäväksi kaikkia mahdollisia kombinaatioita. Monet tuloksista saavat lopullisen mielenkiintonsa vasta kun niitä voidaan vertailla muista korpuksista saatuihin tuloksiin. Tällöin selviää nimittäin viime kädessä se, missä määrin nyt tutkitut ominaisuudet ovat yleispäteviä suomen kielen piirteitä, mitkä ehkä yleistettävissä yli kielirajankin, mitkä puolestaan osoittavat vain korpuksenamme olleen asiaproosan ominaisuuksia.

*Lauseen virkeasema.* Lauseen virkeasemaa ja tyyppiä koskevia tietoja on koottu kahteen muuttujaan (6) ja (8). Näistä edellinen paitsi erottaa toisistaan pää- ja sivulauseet, selvittää myös eräitä päälauseen ominaisuuksia: muodostaako sen yksi lause, onko päälause virkkeen alussa, keskellä vai lopussa. Tämä muuttuja on tuonnempana runsaassa käytössä taustamuuttujana, koska halutaan selvittää missä määrin virkeasema säätelee muita lauseessa mahdollisia tekstuaalisia valintoja. Muuttuja (8) puolestaan kokoaa tarkempia tietoja sivulauseityypeistä.

TAULUKKO 14. *Lauseen paikka virkkeessä* (6)

	%	(% sivul.)
päälause yksin t. ensimmäisenä	3475	34
päälause rinnasteisena alussa	1191	12
päälause keskellä	776	8
päälause lopussa	1213	12
sivulause päälauseensa edellä	489	5 (14)
sivulause päälauseensa sisällä	179	2 (5)
sivulause päälauseensa jäljessä	2832	28 (81)
<hr/>		
yht.	10149	

Taulukosta käy ilmi, että asiateksteissä päälauseita on n. 2/3 lauseista ja että vallitseva virkekonstruktio on sellainen, jossa päälause alkaa virkkeen (kategoriat 1 ja 2 = 46 %). Se, että sivulauseista vain viidennes sijaitsee muualla kuin päälauseensa jäljessä, riippuu oletettavasti tekstilajista. Sekä esitutkimuksessamme olleessa kertomustekstissä että Mirja Saaren (1975:37) ruotsin puhekielen tutkimuksessa päälauseiden määrä oli 80 %. Oletettavaa on edelleen, että kertovissa teksteissä olisi runsaammin sivulauseita päälauseen edellä, koska sivulauseet ovat useammin temporaalisia kuin tiedottavassa tekstissä.

TAULUKKO 15. *Lausetyyppi* (8)

		%	(% sivul.)
päälause	5336	53	
konj. alk. päälause	1099	11	
sitaatti	356	4	
relatiivilause	1028	10	(31)
<i>että</i> -lause	891	9	(27)
epäsuora kysymys	192	2	(6)
adverbiaalilause	676	7	(20)
ei-finiittiset "lauseet"	450	4	(13)
vertailulauseet	117	1	(4)
yht. 10145			

Sivulauseiden keskinäinen jakauma lienee tekstilajikohtainen. Taulukossa 15 on ilmään pistävää relatiivilauseiden runsaus ja adverbiaalilauseiden vähäisyys. Käsitteellisen kielen merkinä voi pitää paitsi relatiivilauseiden runsautta myös epäsuorien kysymysten ja ei-finiittisten lauseenvastikkeiden tuntuva määrä. Rinnasteisia on päälauseista noin kuudennes; tämän tiedon merkitys selvenee kun saadaan vertailukohteita toisista tekstilajeista. Seuraavaksi katsotaan, miten eri lausetyypit sijoittuvat virkeasemiin. Tämä selviää ristiintaulukoimalla muuttujat (6) ja (8). Katsokaamme ensin päälauseita.

TAULUKKO 16a. *Virkeasema ja päälausetyyppi* (6; 8:1-4)

	pelkkä pl.		konj. alk. pl.		sit.		yht.
		%		%		%	
pl. yksin	3294	62	114	10	66	38	3474
pl. rinn:n alussa	1111	20	32	3	35	20	1166
pl. keskellä	420	8	306	28	38	22	764
pl. lopussa	511	10	647	59	36	20	1194
yht. 5336			1099		175		

Kiinnostavin luku taulukossa 16a on se että peräti 13 % konjunktiolla alkavista päälauseista on virkkeenalkuisia, käytetään siis rinnastuskonjunktioita (*mutta, ja*) kytkemään virkettä johonkin edellä olevaan tekstijaksoon (vrt. esim. Teleman (1974)) eikä virkkeensisäisiä liittimiä.

TAULUKKO 16b. *Virkeasema ja sivulausetyyppi (6; 8:4-E)*

	rel.s		rel.o		rel.m		<i>että</i>		e-kys.		advl.		(jatkuu)
	%		%		%		%		%		%		
pl:n ed.	6	9	0	0	0	0	1	0	5	3	261	39	
pl:n sis.	63	19	11	5	24	5	13	2	7	4	2	3	
pl:n jälj.	241	72	213	87	390	7	787	88	155	81	363	54	
lopussa	26	8	20	8	34	8	94	11	25	13	30	4	
yht.	336		244		448		895		192		676		
	partis.		temp.		mod.		inf.subj.						
	%		%		%		%						
pl:n ed.	0	0	63	37	1	5	0	0					
pl:n sis.	0	0	30	18	0	0	2	2					
pl:n jälj.	131	96	72	42	16	80	109	90					
lopussa	6	4	7	4	3	15	10						
yht.	137		172		20		121						

Tätä taulukkoa varten jätettiin kolme relatiivilauseiden ryhmää erilleen. Subjektissa olevista relatiivilauseista (rel.s ) peräti 28 % jäi virkkeen alkuun tai keskelle, kun sen sijaan vain 5 % objektissa (rel.o ) tai muissa lauseenjäsenissä (rel.m) olevista relatiivilauseista oli edes päälauseen sisällä. Kiinnostava on myös tieto, että vain 8 % kustakin relatiivilausekategorista esiintyi virkkeen absoluuttisessa lopussa. Lähes aina virke on jatkunut siitakin huolimatta, että relatiivilause on jo heittänyt päälauseensa jälkeen. Vaikeasti tulkittavat päälauseensa edellä sijaitsevat *että*- ja kysymyslauseet ovat nekin tässä tekstilajissa harvinaisuuksia; luontevasti päälauseen edellä sijaitsevat vain adverbiaalilauseet ja temporaaliset lauseenvastikkeet.

*Lauseen modaalisuus.* Modaalisuutta on kooditettu muuttujiin (9: lausemuoto), (10: leksikaalis-morfologinen modaalisuus), (11: tekstifunktio) ja (14: kielto). Tarkastellaan ensin syntaktista ja morfologis-leksikaalista modusta.

TAULUKKO 17. *Lausemuoto (9)*

		%
väitelause	9772	96
-kO-kysymys	121	1
hakukysymys	200	2
käsky	41	0
huudahdus	15	0
yht.	10149	

TAULUKKO 18. *Tekstifunktio* (11)

		%
tiedotus, toteamus, kuvaus	4078	40
selitys, tarkennus, esimerkki	957	9
päätelmä	102	1
arvio, arvoarvostelma	1303	13
käsky, kierto, ohje	125	1
vetoamus, toivomus	185	2
kysymys	97	1
vastaus, responsi	64	1
metateksti	98	1
i	3140	31
yht.		10149

TAULUKKO 19. *Modaalisuus* (10)

		%
verbimodus	391	29
modaaliverbi	768	58
adverbi	143	11
predikaatti	27	2
yht.		1329

TAULUKKO 20. *Lausemuoto ja modaalisuus* (9,10)

	v-modus		mod-verbi		adverbi		pred.		yht.
		%		%		%		%	
väite	368	29	730	58	141	11	26	2	1265
kysymys	17	29	38	66	2	3	1	2	58
käsky	5	100	-	-	-	-	-	-	-
huudahdus	1	100	-	-	-	-	-	-	-

Syntaktista modusta, lausemuotoa vaihdellaan vähän sellaisessa asiatekstissä, jossa dialogia on tuskin lainkaan. Huomattakoon, että muodoltaan kysymyksiä on aineistossa 321, mutta koodittaja on pitänyt näistä aitoina, funktioltaan kysymyksiä vain 97: tavallisemmin interrogatiivista lausetta käytetäänkin retorisenä kysymyksenä, jonka tekstifunktiona on vetoamus, toivomus taikka mahdollisesti käsky. Vaikka tekstifunktion määrittämistä voikin pitää vain alustavanluonteisena, se vaikuttaa ainakin sikäli luotettavalta, että irrelevantiksi tämä on osoittautunut n. 1/3:n lauseista kohdalla. Toisin sanoen lähinnä päälauseelle on määritettävissä itsenäinen tekstifunktio. Oireellisenä tekstilajin kannalta voinee pitää sitä, että "arvioiden" luokka on toteavien lauseiden jälkeen toiseksi yleisin ja että "selityksiä" on niin ikään runsaasti. Sen sijaan päätelmiä, joita odotettaisiin joko tieteellisessä tai argumentoivassa tekstissä olevan runsaasti, on tiedottavassa tekstissä vähänlaisesti. Lausemuotoisen metatekstin vähäisyys on ehkä odottamaton - sen tarve saattaa olla suurempi vaikeatajuisemmissa teksteissä kuin nämä ovat olleet.

Leksikaalis-morfologisin keinoin ilmaistu modaalisuus on n. neljä kertaa yleisempää kuin syntaktinen modaalisuus. Yleisin keino on modaaliverbin käyttö (53 %), mitä tosin ei esiinny kuin väitteissä ja kysymyksissä. Verbimodus, toiseksi yleisin käyttö, esiintyy kaikissa lausetyypeissä, kun taas adverbilla ilmaistu modaalisuus on lähes tyystin väitelauseen keino (vain 11 %). - Muut tarkemmat ristiintaulukoinnit puuttuvat toistaiseksi. Tarkastelkaamme seuraavaksi kieltoa.

TAULUKKO 21. *Kielto* (14)

		%
myönteinen	9222	91
kieltosanakielto	904	9
<i>tuskin</i> yms.	23	0

---

yht. 10149

TAULUKKO 22. *Kielto ja lausemuoto* (14,9)

	myönteinen		kieltosanakielto		yht.
		%		%	
väite	8864	91	908	9	9772
-kO-kysymys	108	89	13	11	121
hakukysymys	197	99	3	2	200
käsky	38	93	3	7	41
huudahdus	15	100	-	-	15

TAULUKKO 23. *Kielto ja lausetyyppi* (14,8)

	myönteinen		kieltosanakielto		yht.
		%		%	
pl. yht.	6178	91	613	9	6791
rel.s	298	89	38	11	336
rel.o	235	96	9	4	244
rel.m	413	92	35	8	448
<i>että</i>	744	84	147	17	891
e-kys.	187	97	5	3	192
advl.	600	89	75	11	676
muut	563	99	4	1	567

---

sivulauseet yht. 3354

Kieltolauseiden osuus liikkuu 10 %:n molemmin puolin, vrt. esim. Savijärvi (1977). Havaitsemme, että tähän lukemaan pääsevät modaalisisista lausetyypeistä väitelause ja joko/tai -kysymys, kun sen sijaan hakukysymyksissä (*mikä-* ja *kuka-*kysymyksissä) ja käskyissä kielteisyyttä on vähäisempää. Pää- ja sivulauseiden välillä ei ole yleisesti ottaen eroja, mutta eri sivulauseissa kiellon määrä vaihtelee selvästi. Peräti 16,5 % *että*-lauseista on kielteisiä, kun taas

epäsuorissa kysymyslauseissa kielteisiä on vain 2,6 %. Lähimmäs päälauseiden arvoja tulevat adverbiaalilauseet. Relatiivilauseiden kesken on myös yllättäviä eroja: objektina olevissa relatiivilauseissa on alle 4 % kielteisiä, tieto jota on vaikea selittää.

Lopuksi esitämme tiedot puhekielessä yleisten pragmaattisten partikkeleiden käytöstä: aineistomme lauseista niitä esiintyy vain 5 %:ssa.

#### TAULUKKO 24. *Pragmaattiset elementit* (13)

		%
-hAn	88	15
verbi + -kin	149	26
liikkuva -kin alussa	135	23
liikkuva -kin lopussa	111	19
-pA	20	3
yhdistelmiä	22	4
<i>kyllä</i>	31	5
<i>sitä</i>	9	2
<i>toki</i>	16	3

---

yht. 581

Kirjakielissä yleisin on, kuten tunnettua, tuskin lainkaan modaalinen -kin, tosin näistä runsas kolmannes on verbiin liittyneinä eräänlaisia syntaktisen moduksen keinoja (myönnytyksenä: *Hän onkin siinä onnistunut*). Yhteensä 23 % partikkelitapauksista muodostuu vakuuttelua ja perustelua ilmentävistä keinoista, joita ovat -hAn, *kyllä* ja *toki*. Liitepartikkeli -pA lienee kirjoitetussa kielessä lähinnä argumentoivaan tyyliin sidoksissa, ja partikkelistunut *sitä* on korkeintaan pakinoissa esiintyvä puhekielimäisyys.

#### 4.2. *Konstituenttien rakenteita koskevat muuttujat*

Tässä jaksossa käsitellään subjektin, objektin, predikatiivin sekä alku- ja loppuadverbiaalin konstituenttirakennetta kuvaavia muuttujia (19,27,33,38,45) sekä subjektin, objektin ja predikatiivin sijaanmerkintää (20,28,34) ja predikatiivin lukukongruenssia (35). Lisäksi käsitellään adverbiaalien syntaktisia tyyppisiä (39,46) ja semanttisia lajeja (40,47), perustavaa lauserakennetta eli etenkin transitiivisuuden ja intransitiivisuuden suhteita (15) ynnä vielä verbiryhmien rakennetta (25) ja niiden kongruenssisuhteita (26).

Taulukossa 25 vertaillaan ensin subjektin ja objektin konstituenttirakennetta. Havainnollisuuden lisäämiseksi alakategoriat on ryhmitelty osittain uudelleen; näin menetellään jatkossakin.

TAULUKKO 25. *Subjektin ja objektin konstituenttirakenne* (19,27)

	subjekti	objekti
	%	%
N	3215 40 (47)	1520 33 (34)
määrite + N	2187 27 (32)	1457 32 (33)
N + app./advl.	381 5 (6)	236 5 (5)
N + lause	363 5 (5)	286 6 (6)
lause	99 1 (1)	599 13 (14)
infinitiivi	110 1 (2)	(vrt. muuttuja (25:3,4))
nominaalitus	208 3 (3)	137 3 (3)
rinnasteinen	303 4 (4)	182 4 (4)
inkorporoitu	403 5 -	(ei koske objektia)
geneerinen 0	160 2 -	32 1 -
ellipsi	593 7 -	128 3 -
yht.	8022	4577

Subjektin ja objektin lukumäärät eivät tällaisinaan ole suoraan verrannollisia. Subjektien yhteismäärään 8022 kuuluu 1056 inkorporoitua, geneeristä ja ellipsitapausta (tule+n; jos 0 tulee; ... ja 0 tulee), jotka olisi vähennettävä kun verrataan "materiaalisia" subjekteja ja objekteja. Vastaavia objektitapauksia on 160. Kun nämä vähennetään, saadaan subjektien määräksi 6866 ja objektien määräksi 4417. Sulkeissa olevat prosenttiluvut ovat suhteessa näihin kokonaismääriin, ja niitä käytetään jatkossa.

Subjektin ja objektin konstituenttirakenteiden frekvenssit eroavat huomattavasti toisistaan. Yleistendenssi on se, että subjekti on rakenteellisesti yksinkertaisempi (ja sanamäärältään lyhyempi, vrt. taulukkoa 50 jaksossa 4.3.). Tämä ilmenee erityisen selvästi määritteettömien substantiivien sekä lauseiden esiintymisfrekvensseissä. Pelkkiä substantiiveja on subjektina 47 %, objektina vain 34 %.

Attribuutillisia substantiivilausekkeita sen sijaan on subjektina ja objektina suurin piirtein yhtä paljon, etuattribuutillisia 33 % vs. 32 % ja jälkiattribuutillisia (lauseet mukaan lukien) 11 % kummassakin. Huomattava ero löytyy pelkkien lauseiden kohdalla, joita subjektina on vain 1 % mutta objektina peräti 14 %. Enemmistö objektilauseista on luonnollisestikin sanomis- ja mentaalisten verbien *että*-lauseobjekteja tai partisiippirakenteita.

Mainitut kaksi subjektin ja objektin eroa voidaan yhdistää informaatorakenteen tendensseihin sekä viskurilakiin. Subjektin tunnusmerkitön paikka on verbinetinen ja siksi se on kernaimmin rakenteellisesti yksinkertainen ja lyhyt, objektin suosikkipaikka taas on verbinjälkeinen ja siksi se voi helpommin olla kompleksinen ja pitempi.

Subjektin osalta on syytä kiinnittää huomiota 0-edustumien huomattavaan määrään, etenkin inkorporoituihin subjekteihin (5 %; esim. tule+n) ja geneerisiin 0-subjekteihin (2 %; esim. Jos 0 tulee, ...). Jos vielä ajattelee hyvin yleisiä subjektittomia passiivilauseita, käynee selväksi, että pintasubjektittomien lauseiden olemassaolo on tosiseikka, jolle tulisi suoda sille kuuluva asema suomen lauseopin kuvauksessa. Subjekti ei ole suomessa siinä määrin "lauseen toinen

pakollinen pääjäsen" kuin se on indoeur. kielissä, joskin se on primaarinen nominaalijäsen (vrt. Hakulinen & Karlsson (1979:165-)).

Taulukon 25 osalta on vielä muistutettava siitä, että olemme tulkinneet perinteiset inf.objektit verbiketjun osiksi eivätkä ne siis tule esiin muuttujassa (27) ollenkaan. Verbiketjujen lukumäärä käy ilmi taulukosta 43. Vrt. muuttujan (25, etenkin alakategorian 3) esittelyä jaksossa 2.5.

Subjektin ja objektin konstituenttirakenne ovat muuttujia, jotka ovat hyödyllisiä monissa ristiintaulukoinneissa. Taulukossa 26 osoitetaan, miten subjektin konsituenttirakenne ja sijainti ovat sidoksissa toisiinsa. Tämä taulukko täydentää lausekesubjektin sanamäärää ja sijaintia esittävää taulukkoa 51. Sanajärjestysmuuttuja (18) on pelkistetty kahdeksi peruskategoriaksi, SV ja VS, ja subjektin konstituenttirakenteessakin (19) on suoritettu eräitä uudelleen ryhmittelyjä.

TAULUKKO 26. *Subjektin konstituenttirakenne ja sijainti (19,18)*

	SV		VS		i	
		%		%		%
N	2844	89	352	11	13	0
määrite + N	1541	71	599	27	46	2
N + app./advl.	227	60	147	39	7	2
rinnasteinen	158	52	131	43	14	5
(N +) S, inf., nom.	236	30	521	67	18	2
	<hr/>					
yht.	6000	75	1763	22	219	3

Subjektin rakenteen vaikutus sen sijaintiin on aivan ilmeinen: rakenteellinen kompleksisuus lisää viskurilain mukaisesti subjektin alttiutta esiintyä verbinjälkeisessä asemassa. Määritteettömistä NP-subjekteista peräti 89 % on verbinetisiä, lausemaisista (N + S, S, infinitiivit ja nominaalistukset) enää vain 30 %, ja tähän väliin sijoittuvat laskevassa järjestyksessä etuattribuutilliset, jälkiattribuutilliset ja rinnasteiset subjektit. Lause- ja infinitiivisubjektien kohdalla voi jo melkein puhua verbinjälkeisyyden säännöstä, tai vähintään voimakkaasta tendenssistä.

Toinen päätelmä voidaan tehdä SV- ja VS-järjestysten kokonaisluvuista: peräti neljännes kaikista subjekteista on verbinjälkeisiä. Väitteeseen, että suomen perussanajärjestys on SVX, on siis tehtävä eräitä tärkeitä täsmennyksiä. Subjektin sijoittumiseen verbin jälkeen vaikuttavat paitsi viskurilaki myös tekstuaaliset siirrot; vrt. Hakulinen & Karlsson (1979:305-) sekä jaksoa 4.4.

Seuraavaksi tarkastellaan predikatiivin sijan määräytymistä subjektin rakenteen mukaan.

TAULUKKO 27. *Subjektin konstituenttirakenne ja predikatiivin sija (19,34)*

	nominatiivi		partitiivi		yht.
		%		%	
N	322	61	203	39	525
määrite + N	219	64	119	35	338
N + app./advl.	47	68	22	32	69
rinnasteinen	14	27	38	73	52
lause	21	32	45	68	66
N + lause	34	53	30	47	64
infinitiivi	35	44	45	56	80
nominaalistus	31	48	33	52	64

Kun substantiivilauseke on subjektina, predikatiivin sijana on useammin nominatiivi kuin partitiivi (61 - 68 % vs. 39 - 32 %). Nominatiivi on harvinaisempi, kun subjektina on N + S - rakenne, mutta silti yleisempi kuin partitiivi (53 % vs. 47 %). Partitiivipredikatiivi dominoi (52 - 56 %), kun subjektina on nominaalistettu rakenne tai infinitiivi, ja partitiivin yliote käy vielä selvemäksi lausesubjektien (68 %) ja rinnastettujen subjektien (73 %) kohdalla.

