

Suomen substantiivilausekkeen rakenne¹

KRISTIINA JOKINEN

1. Johdanto

Suomen kielen substantiivilausekkeen ominaisuuksia ja rakennepiirteitä on tutkittu melko paljon, mutta sen tarkastelu lausekerakennekielioppien (Phrase Structure Grammars) kannalta on jäänyt vähemmälle (vrt. kuitenkin Hakulinen & Karlsson 1979). Tässä artikkelissa käsittelen substantiivilausekkeen rakennetta X'-teorian² pohjalta ja esittelen tietokoneelle tehtyä toteutusta kuvauksestani.

X'-teorian perusteos on Jackendoff (1977). Teorian pääpiirteet on esittänyt jo Chomsky (1970), enkä tässä käy niitä enemmälti esittelemään. Kuvaukseni on saanut vaikutteita myös yleistetystä lausekerakennekieliopista (GPSG, Generalized Phrase Structure Grammar, Gazdar 1982, Gazdar ym. 1985), joka korostaa lausekerakennesääntöjen ja syntaktisen piirteistön käsittelyssä teorian perusteiden matemaattista hallintaa. GPSG:stä on hyvä yleisesitys Vilkunan (1986) artikkelissa. Selkeä johdatus GPSG:hen ja myös muihin nykyisiin kielioppimalleihin on Sells (1986).

Artikkelini rakenne on seuraava. Määritelmän jälkeen tarkastelen lyhyesti rakenteen hajoamista (luku 3). Sitten käsittelen substantiivilausekkeen rakennetta yleisesti (luku 4) ja kiinnitän erityistä huomiota genetiiviattribuuttiin (luku 5) sekä numeraaleja sisältäviin lausekkeisiin (luku 6). Lopuksi esitän lausekerakennesäännöt, jotka olen implementoinut tietokoneelle, ja annan esimerkkejä tietokoneanalyysista.

¹ Kirjoitus perustuu substantiivilausekkeen rakennetta käsittelevään opinnäytteeseeni (Jokinen 1986), joka muutamissa vähäisissä kohdissa poikkeaa artikkelin esityksestä. Tutkimus on tehty Helsingin yliopiston tietokone-lingvistiikan tutkimusyksikössä, jossa myös on toteutettu kuvauksen tietokoneimplementaatio.

² X'-teoria eli X-bar-teoria on saanut nimensä leksikaalisen kategorian projektiotasoihin viittaavista viivoista (engl. *bar*), jotka merkitään kategoriasymbolin yläpuolelle. Esim. \bar{N} tarkoittaa ykköstason ja $\bar{\bar{N}}$ kakkostason substantiivilauseketta. Nykyisin viivat on typografisista syistä korvattu joko pilkuilla (N', N'') tai numeroilla (N1, N2), ja teorian nimi on vastaavasti muutettu X'-teoriaksi.

2. Substantiivilausekkeen määritelmä

Substantiivilausekkeella (*SbP*) tarkoitan syntaktista lauseketta, jonka pääsanana on substantiivi. Lausekkeen pääsanana ei aina suinkaan ole selvä, ja sen yleistä määrittämistä on pohdittu paljonkin.³ X'-teorian pääsanakäsite on periaatteessa erittäin selvä: leksikaalinen pääkategoria toimii pääsanana ensimmäisellä projektiotasolla ja sen syntaktiset projektiot muilla tasoilla. Toisaalta Muysken (1982) on tutkittuaan yhdeksäntoista X'-syntaksiin pohjautuvan teorian ehdotuksia pääsanakategorioiden määrittelemiseksi todennut mielenkiintoisen seikan: mitkään kaksi pääsanamäärittelyä eivät olleet keskenään samanlaisia huolimatta teorioiden melko voimakkaistakin yhteisistä perusoletuksista. Lausekkeen syntaktiseen käyttäytymiseen liittyvien ilmiöiden kuvauksessa pääsanana on joka tapauksessa tärkeä: se määrää lausekkeen kategorian, sisäisen rakenteen, suhteet muihin syntaktisiin elementteihin, ja se toimii myös lausekkeen semanttisen tulkinnan perustana. Substantiivilausekkeessa pääsanana-määrite -suhde on melko ongelmaton, tosin esim. numeraalin ja substantiivin suhde on tulkinnanvarainen (luku 6).

Teen eron substantiivilausekkeen ja nominaalilausekkeen (*NP*) välille, joita yleensä on pidetty enemmän tai vähemmän synonyymeinä. *NP* on alaltaan laajempi yleistermi, joka viittaa kaikkiin niihin konstituentteihin, jotka voivat toimia lauseen subjektina ja objektina, ts. verbin pakollisina komplementteina. *NP* ei siis ole varsinainen X'-teorian mukainen syntaktinen kategoria kuten *SbP* vaan funktionaalinen luokka. Kaikki kursivoidut konstituentit esimerkkilauseissa (1) voidaan näin ollen tulkita *NP*:iksi (nominaalilausekkeiksi), mutta mikään niistä ei ole edellä annetun määritelmän mukainen *SbP* (substantiivilauseke).