TAULUKKO 28. *Predikatiivin konstituenttirakenne (33)*

		%
määrite + N, N + määrite	549	31
A	471	27
määrite + A, A + määrite	286	16
N	181	10
N + lause	102	6
A + A	49	3
A + lause	37	2
nominaalistus	27	2
N + N	32	2
vertailu ( <i>kuin x</i> )	13	1
ellipsi	17	1

---

yht. 1767

Predikatiivin konstituenttirakenne poikkeaa selvästi subjektista yhtä hyvin kuin objektistakin. Ensi silmäyksellä yllättävää on se, että (toisin kuin subjektissa ja objektissa) yleisin konstituenttirakenne ei ole pelkkä substantiivi tai adjektiivi. Pelkkiä adjektiiveja on tosin 27 % ja substantiiveja 10 %, yhteensä 37 %. Näitä minimirakenteita huomattavasti enemmän on kuitenkin määritteellisiä konstituentteja, määritteellisiä substantiivirakenteita 31 % ja määritteellisiä adjektiivijonoja 16 %, yhteensä 47 %. Predikatiivin suosikkirakenne on toisaalta "Tämä on hyvä auto" (31 %), toisaalta "Tämä auto on hyvä" (27 %). Predikatiivin suhteellinen pidemmys subjektiin ja objektiin verrattuna on ymmärrettävissä, kun ajattelee predikatiivin funktiota. Useimmat predikatiivithan ovat subjektin tarkoitetta luonnehtivia

luokituspredikatiiveja, ja luonnehtiminen edellyttää melkein määritelmänomaisesti attribuuttirakenteen käyttöä.

Substantiivi- ja adjektiivipredikatiivit näyttävät käytännössä olevan yhtä yleisiä. Substantiivirakenteiden osuus on 49 % (määrite + N, N + määrite, N, N + S, N + N), adjektiivirakenteiden 48 % (A, määrite + A, A + määrite, A + A, A + S).

Seuraavaksi käsitellään subjektin ja objektin sijanmerkintää. Nämä muuttujat on ryhmitelty alakategorioiksi siten, että tietyt kvanttorin sisältävät tapaukset (esim. *kaksi* naista; *paljon* kahvia; *kaikki* naiset; *eräs* mies) on sijaan katsomatta kooditettu erilleen kvanttorittomista subjekteista. Nämä kvanttorin sisältävät subjektit ja objektit on jätetty pois taulukosta 29: tällaisia subjekteja oli 368 ja objekteja 251. Tuskin on syytä olettaa, että sijojen jakauma on näissä ratkaisevasti erilainen kuin kvanttorittomissa lausekkeissa, ja siksi niiden poisjätö ei vaikuttane taulukon lukujen luotettavuuteen.

Lohjenneet partitiivisubjektit (N 86) ja objektit (N 38) on sisällytetty ao. partitiivien osuuksiin. Nominatiivilla tarkoitamme objektin kohdalla päätetöntä akkusatiivia (sekä yksikössä että monikossa), genetiivillä *-n*-päätteistä akkusatiivia.

TAULUKKO 29. *Subjektin ja objektin sija* (20,28)

	subjekti		objekti	
		%		%
nominatiivi	5586	89	761	21
genetiivi	241	4	718	20
<i>-t</i> -akk.	-	-	29	1
partitiivi	466	7	2049	58
	yht. 6293		3557	

Kieliopillisten sijojen frekvenssijakauma on subjektissa aivan toisenlainen kuin objektissa. Keskeisin yleistys on se, että nominatiivi on subjektin suosikkisija (89 %), partitiivi objektin (58 %). Näiden lukujen valossa on ilmeisen harhaanjohtavaa sanoa (niin kuin monissa kieliopeissa te#hdään), että subjektin sijat ovat nominatiivi ja partitiivi, objektin sijat akkusatiivi ja partitiivi. Partitiivin sekundaarisuus subjektin sijana olisi tuotava selvästi julki, samaten sen ensisijaisuus objektin sijana.

Subjektin sijojen frekvenssijakauma osoittaa senkin, että ns. eksistentiaalilauseet eivät ole suomen kielen rakenteessa niin keskeinen kategoria kuin usein on annettu ymmärtää. Onhan part.subjektien osuus lausekesubjekteista vain 7 %, ja juuri part.subjektisia lauseitahan on pidetty e-lauseiden prototyypinä; vrt. Ikola (1964) ja Hakanen (1972) sekä myös Hakulinen & Karlsson (1979:95-).

Päätteetön ja *-n*-päätteinen akkusatiivi ovat melkein yhtä yleisiä, 21 % ja 20 %. *-t*-akkusatiivi on odotetusti harvinainen, 1 %.

Taulukko 30 osoittaa, miten subjektin sijat ja perustavat sanajärjestystyyppit korreloivat keskenään. Kvanttorin sisältävät subjektit eivät taaskaan ole mukana; taulukon partitiivin osuudesta puuttuvat myös lohjenneet part. subjektit (20:8), joita emme tässä käsittele.

TAULUKKO 30. *Subjektin sija ja sijainti (20,18)*

	SV		VS		i		yht.
		%		%		%	
nominatiivi	4421	79	1092	20	68	1	5586
partitiivi	87	23	278	73	12	3	380
genetiivi	233	97	8	3	0		241

Nominatiivi- ja partitiivisubjektien tunnusmerkittömät paikat ovat tasan päinvastaiset. Nom.subjekti on yleensä verbinetinen (79 %), part.subjekti verbinjälkeinen (73 %). Mutta silti ei sovi unohtaa sitäkään, että kummassakin tapauksessa 20 - 25 % esiintymistä on siirtosääntöjen myötä joutunut tunnusmerkikseen paikkaan, nom.subj. verbin jälkeen ja part.subj sen eteen. Nom.subjekteista 20 % on sirretty loppuun joko viskurilain vaatimusten taikka subjektin tarkoitteen uutuuden takia; vrt. Hakulinen & Karlsson (1979:306-). Tällaisissa lauseissa on lisäksi usein topikaalistettu jokin muu ydinkonstituentti, yleensä objekti.

Part.subjekteista 23 % on topikaalistettu verbinetiseen asemaan, yleensä siksi, että subjektina on anaforinen pronomini (*Niitä* on täälläkin) taikka siksi, että siinä on toinen tekstuaalinen sidoskeino, esim. anaforinen attribuutti tai possessiivisuffiksi. Olettamamme genetiivisubjekti on melkein järjestään verbinetinen.

TAULUKKO 31. *Kielto ja subjektin sija (15,20)*

	nominatiivi		partitiivi		genetiivi	
		%		%		%
myönt.	5178	93	349	75	237	98
kieltosana	402	7	116	25	4	2
<i>tuskin</i>	6	0	1	0	-	-
yht.	5586		466		241	

Nominatiivi- ja partitiivisubjektien käyttäytyminen kiellon vaikutusalueella on odotettavasti melko erilaista, koska mm. trans.verbien subjektit ovat kiellon partitiivistamissäännön ulkopuolella. Part.subjekteista neljännes esiintyykin kieltolauseissa, nom.subjekteista vain 7 %. Olettamamme gen.subjektit suosivat selvästi myöntölauseita.

Näitä subjektin lukuja voi verrata objektin vastaaviin lukuihin, jotka ilmenevät taulukosta 32.

TAULUKKO 32. *Kielto ja objektin sija (15,28)*

	nominatiivi		genetiivi		-t-akkusatiivi		partitiivi		yht.
		%		%		%		%	
myönteinen	757	23	717	22	29	1	1740	54	3243
kieltosana	4	1	1	-	-	-	301	98	306
<i>tuskin</i>	-	-	-	-	-	-	6	100	6

Kieltolauseen part.objekteja on enemmän (N 301) kuin kieltolauseen part.subjekteja (N 116), vaikka subjekti muuten on selvästi yleisempi lauseenjäsen kuin objekti. Partitiivin kieltosääntö on siis nimenomaan objektin sääntö. Toisaalta on silmäänpistävä, että kielteisiä part.objekteja on vähän verrattuna myöntölauseiden irresultatiivista aspektia ja epämääräistä paljoutta ilmaiseviin partitiiveihin, joita on peräti 1740 eli runsaasti 5 kertaa enemmän kuin kieltopartitiiveja.

Taulukko 33 näyttää, miten subjektin sija suhtautuu sen tarkoitteen tunnettuuteen. Etukäteen voisi tietysti olettaa, että nominatiivisubjektien enemmistö on tavalla tai toisella tuttuja, part.subjektien enemmistö taas uusia, sikäli kuin e-lauseiden prototyypissä eli loppuasemaisen partitiivisubjektin sisältävässä lauseessa todella on kyse uuden tekstireferentin esittelemisestä. Kvanttorin sisältävät subjektit ja partitiivisubjektin lohkeamat eivät taulukkoon sisälly.

TAULUKKO 33. *Subjektin sija ja sen referentin tunnettuus (20,21)*

	mainittu		"heik."		asiain-		impli-		edelly-		uusi		i	yht.	
		%		%		%		%		%		%			
nom.	2303	41	655	12	315	6	873	16	279	5	923	17	238	4	5586
part.	30	8	80	21	7	2	72	19	4	1	137	36	50	13	380
gen.	116	48	39	16	11	5	37	15	12	5	18	8	8	3	241

On selviä eroja, vaikka ne eivät ole niin jyrkkiä kuin saattaisi odottaa. Nom. subjekteista peräti 41 % viittaa aikaisemmin mainittuun tarkoitteeseen, part. subjekteista vain 8 %. "Heikko" mainittuus eli samamerkityksisyys näyttää olevan leimallisemmin part.subjektin ominaisuus, 21 % vs. nom.subjektin 12 %. Odotuksenmukaista on, että part.subjektin tarkoite on paljon useammin uusi kuin nom.subjektin, 36 % vs. 17 %, mutta tämä ero ei ole kovin suuri. Part. subjekteista peräti 2/3 on eri tavoin tuttuja, mikä on ristiriidassa perinteisen e-lausedeorian kanssa, jonka mukaan e-lauseiden subjekti e i edellyttää tarkoitetta olemassa olevaksi tai tutuksi tai mainituksi. E-lauseiden tarkastelu konkreettisissa tekstiyhteyksissä osoittaa siis kuitenkin toista. Nom. subjektin tarkoite on kielenulkoisen tiedon perusteella tiedetty 5 kertaa yleisemmin kuin part.subjektin, 5 % vs. 1 %. Yleistävästi voi todeta, että nom.subjekteista on eri tavoin tuttuja (mainittuja, "heikosti" mainittuja, implikoituja tai edellytetyjä) 3/4, part.subjekteista puolet.

Taulukko 34 osoittaa, miten topikaalistus ja lausekeobjektin sija (28:1-4) suhtautuvat toisiinsa.

TAULUKKO 34. *Lausekeobjektin sija ja topikaalistus* (28:1-4; 54:1)

	topikaalistettu lausekeobjekti		kaikki objektit	
	%		%	
nominatiivi	247	32	761	21
genetiivi	79	10	718	20
-t-akkusatiivi	6	1	29	1
partitiivi	442	57	2049	58
yht.	774		3557	

Selvästi näkyy, miten objektin topikaalistus suosii päätteettömiä akkusatiiveja ("nominatiiveja"). Selitystä ei ole vaikea löytää: päätteettömän akkusatiivin suosinta johtuu subjektittomien passiivilauseiden topikaalistustarpeesta. Pinnan kanoninen tavoiterakenne "toivoo" lauseenalkuista nominaalista lauseketta, ja tällaiseksi joutuu passiivilauseissa hanakasti päätteettömässä akkusatiivissa oleva objekti. Tätä siirtosääntöä on nimitetty kompensatoriseksi tematisoinniksi, ks. Hakulinen & Karlsson (1979:304). Pass.lauseiden kompensatorinen tematisointi tapahtuu nimenomaan *-n*-päätteisen akkusatiivin kustannuksella; part.objektin topikaalistumisfrekvenssissä ei ole mainittavaa eroa koko objektien joukkoon verrattuna.

Siirrymme käsittelemään predikatiivin sijanmerkintää. Kooditusjärjestelmässämme myös predikatiiviadverbiaalit on otettu mukaan tähän muuttujaan (vrt. jaksoa 2.5.). Sulkeissa olevat prosenttiluvut nominatiivin, partitiivin ja genetiivin kohdalla ovat suhteessa perinnäisesti määriteltyn predikatiiviin, johon vain nämä kolme kuuluvat.

TAULUKKO 35. *Predikatiivin sija* (34)

		%
nominatiivi	835	47 (58)
partitiivi	592	34 (41)
genetiivi	6	0
essiivi	144	8
translatiivi	102	6
kvanttori (+ N)	68	4
i	20	1
yht.	1767	

Nominatiivi on predikatiivin perussija niin kuin se on subjektinkin. Partitiivi on huomattavasti yleisempi predikatiivin kuin subjektin sijana. Nominatiivipredikatiivin osuus on käytännössä hieman suurempikin kuin 47 (58) %, koska kvanttori (+ N) -alakategorian esiintymät ovat (melkein?) järjestään nominatiivinmuotoisia, esim. Nämä ovat *kolme* autoa. Genetiivipredikatiivin harvinaisuus on intuitiivisestikin tiedettävissä. Predikatiiviadverbiaaleista essiivinmuotoinen on yleisempi kuin translatiivinmuotoinen.

Subjektin konstituenttirakenteen heijastumista predikatiivin sijaanmerkintään on selvitelty taulukossa 27.

TAULUKKO 36. *Predikatiivin lukukongruenssi (35)*

SUBJEKTI	PREDIKATIIVI		%
yksikkö -	yksikkö	1188	67 (76)
yksikkö -	monikko	30	2 (2)
monikko -	yksikkö	45	3 (3)
monikko -	monikko	309	17 (20)
irrelevantti		195	11
yht.		1767	

Sulkeissa oleva prosenttiluvut on laskettu suhteessa predikatiivien lukumäärään 1572, josta irrelevantit tapaukset on jätetty pois. Odotuksenmukaisesti valtaenemmistössä tapauksia yksiköllistä subjektia seuraa yksiköllinen predikatiivi, 67 %. Monikollisia predikatiiveja monikollisen subjektin ohessa on 17 %. Monikko-yksikkö -yhdistelmä (*Nämä suunnitelmat ovat mainio ajatus*) on jonkin verran yleisempi kuin yksikkö-monikko -yhdistelmä (*Tämä on parhaita suunnitelmia*), 3 % vs. 2 %.

Adverbiaalien osalta kooditusjärjestelmämme rajoittuu lauseen alku- ja loppuadverbiaaleihin, keskiasemaisia adverbiaaleja ei analysoida tarkemmin. Käsittelemme seuraavaksi rinnakkain alku- ja loppuadverbiaalin konstituenttirakennetta, syntaktista tyyppiä (mm. kiinteysastetta suhteessa verbiin) sekä merkitystä.

TAULUKKO 37. *Alku- ja loppuadverbiaalin konstituenttirakenne (38 45)*

	alkuadverbiaali		loppuadverbiaali	
		%		%
yksi sana	1667	50	1094	25
lauseke	891	27	1497	35
PP	283	9	316	7
3. inf., nominaalitus	5	2	307	7
lauseenvastike	60	2	74	2
pääsana + lause	50	2	539	12
lause	215	7	320	7
rinnasteinen	28	1	163	4
ellipsi	52	2	17	0
yht.	3304		4327	

Konstituenttirakenteeltaan alku- ja loppuadverbiaali eroavat toisistaan monin tavoin. Alkuadverbiaaleista peräti puolet on lyhyitä, yhden sanan mittaisia, ja näistä huomattava osa on kehysadverbiaaleja tyyppiä *nyt, tänään, siellä*; vrt. taulukkoa 38. Loppuadverbiaaleista vain

25 % on yhden sanan pituisia. Toinen huomattava ero ilmenee alakategoriassa pääsana + lause, joita alussa on vain 2 % mutta lopussa 12 %; nämä adverbiaalit ovat tyyppiä *siellä, missä sitruunapuut kukkivat*. Tämäkin ero on edellisen tapaan viskurilain mukainen. Adverbiaalilauseiden esiintymisessä sen sijaan ei ole eroa alku- ja loppuaseman välillä, 7 % kummassakin tapauksessa. Viskurilain mukaista on sekin, että ei-minimaalisen lausekkeen muotoisia adverbiaaleja on enemmän lopussa kuin alussa, 35 % vs. 27 %. Jonkin verran yllättävää on PP-lausekkeiden alkuasemainen suosittuus, 9 % vs. 7 %. Verbisyntaksin kanoniset rakenteet riittävät selittämään, miksi lopussa on enemmän 3. infinitiivejä kuin alussa. Ellipsi on adverbiaaleissa (alussa 2 %, lopussa 0 %) harvinaisempaa kuin subjektissa (7 %) ja objektissa (3 %), vrt. taulukkoa 25.

TAULUKKO 38. Alku- ja loppuadverbiaalin tyyppi (39,46)

	alkuadverbiaali		loppuadverbiaali	
		%		%
valenssi	936	28	1788	41
kehys	765	23	374	9
kommentti	106	3	33	1
konnektiivi	272	8	13	0
pakollinen pred.advl.	36	1	206	5
irrationaalinen pred.advl.	52	2	123	3
muu	1085	33	1773	41
i	52	2	17	0
yht.		3304	4327	

Muuttujien (39,46) alakategoriajako ei voi pitää onnistuneena jo siitä syystä, että jäännöskategoriaan "muu" on joutunut liian suuri osa tapauksista. Valenssiadverbiaalien erottaminen vapaista määritteistä on usein makukysymys. Silti eräät suuntaukset ovat selviä. Kehysadverbiaalin suosikkipaikka on alussa, samaten konnektiivin. Nämä tosiseikat seuraavat viskurilaista ja tekstinrakentelun yleisperiaatteista. Predikatiiviadverbiaalit ovat paljon useammin lauseen lopussa kuin alussa. Valenssiadverbiaali käyttäytyy objektin tavoin. Normaalisti se on verbinjälkeinen ja nimenomaan loppuasemainen, mutta se voidaan myös helposti topikaalista tekstuaalisista tai muista syistä.

TAULUKKO 39. Alku- ja loppuadverbiaalin merkitys (40,47)

	alkuadverbiaali		loppuadverbiaali	
		%		%
aika	554	17	380	9
paikka	307	9	360	8
habitiivi	248	8	138	3
tapa	132	4	342	8
muu materiaallinen	394	12	657	15
OSMA	30	1	72	2
modaalinen	41	1	9	0
muu	1546	47	2352	54
i	52	2	17	0
yht.	3304		4327	

Muuttujia (40 47) on syytä arvostella samasta syystä kuin edellisiä. Niistä saa selville vain muutaman merkityskategorian frekvenssijakauman ja liian suuri osa tapauksista, noin puolet, jää jäännöskategoriaan "muu". Huomiota kiinnittää kuitenkin, että lukumääräisesti paikan adverbiaaleja on jokseenkin yhtä paljon alussa kuin lopussakin, kun sen sijaan ajan adverbiaalit ovat selvemmin kasautuneet alkuasemaan. Habitiivit ovat yllättävän yleisiä loppuasemassa; omistusrakenteen kiteytyneisyyden takia niitä odottaisi olevan enemmänkin alkuasemassa. Modaaliset adverbiaalit osuvat paljon harvemmin loppu- kuin alkuasemaan, OSMA:t taas suosivat loppuasemaa.

TAULUKKO 40. Lauserakenne (15)

		%
aktiivinen transitiivinen	3049	30
aktiivinen intransitiivinen	4754	47
passiivinen transitiivinen	1434	14
passiivinen intransitiivinen	228	2
omistusrakenne	232	2
<i>pojasta tuli mies</i>	38	0
sää, olotila	61	1
psykkinen kausatiivi	11	0
idiomi, kangistuma	90	1
i fragmentti	252	2
yht.	10149	

Aktiivisten intransitiivilauseiden osuus on noin puolet kaikista lauseista, jos ottaa huomioon, että harvinaisista alakategorioista (omistusrakenne, *pojasta tuli mies* -tyyppi eli tulorakenne, säälauseet jne.) useimmat ovat intransitiivilauseiden kiteytyneitä erikoistapa#uksia. On melkein itsestään selvää, että aktiivilauseet ovat passiivilauseita huomattavasti yleisempiä. Onhan

suomen passiivi yksi kieliopillinen persoona ja aktiivilauseiden kategoriaan sisältyy kolme persoona, ks. Hakulinen & Karlsson (1979:254-).

Mielenkiintoisempaa on todeta, että passiivin jakauma on aivan erilainen transitiivi- ja intransitiivilauseissa. Kahta aktiivista transitiivilauseetta kohden esiintyy yksi passiivinen trans.lause, kun sen sijaan kahtakymmentä aktiivista intrans.lauseetta kohden esiintyy yksi passiivinen intrans.lause. Akt. intrans.lauseita on puolitoista kertaa enemmän kuin akt. trans.-lauseita, mutta pass. trans.lauseita on kuusi kertaa enemmän kuin pass. intrans.-lauseita.

Passiivi suosii siis mitä ilmeisimmin trans.verbejä, ja tästä on pääteltävä, että suomen passiivin päätehtävä on sama kuin indoeur. passiivirakenteen yksi päätehtävä, nimittäin primaarisen subjektin poistuminen objektin tieltä silloin, kun objektin rankia halutaan korottaa.

Lauserakennemuuttuja (15) on hyödyllinen monissa ristiintaulukoinneissa. Taulukko 41 suhteuttaa akt. trans.- ja intrans.lauseet subjektin konstituenttirakenteeseen.

TAULUKKO 41. Aktiivisten lauseiden lauserakenne ja subjektin konstituenttirakenne (15:1,2; 19)

	rinn. lause %	pääs. + lause %	infin. %	nomin. %	N + määrite %	määrite + N %	(jatkuu)
akt.	91 3	65 2	8 0	2 0	58 2	76 3	641 22
trans- akt.	194 4	270 5	86 2	104 2	147 3	252 5	1393 30
intrans.							
	N %	inkorp. %	gen. 0 %	ellipsi %	yht.		
akt.	1293 44	268 9	116 4	305 10	2923		
trans. akt.	1834 39	134 3	42 1	257 6	4713		
intrans.							

On helppo nähdä, että akt. intrans.verbien subjekti on rakenteellisesti hieman kompleksisempi kuin akt. trans.verbien. Esim. pääsana + lause-, lause-, infinitiivi- ja nominaalistettujen subjektien osuudet ovat intransitiivisillä verbeillä 5 %, 2 %, 2 % ja 3 %, kun transitiiviverbien vastaavat luvut ovat 2 %, 0 %, 0 % ja 2 %. Erityisen selvänä ero ilmenee määrite + N -subjekteissa, jonkalaisia intrans.verbeihin liittyy 30 %, mutta trans.verbeihin vain 22 %. Toisaalta trans.verbeillä on enemmän lyhyitä subjekteja ja nollasubjekteja kuin intrans.verbeillä: pelkkä substantiivi 44 % vs. 39 %, verbiin inkorporoitu (syvä)subjekti 9 % vs. 3 %, geneerinen 0-subjekti 4 % vs. 1 % ja elliptinen subjekti 10 % vs. 6 %. Transitiivisuuden kompleksisuuden ja subjektin konstituenttirakenteen kompleksisuuden kielteinen korrelaatio ("yksinkertaisempi VP - kompleksisempi subjekti-NP") on luultavasti lähinnä osoitus siitä, että intransitiiviverbeihin liittyy enemmän abstraktisia primaariaktantteja, etenkin propositioita ilmaisevia, kun taas transitiiviverbien subjektit useammin ilmaisevat konkreettisia agenttiiviolioita.

Taulukosta 42 käy ilmi, miten keskeisimpien konstituenttien topikaalitus korreloi lauserakenteen perustavien alakategorioiden kanssa.

TAULUKKO 42. *Lauserakenne ja topikaalitus* (15:1-4; 54:1,2,4,6,8,C,D)

	lauseke- obj.		lause- obj.		predika- tiivinen		advl.		advl.+obj.		advl.+advl.		NP laus.- vast:sta	
	%		%		%		%		%		%		%	
akt. trans.	320	33	90	9	-	-	497	51	16	2	14	1	32	3
akt. intrans.	-	-	-	-	277	16	1307	77	-	-	82	5	24	1
pass. trans.	508	45	-	-	-	-	450	40	125	11	49	4	2	0
pass. intrans.	-	-	-	-	1	0	134	82	-	-	28	17	-	-

Tämä taulukko oikeuttaa useihin päätelmiin. Lausekeobjektin topikaalitus on paljon yleisempää passiivikuin aktiivilauseissa, 45 % vs. 33 %, samaten adverbiaalin ja objektin yhtäaikainen topikaalitus, 11 % vs. 2 %. Tämä johtuu kuten todettua pintarakenteen pyrkimyksestä nominaali- eikä verbialkuisuuteen, ks. Hakulinen & Karlsson (1979:303-) ja vrt. taulukkoa 34.

Lauseobjekti sen sijaan ei näytä passiivilauseissa topikaalistuvan ollenkaan, mikä johtunee siitä, että lauseesta subjektittomana tulisi liian kokkapainoinen ja viskurilain vastainen. Predikatiivin topikaalistuminen on pass. intrans.lauseissa olematonta, kun se taas on melko yleistä aktiivilauseissa, 16 %. Myös adverbiaalin topikaalitus vaihtelee lauserakenteesta riippuen. Yleisintä tämä on odotettavasti intrans.lauseissa. Akt. intrans.- ja trans. lauseiden suhdeluku on 1.55 (4754/3049, taulukko 40), mutta adverbiaalien topikaalituksen suhdeluku samoissa rakenteissa peräti 2.63 (1307/497).

TAULUKKO 43. *Verbiryhmän rakenne* (25)

		%
yksi sana	6234	63 (66)
liittomuoto	1817	18 (19)
verbiketju	851	9 (9)
verbiketju liittomuod.	265	3 (3)
<i>on aika</i> + inf., <i>katsoa</i>	30	0 (0)
<i>hyväkseen</i> + inf. edellinen		
liittomuodossa	10	0 (0)
phrasal verb	145	1 (2)
edellinen liittomuodossa	59	1 (1)
sanapari	8	0 (0)
edellinen liittomuodossa	2	0 (0)
finiittiverbin ellipsi	519	5
yht.		9940

Sulkeissa olevat prosenttiluvut on laskettu suhteessa konkreettisten verbiryhmien lukumäärään 9421, jossa finiittiverbien ellipsit eivät ole mukana. Luonnollisestikin yksisanaiset verbit ovat yleisimpiä, 66 %. Kooditusjärjestelmämme ei tässä kohden erottele presens- ja imperfektimuotoja toisistaan.

Liittomuotojen eli perfektin, pluskvamperfektin ja kieltomuotojen osuus on 22 %, kun mukaan lasketaan myös verbiketjujen ja fraasiverbien liittomuodot. Liittomuotojen kokonaismäärästä 2153 vajaa puolet eli 904 on kieltomuotoja (muuttuja 14:2). Finiittiverbin ellipsi on uomattavan yleinen ilmiö, sen lukumäärää 519 sopii verrata subjektin ellipsiin, jota on vain hieman enemmän (N 593).

TAULUKKO 44. *Verbin monikkokongruenssi (26)*

		%
kongruoi	1499	15
ei kongruoi	60	1
irrelevantti	8381	84
<hr/>		
yht.	9940	

Aineistossamme on vain 60 tapausta, joissa verbin monikkokongruenssi ei ole toteutunut. Kooditusjärjestelmämme ei erottele toisistaan tyyppjä *Säkissä on saappaat* ja *Ne menee pois*, joten ei voi päätellä, missä määrin jälkimmäinen puhekielelle ominainen inkongruenssityyppi on levinnyt kirjoitettuun kieleenkin.

### 4.3. *Lause- ja sanamääriä koskevat muuttujat*

Tässä jaksossa käsitellään kaikkein mekaanisimpia muuttujiamme, niitä joissa on kyse lauseiden, konstituenttien taikka sanojen määristä. Nämä muuttujat ovat (4) lauseiden lukumäärä virkkeessä, (16,17) lauseen sana- ja konstituenttimäärä, subjektin (24), objektin (32), predikatiivin (37), alkuadverbiaalin (44) ja loppuadverbiaalin (51) sanamäärä sekä adverbialien määrä lauseessa (52). Lopuksi esitämme myös eräitä keskeisiä vertailevia frekvenssitietoja tärkeimpien nominaalisten lauseenjäsenten yleisyydestä.

TAULUKKO 46. *Lauseiden lukumäärä virkkeessä (4)*

lauseet		%
1	1929	38
2	1754	35
3	851	17
4	330	7
5	103	2
6	31	1
7	10	0
8	3	0
9	3	0
10	1	0
11	1	0
yht. 5016		

Virkkeitä aineistossamme on 5016 ja lauseita 10149; virkettä kohti on siis melko tasan 2 lausetta. Tämän suhteen arvioimisessa on tosin pidettävä mielessä, että "lauseen" määritelmämme poikkeaa hieman perinteisestä, koska pidämme lauseina myös eräitä lauseenvastikkeita sekä poistosääntöjen tuottamia konjunkteja. Sellainen jakso kuin *Kalle syö ja juo* on siis meidän järjestelmässämme kaksilauseinen; ks. jaksoa 2.4. Lauseina pidettyjä lauseenvastikkeita aineistossa on 459 (muuttuja 8:B-E), usein voimakkaasti redukoituneita komparatiivilauseita 117 (8:A). Lauseita, joiden subjekti on elliptisesti poistettu (19:D), on 593; näistä melkoinen osa on konjunkteja. Jos nämä kaikki vähennetään, lauseiden määräksi saadaan 8980, mikä tarkoittaisi, että virkettä kohti on noin 1,8 lausetta. Virkkeiden lausemäärän suhdeluku 1,8 on täsmälleen sama, minkä Leskinen & Savijärvi & Särkkä (1974:24) ovat todenneet keskimääräisesti vallitsevan oppikoululaisten aineissa (II, V ja VIII luokan keskiarvo). Konjunktitulkitamme takia suhdelukumme 1,8 on kuitenkin vähän epätarkka, ja saattaa olla lähempänä 1,9.

On ehkä yllättävää, että yksilauseisten virkkeiden osuus on niinkin alhainen kuin 38 %. Esim. tunnusmerkkisyysteorian näkökulmasta saattaisi hyvinkin ennustaa, että niiden osuuden pitäisi olla suurempi. Joka tapauksessa tämä luku osoittaa, että *yhdyslauseiden* tekstuaalisten ja lauseopillisten piirteiden selvittämiseksi pitäisi omistaa entistä enemmän huomiota.

TAULUKKO 47. *Lauseen sanamäärä (16)*

sanat		%
1	138	1
2	590	6
3	954	9
4	1297	13
5	1381	14
6	1396	14
7	1061	10
8	962	9
9	665	7
10	518	5
11	385	4
12	217	2
13	183	2
14	126	1
15	70	1
16	57	1
17	44	0
18	30	0
19	18	0
20	16	0
21	9	0
22	8	0
23	7	0
24	4	0
25	1	0
26	1	0
27	4	0
28	1	0
29	2	0
31	1	0
32	1	0
40	1	0
50	1	0

---

yht. 10149

5-6 -sanaiset lauseet ovat yleisimpiä, kumpaakin on 14 %. Kolme neljännestä (76 %) kaikista lauseista on 8-sanaisia tai lyhyempiä. Näiden lukujen arvioinnissa on jälleen pidettävä mielessä, että "lauseen" määritelmästä johtuen lauseemme ovat keskimäärin hieman lyhyempiä kuin mihin perinteinen yksinkertaisen lauseen määritelmä johtaisi.

Aineistomme sanojen (= saneiden) kokonaismäärä on 66851. Lausetta kohti sanoja on siten 6,6. Tätä lukua sopii verrata Leskisen & Savijärven & Särkän (1974:29) tuloksiin, joiden mukaan oppikoulun VIII-luokkalaisten lauseiden keskimääräinen sanepituus on 7,2 ja kaikkien kolmen

tutkitun luokan 6,1. Hyvällä syyllä voi siis väittää, että asiaproosan lauseiden keskimääräinen pituus on 6-7 sanaa.

Kooditusjärjestelmämme teknisten muuttujien avulla (2,3,5) olemme tulostaneet myös virkkeiden sanamääräjakauman, joka on seuraavankaltainen.

TAULUKKO 48. *Virkkeen sanamäärä*

sanat		%
1	3	0
2	20	0
3	65	1
4	123	2
5	218	4
6	280	6
7	266	5
8	335	7
9	345	7
10	345	7
11	357	7
12	279	6
13	293	6
14	290	6
15	232	5
16	200	4
17	202	4
18	167	3
19	153	3
20	133	3
21	117	2
22	99	2
23	79	2
24	67	1
25	52	1
26	51	1
27	35	1
28	33	1
29	25	0
30	25	0
31 - 73	123	3
<hr/>		
yht.	5016	

Virkettä kohti on keskimäärin 13,3 sanaa. Tässä kohden voidaan verrata Westmanin (1974:56) ruotsin vastaavaa tekstilajia ("bruksprosa") koskeviin tuloksiin. Hänen mukaansa ruotsin virkkeen (= grafisk mening) keskipituus on 14,4 sanaa. Luonnollisestikin suomen lauseessa ja

virkkeessä on vähemmän sanoja kuin ruotsin, koska kielten morfologisessa rakenteessa on huomattavia eroja. Silti suomen ja ruotsin ero tuntuu yllättävän pieneltä.

TAULUKKO 49. *Lauseen konstituenttimäärä (17)*

konstituentit		%
1	433	4
2	1331	13
3	4012	40
4	3158	31
5	1004	10
6	182	2
7	26	0
8	3	0
<hr/>		
yht.	10149	

3 ja 4 konstituentin lauseet ovat yleisimpiä; "konstituentin" määritelmästämmme ks. muuttujan (17) esittelyä jaksossa 2.6. Yksikonstituenttisia lauseita on vähän, mutta silti 4 %. Vain poikkeustapauksissa lause sisältää enemmän kuin 6 konstituenttia. Korpuksemme konstituenttien kokonaismäärä on 34081, joten konstituentti sisältää keskimäärin melko tarkalleen 2 sanaa. Huomautettakoon vielä, että äskeisissä luvuissa ovat mukana konstituenttina esiintyvät lauseetkin. Ellegårdin (1978:44) mukaan englannin ei-lausemaisten konstituenttien pituus on 2,2 sanetta. Morfologisten erojen takia on selvää, että indoeur. kielten konstituentit ovat pitempiä kuin suomen: jo prepositioiden ja artikkeleiden olemassaolo kasvattaa indoeur. kielissä huomattavasti konstituenttien pituutta.

TAULUKKO 50. *Subjektin, objektin ja predikatiivin sanamäärät (24,32,37)*

sanoja	subjekti		objekti		predikatiivi	
		%		%		%
1	3374	42 (51)	1637	36 (43)	676	38
2	1457	18 (22)	984	21 (26)	501	28
3	695	9 (10)	486	11 (13)	256	14
4	409	5 (6)	270	6 (7)	147	8
5	280	3 (4)	165	4 (4)	66	4
6	157	2 (2)	118	3 (3)	37	2
7	103	1 (2)	43	1 (1)	26	1
8	63	1 (1)	37	1 (1)	21	1
9-17	127	2 (2)	69	2 (2)	17	1
i	1357	17	768	17	20	1
<hr/>						
yht.	8022		4577		1767	

Tämän taulukon sanamäärät koskevat muita subjekteja, objekteja ja predikatiiveja kuin lauseenmuotoisia. Kategorian "irrelevantti" suuri osuus subjektin ja objektin kohdalla, 17 %, johtuu siitä, että siihen on sisällytetty lauseenmuotoisten lisäksi myös inkorporoidut, elliptiset ja geneeriset nollasubjektitapaukset (vrt. muuttujia (19:B-D; 27:A,B; 33:C). Kun subjektin ja objektin irrelevantteiksi katsotut tapaukset jätetään huomiotta, saadaan lausekesubjektien summaksi 6665 ja lausekeobjektien 3809. Sulkeissa olevat prosenttiluvut on laskettu suhteessa näihin lukuihin.

Subjekti on tunnusmerkittävästi verbin edessä, objekti ja predikatiivi taas verbin jäljessä. Samaten subjekti on yleensä temaattinen lauseenjäsen ilmaisten puheenaiheen, joka vielä on usein tunnettu, kun objekti ja predikatiivi useammin ovat remaattisia kommentteja, joiden avulla ilmaistaan uutta informaatiota. Nämä tutut seikat heijastuvat hyvinkin tuntuvasti konstituenttien pituuteen. Yksisanaiset subjektit ovat paljon yleisempiä kuin yksisanaiset objektit ja predikatiivit, 51 % vs. 43 % vs. 38 %. Näistä subjekteista suurin osa on erisnimiä tai anaforisia pronomineja. Pitempien primaaristen lausekkeiden kohdalla pätee seuraava poikkeukseton tendenssi: predikatiivi on aina pitempi kuin objekti, joka on aina pitempi kuin subjekti. Kaksisanaisia predikatiiveja on 28 %, objekteja 26 % ja subjekteja 22 %, kolmisanasia predikatiiveja 14 %, objekteja 13 % ja subjekteja 10 % jne. Predikatiivi on siis luonnehtiva lauseenjäsen par préférence, se sisältää enemmän attribuutteja ja muita alemman rangin määritteitä kuin subjekti ja objekti; vrt. myös taulukkoa 28.

Lausesubjektien, -objektien ja -predikatiivien status ja yleisyys käy ilmi taulukoista 25 ja 28.

Sanamäärämuuttujia voidaan käyttää monissa ristiintaulukoinneissa. Seuraavassa taulukossa osoitetaan, miten lausekesubjektin sanamäärä korreloi sen sijainnin kanssa. Sanajärjestysmuuttujia (18) on yhdistetty kahdeksi subjektin sisältäväksi alakategoriaksi, SV ja VS, subjektin sanamäärästä taas 7-17 sanan mittaiset on yhdistetty.

TAULUKKO 51. *Lausekesubjektin sanamäärä ja sijainti (24, 18)*

sanat	SV		VS		i	
		%		%		%
1	2889	86	460	14	19	1
2	1024	70	406	28	26	2
3	470	68	217	31	8	1
4	231	57	165	40	12	3
5	164	59	108	39	8	3
6	82	52	70	45	5	3
7-17	140	48	137	47	16	6
yht.	6000	75	1763	22	219	3

Subjekttilausekkeen pituus vaikuttaa ilmiselvästi sen sijaintiin. Yksisanaisista subjekteista peräti 86 % on verbinetisiä, kaksi- ja kolmisanasista enää noin 70 %, neli- ja viisisanasista vajaat 60 % ja 6-17 -sanaisista noin puolet. Tämä taulukko on subjektin osalta eräänlainen viskurilain mitta. Ei ole aivan oikein sanoa ehdottomasti, että suomen perussanajärjestys on SVO ottamatta

huomioon subjektikonstituentin pituutta: todennäköisyys, että pitkä subjekti on verbinjälkeinen on kuten todettua noin 50 %. - Konstituenttirakenteen vaikutuksesta subjektin paikkaan ks. taulukkoa 26, jossa ovat mukana myös lauseenmuotoiset subjektit: nämä puuttuvat äskeisestä taulukosta.

Seuraavaksi pohditaan, miten subjektin sanamäärä vaikuttaa topikaalistukseen heittyvätkö topikaalistettavat lausekeobjektit ja adverbiaalit esim. helpommin pitkän kuin lyhyen subjektin yli?

TAULUKKO 52. *Subjektin sanamäärä sekä lausekeobjektin ja adverbiaalilin topikaalistus (24, 54:1,6)*

sanat	topikaalistettu lausekeobjekti		topikaalistettu adverbiaali		kaikki subjektit	
		%		%		%
1	119	36	828	42	3374	42
2	61	19	412	21	1457	18
3	13	4	208	10	695	9
4	15	5	129	7	409	5
5	19	6	73	4	280	3
6	5	2	48	2	157	2
7	4	1	27	1	103	1
8	3	1	15	1	63	1
9	4	1	6	0	39	0
10	3	1	6	0	24	0
11-17	4	1	24	2	64	1
i	77	24	218	11	1357	17
yht.	327		1994		8022	

Adverbiaalilin topikaalistuksen ja subjektin sanamäärän välillä ei näytä olevan minkäänlaista riippuvuutta, adverbiaali siirtyy alkuun subjektin pituudesta riippumattomasti. Lausekeobjekti sen sijaan ei näytä topikaalistuvan yhtä riippumattomasti. Se siirtyy jonkin verran hanakammin alkuun, jos subjekti on pitkä kuin jos se on lyhyt. Tämä on sikäli odotuksenmukaista, että pitkä subjekti pyrkii viskurilain mukaisesti verbin jälkeen, ja tämä edistää objektin topikaalistusta, jolloin järjestys SVO vaihtuu OVS:ksi. Objekti topikaalistuu siis herkemmin, jos pitkä subjekti samanaikaisesti siirtyy verbin jälkeen.

Seuraava taulukko osoittaa, miten objektin sanamäärä vaikuttaa sen topikaalistumisalttiuteen.

TAULUKKO 53. *Lausekeobjektin sanamäärä- ja topikaalistus (32,54:1)*

sanat	lausekeobjekti topikaalistettu		vrt. kaikki lausekeobjektit	
		%		%
1	427	51	1637	43
2	186	22	984	26
3	88	11	486	13
4	51	6	270	7
5	26	3	165	4
6	15	2	118	3
7-16	16	2	149	4
i	27	3	-	-
<hr/>				
yht.	836		3809	

Taulukko 53 osoittaa, miten muun kuin lauseenmuotoisen objektin topikaalistumisherkkyys on sidoksissa sanamäärään; oikealla olevat luvut ovat kaikkien lausekeobjektien vertailulukuja. Yksisanaiset objektit topikaalistuvat herkimmin. Niiden osuus topikaalistetuista on 51 %. Yksisanaisia objekteja on kaikista vain 43 %, joten tendenssi on ilmeinen. 2-16 -sanaiset objektit topikaalistuvat vastahakoisesti, näiden tapausten prosenttiosuudet ovat järjestään pienempiä kuin vastaavat kaikkien lausekeobjektien prosenttiosuudet. Yksisanaisten objektien topikaalistumisherkkyys on ymmärrettävissä, koska pitkien konstituenttien topikaalistaminen on ristiriidassa viskurilain päinvastaisen pyrkimyksen kanssa ("Gesetz der wachsenden Glieder", "fat things last").