- | | |
|----------------------------------------|---------------------------------|
| (1) a. <i>Hän</i> astui ovesta sisään. | (pronomini subjektina) |
| b. <i>Uusi</i> on aina parempi. | (adjektiivi subjektina) |
| c. Tiedän <i>kaiken</i> . | (kvanttori objektina) |
| d. Haluan <i>lähteä</i> . | (infinitiivi objektina) |
| e. <i>Se, että tulit</i> on mukavaa. | (pronomini ja lause subjektina) |

Myös Hakulinen & Karlsson (1979:112) tekevät tämän eron perustellessaan *NP*:n nimeämistä nominaalilausekkeeksi, vaikka he käyttävätkin lyhennettä

³ Bloomfieldin jakoa endo- ja eksosentrisiin konstruktioihin ja sen mukaista distributionaalista määritelmää on kritisoinut mm. Lyons (1977). Toisaalta pääsanana ei ole pidetty primitiivikäsitteenä lainkaan (Vennemann & Harlow 1977, Hawkins 1984, Hoeksema 1984). Zwicky (1985) taas korostaa kieliopin kokonaisvaltaisuutta: määrittelyssä on otettava huomioon se, millaisia pääsanakäsitteitä kieliopissa ylipäättään tarvitaan, jotta hallitsemis- ja määrittäyssuhteet sekä syntaksin vuorovaikutus semantiikan, leksikon ja morfologian kanssa saataisiin kuvattua.

NP sekä substantiivi- että nominaalilausekkeesta. Koska haluan korostaa, että kuvaukseni kohde on nimenomaan substantiivin sisältävä substantiivilauseke, käytän substantiivista lyhennettä *Sb* ja substantiivilausekkeesta lyhennettä *SbP*.⁴

Artikkelissani keskityn käsittelemään prototyyppejä substantiivilausekkeita, ts. lausekkeita joiden pääsanana on johtamaton substantiivi. Nominaaliset, erityisesti *minen*-johdokset ja muut teonnimet, ovat tutkimuksen kohteena jatkotyössäni. Toisaalta raja johdoksen ja leksikaalistuneen substantiivin välillä on epäselvä. Etenkin jälkiattribuutteja tarkastellessa rajaus prototyyppeihin pääsanasubstantiiveihin ei aina liene tarpeellinen, sillä suuri osa alakategorioivista substantiiveista on juuri verbijohdoksia (*usko/uskoa johonkin*). Deverbaalisubstantiivien kuvauksessa on kuitenkin otettava huomioon mm. kantaverbin argumenttirakenne, minkä tutkiminen menee artikkelini aiheen ulkopuolelle.

Kuvaamani SbP:t ovat kirjoitetun yleiskielen mukaisia, ja niitä kuvaavat säännöt on tarkoitettu automaattisen lauseenjäsennyksen perustaksi. Lausekerakennesääntöni eivät sovi esim. runomuotoisen tekstin kuvaukseen eivätkä sellaisenaan myöskään puhekielen kuvaukseen (puhekielen rakenteet eivät välttämättä muistuta kirjoitetun kielen rakenteita, vrt. Hakulinen & Sorjonen 1986). Toisaalta ne syntaktis-semanttiset periaatteet, jotka vaikuttavat kirjoitetun yleiskielen rakenteisiin, koskevat kaikkea kielenkäyttöä, enkä näin ollen halua vetää liian jyrkkää rajaa eri kielimuotojen välille.

3. SbP hajoavana rakenteena

Substantiivilauseke on intuitiivisesti varsin selkeä kokonaisuus, onhan se sellainen kielen rakenne, jonka avulla puhuja koodaa viittaukset diskurssimaailman olioihin (ks. Hopper & Thompson 1984 sanaluokkaerottelun, erityisesti verbien ja substantiivien pragmaattis-semanttisista perusteluista). Tämä heijastuu myös SbP:n rakenteessa: se on kiinteä konstituentti, jonka määritteet muutamia poikkeuksia lukuunottamatta pysyvät tiukasti pääsanan yhteydessä. Kiinteyttä lisää vielä etuattribuuttien kongruenssi (F. Karlsson 1977).