TAULUKKO 54. *Alku- ja loppuadverbiaalin sanamäärät (44,51)*

sanat	alkuadverbiaali		loppuadverbiaali	
		%		%
1	1697	51	1366	32
2	807	24	1133	26
3	252	8	636	15
4	114	3	303	7
5	47	1	178	4
6	27	1	125	3
7	11	0	55	1
8	11	0	46	1
9	2	0	24	1
10-17	8	0	58	1
i	328	10	403	9
<hr/>				
yht.	3304		4327	

Alku- ja loppuadverbiaalin pituudessa on melkoisia eroja. Huomattavin on alkuadverbiaalin pyrkimys yksisanaisuuteen, enemmän kuin puolet on yksisanaisia, loppuadverbiaaleista vain

noin kolmannes. Yksisanaisten alkuadverbiaalien enemmistö on ajan ja paikan kehysadverbiaaleja (*täällä, nyt, kotona, eilen* jne.). Kaksisanaisia alku- ja loppuadverbiaaleja on suurin piirtein yhtä paljon, mutta kolmisanaisia ja pitempiä loppuadverbiaaleja on ainakin kaksi kertaa enemmän kuin vastaavia alkuadverbiaaleja. Tässäkin viskurilain vaikutus siis näkyy yhtä lailla kuin subjektin, objektin ja predikaatiivin kohdalla taulukossa 50.

TAULUKKO 55. *Adverbiaalien määrä lauseessa (52)*

adverbiaalit		%	yht. adverbiaaleja
1	4322	59	4322
2	2311	31	4622
3	616	8	1848
4	99	1	396
5	12	0	60
6	7	0	42
<hr/>			
yht.	7367		11290

Yksi- ja kaksiadverbiaaliset lauseet muodostavat adverbiaalin sisältävistä lauseista peräti 90 %. Kolme tai useampia adverbiaaleja lauseessa on vain poikkeustapauksissa, yhteensä noin 1 %:ssa kaikista lauseista. Korpuksemme 10149 lauseesta 7367 eli 73 % sisältää ainakin yhden adverbiaalin; noin neljäsos lauseista on silti adverbiaalittomia. Taulukosta käy ilmi, että alkuadverbiaalien lukumäärä on 3304 ja loppuadverbiaalien 4327, yhteensä 7631. Alku- tai loppuadverbiaaleja, jotka kooditusjärjestelmämme kattaa, on näin ollen 68 % koko aineiston 11290 adverbiaalista. Kolmannes kaikista adverbiaaleista jää siis meidän tarkemmalta analyysiltämme tavoittamatta.

Keskiasemaisia adverbiaaleja on 3659. Pelkkien lukumäärien perusteella näyttää siltä, että suomen adverbiaaleilla ei ole mitään ehdotonta suosikkiasemaa. Loppuasemaisia adverbiaaleja on kuitenkin eniten eli 38 %, alkuasemaisia 29 % ja keskiasemaisia 32 %. Nämä suhteet antavat lievää tukea Hakulisen & Karlssonin (1979:304) oletukselle adverbiaalien loppuasemaisuuden tunnusmerkittömyydestä. Tämä on silti vain heikko suuntaus.

Lopuksi esitämme muutamia tärkeitä nominaalisten lauseenjäsenten yleisyyttä koskevia tietoja, adverbiaalit mukaan lukien. Luvut on laskettu jakson 4.2. subjektin, objektin ja predikaatiivin konstituenttirakennetta käsittelevistä taulukoista 25 ja 28 niin, että kaikki 0-jäsenet on jätetty pois, siis ellipsisit ynnä inkorporoidut ja generiset 0-subjektit. Adverbiaaleista ks. taulukkoa 55.

TAULUKKO 56. *Nominaalisten lauseenjäsenten lukumäärä*

		%
subjekti	6866	28
objekti	4417	18
predikaatiivi	1750	7
adverbiaali	11290	46
<hr/>		
yht.	24323	

On selvää, miksi adverbiaaleja on eniten: niitä toisin kuin subjekteja, objekteja tai predikatiiveja voi yhdessä yksinkertaisessa lauseessa olla useitakin, subjekteja jne. sen sijaan vain yksi. Monien adverbiaalien valinnaisuuden takia taulukon 56 adverbiaalien lukumäärä ei ole suoraan verrannollinen muiden nominaalisten jäsenten lukuihin. Koska yksityiskohtainen kooditusjärjestelmämme kattaa vain alku- ja loppuadverbiaalit, emme voi tarkalleen ilmoittaa, montako pakollista adverbiaalia lukuun 11290 sisältyy, mutta alku- ja loppuasemaisia valenssiadverbiaaleja on yhteensä 2724, johon voidaan vielä lisätä pakollisten predikatiiviadverbiaalien osuus 242; pakollisten alku- ja loppuadverbiaalien lukumäärä olisi täten 2966, mikä on 39 % kaikista alku- ja loppuadverbiaaleista (taulukko 38) Jos tämä luku yleistetään kaikkiin adverbiaaleihin (N 11290), arvioimme pakollisten valenssinmukaisten adverbiaalien ja predikatiiviadverbiaalien lukumääräksi noin 4400 (4403), mikä luku tuntuu sikäli luonnolliselta, että se on melkein sama kuin objektien määrä 4417.

Jos taulukko 56 laaditaan uuteen muotoon niin, että äskeisen arvion mukaisesti verrataan vain pakollisten nominaalijäsenten yleisyyksiä, saadaan seuraavat tulokset, jotka myös kuvastavat näiden jäsenten *primaarisuushierarkiaa*.

TAULUKKO 57. *Pakollisten nominaalijäsenten lukumäärä*

		%
subjekti	6866	39
objekti	4417	25
valenssiadverbiaali	4403	25
predikatiivi	1750	10
<hr/>		
yht.	17436	

Subjekti on primaarinen nominaalijäsen, niitä on puolitoistakertaisesti objekteihin ja adverbiaaleihin verrattuna; vrt. myös Hakulinen & Karlsson (1979:165-). Jos pakollisista adverbiaaleista erotetaan predikatiiviadverbiaalit omaksi ryhmäkseen (vrt. ed.), päädytään siihen, että objekti sijoittuu primaarisuushierarkiassa valenssiadverbiaalien edelle. A priori tekeekin mieli arvata, että objektillisiä transitiiviverbejä on enemmän kuin valenssiadverbiaalin vaativia intransitiiviverbejä eli ns. semitransitiivisiä verbejä; ks. Hakulinen & Karlsson (1979:94).

Predikatiivi jää nominaalijäsenten primaarisuushierarkiassa viimeiseksi (yhdessä predikatiiviadverbiaalien kanssa).

#### 4.4. *Sanajärjestykseen vaikuttavista tekijöistä*

Tässä jaksossa käsitellään etenkin sanajärjestysmuuttujaa (18) sekä eräitä olettamiamme siirtosääntöjä (54-58), jotka täydentävät muuttujaa (18). Ristiintaulukointien avulla on selvitetty edelleen eräitä sanajärjestykseen vaikuttavia tekijöitä, joista tässä yhteydessä tarkastellaan lauseen virkeasemaa sekä sivulauseen tyyppiä.

TAULUKKO 58a. *Pintajärjestys* (18)

		%
SVX	4928	49
SXV	60	1
VSX	177	2
VXS	233	2
XSV	348	3
XVS	1139	11
SV	746	7
VS	218	2
VX	970	10
XV	861	8
V	36	0
i	433	4
<hr/>		
yht.	10149	

TAULUKKO 58b. *Yhdistetty pintajärjestys* (18)

		%
SV	6082	61
VS	1590	16
V	1867	19
i	433	4
<hr/>		
yht.	10149	

Pintajärjestyksessä huomio kiinnitettiin vain pääjäseniin: subjettiin (S), finiittiverbiin (V) ja tämän toiseen kiinteään aktanttiin (X), oli se sitten objekti, predikatiivi tai valenssiadverbiaali. Havaitsemme, että suomen sanajärjestyksen ns. vapaudesta huolimatta tekstilauseissa on selvä SV-järjestyksen ylivoima. Edelleen näkyy, että toisen nominaalisen pääjäsenen sijoittuminen subjektin ja verbin väliin on harvinaista: tyyppi SXV (*Minä häntä silittelin*) esiintyy vain 1 %:ssa, tyyppi VXS (*Aiheutuu tästä ongelmia puolestaan*) 2 %:ssa lauseista. Taulukon 58a esittämä kuva sanajärjestyksyyppien yleisyys-suhteista selvenee, kun se yleistetään muotoon 58b. Lauseet, joissa subjekti edeltää verbiä, muodostavat kaikkiaan 61 % korpuksen tekstilauseista. Merkille pantavaa on edelleen sekin, että subjektittomia lauseita on niinkin paljon kuin 19 % ja subjektiloppuisia lauseita 16 %. Kaksi viimeksi mainittua seikkaa ovat "unohtuneet" monista kieliopeista, subjektittomat lauseet siksi, että on seurattu klassisten kielten kielioppijärjestelmää liian orjallisesti ja ajateltu että suomessakin subjekti on predikaatin ohella lauseen toinen pääjäsen aivan kuten klassisissa kielissä, subjektiloppuiset lauseet taas lähinnä siksi, että yleensä ei ole tarkasteltu tekstilauseita, vaan kontekstittomia, kieliopintekijän keksimiä lauseita.

Taulukossa 58c tarkastellaan verbin sijoittumista suhteessa kahteen nominaaliseen pääjäseneseen. Tämäkin taulukko perustuu taulukkoon 58a.

TAULUKKO 58c. *Verbin sijainti suhteessa nominaalisiin pääjäseniin* (18)

		%	
verbi alussa	1598	16	(alakat. 3,4,8,9)
verbi keskellä	6067	60	(1,6)
verbi jäljessä	2015	20	(2,5,7,A)
i	469	4	(B,i)
<hr/>			
yht.	10149		

Kun kaikentyyppiset lauseet ovat yhdessä, verbiloppuisuus on hieman yleisempää kuin verbialkuisuus. Tämä tulos osoittaa, miten suomi eroaa naapurikielestä ruotsista, jossa nimenomaan verbin paikka väitelauseessa toisena konstituenttina on sanajärjestyksen tiukoin vaatimus. On edelleen muistettava, että aineistomme lauseista 96 % on väitelauseita, joten lauseen syntaktinen modus ei sanajärjestystuloksiin juuri vaikuta.

Katsokaamme seuraavaksi, miten lauseen virkeasema vaikuttaa sen sanajärjestyksen muotoutumiseen. Taulukossa 60a ovat mielenkiintoisimmat pintajärjestykset virkeasemittain. Taulukot 60b,c noudattavat samaa ryhmittelyä kuin taulukot 58b,c.

TAULUKKO 60a. *Virkeasema ja sanajärjestys* (6; 18:1,5,6,7)

	SVX		XSV		XVS		SV		yht.	(jatkuu)
päälauseet	%		%		%		%		(rivisummista)	
yksin t. al.	1681	48	78	2	524	15	154	4	2437	
rinn. al.	647	54	22	2	129	11	103	9	901	
keskellä	341	44	15	2	70	9	66	9	492	
lopussa	516	43	67	6	138	11	86	7	807	
<hr/>										
yht.	3187	47	182	3	861	13	406	6	4637	
<hr/>										
sivulauseet:	SVX		XSV		XVS		SV		yht.	
pl:n edellä	%		%		%		%			
pl:n sisällä	217	44	12	3	31	6	75	15	335	
pl:n jäljessä	77	45	15	9	12	7	22	13	126	
lopussa	1307	51	132	5	222	9	219	9	1880	
	142	54	7	3	13	5	21	8	183	
<hr/>										
yht.	1743	50	166	5	278	8	337	10	2524	

Näiden poimintojen perusteella havaitaan, että tunnusmerkitön SVX-sanajärjestys on yleisin virkkeenalkuisissa päälauseissa ja virkkeenloppuisissa sivulauseissa (kummassakin 54 %), kun

sen sijaan päinvastaisissa virkeasemissa tätä tyyppiä on vähiten. Tässä siis virkeasema vaikuttaa pää- ja sivulauseissa päinvastaisella tavalla. Verbiloppuinen sanajärjestys, jossa subjektin eteen on topikaalistettu jokin lauseen ydinosa, on vain virkkeenloppuisissa päälauseissa keskimääräistä yleisempi, sivulauseista taas virkkeensisäisissä. Subjektiloppuinen tyyppi XVS (e-lauseen normaalimuoto) on sekä absoluuttisesti että suhteellisesti yleisin virkeenalkuisissa päälauseissa. Sivulauseista vain päälauseensa jäljessä olevissa se on merkittävän runsaassa käytössä. Intransitiivista lauserakennetta SV käytetään suhteellisesti enemmän sivu- kuin päälauseissa; suhteessa harvinaisin tämä rakenne on virkeenalkuisissa päälauseissa.

TAULUKKO 60b. *Virkeasema ja sanajärjestys* (6; 18:7-i)

	SV		VS		(X)V(X)		i		yht.
		%		%		%	%		
päälauseet:		%		%		%			
yksin t. al.	1913	55	822	22	653	19	77	4	3475
rinn. al.	777	65	177	15	207	17	30	3	1191
keskellä	425	54	142	18	130	17	79	11	776
lopussa	672	55	221	18	186	15	134	11	1213
<hr/>									
yht.	3989	58	1362	20	1176	17	320	5	6845
<hr/>									
sivulauseet:									
pl:n edellä	308	63	60	12	99	20	22	5	489
pl:n sisällä	119	66	13	7	38	21	3	5	179
pl:n jälj.	1687	65	312	12	510	20	61	3	2570
lopussa	171	65	20	9	44	16	27	10	262
<hr/>									
yht.	2285	65	405	12	691	20	113	3	3494

Tämä yhdistetty taulukko viittaa siihen, että SV-järjestys on jonkin verran yleisempää sivu- kuin päälauseissa; sivulauseen tasolle yltävät vain rinnastuksen alkuosat, siis juuri ne lauseet, joiden jälkiosiin ellipsi useimmiten soveltuu: ellipsi tarvitsee mahdollisimman yksiselitteisen vertauskohdan. Verbialkuisuus on puolestaan varsin selvästi yleisempää pää- kuin sivulauseissa. Sitä esiintyy lähes viidennes muissa kuin rinnastuksenalkuisissa päälauseissa, kun sen sijaan sivulauseissa lukemat ovat IO:n molemmin puolin. Päälauseen sisällä oleva sivulause erottuu jossain määrin omaksi tyyppikseen. Katsokaamme vielä verbin sijaintia eri virkeasemissa.

TAULUKKO 60c. *Virkeasema ja verbin sijainti (6,18)*

	V-		-V-		-V		i		yht.
		%		%		%		%	
päälauseet:									
yksin t. al.	689	20	2205	63	496	14	85	2	3475
rinn. al.	140	12	776	65	239	20	36	3	1191
keskellä	153	20	411	53	127	16	85	11	776
lopussa	184	15	654	54	236	19	139	11	1213
yht.	1166	18	4046	61	1098	16	345	5	6655
sivulauseet:		%		%		%		%	
pl:n edellä	69	14	248	51	146	30	26	5	489
pl:n sisällä	15	8	99	55	66	37	3	2	179
pl:n jälj.	319	12	1529	59	654	25	68	3	2570
lopussa	29	11	155	59	51	19	27	10	262
yht.	432	12	2031	58	917	26	124	4	3504

Verbialkuisia on viidennes eräistä päälausetyppeistä; vähiten tätä sanajärjestystä esiintyy "konservatiivisessa" rinnastuksenalkuisessa lauseessa. Sivulauseista virkkeenalkuinen tulee lähimmäs päälauseen käytäntöä, päälauseen sisälle jäävä sivulause on tästä etäimmällä. Verbiloppuisuus on etenkin päälauseensisäisen ja -etisen sivulauseen ominaisuus, jota tavallisimmassa päälausetypissä esiintyy huomattavan vähän. Tällaisena heijastuu siis nykykielen kirjoitettuun muotoon sanajärjestystyyppi, jonka on oletettu olleen suomen kielen esivaiheissa vallitsevana: jäljellä on lievä tilastollinen suuntaus.

TAULUKKO 61a. *Sanajärjestys sivulauseissa (8:4-E; 18:1,3,5,6,7)*

	SVX		VSX		XSV		XVS		SV		yht.
		%		%		%		%		%	
rel.	539	52	1	0	97	9	111	11	61	6	809
<i>että</i>	479	54	3	0	4	0	85	10	79	9	650
e-kys.	50	26	21	11	30	16	11	6	6	3	118
advl.	329	49	2	0	6	1	51	8	87	13	475
partis.	96	70	-	-	-	-	1	1	23	17	120
temp.	75	44	-	-	16	9	-	-	48	28	139
inf.S	66	55	-	-	2	2	-	-	3	3	71
yht.	1634		27		155		259		307		2382

TAULUKKO 61b. *Sanajärjestys sivulauseissa* (8:4-E; 18:7-i)

	SV	VS	(X)V(X)	i	yht.
	%	%	%	%	
rel.	712 69	129 13	173 17	15 2	1029
<i>että</i>	567 64	113 13	189 21	19 2	888
e-kys.	90 53	21 12	41 24	19 11	171
advl.	429 64	69 10	162 24	14 2	674
partis.	125 91	2 2	8 6	2 2	137
temp.	139 80	1 1	30 17	2 1	172
inf.S	71 59	1 1	48 40	1 1	121
yht.	2133 67	336 11	651 20	72 2	3192

TAULUKKO 61c. *Verbin sijainti sivulauseissa* (8:4-E; 18)

	V-	-V-	-V	i	yht.
	%	%	%	%	
rel.	82 8	650 63	280 27	16 2	1029
<i>että</i>	119 13	564 63	187 21	21 2	891
e-kys.	44 23	61 32	68 35	19 10	192
advl.	88 13	380 56	189 28	19 3	676
partis.	7 5	97 71	31 23	2 2	137
temp.	13 8	75 44	82 48	2 1	172
inf.S	36 30	66 56	16 13	1 1	119
yht.	389	1893	838	80	3200

Taulukossa 61a on ensinnäkin huomattava partisiippirakenteen vähäinen sanajärjestyksen liikkuvuus. Toiseksi huomiota kiinnittää, että epäsuorat kysymykset ja relatiivilauseet poikkeavat toisistaan jonkin verran: kysymyslauseessa on yleisempää, että subjekti jää lykkäämättä loppuun kun kysymystransformaatio siirtää toisen konstituentin lauseenalkuiseksi (En tiedä, *missä hän on*), relatiivilauseessa verbioppisuus ja lykätty subjekti ovat yhtä yleisiä. Subjektin sijaintia kuvaavasta taulukosta 60b näkyy, että erot päälauseeseen nähden ovat lausetyyppikohtaisia: *että*-lause ja adverbiaalilause eivät juuri poikkea konservatiivisista päälauseista muuten kuin verbioppisuutensa suhteen; relatiivilause poikkeaa selvemmin ja lauseenvastikkeet tyystin. Emme siis voi puhua mistään sivulauseelle tyypillisestä sanajärjestyksestä ainakaan subjektin sijainnin suhteen. Verbin sijainnin suhteen (taulukko 61c) pätee jokseenkin sama. Verbioppisuus on sivulauseissa tosin tavallisempaa kuin päälauseissa, mutta ero *että*-lauseiden ja rinnastuksenalkuisten päälauseiden välillä on mitätön; epäsuorat kysymyslauseet poikkeavat jo kuvasta selvästi, ja relatiivi- ja adverbiaalilauseillakin verbioppisuus lähenee 30 %:a.

Näitä lukuja olisi kiinnostavaa verrata vastaaviin puhutusta kielestä tai romaanien dialogista saatuihin, jotta nähtäisiin, missä määrin pää- ja sivulauseiden sanajärjestyserojen tasoittuminen on kirjoitetun kielen ominaisuus.

Tarkastelemme seuraavaksi muuttujan (54) antamaa hieman tarkempaa kuvaa sanajärjestyksen käyttäytymisestä. Tämän muuttujan lähtökohtanahan oli se, että jokaisen lauseen oletettu perusjärjestys on SVX, ja tästä poikkeavat järjestykset ovat siirtosääntöjen tulosta. Tällä tavoin saamme äskeistä tarkemman kuvan siitä, missä maarin muut nominaaliset ainekset sijoittuvat subjektin edelle tai sen paikalle.

TAULUKKO 62. *Topikaalistukset (54)*

		%
lausekeobjekti	840	19
lauseobjekti t. sen osa	96	2
predikatiivi	285	6
adverbiaali t. sen osa	2667	60
adverbiaali ja objekti	143	3
objekti ja adverbiaali	50	1
adverbiaali ja predikatiivi	17	0
adverbiaali ja adverbiaali	223	5
NP lauseenvastikkeen yli	66	1
3 konstituenttia	23	1
yht.		4410

Näin tulkittuna peräti 43 %:ssa lauseista on jokin topikaalistettu konstituentti. Ylivoimaisesti yleisintä on adverbiaalin sijoittuminen alkuun, mutta typologisesti kiinnostavaa on myös objektin herkkä siirtyminen.

Jos lauseen alussa on kaksi nominaalista konstituenttia, joista kumpikaan ei ole subjekti, ne ovat tavallisimmin kaksi adverbiaalia (N 223) taikka adverbiaali ja objekti (N 143). Kolmen nominaalisen konstituentin sijoittuminen alkuun on aivan marginaalista (N 23). Verrattaessa toisiinsa eri konstituenttien sijoittumista lauseen alkuun, on verrattava lauseenalkuisia tapauksia vastaavan konstituentin kokonaismäärään aineistossamme. Näistä luvuista olemme kuitenkin vähentäneet kaikki 0-edustumat, vrt. taulukoita 25, 28 ja 55.

TAULUKKO 63. *Topikaalistettujen konstituenttien määrä*

	kokonaismäärä	topikaalistetut	%
objektit	4417	1129	26
predikatiivit	1750	302	17
adverbiaalit	11290	3323	29

Konstituenttien kokonaismääriin suhteutettunakin ero säilyy, mutta adverbiaalin ja objektin ero tasoittuu silti huomattavasti. Alkuperäisessä taulukossamme ovat lauseke- ja lauseobjektit erillään; on valaisevaa katsoa tässä vaiheessa, millainen ero niiden käyttäytymisessä on, kun luvut suhteutetaan kokonaismääriin. Lausekeobjekteja on kaiken kaikkiaan 3395 (taulukko 25; alakat. N, määrite + N, N + app./ advl., rinnasteinen), joista topikaalistettuja 1033 eli 30 %. Lauseobjekteja ja sen tapaisia (alakat. lause, N + lause, nominaalistus) on 1022, joista topikaalistettuja 96 eli 9 %. Noin neljäsosa eli 23 % "materiaalisista" objekteistamme on

lausemaisista (taulukko 25), mutta topikaalistetuista vain 9 % on S-muotoisia, mikä tietysti on viskurilain mukaista: raskaat rakenteet siirtyvät vain vastahakoisesti alkuun. Kun lauseobjektit poistetaan objektien kokonaismäärästä, havaitaan, että peräti kolmasosa muista objekteista esiintyy aineistossamme verbinetisesti.

Muuttujaan (55) koodittaja on merkinnyt intuitiivisen käsityksensä topikaalistuksen syistä; samaan suuntaanhan voi olla vaikuttamassa useita syitä.

TAULUKKO 64. *Topikaalistuksen todennäköinen syy (55)*

		%
kompensatorinen	1313	30
sidonta	645	15
emfaasi	140	3
kontrasti	196	4
syntaktinen yhdenmukaisuus	18	0
ei syytä? tunnusmerkitön	529	12
kompensatorinen + sidonta	675	15
kompensatorinen + emfaasi	80	2
kompensatorinen + kontrasti	92	2
sidonta + emfaasi	18	0
sidonta + kontrasti	20	0
tunnusmerkitön + kompensatorinen	149	3
tunnusmerkitön + sidonta	74	2
tunnusmerkitön + kontrasti	14	0
tunnusmerkitön + emfaasi	7	0
emfaasi+ kontrasti	3	0
mahdoton ratkaista	437	10
yht.		4410

Tulokset ovat tulkintojen subjektiivisuudenkin huomioon ottaen kiinnostavia. Havaitsemme, että asiaproosassa sanajärjestyksen vaihteluun vaikuttavat syyt ovat voittopuolisesti koheesioon liittyviä, ja emfaasi sekä kontrasti ovat vaikuttamassa vajaassa 10 %:ssa tapauksista. Se, että "kompensatoriselta" tuntuva siirto on yleisempi kuin sidos, kertoo omalla tavallaan sen, mikä on tullut muistakin muuttujista esiin. On lukuisia syitä siihen, että subjektia ei ole lauseen alussa. Subjektittomia lauseita on eri sanajärjestyksityypeissä (muuttuja 18) 19 %, mihin lukuun tulevat lisäksi ellipsi-, inkorporointija generiset 0-subjektitapaukset, joten itse asiassa lausennalkuinen subjekti puuttuu n. 30 %:ssa lauseista. Kun subjekti on syystä tai toisesta lykätty (vrt. taulukkoa 65), verbin eteen kaivataan niin ikään jokin konstituentti "kompensoimaan", saattamaan lause tasapainoon (verbialkuisuuttahan esiintyy vain 16 %:ssa tapauksista). Kaikki kompensatoriset alakategoriat yhteenlaskettuina (1,7,8,9,C) antavat prosenttiluvuksi 52, mikä on kaiken edellä mainitunkin huomioon ottaen sangen korkea luku. Huomattakoon kuitenkin, että tästä luvusta melkein puolet koostuu tapauksista, joissa vaikuttavia tekijöitä on ollut muitakin kuin kompensointi. Tekstuaalinen sidos on yhtä usein (15 %) ensisijaisesti kuin yhdessä kompensoinnin kanssa vaikuttavana tekijänä tulkittu topikaalistuksen syyksi.

12 %:ssa tapauksista koodittajan käsitys on ollut, että ei ole topikaalistettukaan, vaan kyseessä on tunnusmerkitön sanajärjestys, vaikka se poikkeaa SVX-kaavasta. Tämä tukisi sitä Hakulisen & Karlssonin (1979:97) käsitystä, että perussanajärjestyksiä ei ole suomessa yksi vaan kenties kaksi tai kolme. Luku on itse asiassa tätäkin korkeampi; jos siihen lisätään alakategoriat C-F saadaan tunnusmerkittömien osuudeksi 17 %.

TAULUKKO 65. *Subjektin lykkäys (56)*

		%	kokonaismäärä	lykkättyjen subjektien %- osuus
intransitiivisubjekti	1100	71	3823	28
transitiivisubjekti	212	14	2159	10
lausesubjekti	89	6	462	20
infinitiivisubjekti	97	6	110	88
subjektin osa	5	0		
kvanttorin siirto	40	3		
<hr/>				
yht.	1543			

Subjekti on lykkäty noin 25 %:ssa sellaisista lauseista, joissa on ilmisubjekti. Yleisintä tämä on intransitiivilauseissa, joihin ns. eksistentiaalilauseet on tässä luettu. Kenties yllättävä tulos on se, että infiniivisubjekti ei välttämättä sijaitse verbinsä jäljessä ja että finiittinen lause päättyy vain joka viidennessä tapauksessa ekstrapositioon. Nämä ovat piirteitä, joita alustavasti voisi pitää vaikealukuisen tekstin merkkeinä.

TAULUKKO 66. *Verbin siirto (57)*

		%
alkuun	152	55
loppuun	96	35
molemmat	26	9
infinitiivi verbiketjusta alkuun	2	1
<hr/>		
yht.	276	

Vain vajaassa 3 %:ssa aineiston lauseista oli tulkittavissa tapahtuneeksi verbin siirto joko lauseen alkuun tai loppuun. Vähäinen määrä on oletettavasti tekstilajin tunnuspiirre. Onhan alkuun siirto usein dialogissa käytetty keino (inttäminen tai myönnyttely - joka tapauksessa reagointi toisen puheeseen). Vastaavasti siirto loppuun aiheuttaa emfaattisuutta, jota sitäkin on tässä kirjoitetussa aineistossa vähän. Kuten aikaisemmin on huomautettu, verbialkuisia ja -loppuisia rakenteita on aineistossa kuitenkin enemmän kuin nämä 276 tapausta (vrt. taulukkoja 61a-c), mikä johtuu siitä, että tällaisiin rakenteisiin päädytään myös nominaalisia konstituentteja siirtämällä.

Virkeasema ja lausetyyppi vaikuttavat osaltaan siihen, kuinka vapaasti topikaalistuksia la#useessa käytetään. Taulukkoa 67 on yksinkertaistettu siten, että kaikki kahden konstituentin topikaalistukset on yhdistetty. Huomio kiinnitetään vain virkkeenalkuisiin ja -loppuisiin lauseisiin polaaristumisen selventämiseksi. Yhteissummiin on lisätty myös virkkeiden keskellä olevien pää- ja sivulauseiden määrät.

TAULUKKO 67. *Virkeaseman ja lausetyypin vaikutus topikaalistukseen (6,8,54)*

	obj.		obj <sub>i</sub>		predvi		advl.		2 konst		lv		3 konst	
		%		%		%		%		%		%		%
päälauseet:														
yksin t. al.	378	17	4	0	206	9	1426	63	228	10	1	0	15	1
lopussa	80	15	74	14	19	4	309	58	48	9	1	0	1	0
kaikki pl.	505	16	91	3	239	8	926	62	290	9	2	0	18	1
sivulauseet:														
alussa	47	31	-	-	2	1	94	61	10	7	-	-	-	-
lopussa	266	24	2	0	45	4	613	55	122	11	64	6	4	0
kaikki sl.	535	35	2	0	48	3	740	48	143	9	64	4	5	0

Huomiota kiinnittää ensimmäiseksi se, että objektin topikaalistuksia on sekä absoluuttisesti että suhteellisesti enemmän sivu- kuin päälauseissa. Koska suuri osa aineiston sivulauseista on päälauseensa jäljessä tai keskellä, niiden objektin voi olettaa sisältävän usein tekstuaalisesti päälauseeseen sidoksissa olevaa ainesta. Miksi tämä sitten sisältyy juuri objektiin eikä niinkään predikatiiviin tai edes adverbialiin, on avoimeksi jäävä kysymys. Adverbiaalien kohdalla pää- ja sivulauseiden eroa ei ole; oletettavasti ero näkyisi, jos voisimme tarkastella adverbiaalien eri alatyypien topikaalistuksia. Tämä ei kuitenkaan meidän kooditusjärjestelmässämme ole mahdollista, koska lauseen kaikkia adverbiaaleja ei ole kooditettu tyyppin mukaan.

Eri sivulauseissa topikaalistukset vaihtelevat suuresti, kuten taulukosta 68 käy ilmi.

TAULUKKO 68. *Topikaalistus sivulauseissa (54; 8:4-E)*

	obj.		p:vi		advl.		A + X		A + A		lv		3 k.		yht.
		%		%		%		%		%		%		%	
rel.s.	24	16	6	4	94	64	10	7	11	8	-	-	1	-	146
rel.o.	34	28	1	1	70	58	9	7	6	5	-	-	1	-	121
rel.m.	60	24	7	3	142	58	23	9	13	5	-	-	-	-	245
<että>	80	29	12	4	180	59	10	3	15	5	-	-	1	-	306
e-kys.	15	13	12	10	65	56	17	15	5	4	-	-	2	-	116
advl.	77	38	7	4	104	52	6	3	7	4	-	-	-	-	201
partis.	2	4	-	-	6	13	-	-	-	-	39	83	-	-	47
temp.	24	60	1	3	12	30	3	8	-	-	-	-	-	-	40
inf.S	-	-	-	-	1	4	-	-	-	-	25	96	-	-	26

Se, että adverbiaalilauseissa on eniten topikaalistettuja objekteja, ei käy yksiin sen edellä esitetyn oletuksen kanssa, että sivulauseissa olisi nimenomaisesti tekstisidonnaisia, päälauseeseen viittaavia objekteja paljon, sillä adverbiaalilauseethan esiintyvät sivulauseista eniten päälauseen edellä, mutta subjektikonstituentissa esiintyvien relatiivilauseiden kohdalla tämä virkeasema selittäisi topikaalistusten vähäisyyden. Temporaaalisissa lauseenvastikkeissa vältellään verbialkuisuutta, mistä syystä objekti yleensä topikaalistuu, esim. *ruoan syötyään*, ei *syötyään ruoan*. Miksi subjektissa oleva relatiivilause jälleen käyttäytyy toisin kuin muissa lauseenjäsenissä olevat, on kysymys, joka ansaitsisi lähemmän tutkimuksen. Kommentoinnin arvoinen on vielä lauseenvastikkeiden sarake, josta näkyy, miten runsasta on partisiippirakenteista ja infinitiivisubjekteista topikaalistaminen yli päälauseen koko virkkeen alkuun. 28 %:ssa partisiippirakenteita on tällainen topikaalistus ja 20 %:ssa inf.subjekteista. Näillä lauseenvastikkeilla ei sisäistä sanajärjestyksen vaihtelua juuri olekaan, mikä näkyy tästäkin taulukosta selvästi.

#### 4.5. Tekstuaalisten tekijöiden yhteispeli

Kuudesta keskeisestä nominaalijäsenestä on kerätty tiedot niiden tekstisidonnaisuudesta sekä niiden sisältämän informaation tunnettuudesta. Tässä luvussa esitellään näin saatuja tuloksia sekä tarkastellaan lisäksi sitä, millä tavoin eri sidoskeinoja käytetään eri lauseenjäsenissä tai miten tunnettua informaatiota sisältävä konstituentti sijoittuu eri lauseenjäseniin kun erilaiset virkeasemat taikka sanajärjestykseltään erilaiset lauseet ovat kyseessä. Sovitamme siis yhteen myös edellisissä luvuissa esiteltyä tietoa keskeisistä tekstimuuttujista saatavan tiedon kanssa.

TAULUKKO 69. Sidiskeinojen yleisyys eri konstituenteissa

	ellipsi		anafora		katafora		toisto		kompar.		synon.		(jatkuu)
	%		%		%		%		%		%		
subjekti	83	1	1351	17	109	1	2237	28	93	1	511	6	
objekti	34	1	506	11	97	2	1101	24	60	1	289	6	
predik.	16	1	59	3	63	4	230	13	67	4	30	2	
alkuadvl.	14	0	860	26	27	1	558	17	277	8	130	4	
loppuadvl.	14	0	102	2	294	7	760	18	96	2	189	4	
yht.	161	2878	590	4886	593	1149							
	leksik.		eksofora		vapaa		i						
subjekti	1218	15	84	1	958	12	1378	17					
objekti	915	20	25	1	764	17	786	17					
predik.	416	24	3	0	849	48	34	2					
alkuadvl.	314	10	165	5	401	12	558	17					
loppuadvl.	891	21	76	2	1410	33	495	11					
yht.	3754		353		4382		3251						

Yhteenlasketuista luvuista näkyy, että vapaat ja toistetut, siis vailla ilmipantua sidoskeinoja olevat konstituentit muodostavat lähes puolet tarkasteltujen konstituenttien määrästä. Leksikaaliset sidokset, siis samassa merkityskentässä pysyttelevät, ovat yleisempiä kuin selvät anaforiset pronominit. Anaforien ja synonyymien ynnä parafrasien suhde on luku, jonka voi olettaa olevan tekstilajikohtainen. Huomiota kiinnittää "vahvimman" sidoskeinoon, ellipsin, ylen vähäinen käyttö; tämäkin saattaa olla tekstilajikohtaista.

Eri konstituenttien väliset erot ovat monessa suhteessa huomionarvoiset. Ellipsi on voittopuolisesti subjektin ja objektin keino, kun taas anafora esiintyy eniten alkuasemassa: predikatiivissa ja loppuadverbiaalissa se on marginaalista, joskaan se ei niistäkään puutu tyystin. On odotuksenmukaista, että kataforaa, eteenpäin viittaavaa sidosta, esiintyy lähes kolme kertaa enemmän lauseenloppuisissa kuin -alkuisissa konstituenteissa. Peräti 28 %:ssa subjekteista toistetaan jotakin edellä esiintynyttä leksikaalista ainesta, objekteistakin neljäosassa. Huomattavaa on, että toistot käyttäytyvät toisin kuin kieliopilliset sidokset: alkuadverbiaaleissa oli pronominisidos yleisintä, kun taas toiston kohdalla alku- ja loppuadverbiaalit käyttäytyvät keskenään samoin. Sarake nimeltä "komparat.", johon kuuluvat komparaation lisäksi muut kieliopilliset sidokset, käyttäytyy jälleen samoin kuin anaforasarake. Synonyymien ja parafrasien käyttö on asiaproosassa kauttaaltaan vähäistä, eivätkä konstituenttien väliset erot ole huomattavia. Samaan merkityskenttään kuuluvan sanan käyttäminen on suhteellisesti yleisempää muissa konstituenteissa kuin subjektissa ja alkuadverbiaalissa - tulos on käänteinen kieliopillisten sidoskeinojen jakaumalle. Kolme viimeistä saraketta, nimittäin eksoforat, vapaat ja irrelevantit, edustavat konstituentteja, joissa ei ole tekstuaalista sidosta. Irrelevantti (i) tarkoittaa eri konstituenteissa eri asioita: subjektissa ja objektissa etenkin lausetta ja lauseenvastiketta, adverbiaalissa mm. eräitä adverbeja (*aina, ehkä, joskus*). Lähes puolet predikatiiveista on tekstuaalisesti vapaita, loppuadverbiaaleistakin kolmannes. Lausekeobjekteista, joissa sidos voisi tulla kyseeseen, vapaita on vain 17 % - ei kovinkaan paljon enempää kuin subjekteissa ja alkuadverbiaaleissakaan. Irrelevanttien konstituenttien suhteellinen osuus näissä konstituenteissa on sama.

Taulukko osoittaa, että ainakaan sidoskeinoja ei voida käyttää minkään tekstuaalisten kategorioiden kuten teeman postuloimisen pohjana - niitä esiintyy vähäisemmässä tai suuremmissa määrin kaikissa nominaalisissa lausekkeissa.

Taulukko 69 sisältää pohjatietoa sekä tyyli- että kieltenvälisiä vertailuja varten. Voidaan esittää esim. seuraavanlaisia hypoteeseja. Toiston suhteellinen osuus kasvaa mitä yksinkertaisemmasta tyylistä on kysymys; painollinen toisto lienee tämän lisäksi yksi kaunokirjallinen tehokeino. Synonyymien ja parafrasien määrä on myös suurempi kaunokirjallisessa ja esseetyylissä kuin nyt tarkasteltavana olevassa asiatyyliä; edelleen sen voi olettaa olevan runsaampaa pitemmän kulttuuriperinteen omaavissa kielissä (esim. ranskassa ja englannissa) kuin suomessa, vaikka tyyliolisi samakin. Mitä suurempi osuus konstituenteista on tekstuaalisesti vapaita, sitä informatiivisempaa tai mahdollisesti raskaslukuisempaa teksti on. Hypoteeseja voi todennäköisesti esittää tämän aineiston nojalla lisääkin.

Ellipsi on kooditettu sekä yksittäisten konstituenttien rakenteita kartoittaviin muuttujiin että erikseen muuttuun (59). Kokoamme tiedot eri muuttujista tähän.

TAULUKKO 70a. *Ellipsi eri konstituenteissa*

	ellipsi	ellipsi sidoskeinona	yht.
		%	
subjekti	593	83	8022
objekti	128	34	4577
predik.	17	16	1767
alkuadvl.	52	14	3304
loppuadvl.	17	14	4327
fin.verbi	519	-	9940

Havaitsemme, että subjektin ja objektin, yllättäen myös alkuadverbiaalin ellipsiä on kahta lajia; ellipsi sidoskeinona muodostaa deletioiden pienen alaryhmän, kun taas predikatiivin ja loppuadverbiaalin ellipsi on yleensä tekstuaalista. Katsokaamme seuraavaksi verbin ellipsiä lähemmin.

TAULUKKO 70b. *Verbin ja verbiryhmän ellipsi*

pelkkä finiittiverbi	348
verbi + objekti	47
verbi + adverbiaali	45
verbi + muuta	7
VP:n ellipsi	87
<i>niin ... -kin</i>	2
<hr/>	
yht.	536

Noin kaksi kolmasosaa verbin ellipsisistä kohdistuu yksinomaisesti verbiin. Ristiintaulukointi (jota ei ole mukana tässä raportissa) tulee osoittamaan, kuinka usein tämä osuu yksiin subjektin poiston kanssa. Suomessa finiittiverbin poisto (ns. aukko) on yli viisi kertaa yleisempää kuin ns. VP:n poisto, jossa pyyhitään verbin ei-finiittistä ainesta. Tämä suhde on todennäköisesti myös kielitypologisesti kiinnostava vertailukohde. Verbiaineksen ellipsiä esiintyy runsaammin kuin pelkkien nominaalisten konstituenttien ellipsiä: sidoskeinoista nominaalijäsenten poistoa on yhteensä vain n. 150 tapausta eli vajaa puolet finiittiverbin ellipsin määrästä. Se, että eri taulukoiden verbin ellipsiä koskevat kokonaisluvut poikkeavat hieman toisistaan (519 ja 536) johtunee siitä, että ns. NP:n ellipsisissä voi olla kyse myös pelkän ei-verbaalisen aineksen poistamisesta (vrt. koodituksen esittelyä luvussa 2).

Katsokamme sitten, vaikuttaako jokin aikaisemmin käsitellyistä taustamuuttujista siihen, miten yleisesti mitään sidoskeinoja on käytetty. Kolme tässä kyseeseen tulevaa muuttujaa ovat virkeasema, sanajärjestys ja topikaalistaminen.

TAULUKKO 71. *Lauseen virkeasema ja subjektin sidoskeino* (% sarakesummista)

	pl. alussa		pl. lopussa		sl.alussa		sl. lopussa		subj:lla yl.
		%		%		%		%	%
ellipsi	25	1	34	2	2	0	22	1	1
anafora	459	12	229	15	75	15	588	27	17
katafora	83	2	9	1	6	1	11	1	1
toisto	1280	34	329	21	140	27	488	22	28
kompar.	54	1	12	1	2	0	25	1	1
synon.	283	8	76	5	39	8	113	5	6
leksik.	614	16	250	16	60	12	294	13	15
vapaa	523	14	149	10	51	10	235	11	12
eksofora	38	1	19	1	13	3	14	1	1
i	384	10	459	29	127	25	408	19	17

Anafora ja toisto käyttäytyvät keskenään päinvastaisella tavalla. Virkkeen lopussa olevan sivulauseen subjekti on yli kaksi kertaa useammin anaforinen kuin virkkeen alussa olevan päälauseen, kun taas toisto on virkkeenalkuisen päälauseen subjektissa kaikista lauseasemista yleisintä. Utta virkettä aloitettaessa on varmistettava kuulija/lukija siitä, että subjektin tarkoite on entinen, ja tämä käy parhaiten päinsä toistamalla leksikaalinen aines, virkkeen sisäpuolella pronomini taas yksiselitteisemmin on tulkittavissa. Huomattavaa on, että virkkeenloppuinen päälause käyttää yhtä paljon anaforaa kuin virkkeenalkuinen sivulause: pilkun yli päälauseeseen siirrytään vaikeammin kuin pilkun yli sivulauseeseen ja yhtä helposti kuin pisteen yli sivulauseeseen. Muut leksikaaliset keinot kuin toisto ovat yleisempiä pää- kuin sivulauseissa. Virkkeenalkuisen päälauseen subjekti on tekstuaalisesti vapaa useammin kuin keskimääräinen alkuadverbiaali eikä jää kauas keskimääräisen objektin lukemista (vrt. taulukkoa 69).