Toisaalta substantiivilla on irtautuvia määritteitä, jotka osoittavat, että SbP:kin voi hajota muun lauseen sekaan. Tällaisiin määritteisiin kuuluvat

⁴ Luonnollisesti lyhenteet on mahdollista korvata »kansainvälisemmillä» lyhenteillä *N* ja *NP* (ja nominaalilausekkeen lyhenne esim. *NomP*:llä), sillä kyse ei ole syvällisestä teoreettisesta ongelmasta vaan ainoastaan symbolien valinnasta.

mm. ajelehtivat kvanttorit (floating quantifiers). Termi on peräisin logiikasta (ks. esim. Barwise & Cooper 1981), ja sillä viitataan kvanttoireihin, jotka syntaktiselta käyttäytymiseltään muistuttavat lauseen pääjäsenten seassa vapaasti esiintyviä lauseadverbeja. Ajelehtivia kvanttoireita ovat esim. *kaikki*, *molemmat*, *kumpikin* ja *kumpikaan*:

- (2) a. *Molemmat/kummatkin/kaikki/myös* nämä seikat ovat tärkeitä.
 b. Nämä *molemmat/kummatkin/kaikki/*myös* seikat ovat tärkeitä.
 c. Nämä seikat *molemmat/kummatkin/kaikki/myös* ovat tärkeitä.
 d. Nämä seikat ovat *molemmat/kummatkin/kaikki/myös* tärkeitä.
 e. Nämä seikat ovat tärkeitä *molemmat/kummatkin/kaikki/myös*.

Myös rajoittavat relatiivilauseet voivat lohjeta pääsanansa yhteydestä. Ilmiö on sukua viskurilaille eli raskaan substantiivilausekkeen siirrolle (heavy NP-shift), ja se on mahdollista, jos pääsana-substantiivia määrittää demonstratiivi. Demonstratiivi tekee SbP:stä definiittisen, ja lisäksi se merkitsee relatiivilauseen pääsanana, vrt. (3a–b). Pelkkä demonstratiivi ei kuitenkaan riitä disambiguoimaan relatiivilauseen pääsanana, vaan analyysiin vaikuttavat erityisesti intonaatio ja painotus (3c).

- (3) a. Se opiskelija saa lainan, jolla on tarpeeksi opintosuorituksia.
 b. ?Opiskelija saa lainan, jolla on tarpeeksi opintosuorituksia.
 c. Se laatikko sisältää purkin, jossa on kansi.

Lohkeamat, joissa lauseattribuutin pääsanana ei ole substantiivi vaan pronomini (*Se saa lainan, jolla on tarpeeksi opintosuorituksia*), eivät SbP:n määritelmän perusteella kuulu substantiivilausekkeen lohkeamiin. Sama pätee myös lohjenneisiin *että*-lauseisiin (*Se on varmaa, että Marsissa asuu pieniä vihreitä miehiä*). Substantiivia määrittävät *että*-lauseet taas kuuluvat substantiivin alakategorioimiin määritteisiin, eivätkä ne siksi yleensä irtoa: **Ajatus on mieleton, että Marsissa asuu pieniä vihreitä miehiä*.

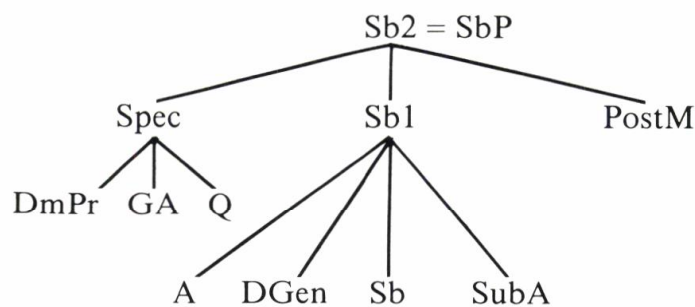
Kolmas irtoavien määritteiden ryhmä muodostuu ei-satunnaista omistusta ilmaisevasta genetiivistä: *Aimon on pää kipeä*. Aina ei kuitenkaan ole selvää, onko kyse irtautuneesta genetiivistä vai verbin sidonnaisesta adverbialista (Penttilä 1963:346). Nykykielessä jälkimmäinen tulkinta on yleisin, mitä todistaa genetiivin korvautuminen ulkopaikallissijalla. Toisaalta myös omistuliitteen puuttuminen auttaa hyväksyttävyydessä (*Mun on pää kipeä* vs. ?*Minnun on pääni kipeä*), mikä tukee adverbialitulkintaa: genetiivi on menettänyt possessiivisen merkityksensä. Lohkeamisessa kyse näyttääkin olevan siitä, että tilanteen toteamisen sijasta korostetaan jotain tapahtuvaksi jollekin (*Mun pää on kipeä* vs. *Mulla on pää kipeä* tai *Mun juna lähtee neljältä* vs. *Multa lähtee juna neljältä*, vrt. Kangasmaa-Minn 1980).