TAULUKKO 72. *Subjektin sanajärjestys ja sidoskeino*

*Objektin sanajärjestys ja sidoskeino*

	SV		VS		i		VX		XV		kokon.- %
		%		%		%		%		%	
ellipsi	54	1	10	1	16	1	18	1	6	1	1
anafora	1276	21	33	2	5	17	163	5	342	28	11
katafora	24	0	80	5	3	1	86	3	9	1	2
toisto	1844	31	319	20	19	28	763	23	330	27	24
kompar.	57	1	33	2	2	1	37	1	22	2	1
synon.	433	7	60	4	3	6	192	6	95	8	6
leksik.	765	13	396	25	30	15	700	22	205	17	20
vapaa	461	8	454	29	16	12	697	21	51	4	17
eksofora	69	1	6	0	0	1	22	1	2	0	1
i	1017	17	196	12	339	17	575	18	158	13	17

Vertailun helpottamiseksi on taulukkoon lisätty prosenttiluvut, jotka osoittavat, mikä määrä kutakin sidoskeinoa on koko subjektin ja objektin määrästä: niinpä esim. ellipsiä on vain 1 %

koko subjektien lukumäärästä, kun taas toistoa on peräti 28 %:ssa subjekteista. Taulukosta havaitaan, että sidoskeinoon valintaan vaikuttaa konstituentin sijainti lauseessa enemmän kuin sen lauseenjäsenyys. Anaforaa on 28 % verbinetisistä objekteista, 21 % verbinetisistä subjekteista ja kummassakin lauseenjäsenessä sangen vähän (2% ja 5 %) silloin kun ne sijaitsevat verbin jäljessä. Toistossa sijainti vaikuttaa samalla tavoin, mutta muissa leksikaalisissa keinoissa tilanne ei ole yhtä yksiselitteinen, sillä tätä keinoa näyttää suosivan verbinjälkeinen subjekti, kun taas objektien välinen ero on paljon vähäisempi. Kaikkein useimmin tekstuaalisesti vapaa on verbinjälkeinen subjekti, mikä on ymmärrettävää, koska subjektin verbinjälkeisyyteen tarvitaan erityisyys, kun taas objekti voi sijaita verbin jäljessä vaikka se olisi tekstuaalisesti sidoksenkin, koska tämä on objektin tunnusmerkitön paikka. Huomattava kuitenkin on, että 8 % lauseenalkuisistakin subjekteista on tekstuaalisesti vapaita. Tämä on siis suomessa mahdollista, vaikka suhteellisen harvinaista.

Tarkastelimme myös sitä, miten eri lauseenjäsenissä esiintyvä sidoskeino vaikuttaa topikaalistumiseen. Havaittiin, että subjektin, tekstuaalinen vapaus ei lisännyt muiden konstituenttien topikaalistumisherkkyttä. Objektit pikemminkin siirtyivät alkuun tavallista vähemmän, jos subjekti oli tekstuaalisesti vapaa. Missä määrin tekstisidonnainen objekti ja predikatiivi sitten pyrkivät siirtymään lauseen alkuun, näkyy seuraavasta taulukosta.

#### TAULUKKO 73. *Topikaalistus ja sidoskeino*

	topik. obj.	% sarake- summasta	% sidoks. obj:sta	topik. predik.	% sarake- summasta	% sidoks. predik:sta
ellipsi	5	1	14	4	2	25
anafora	247	30	49	38	14	64
katafora	4	1	4	2	1	3
toisto	246	30	22	65	23	28
kompar.	21	3	35	28	10	42
synon.	66	8	23	10	4	33
leksik.	159	0	17	77	27	19
vapaa	45	6	6	46	16	5
eksofora	2	0	8	2	1	66
i	41	7	5	9	3	
yht.	816			281		

Taulukossa olevista prosenttilukusarakkeista ensimmäinen osoittaa esiintymien %-osuutta topikaalistetuista objekteista ja predikatiiveista, toinen sarake sitä, mikä osa kutakin sidoskeinoa sisältävästä objektista tai predikatiivista on topikaalistunut. Kun vertaa lukuja taulukkoon 69, huomaa, että esim. kolmasosa topikaalistuneista objekteista sisältää anaforan, kun koko objektien kohdalla luku on vain 11 %. Vastaavasti näkyy, että lähes puolet anaforisista objekteista on topikaalistunut. Synonyymit ja leksikaaliset sidokset eivät vaihtelee sanottavasti, mutta vapaiden objektien määrä lauseenalkuisissa on huomattavasti vähäisempi kuin kokonaismäärässä. Peräti 64 % anaforan sisältävistä predikatiiveista on siirtynyt lauseen alkuun, ja niinpä topikaalistuneiden predikatiivien määrästä anaforisia on 14 %, kun kokonaismäärästä niitä on vain 3 %. Vastaavasti topikaalistuneiden predikatiivien määrästä vain 16 % on vapaita (topikaalistuksen syynä tällöin esim. kontrastointi), kun niitä

kokonaismäärästä on noin puolet. Kiinnostavaa on, että noinkin runsaasti vapaat predikatiivit siirtyvät lauseen alkuun myös kirjoitetussa kielessä; puheessa voi olettaa tämän olevan vielä yleisempää.

Siirrymme nyt tarkastelemaan tekstikoheesion semanttista puolta eli sitä, millä tavoin tuttu ja uusi tieto jakautuu tutkittavien konstituenttien kesken. Katsotaan, miten tähän sijoittumiseen vaikuttaa lauseen virkeasema, sanajärjestys sekä lopuksi verrataan sitä, miten korrelaatin suhteellinen etäisyys vaikuttaa sidoskeinon valintaan.

TAULUKKO 74. *Tunnettuus eri lauseenjäsenissä*

	mainittu		"heik." mainittu		asiain- tila		implikoitu		edelly- tetty		uusi		i	
	%		%		%		%		%		%		%	
subjekti	2493	31	940	12	341	4	1060	4	304	4	1166	15	1718	21
objekti	861	19	617	13	201	4	738	16	70	2	1004	22	1086	24
alkuadvl.	888	27	174	5	189	6	248	8	95	3	137	4	1573	48
loppuadvl.	354	8	320	7	82	2	464	11	161	4	640	15	2306	53
yht.	4596		2051		813		2510		630		2947		6683	

Samatarkoitteisuutta (sarake "mainittu") on suhteellisesti eniten subjekteissa ja alkuadverbiaaleissa. Lähes kolmasosa näistä on mainittu aiemmin tekstissä; objekteista enää viidennes ja loppuadverbiaaleista vajaa 10 %. Tarkoitteen tuttuus erottaa siis jyrkemmin konstituentteja toisistaan kuin pelkkä sidoskeino. Samamerkityksisyys ("heikosti mainittu") ei erottele yhtä lailla lauseenjäseniä toisistaan. Subjekti ja objekti käyttäytyvät tämän ilmiön suhteen samalla tavoin, adverbiaalit keskenään samalla tavoin. Tämän piirteen kohdalla ei lauseasema ole ratkaiseva. Kun korrelaattina on koko lause, keskimäärin n. 5 %:ssa tapauksista, eroja ei muodostu lauseenjäsenten välille. Implikoidusti tunnettuja, siis esillä olleeseen merkityskenttään liittyviä tarkoitteita, on eniten objekteissa, vähiten alkuadverbiaaleissa. Tässä siis lauseasema on vaikuttavana tekijänä. Luku käynee ymmärrettävämmäksi, kun saadaan vertailukelpoista tietoa toisentyypisistä teksteistä. Samansuuntaisesti jakautuvat uudet tarkoitteet: näitä on suhteellisesti eniten objekteista, subjekti ja loppuadverbiaali käyttäytyvät samoin, ja alkuadverbiaalissa uudet tarkoitteet ovat harvinaisia. Nämä luvut ovat jälleen tärkeitä perustietoja tutkittaessa erityyppisten tekstien ja tyylien informaatiotiheyttä. Noin puolet alku- ja loppuadverbiaaleista ovat tunnettuuden suhteen irrelevantteja, ts. koko kysymys ei ole niiden kohdalla mielekäs (modaaliset adverbis, "OSMA:t" jne.). Subjektilla ja objektilla irrelevanttius on kaksi kertaa harvinaisempaa kuin adverbiaaleilla. Jos katsotaan absoluuttisia lukuja, havaitaan kuitenkin, että luvut eivät kovin suuresti poikkea toisistaan.

Uusia tarkoitteita on tullut aineistomme teksteihin neljän tarkasteltavana olleen konstituentin (subjektin, objektin, alku- ja loppuadverbiaalin) kautta yhteensä 2947. Muuttujasta (53) saatiin tiedot siitä, kuinka monta uutta tarkoitetta lausetta kohden teksteissä oli sekä uusien tarkoitteiden kokonaismäärä.

TAULUKKO 75. *Uusien tarkoitteiden määrä lauseessa*

	lauseita	tarkoitteita
1	3848	3848
2	1191	2348
3	259	777
4	73	292
5	29	145
6	16	96
7	4	28
9	2	18
<hr/>		
yht.		7588

Uusia tarkoitteita on lähes puolessa lauseista vain yksi tai kaksi. Mutta aineiston tarkoitteiden kokonaismäärä on 7588, ja näistä vain 39 % esiintyy nyt tutkituissa lauseenjäsenissä. Vastaisuudessa olisikin kiinnostavaa selvittää, mitä muita teitä uusia tarkoitteita tekstiin tulee. Mahdollisuuksina ovat predikatiivi, nominaalistumien sisäiset attribuutit, lauseen keskellä olevat adverbiaalit

Tutkimme myös, miten etäällä korrelaatti oli mainittu niissä tapauksissa, joissa tarkoite oli selvärajaisesti ollut aikaisemmin esillä. Kyseeseen tulivat lähinnä "mainittuus"-muuttujan alakategoriat 1-3. Tulokset näkyvät seuraavasta taulukosta.

TAULUKKO 76. *Etäisyys tarkoitteen edellisestä maininnasta*

	ed. lauseessa		ed. virkkeessä		samassa yht.		kauemp.	yht.
	%		%		%		%	
subjekti	901	24	1328	24	1517	28	31 1	3737
objekti	391	23	505	30	761	45	22 1	1669
alkuadvl.	458	36	356	28	426	33	15 1	1255
loppuadvl.	68	9	176	23	498	65	17 2 7	59

Kuten kooditussuunnitelman esittelyssä mainittiin, jaottelu ei tässä muuttujassa ollut täysin onnistunut. Eräitä mielenkiintoisia seikkoja näkyy silti saaduista tuloksista, joskin tämä muuttuja tulee kiinnostavammaksi vasta kun sen tiedot ristiintaulukoidaan muiden muuttujien kanssa. Edellisessä lauseessa oli suhteellisesti useimmin alkuadverbiaalin korrelaatti, loppuadverbiaalilla tämä oli harvinaista. Ehkä yllättävää on, että peräti 52 %:lla subjekteista korrelaatti oli virkerajan takana, objekteista tätäkin useampi eli 75 %:ssa tapauksista. Lähes puolet objekteista, joiden tarkoite oli tekstistä tuttu, oli esitelty edellistä virkettä kauempana. Korrelaatin esiintyminen kauempana kuin "samassa yhteydessä" näyttää olevan sangen vähäistä tämän tyyppisessä asiatekstissä. Lauseasema on tärkeämpi kuin kieliopillinen funktio, mikä sekään ei ole kokonaan vaikuttamatta suuntaukseen. Tästä selittyy "edellisen virkkeen" jakauma, jossa loppuadverbiaali ja subjekti käyttäytyvät samoin, toisaalta objekti ja alkuadverbiaali samansuuntaisesti.

Katsotaan seuraavaksi, miten lauseen virkeasema vaikuttaa subjektin tunnettuuteen; muistamme, että virkeasema vaikutti selvästi sidoskeinon jakaumaan.

TAULUKKO 77. *Lauseen virkeasema ja subjektin tunnettuus*

	mainittu		"heik." mainittu		asiain- tila		impli- koitu		edelly- tetty		(jatkuu)
		%		%		%		%		%	
pl. yksin	831	30	320	11	169	6	397	14	132	5	
pl.rinn:n alussa	313	33	146	15	45	5	147	15	47	5	
pl. keskellä	137	23	55	9	9	2	63	11	15	3	
pl. lopussa	274	28	100	10	30	3	119	12	26	3	
yht.	1545	24	621	10	253	4	726	11	220	4	
sl. pl:n ed.	105	28	42	11	18	5	43	11	15	4	
sl. pl:n sis.	68	51	7	5	1	1	13	10	1	1	
sl. pl:n jälj.	744	37	254	13	67	3	255	13	64	3	
sl. pl:n lopussa	31	15	16	8	2	1	23	12	4	2	
yht.	948	34	319	12	88	3	334	2	84	3	
		uusi		i			yht.				
		%		%							
pl yksin		503	18	441	16		2783				
pl. rinn:n alussa		143	15	119	12		960				
pl. keskellä		92	15	228	38		599				
pl. lopussa		125	13	293	30		967				
yht.		863	14	1081	17		5309				
sl. pl:n edellä		34	9	127	33		384				
sl. pl:n sisällä		17	13	27	20		134				
sl. pl:n jälj.		233	12	378	19		1995				
sl. pl:n lopussa		19	9	108	58		203				
yht.		303	11	640	24		2716				

Selvimmät erot muodostuvat eri virkeasemissa tunnettujen (ensimmäinen sarake) ja kokonaan uusien konstituenttien välille. Suhteellisesti eniten tunnettua, aiemmin mainittua tietoa sisältäviä subjekteja on päälauseen sisällä olevissa sivulauseissa (puolet), kun taas virkkeen loppua kohden tunnettujen subjektien määrä vähenee, oli sitten kyseessä pää- tai sivulause. Kun myös kokonaan uusia subjekteja on vähän virkkeen loppua kohti, on pohdittava, mitä virkkeen viimeisten lauseiden subjektit oikein ovat. Varsin suuri määrä niistä on infinitiivi- ja lausesubjekteja, mikä tekee niistä tunnettuuden kannalta irrelevantin (vrt. saraketta i). Selvänä

tuloksena on myös pidettävä sitä, että pragmaattisesti, siis tilanteesta käsin tunnetut ja epäsuorasti, implikoidusti tunnetut konstituentit eivät näytä lainkaan olevan riippuvaisia virkeasemasta.

TAULUKKO 78. *Sanajärjestys ja subjektin tunnettuus*

	mainittu		"heik." mainittu		asiain- tila		impli- koitu		edelly- tetty		uusi		i		yht.
	%		%		%		%		%		%		%		
SVX	1907	39	535	11	250	5	612	13	218	5	391	8	939	20	4852
SXV	17	28	12	20	-		7	12	-		7	12	17	28	60
VXS	60	34	23	13	13	7	28	16	6	3	19	11	27	15	176
VXS	4	2	12	5	1	0	12	5	1	0	48	21	155	67	233
XSV	179	52	17	5	15	4	13	4	5	1	22	6	94	27	345
XVS	82	7	148	13	6	1	191	17	41	4	482	43	187	16	1137
SV	213	29	116	16	47	6	142	19	28	4	83	11	114	15	743
VS	20	9	50	23	6	3	40	18	5	2	74	34	22	10	217
yl-S:lla	31		12		4		13		4		15		21		

Huomiota kiinnittää ensiksi sanajärjestystyyppi XSV (*Rahaa huoltaja tarvitsee.*) Subjekteista peräti 52 % on tunnettuja, ja jos jätetään irrelevantit huomiotta, luku nousee 70 %:iin. Vastaavasti uusia tarkoitteita on tässä lauseasemassa vain 22 kplettä, mikä on 0,3 % aineiston subjektin koko määrästä. Tässä tilastollinen arvo lähestyy lukemia, joiden nojalla voidaan puhua säännöstä: jos subjektin edelle on topikaalistettu toinen ydinkonstituentti, subjekti ei yleensä sisällä uutta informaatiota. Vastaavasti tyyppin VXS, jossa niin ikään subjektia edeltää toinen ydinkonstituentti mutta subjekti on lauseenloppuinen, tunnetut tarkoitteet ovat äärimmäisen harvinaisia (aineistossamme vain neljä esimerkkiä). Vaikka tyyppin SVX subjektikin sisältää suhteellisen harvoin uutta informaatiota, tämä luku on kuitenkin absoluuttisesti merkittävän suuri ja muodostaa aineiston kaikista subjekteista viitisen prosenttia. Ns. eksistentiaalilauseen perussanajärjestystyyppissä XVS subjekti sisältää uutta informaatiota 43 %:ssa tapauksista. Jos vähennämme kokonaismäärästä jälleen irrelevantit (siis infinitiivi- ja lausesubjektit), osuus nousee 50 %:iin, mitä voi pitää odottamattoman alhaisena. Kiinnostava on myös lauseenloppuisten ja intransitiivilauseiden subjektien muita korkeampi leksikaalisten sidoskeinojen käyttö, vaikka tätä onkin vaikea selittää.

Edellä todettiin, että suhteellisesti eniten tunnettuja subjekteja on sellaisessa sanajärjestystyyppissä, jossa subjektin eteen on topikaalistettu jokin toinen ydinjäsen. Seuraava taulukko osoittaa, miten objekti, predikatiivi ja adverbiaali eroavat tässä suhteessa toisistaan.

TAULUKKO 79. *Subjektin tunnettuus, objektin, predikatiivin ja adverbiaalinen topikaalistus*

subjekti on	objekti topikaal.		predik. topikaal.		advl. topikaal.		muuta		yht.
		%		%		%		%	
mainittu	97	12	27	3	559	71	31	4	784
"heik." main.	26	8	24	7	267	77	27	8	347
asiointi	3	4	9	11	70	81	4	5	86
implikoitu	20	5	45	12	276	73	37	10	379
edellytetty	13	14	9	10	58	64	10	11	90
uusi	17	11	91	13	426	62	73	11	684
i	91	16	72	13	341	61	47	8	556
topik. yleensä		19		6		60		16	

Tulokset ovat yllättäviä. Vain adverbiaalit topikaalistuvat herkästi kun subjektin tarkoite on tunnettu, objekti ja predikatiivi sen sijaan keskimääräistä vähemmän. Lisäselvyyttä tähän asiaan saataneen, kun loput ristiintaulukoinneista ovat käytettävissä.

Viimeisessä taulukossa katsotaan, mikä sidoskeino on ollut suhteellisesti yleisin etäisyyden korrelaatiosta vaihdellessa. Taulukossa on otettu huomioon vain selvät anaforiset sidoskeinot ja jätetty pois sarakkeet "katafora", "vapaa" ja "eksofora". Ne on kuitenkin laskettu kokonaismäärään mukaan.

TAULUKKO 80. *Korrelaatiin etäisyys ja konstituentin sidoskeino*

SUBJEKTI	ellipsi		anafora		toisto		komp.		synon.		yht.
		%		%		%		%		%	
ed. lause	46	5	708	79	95	11	13	1	38	4	901
ed. virke	34	3	444	33	640	48	22	2	178	13	1384
sama yht.	-	-	68	5	1158	76	18	1	264	17	1517
kauemp.	-	-	1	3	27	87	-	-	2	7	31
i	1	0	8	1	9	1	10	1	3	0	1715
yhteensä	81	1	1229	22	1929	35	63	1	485	9	5492
(koko S)	83	1	1351	17	2237	28	93	1	511	6	8022

  

OBJEKTI	ellipsi		anafora		toisto		komp.		synon.		yht.
		%		%		%		%		%	
ed. lause	21	5	284	73	46	12	5	1	32	8	391
ed. virke	7	1	133	26	248	49	16	3	95	19	505
sama yht.	2	0	24	3	574	75	10	1	139	18	761
kauemp.	-	-	-	-	18	82	-	-	4	18	22
i	1	0	7	1	17	2	12	1	1	0	1086
yhteensä	31	1	448	16	903	33	43	2	271	10	2765
(koko O)	34	1	506	11	1101	24	60	1	289	6	4577

	ellipsi		anafora		toisto		komp.		synon.		yht.
<i>ALKUADV.L.</i>											
ed. lause	8	2	419	92	19	4	3	1	9	2	458
ed. virke	4	1	173	49	124	35	9	3	40	11	356
sama yht.	1	0	38	9	313	74	6	1	66	16	426
kauemp.	-	-	-	-	11	73	-	-	3	20	15
i	1	0	193	12	46	3	252	16	4	0	1569
yhteensä	14	1	823	29	513	18	270	10	122	4	2824
(koko A-A)	14	0	860	26	558	17	277	8	289	6	3304

	ellipsi		anafora		toisto		komp.		synon.		yht.
<i>LOPPUADV.L.</i>											
ed. lause	4	6	26	38	23	34	3	4	10	15	68
ed. virke	1	1	21	12	96	55	4	2	45	26	176
sama yht.	1	0	8	2	377	76	2	0	101	20	498
kauemp.	-	-	1	6	12	70	2	12	2	12	17
i	3	0	28	1	142	6	73	3	12	1	2303
yhteensä	9	0	84	3	650	21	84	3	170	6	3062
(koko L-A)	14	0	102	2	760	18	96	2	189	4	4327

Ellipsi on selvimmin sidoskeino, jota käytetään vain vierekkäisten lauseiden kesken, joko virkerajan yli tai virkkeen rajoissa. Anaforinen pronominiikin esiintyy vain hajatapauksissa korrelaatistaan tätä etäämmällä; useimmiten tällainen on mahdollista subjektissa ja alkuadverbiaalissa. Toistoa käytetään vähän vierekkäisten lauseiden välillä; vähintään edellytetään tällöin virkerajaa, jotta sama elementti voidaan toistaa. Tavallisimmin kuitenkin korrelaatti on esiintynyt samassa yhteydessä mutta kauempana. Synonymia asettuu anaforan ja toiston välimaille: sitä käytetään kun korrelaatti on edellisessä virkkeessä taikka samassa yhteydessä, mutta vähän jos korrelaatti on joko edellisessä lauseessa taikka kauempana tekstissä. Lauseenjäsenten välillä ei ole erityisen suuria eroja. Edellä jo todettiin, että subjekti ja alkuadverbiaali sallivat enemmän anaforan käyttöä kuin etäisyys on enemmän kuin yksi virke korrelaattiin. Vastaavasti loppuadverbiaalissa on runsaammin toistoa. Sikäli lauseenjäsenten erot toki säilyvät, että anaforaa on paljon runsaammin (92 % sidoksista) alku- kuin loppuadverbiaalissa (38 %); vastaavat luvut näiden kokonaismääristä ovat 26 % ja 2 %. Voidaan lopuksi olettaa, että tässä tavatut suuntaukset ovat yleispäteviä, tekstilajista toiseen yleistettävissä, kun sen sijaan jonkin sidoskeinon runsaus saattaa vaihdella tyyllilajista toiseen, siis esimerkiksi toiston käyttö joko tekstiä selventämässä taikka tyyllittelyn keinona, synonyymien käyttö kaunokirjallisiin tarkoituksiin jne.

## KIRJALLISUUTTA

Aarts, J. 1978. Syntactic Coding of a Computer Corpus. Paper read at the 8th AILA Congress, Montreal, 1978.

Allen, Sture 1970. *Nusvensk frekvensordbok*. Almqvist & Wicksell, Stockholm.

Allwood, Jens 1977. A Critical Look at Speech Act Theory. Ks. Dahl (1977, toim.).

Andersson, Erik 1978, toim. *Working Papers on Computer Processing of Syntactic Data*. Publications of the Åbo Akademi Research Foundation, Nr. 3 . Åbo.

Andersson, Erik & Östman, Jan-Ola 1978. Computer Processing of Swedish Syntactic Data: Some Preliminaries and Tentative Results. Ks. Andersson (1978, toim.).

Anward, Jan 1976. Vem har ordet och vad innebär det? Ks. Karlsson (1976, toim.).

Bergenholtz, H. & Schaefer, B. 1978, toim. *Text-Corpora: Materialien für eine empirische Sprach- und Literaturwissenschaft*. Kronberg, Taunus.

Cedergren, Henrietta & Sankoff, David 1974. Variable Rules: Performance as a Statistical Reflection of Competence. *Language* 50:2, 333-55.

Dahl, Östen 1977, toim. *Logic, Pragmatics and Grammar*. University of Göteborg, Department of Linguistics, Göteborg.

van Dijk, Teun A. 1977. *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. Longman, London.

Einarsson, Jan 1978. *Talad och skriven svenska*. Studentlitteratur, Lund.

Ellegård, Alvar 1978. *The Syntactic Structure of English Texts. A Computer-Based Study of Four Kinds of Text in the Brown University Corpus*. Acta Universitatis Gothoburgensis. Gothenburg Studies in English 43. Göteborg.

Enkvist, Nils Erik 1975. *Tekstilngvistiikan peruskäsitteitä*. Gaudeamus, Helsinki.

Enkvist, Nils Erik & von Wright, Marianne 1978. Problems in the Study of Textual Factors in Topicalization. Ks. Andersson (1978, toim.).

Gellerstam, Martin 1970. *Jämförande vokabulärstudier i svensk morgonpress 1965*. Göteborgs universitet, Inst för nordiska språk. Moniste.

Guiraud, Pierre 1954. *Les caractères statistiques du vocabulaire*. Paris  
--- 1959. *Problèmes et méthodes de la statistique linguistique*. Dordrecht, Holland.

- Hakanen, Aimo 1972. Normaalilause ja eksistentiaalilause. *Sananjalka* 14:36-76.
- Hakulinen, Auli 1973. Semanttisia huomioita lauseenvastikkeista. *Sananjalka* 15:38-68.
- 1976. *Reports on Text Linguistics: Suomen kielen generatiivista lauseoppia*. Publications of the Åbo Akademi Research Foundation, Nr. 7. Åbo.
- 1978. Nollien syntaksia. Ks. *Rakenteita. Juhlakirja Osmo Ikolan 60-vuotispäiväksi 8.2.1978*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 6. Turku.
- Hakulinen, Auli & Karlsson, Fred 1979. *Nykysuomen lauseoppia*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- 1980. Finnish Syntax in Text: Methodology and Some Results of a Quantitative Study. *Nordic Journal of Linguistics* 3:2.
- Halliday, M. A. K. & Hasan, Ruqaiya 1976. *Cohesion in English*. Longman, London.
- Hansson, Åke 1973. *Eterspråkets grammatik I. Textkaraktäriserande aspekter*. Lundastudier i nordisk språkvetenskap 6:4-34
- Hooper J. & Thompson S. A. 1973. On the Applicability of Root Transformations. *Linguistic Inquiry* 4:465-97.
- Hultman Tor G. & Westman Margareta 1977. *Gymnasistsvenska*. LiberLäromedel, Lund.
- Ikola, Osmo 1964. *Lauseopin kysymyksiä*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- 1975. Mechanical Treatment of Syntactic Material in Finnish Dialects. Ks. Valmen Hallap, toim., *Congressus Secundus Internationalis Fenno-Ugristarum, Pars I, Acta Linguistica*. Valgus, Tallinn.
- Ikola, Osmo & Karjalainen, Yrjö 1977. Syntactic Archives of Finnish Dialects. Ks. Zampolli & Calzolari (1977, toim.).
- Itkonen, Terho 1974-75. Ergatiivisuutta suomessa I-II. *Virittäjä* 78:379-98 ja *Virittäjä* 79:31-65.
- Jörgensen, Nils 1970. *Om makrosyntagmer i informell och formell stil*. Studentlitteratur, Lund.
- 1976. *Meningsbyggnaden i talad och skriven svenska*. Studentlitteratur, Lund.
- 1977. Studium över textbyggnaden i talad svenska. *Nysvenska studier* 57:31-9.
- Karlsson, Fred 1978. Nominaalilausekkeen tematiikka ja eksistentiaalilauseiden ongelma. Ks. *Rakenteita. Juhlakirja Osmo Ikolan 60-vuotispäiväksi 8.2.1978*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 6. Turku.

Karlsson, Fred 1976 toim. *Papers from the Third Scandinavian Conference of Linguistics*. Text Linguistics Research Group, Turku.

Kiefer, Ferenc 1977. Review of Janos S. Petöfi and Hannes Rieser, toim., *Studies in Text Grammar*. *Journal of Pragmatics* 1:177-92.

Kohonen, Viljo 1976. A Note on Factors Affecting the Position of Accusative Objects and Complements in Aelfric's *Catholic Homilies* I. Ks. Enkvist, Nils Erik & Kohonen, Viljo, toim., *Reports on Text Linguistics: Approaches to Word Order*. Publications of the Research Institute of the Åbo Akademi Foundation, Nr. 8. Åbo.

Kohonen, Viljo & Salmela, Jussi 1978. Aineiston valinnan ja automaattisen tietojenkäsittelyn ongelmia kielitieteellisessä tutkimuksessa. Ks. Andersson (1978, toim.).

Kucera, Henry & Francis, W. Nelson 1967. *Computational Analysis of Present-Day Standard English*. Brown University Press, Providence, Rhode Island.

Labov, William 1972. *Sociolinguistic Patterns*. Basil Blackwell, Oxford.

Leskinen, Heikki & Savijärvi, Ilkka & Särkkä, Tauno 1974. *Koululaisten kirjallisen ilmaisutaidon kehityksestä*. Jyväskylän yliopiston suomen kielen laitoksen julkaisuja 8. Jyväskylä.

Loman, Bengt & Jörgensen, Nils 1971. *Manual för analys och beskrivning av makrosyntagmer*. Studentlitteratur, Lund.

Lyons, John 1977. *Semantics* I-II. Cambridge University Press, Cambridge.

Matihaldi, Hilikka-Liisa 1979. *Nykysuomen modukset I. Kvalitatiivinen analyysi*. Acta Universitatis Ouluensis. Series B. Humaniora N:o 7. Philologica N:o 2. Oulu.

Platzack, Christer 1974. *Språket och läsbarheten*. Gleerup, Lund.

Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan 1972: *A Grammar of Contemporary English*. Longman, London.

Quirk Randolph & Svartvik Jan 1978. *A Corpus of Modern English*. Ks. Bergenholtz & Schaeder (1978, toim.).

Saari Mirja 1975 *Talsvenska Studier i nordisk filologi* 60. Helsingfors.

Salmela, Jussi & Kohonen, Viljo 1977. CHITAB - A "Poor Man's" Shortcut to Computer Processing of Linguistic Data. Ks. Gellerstam, Martin, toim., *Nordiska datalingvistikdagar 1977*. Rapporter från Språkdata, Göteborgs universitet, 3/1977.

Saukkonen, Pauli 1976. *Kielen mikrosysteemit*. Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja 1. Oulu.

--- 1977. *Nykysuomen saneiston yleisyytilastoa saneenloppuisessa aakkosjärjestyksessä*. Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja 9. Oulu.

Savijärvi, Ilkka 1977. *Redundanssi ja kieltoverbin ellipsi suomen kielen negaatiojärjestelmässä*. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 14. Jyväskylä.

Strand, Hans 1977. Projektet Tidningsspråk. *Meddelanden från Institutionen för nordiska språk vid Stockholms universitet* 1:40-51.

Svartvik, Jan 1980. Computer-Aided Grammatical Tagging of Spoken English. Survey of Spoken English, University of Lund.

Teleman, Ulf 1974. *Manual för grammatisk beskrivning av talad och skriven svenska*. Studentlitteratur, Lund.

Tuomikoski, Risto 1978. Objektinsijaisista määrän adverbialleista. *Virittäjä* 82: 26-51.

Urmson, J. O. 1963. Parenthetical verbs. Ks. Caton, J. toim., *Philosophy and Ordinary Language*. University of Illinois Press, Urbana.

Westman, Margareta 1974. *Bruksprosa. En funktionell stilanalys med kvantitativ metod*. LiberLäromedel, Lund.

Vilkuna, Maria 1980. Geneerisyys. Käsikirjoitus.

Virkkunen, Hilikka 1977. *Essiivi, translatiivi, abess#iivi, komitatiivi ja <uinstruktiivi 1960-luvun suomen yleiskielessä*. Oulun yliopiston suomen ja saamen kielen laitoksen tutkimusraportteja 1-1. Oulu.

Yli-Vakkuri, Valma 1972. Muoto- ja lauseopillinen analysointiohjelma tietokoneaineistoa varten. Moniste. Tampereen yliopiston suomen kielen laitos.

Zampolli, A. & Calzolari, N. 1977, toim. *Computational and Mathematical Linguistics*. Proceedings of the International Conference on Computational Linguistics, Pisa 27.8. - 1.9.1976. Firenze.

## LIITE 1: Kooditus suunnitelma

### MUUTTUJA ALAKATEGORIAT

- 1 tekstin numero
- 2 virkkeen numero tekstissä
- 3 lauseen numero virkkeessä
- 4 lauseiden lukumäärä virkkeessä
- 5 sivulauseen matriisilauseen numero
  
- 6 lauseen paikka virkkeessä
  - 1 päälause yksin tai ensimmäisenä
  - 2 päälause rinnastuksen alussa
  - 3 päälause keskellä
  - 4 päälause lopussa
  - 5 sivulause päälauseensa edellä
  - 6 sivulause päälauseensa sisällä
  - 7 sivulause päälauseensa jäljessä
  - 8 sivulause virkkeen lopussa
  
- 7 konstruktio tyyppi
  - 1 yksinkertainen
  - 2 kompleksinen
  - 3 rinnasteisen konstruktion yksinkertainen alkuosa
  - 4 rinnasteisen konstruktion yksinkertainen keskiosa
  - 5 rinnasteisen konstruktion yksinkertainen loppuosa
  - 6 rinnasteisen konstruktion kompleksinen alkuosa
  - 7 rinnasteisen konstruktion kompleksinen keskiosa
  - 8 rinnasteisen konstruktion kompleksinen loppuosa
  - 9 irrallinen parenteesi
  
- 8 lausetyyppi
  - 1 päälause
  - 2 konjunktioilla alkava päälause
  - 3 sitaatti (pää- tai sivulause)
  - 4 relatiivilause subjekti-NP:ssä
  - 5 relatiivilause objekti-NP:ssä
  - 6 relatiivilause muussa NP:ssä
  - 7 *että*-lause
  - 8 epäsuora kysymyslause
  - 9 adverbiaalilause
  - A komparatiivilause (pää- tai sivulause)
  - B partisiippirakenne
  - C temporaalirakenne
  - D "väärin" käytetty modaalirakenne
  - E subjektina oleva infinitiivirakenne

## 9 lausemuoto

- 1 väitelause
- 2 *-kO*-kysymys
- 3 hakukysymys
- 4 käskylause
- 5 huudahduslause

## 10 modaalisuus

- 1 verbimodus (*-isi-*, *-ne-*)
- 2 modaalinen apuverbi verbiketjussa
- 3 modaalinen apuverbi
- 4 modaalinen predikaatti

## 11 tekstifunktio

- 1 tiedotus, toteamus, kuvaus
- 2 selitys, tarkennus, esimerkki
- 3 päätelmä
- 4 arvio, arvoarvostelma
- 5 käsky, kieltä, ohje
- 6 vetoamus, toivomus
- 7 kysymys
- 8 vastaus, responsi
- 9 metateksti

## 12 geneerisyys

- 1 spesifinen lause
- 2 geneerinen lause
- i irrelevantti

## 13 pragmaattiset elementit

- 1 *-hAn*
- 2 verbi + *-kin*
- 3 liikkuva *-kin* alussa
- 4 liikkuva *-kin* lopussa
- 5 *-pA*
- 6 yhdistelmiä (*kyllä ... -hAn* jne.)
- 7 *kyllä*
- 8 *sitä*
- 9 *toki*

## 14 kieltä

- 1 ei kieltä
- 2 kieltosanakielto
- 3 *tuskin*, *tokkopa*

## 15 lauserakenne

- 1 aktiivi transitiiivi
- 2 aktiivi intransitiivi
- 3 passiivi transitiiivi
- 4 passiivi intransitiivi
- 5 omistusrakenne
- 6 *pojasta tuli mies* -tyyppi
- 7 sää, olotila
- 8 psyykkinen kausatiivi
- 9 idiomi, kangistuma
- i fragmentti

## 16 lauseen sanamäärä

## 17 lauseen konstituenttimäärä

## 18 pintajärjestys

- 1 SVX
- 2 SXV
- 3 VSX
- 4 VXS
- 5 XSV
- 6 XVS
- 7 SV
- 8 VS
- 9 VX
- A XV
- B V
- i irrelevantti

## 19 subjektin rakenne

- 1 rinnasteinen
- 2 pääsana + lause
- 3 lause
- 4 infinitiivi
- 5 nominaalistus
- 6 N + selittävä appositio, *kuten*, tms.
- 7 N + adverbiaali tai infinitiivi
- 8 partisiippi + N
- 9 attribuutti/attribuutteja + N
- A N
- B inkorporoitu verbiin
- C geneerinen 0
- D ellipsi

20 subjektin sija ja kvanttori

- 1 nominatiivi
- 2 partitiivi
- 3 genetiivi
- 4 lukusana (+ N)
- 5 *paljon, vähän, liikaa* (+ N)
- 6 *kaikki* (+ N)
- 7 *eräs, harva, joku, moni* (+ N)
- 8 lohjennut partitiivisubjekti
- i irrelevantti

21 subjektin referentin tunnettuus

- 1 mainittu
- 2 "heikosti" mainittu
- 3 asiaintila, lause, juttu
- 4 implikoitu (tutussa merkityskentässä)
- 5 edellytetty (kielenulkoisen tiedon perusteella)
- 6 uusi
- i irrelevantti

22 subjektin referentin mainittuus

- 1 mainittu edellisessä lauseessa
- 2 mainittu edellisessä virkkeessä
- 3 mainittu samassa yhteydessä kauempana
- 4 mainittu eri yhteydessä, hyvin kaukana
- i irrelevantti

23 subjektikonstituentin tekstuaalinen sidoskeino

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparatiivinen tai muu kieliopillinen
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen koheesio
- 8 vapaa
- 9 eksofora
- i irrelevantti

24 subjektin sanamäärä 10A 11B 12C 13D 14E 15F 16G 17H

- i irrelevantti

## 25 verbiryhmän rakenne

- 1 yksi sana
- 2 liittomuoto
- 3 verbiketju
- 4 verbiketju liittomuodossa
- 5 *on aika/määrä + INF, katsoo hyväksi + INF*
- 6 5 liittomuodossa
- 7 fraasiverbi
- 8 7 liittomuodossa
- 9 sanapari
- A 9 liittomuodossa
- B finiittiverbin ellipsi

## 26 verbin monikkokongruenssi

- 1 kongruoi
- 2 ei kongruoi
- i irrelevantti

## 27 objektin rakenne

- 1 rinnasteinen
- 2 pääsana + lause
- 3 lause, lauseenvastike
- 4 nominaalistus
- 5 N + selittävä appositio
- 6 N + adverbiaali tai infinitiivi
- 7 partisiippi + N
- 8 attribuutti/attribuutteja + N
- 9 N
- A geneerinen 0
- B ellipsi

## 28 objektin sija ja kvanttori

- 1 nominatiivi
- 2 partitiivi
- 3 genetiivi
- 4 -t
- 5 lukusana (+ N)
- 6 *paljon, vähän, liikaa (+ N)*
- 7 *kaikki (+ N)*
- 8 *eräs, harva, joku, moni (+ N)*
- 9 lohjennut partitiiviobjekti
- i irrelevantti

29 objektin referentin tunnettuus

- 1 mainittu
- 2 "heikosti" mainittu
- 3 asiaintila, lause, juttu
- 4 implikoitu
- 5 edellytetty
- 6 uusi
- i irrelevantti

30 objektin referentin mainittuus

- 1 mainittu edellisessä lauseessa
- 2 mainittu edellisessä virkkeessä
- 3 mainittu samassa yhteydessä kauempana
- 4 mainittu eri yhteydessä, hyvin kaukana
- i irrelevantti

31 objektikonstituentin tekstuaalinen sidoskeino

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparaatio tai muu kieliopillinen
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen koheesio
- 8 vapaa
- 9 eksofora
- i irrelevantti

32 objektin sanamäärä

33 predikatiivin rakenne

- 1 A
- 2 määrite + A, A + määrite
- 3 A + A...
- 4 A + lause
- 5 vertailu (*kuin x*)
- 6 N
- 7 määrite + N, N + määrite
- 8 N + N...
- 9 N + lause
- A lause
- B nominaalitus
- C ellipsi

34 predikatiivin sija

- 1 nominatiivi
- 2 partitiivi
- 3 genetiivi
- 4 essiivi
- 5 translatiivi
- 6 kvanttori (+ N)
- i irrelevantti

35 predikatiivin lukukongruenssi

- 1 yksikkö - yksikkö
- 2 yksikkö - monikko
- 3 monikko - yksikkö
- 4 monikko - monikko
- i irrelevantti

36 predikatiivin tekstuaalinen sidoskeino

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparaatio
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen koheesio
- 8 vapaa
- 9 eksofora
- i irrelevantti

37 predikatiivin sanamäärä

38 alkuadverbiaalin rakenne

- 1 yksi sana
- 2 lauseke
- 3 pre- tai postpositiolauseke
- 4 III infinitiivi tai nominaalitus
- 5 lauseenvastike
- 6 pääsana + lause
- 7 lause
- 8 rinnasteinen
- 9 ellipsi

39 alkuadverbiaalin tyyppi

- 1 valenssi
- 2 kehys
- 3 kommentti
- 4 konnektiivi
- 5 pakollinen predikatiiviadverbiaali
- 6 irrallinen predikatiiviadverbiaali
- 7 muu
- i irrelevantti

40 alkuadverbiaalin merkitys

- 1 aika
- 2 paikka
- 3 habitiivi
- 4 tapa
- 5 muu materiaallinen
- 6 määrä: OSMA
- 7 modaalinen
- 8 muu
- i irrelevantti

41 alkuadverbiaalin referentin tunnettuus

- 1 mainittu
- 2 "heikosti" mainittu
- 3 asiaintila, lause, juttu
- 4 implikoitu
- 5 edellytetty
- 6 uusi
- i irrelevantti

42 alkuadverbiaalin referentin mainittuus

- 1 mainittu edellisessä lauseessa
- 2 mainittu edellisessä virkkeessä
- 3 mainittu samassa yhteydessä kauempana
- 4 mainittu eri yhteydessä, hyvin kaukana
- i irrelevantti

43 alkuadverbiaalin tekstuaalinen sidoskeino

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparaatio
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen koheesio
- 8 vapaa
- 9 eksofora
- i irrelevantti

44 alkuadverbiaalin sanamäärä

45 loppuadverbiaalin rakenne

- 1 yksi sana
- 2 lauseke
- 3 pre- tai postpositiolauseke
- 4 III infinitiivi tai nominaalistus
- 5 lauseenvastike
- 6 pääsana + lause
- 7 lause
- 8 rinnasteinen
- 9 ellipsi

46 loppuadverbiaalin tyyppi

- 1 valenssi
- 2 kehys
- 3 kommentti
- 4 konnektiivi
- 5 pakollinen predikatiiviadverbiaali
- 6 irrallinen predikatiiviadverbiaali
- 7 muu
- i irrelevantti

47 loppuadverbiaalin merkitys

- 1 aika
- 2 paikka
- 3 habitiivi
- 4 tapa
- 5 muu materiaallinen
- 6 määrä: OSMA
- 7 modaalinen
- 8 muu
- i irrelevantti

48 loppuadverbiaalin referentin tunnettuus

- 1 mainittu
- 2 "heikosti" mainittu
- 3 asiaintila, lause, juttu
- 4 implikoitu
- 5 edellytetty
- 6 uusi
- i irrelevantti

49 loppuadverbiaalin referentin mainittuus

- 1 mainittu edellisessä lauseessa
- 2 mainittu edellisessä virkkeessä
- 3 mainittu samassa yhteydessä kauempana
- 4 mainittu eri yhteydessä, hyvin kaukana
- i irrelevantti

50 loppuadverbiaalin tekstuaalinen sidoskeino

- 1 ellipsi
- 2 anafora
- 3 katafora
- 4 toisto
- 5 komparaatio
- 6 yläkäsite, synonyymi, parafraasi
- 7 leksikaalinen koheesio
- 8 vapaa
- 9 eksofora
- i irrelevantti

51 loppuadverbiaalin sanamäärä

52 adverbiaalien määrä lauseessa

53 uusien referenttien määrä lauseessa

54 topikaalistettu alkukonstituentti

- 1 lausekeobjekti
- 2 lauseobjekti
- 3 osa lauseobjektista
- 4 predikatiivi
- 5 osa predikatiivista
- 6 adverbiaali
- 7 osa adverbiaalista
- 8 adverbiaali + objekti
- 9 objekti + adverbiaali
- A adverbiaali + predikatiivi
- B predikatiivi + adverbiaali
- C adverbiaali + adverbiaali
- D NP (tai kaksi) lauseenvastikkeesta yli hallitsevan lauseen

55 alkukonstituentin topikaalistuksen todennäköinen syy

- 1 kompensatorinen
- 2 sidonta
- 3 emfaasi
- 4 kontrasti
- 5 syntaktinen yhdenmukaisuus
- 6 ei mainittavaa syytä - tunnusmerkitön?
- 7 kompensatorinen + sidonta
- 8 kompensatorinen + emfaasi
- 9 kompensatorinen + kontrasti
- A sidonta + emfaasi
- B sidonta + kontrasti
- C tunnusmerkitön + kompensatorinen
- D tunnusmerkitön + sidonta
- E tunnusmerkitön + kontrasti
- F tunnusmerkitön + emfaasi
- G emfaasi + kontrasti,

i

56 subjektin lykkäys

- 1 intransitiivisubjekti
- 2 transitiivisubjekti
- 3 lausesubjekti
- 4 infinitiivisubjekti
- 5 subjektin osa
- 6 kvanttorin siirto

57 verbin siirto

- 1 alkuun
- 2 loppuun
- 3 molemmat
- 4 infinitiivi verbiketjusta alkuun

58 dislokaatio

- 1 jonkinlainen, oikealle
- 2 jonkinlainen, vasemmalle

59 ellipsit

- 1 aukko: finiittiverbi
- 2 aukko: V + objekti
- 3 aukko: V + adverbiaali
- 4 aukko: V + muuta
- 5 VP:n ellipsi
- 6 VP:n ellipsi + *niin ... -kin*

60 subjektin semanttinen status

- 1 elollinen olio
- 2 esine
- 3 paikka
- 4 aine
- 5 kollektiivi
- 6 teko, tapahtuma, prosessi (II aste)
- 7 abstrakti (III aste)
- i irrelevantti

61 objektin semanttinen status (vrt. (60))

62 alkuadverbiaalin semanttinen status (vrt. (60))

63 loppuadverbiaalin semanttinen status (vrt. (60))

64 korpuksen tekstityypit (*lisätty vuonna 2010*)

- 1 tietokirjat
- 2 Otavan suuri ensyklopedia
- 3 pääkirjoitukset
- 4 kulttuuriartikkelit
- 5 pakinat
- 6 tiedottavat artikkelit

## LIITE 2: Analysoitujen tekstien luettelo

1. TIETOKIRJAT	LAUSEET	VIRKKEET
1. Vygotski tietoisuuden tutkijana. Helkama - Koski-Jännes, Marxilainen sosiaalipsykologia.	46	23
2. Sosiaalistumisen vaiheet. Helkama - Koski-Jännes, Marxilainen sosiaalipsykologia.	108	45
3. Elokuvan tutkimisesta. Olli Alho, teoksessa Aika on aikaa.	73	28
4. Milloin pamfletteja tarvitaan. Pertti Hemanus, teoksessa Eskola-Eskola, Kirjallisuus Suomessa	75	30
5. Musiikki teollistuneissa yhteiskunnissa. Pekka Gronow, teoksessa Aika on aikaa.	94	51
6. Pilasadut, kaskut, huulet. Seppo Knuuttila, teoksessa Folklore tänään.	127	60
7. Kolonialistinen järjestelmä toiminnassa. Timo Melasuo, teoksessa Kehityskaatiedon perusteet.	140	67
8. Suomalainen tutkimus. Nordenstreng (toim.), Joukkotiedotus ja yleisö.	181	71
9. Luonnon ja ympäristömme muuttuminen. Ilmari Hustich, Suomi tänään.	183	95
10. Ongelmallinen täystyöllisyys. Jouko Siipi, Ryysyrannasta hyvinvointivaltioon.	112	58
11. Henkiparannus: alkusanat. Irma Weisen, Henkiparannus.	131	59
	<hr/>	<hr/>
	1270	587
2. ENSYKLOPEDIAT (Otavan Suuri Ensyklopedia)		
12. Afasia	90	40
13. Aivopesu	103	50
14. Akatemit	59	28
15. Anjalan liitto	54	31
16. Apartheid	77	42
17. Delfoi	47	26
18. Don Quijote	93	45
19. Dyynit	114	59
20. Eerik XIV	82	46
21. Eläinyhteisöt	82	38
22. Elämä	122	61
23. Fagerholm, Karl August	68	31
24. Gandhi, Indira	61	32
	<hr/>	<hr/>
	1052	529

### 3. PÄÄKIRJOITUKSET (vuodelta 1977)

25. Tutkalelu kuumentaa tunteita. IS 5.9.	27	15
26. Velvoittava sopimus. HS 23.8.	39	17
27. Muutos eläkepalkkaan. HS 27.8.	32	15
28. Vesakkomyrkyt. HS 29.8.	29	16
29. Jatkokoulutus esiin. HS 4.9.	48	24
30. Vastuu ydinvoimasta. HS 13.9.	19	10
31. Myrsky maitolasissa. HS 18.9.	31	15
32. Rikollista taloutta. HS 19.9.	23	10
33. Maataloustulon tarkistus. KU 25.8.	39	20
34. Itsenäisyyden juhluvuoden teemat ja isänmaallisuus. KU 30.8.	28	15
35. Huippueläkkeiden katto-ongelma. KU 1.9.	40	21
36. Maataloustuloratkaisu. KU 2.9.	33	18
37. Sairas yhteiskunta ja työväenliike. KU 5.9.	42	20
38. Myymälävarkaudet. KU 14.9.	32	19
39. Keinottelua maidon hinnalla. KU 18.9.	39	19
40. Telakoiden tärkeys. Turun Päivälehti 23.8.	32	16
41. Suomen historiaa. Turun Päivälehti 30.8.	30	18
42. Ylieläkkeet osaongelma. Turun Päivälehti 2.9.	51	30
43. Lehtituki tasapuoliseksi. US 4.9.	40	27
44. Hajasijoitus hautaan. US 7.9.	23	13
45. Odotetaan komiteaa. US 7.9.	24	11
46. Bensiinin hinta. US 8.9.	29	18
47. Pankkien voitot. US 11.9.	34	20
48. Liikennetulpat. US 12.9.	33	20
49. Alkutuottajamaa US 13 9	28	17
50. Vesilain uudistus. US 15 9	35	21
51. Kuntien budjetti. US 16.9.	30	17
52. Ansaitsematon kiitos. US 17.9.	63	31
	953	513

### 4. KULTTUURIARTIKKELIT (vuodelta 1977)

53. Tehoa korkeakouluopetukseen. US 8.9.	110	51
54. Kulttuurivientimme kitkat. Erkki Arni, HS 18.9.	214	87
55. Taiteemme katselmus ehdolla Lontooseen. HS 28.8.	76	30
56. Kultamaahan matkaajat. Yrjö Sepänmaa, US 4.9.	93	47
57. Rakkauden aritmetiikka tunteiden labyrintissa. Heikki Eteläpää, US 6.9.	58	27
58. Kun vuodet eivät paina. Heikki Eteläpää, US 13.9.	49	19
59. Lorcan kukkien kieltä Turun teatterissa. Pirjo Vaittinen, KU 14.9.	48	24
60. Turkan Turandot: lahjomattomasti iskevä		

lamanäytelmä. Kirsikka Siikala, HS 21.8.	94	54
61. Liisi Tandefelt Käthe Kollwitzina - kuin herkkä kosketus. Markku Valkonen, HS 23.8.	83	44
62. Bunuelin huumori on yhä julmaa ja huolestunutta. Mirja Bolgär, US 11.9.	113	42
63. Piirpauke paranee yhä. Jorma Tuominen, Turun Päivälehti 24.8.	54	30
64. Rautavaaran konsertto valmis - Sarbun hetki koittaa. Seppo Heikinheimo, HS 23.8.	60	20
65. Salzburgin juhlaviikot IV - Berion Coro puhuu kaikkien puolesta. Pentti Ritolahti, US 12.9.	148	57
66. Asuntojen arkkitehti on yhä ajankohtainen. Leena Maunula, HS 31.8.	86	49
67. Bergelin uhkea Mathis. Olavi Kauko, HS 8.10.	51	24
68. Metrosaarnaaja Siri Derkert. Eeva Siltavuori, HS 18.9.	62	42
69. 1920-luvun taiteilijan vaikein saalis: todellisuus. Markku Valkonen, HS 9.10.	157	88
	<hr/>	
	1556	735

#### 5. PAKINAT (vuodelta 1977)

70. Ihan sinisiä ajatuksia. Simppa, SK 22.7.	69	45
71. Silta. Simppa, SK 8.7.	75	41
72. Välissä. Simppa, SK 18.8.	69	29
73. Kysytään luonnosta. Simppa, SK 25.8.	74	38
74. Kynnelsilmät ohjelmantekoon. Rh. Walta, SK 10.6.	87	58
75. Poliitikka laji sinänsä. Rh. Walta, SK 8.7.	110	58
76. Muuan haastattelu. Rh. Walta, SK 22.7.	108	53
77. Piirileikiksi uhkaa mennä koko elvytyshomma. Rh. Walta, SK 11.8.	97	41
78. Eräitä elvytysvihjeitä. Rh. Walta, SK 25.8.	87	45
79. Junassa. Elina, SK 10.6.	83	26
80. Lukittu Suomi. Elina, SK 17.6.	94	34
81. Žmpäri päässä. Elina, SK 24.6.	82	23
82. Timanttinauha. Elina, SK 12.8.	55	18
83. Keskustelu. Elina, SK 15.7.	91	39
84. Naapuripöydän lasku. Juorkunan Jussi, KU 19.8.	67	32
85. Kansaneläkkeet nousevat. Juorkunan Jussi, KU 20.8.	60	31
86. Varavoima. Juorkunan Jussi, KU 1.9.	62	27
87. Sortumaisillaan. Juorkunan Jussi, KU 3.9.	52	29
88. Tehtaan varjossa. Juorkunan Jussi, KU 18.9.	66	33
89. Mätä maailma. Juorkunan Jussi, KU 28.9.	86	35
90. Sokeritoukka. Anto Leikola, HS 27.8.	61	28
91. Miljoonakaivo. Timo Järviö, HS 26.8.	55	29
	<hr/>	
	1690	792

## 6. TIEDOTTAVAT ARTIKKELIT (vuodelta 1977)

92. Margareta Mickwitz: Pohjoismaiden työ arvostettu Unescossa. Marja Kyllönen, IS 5.9.	78	47
93. Seoslannoitetehtaan suunnittelu Turussa. HS 27.8.	58	46
94. Pugwash-kokous ja aikamme ongelmat. Ilmari Susiluoto, HS 5.9.	88	49
95. Maailman syvimmän salaisuuden edessä. Anto Leikola, HS 18.9.	146	50
96. Partiolaiset keräyttävät jätepaperin liikemiehillä. Reino Ylisipola, HS 28.8.	97	52
97. Pielaveden työttömät marjametsään. HS 20.9.	82	43
98. Näin on marjat. Pentti Sainio, HS 18.9.	194	91
99. Ukrainan ujo ME-mies Jashtshenko. Reetta Meriläinen, HS 26.8.	125	69
100. Ketkä olivat pääosissa kun esitettiin maaherranäytelmä. Janne Virkkunen, HS 4.9.	180	93
101. Valintamyymälä. Pirkko Kolbe, HS 21.5.76	142	76
102. Ankeriaasta arvokala rehevöityviin vesiin. HS 31.8.	77	42
103. Vaatteiden historiaa linnassa. Turun Päivälehti 3.9.	58	38
104. Nuorkauppakamari: Saaristo säilytettävä elävänä. Turun Päivälehti 27.8.	51	28
105. Ei merta edemmäksi... kotimaastakin saa lämminverihavosen. Hanna Toivonen, TS-erikoisnumero 5.9.	81	51
106. Tutkijat todistavat: Lapin ruska on erilainen. KU 30.8.	81	47
107. Presidentti päättää - onko Kyllikki uusi Pekka. Matti Saari, US 10.9.	100	54
108. Pikkupoikien uusin pop: mini trial. Ola Timonen, US 10.9.	60	31
109. Urheilu nuorten suosikki - ja sitä toivotaan lisää. Jyrki Vesikansa, US 4.9.	117	75
110. Neutronit tuottavat syöpää. Jorma K. Miettinen, SK 33/77	51	28
111. Tässä tulevat Sorsan siirtolaiset. Annikki Ukkonen, Oma Markka 7/77	147	68
112. Romahtaako asuntotuotanto vai mikä? Eero Tuomainen, Oma Markka 7/77	126	72
113. Kuntoa sienestä. Rauno Iivonen, Me 24.8.	177	81
114. Valitse huolella koulukengät. Anna-Maija Tähkävuori, Kodin kuvalehti 15/77	80	41
115. Milloinkas lapset tehdään? Aira Huuskonen, Kodin kuvalehti 18/77	113	46
116. Emme ole aggressiivisia, meidät kasvatetaan sellaisiksi. Lumikki Tammi, Kodin kuvalehti 18/77	109	58
117. He aloittivat uuden elämän. Irma Soinio,		

Kodin kuvalehti 18/77	164	68
118. Mikä on paras naimaikä. Anja Niemelä, Kaks plus 9/77	115	58
119. Istuukin jo ihan vähän. Ritva Taskinen, Kaks plus 9/77	124	64
120. Suurperheen aika on ohi. Anja Niemelä, Kaks plus 9/77	164	77
121. Mies miten hoidat ehkäisyn? Ritva Taskinen Kaks plus 9/77	146	80
122. Lue tämä ennenkuin otat kissan. Pirkko Tuominen Kauneus ja terveys 10/77	188	84
123. Alkoholismi on sairaus, josta voidaan toipua. Leo Hilden, Kauneus ja terveys 10/77	104	51
	<hr/>	
	2166	1083
<b>KAIKKI TEKSTIT YHTEENSÄ:</b>	10149	5016

1. Alakorpusten merkitsevyydestit	96
2. Lauseiden lukumäärä virkkeessä	97
3. Lausetyyppi	98
4. Lauseen sanamäärä	99
5a. Pintajärjestys	100
5b. Pintajärjestys pääryhmittäin	100
6. Subjektin konstituenttirakenne	101
7. Alatekstilajien merkitsevyyserot	178
8. Alatekstilajien merkitsevien erojen lukumäärä	103
9. Muuttujien tyylillinen erottelukyky	104
10. Lausetyypit eri alatekstilajeissa	104
11. Objektin sijaanmerkintä eri alatekstilajeissa	105
12. Adverbiaalien määrä lauseessa eri alatekstilajeissa	106
13. Predikatiivin sija ja kvanttori eri alatekstilajeissa	106
14. Lauseen paikka virkkeessä	108
15. Lausetyyppi	109
16a. Virkeasema ja päälausetyyppi	109
16b. Virkeasema ja sivulausetyyppi	110
17. Lausemuoto	110
18. Tekstifunktio	111
19. Modaalisuus	111
20. Lausemuoto ja modaalisuus	111
21. Kielto	112
22. Kielto ja lausemuoto	112
23. Kielto ja lausetyyppi	112
24. Pragmaattiset elementit	113
25. Subjektin ja objektin konstituenttirakenne	114
26. Subjektin konstituenttirakenne ja sijainti	115
27. Subjektin konstituenttirakenne ja predikatiivin sija	116
28. Predikatiivin konstituenttirakenne	116
29. Subjektin ja objektin sija	117
30. Subjektin sija ja sijainti	118
31. Kielto ja subjektin sija	118
32. Kielto ja objektin sija	119
33. Subjektin sija ja sen referentin tunnettuus	119
34. Lausekeobjektin sija ja topikaalistus	120
35. Predikatiivin sija	120
36. Predikatiivin lukukongruenssi	121
37. Alku- ja loppuadverbiaalin konstituenttirakenne	121
38. Alku- ja loppuadverbiaalin tyyppi	122
39. Alku- ja loppuadverbiaalin merkitys	123
40. Lauserakenne	123
41. Aktiivisten lauseiden lauserakenne ja subjektin konstituenttirakenne	124
42. Lauserakenne ja topikaalistus	125
43. Verbiryhmän rakenne	125

44. Verbin monikkokongruenssi	126
46. Lauseiden lukumäärä virkkeessä	127
47. Lauseen sanamäärä	128
48. Virkkeen sanamäärä	129
49. Lauseen konstituenttimäärä	130
50. Subjektin, objektin ja predikaatiivin sanamäärät	130
51. Lausekesubjektin sanamäärä ja sijainti	131
52. Subjektin sanamäärä sekä lausekeobjektin ja adverbiaalin topikaalistus	132
53. Lausekeobjektin sanamäärä ja topikaalistus	133
54. Alku- ja loppuadverbiaalin sanamäärät	133
55. Adverbiaalien määrä lauseessa	134
56. Nominaalisten lauseenjäsenten lukumäärä	134
57. Pakollisten nominaalijäsenten lukumäärä	135
58a. Pintajärjestys	136
58b. Yhdistetty pintajärjestys	136
58c. Verbin sijainti suhteessa nominaalisiin pääjäseniin	137
60a. Virkeasema ja sanajärjestys	137
60b. Virkeasema ja sanajärjestys	138
60c. Virkeasema ja verbin sijainti	139
61a. Sanajärjestys sivulauseissa	139
61b. Sanajärjestys sivulauseissa	140
61c. Verbin sijainti sivulauseissa	140
62. Topikaalistukset	141
63. Topikaalistettujen konstituenttien määrä	141
64. Topikaalistuksen todennäköinen syy	142
65. Subjektin lykkäys	143
66. Verbin siirto	143
67. Virkeaseman ja lausetyypin vaikutus topikaalistukseen	144
68. Topikaalistus sivulauseissa	144

TAULUKKO 7. Alateksilajien merkittävyydet

MUUT- TUJJA	TESTATUT ALATEKSTILAJIT →														
	1/2	1/3	1/4	1/5	1/6	2/3	2/4	2/5	2/6	3/4	3/5	3/6	4/5	4/6	5/6
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	2	11	0	0	1	0	0	0	0	11	0	59	1	12	0
14	4	0	17	2	74	0	0	0	1	0	4	0	28	18	1
16	2	5	80	0	0	77	0	0	0	2	0	0	0	0	0
17	6	16	0	0	0	79	0	0	7	0	0	9	0	0	0
18	0	15	0	0	26	0	0	0	0	0	0	9	8	0	0
19	1	0	1	0	0	1	14	0	26	0	0	0	0	14	0
20	23	96	25	0	13	50	0	0	0	9	0	6	0	35	0
23	0	0	2	0	0	0	0	0	0	5	0	1	0	0	0
24	7	29	45	0	13	8	5	0	0	60	0	0	0	0	0
25	0	0	13	7	32	0	0	0	1	0	0	0	36	5	0
27	0	0	1	0	0	3	8	0	0	0	0	0	0	0	0
28	0	64	13	0	0	0	0	0	0	10	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0
32	1	35	14	0	0	4	0	0	0	6	0	0	0	0	0
33	15	47	22	91	80	6	25	24	38	0	14	17	65	15	69
34	18	11	1	1	0	14	6	6	2	28	75	6	75	93	39
37	2	11	11	63	21	1	5	1	2	6	43	51	11	6	70
38	4	33	23	0	1	11	21	0	0	97	1	57	0	13	1
45	3	2	1	0	0	0	0	0	0	26	5	4	0	0	1
52	50	19	6	0	15	61	41	12	45	98	69	86	69	52	13
54	69	14	34	88	0	7	39	39	0	36	16	0	34	0	0

-10-