Lausekkeen hajoaminen on ilmiö, jossa tarkastelun kohteeksi on otettava koko lauseen rakenne: minkä funktioiset lausekkeet hajoavat, missä lohjennut osa voi esiintyä jne. Ongelmallista on se, millä tavalla toisistaan kaukana olevien osien yhteenkuuluvuus saadaan kuvattua: käytetäänkö ns. aukkokategorioita (slash-categories, GPSG:n esiintuoma käsite), turvaudutaanko semantiikkaan vai pitäydytäänkö siirto-transformaatioihin. Lausekerakennekieliopin kannalta hankala kysymys on myös se, voidaanko hajonneen rakenteen yhteydessä puhua lausekkeesta ylipäätään, sillä lausekkeen määritelmään kuuluu vaatimus jäsenten vierekkäisyydestä. Koska keskityn substantiivilausekkeen sisäiseen rakenteeseen, olen jättänyt hajoamisen lausetason ilmiönä yksityiskohtaisemman tarkastelun ulkopuolelle. Kuvauksen lähtökohta löytyy kuitenkin semanttisten riippuvuussuhteiden määrittämisestä; transformaatioanalyysi ei tule kysymykseen, koska kuvauksessani en oleta syvärakennetta, josta pintataso transformaatioiden avulla johdetaan, ja aukkokategorioiden käyttö lausekerakennesäännöissä taas tuntuu liian vahvalta menetelmältä: jokainen SbP voisi hajota eikä SbP:lle ominaista kiinteyttä (verrattuna esim. verbin ja sen objektin väliseen suhteeseen) saataisi esille.

4. Substantiivilausekkeen rakenne

Suomen kielen substantiivilausekkeen rakenne on kaavamaisesti seuraava:

(4)



Määritteet on jaettu eri tasoille syntaktis-semanttisten kriteerien pohjalta: lineaarinen järjestys, määritesuhteiden kiinteys ja merkitykseen liittyvät seikat ovat olleet tärkeimmät perusteet. Rakenne on kaksitasoinen (ks. tähän liittyvää pohdintaa luvussa 4.3). Ensimmäisen projektiotason konstituentti on *Sb1* ja sen yläpuolella on toisen projektiotason konstituentti *Sb2*, joka on itse substantiivilauseke, *SbP*. *Sb1* sisältää pääsanana *Sb* ohella kuvailuattribuutit, jotka etupäässä ovat adjektiiveja (*A*), kuvailugenetiivin (*DGen*) sekä

alakategorioidut määritteet (*SubA*), joita ovat kvanttoriginaalien substantiivien partitiivisijainen jälkiattribuutti (*pullo olutta*) ja rektiömääritteet (*ehdotus uusiksi tutkintovaatimuksiksi; halu lähteä; ajatus, että Marsissa asuu pieniä vihreitä olioita*).

Toisen tason määritteitä ovat tarkenteet (*Spec*) sekä muut jälkiattribuutit (*PostM*): relatiivilauseet, epäsuorat kysymyslauseet, appositiot ja ne paikallisiaiset substantiivilausekkeet, joita ei voi pitää alakategorioituina määritteinä (*maljakko Kreikasta, puuro lautasella*). Tarkenteet sisältävät demonstratiivipronominit (*DmPr*), kvanttorit (*Q*) ja (omistusta ilmaisevan) genetiiviattribuutin (*GA*). Termi tarkenne vastaa siis suurin piirtein englanninkielistä termiä *specifier*, jolla yleisesti viitataan tämällyyppisiin etumääritteisiin. *Specifier* tosin sisältää myös numeraalit, jotka kuvauksessani muodostavat oman lausekkeensa, *NumP*:n, sekä artikkelit. Olen erottanut kaksi genetiiviattribuuttia, koska heti pääsanana vieressä esiintyvä genetiivi poikkeaa sekä funktioltaan että rakenteeltaan tarkenteeksi luokiteltavasta genetiivistä (luku 5).

Tunnukset *Spec*, *SubA* ja *PostM* ovat lyhenteitä attribuuttien muodostamille heterogeenisille joukoille. Ne eivät viittaa mihinkään erityiseen *X'*-teorian mukaiseen rakenteeseen (kuten esim. *SbP*), vaan toimivat paikanvaraajina: jos substantiivilla on tarkenne, alakategorioitu määrite tai muu jälkiattribuutti, niin se esiintyy tunnuksen osoittamassa paikassa. Varsinaisissa lauserakennesäännöissä tunnuksen sisältämät lausekkeet on eksplisiittisesti kirjoitettu auki. Puukuvaimen ei esimerkiksi tule erillistä *Spec*-solmua vaan kaikki tarkenteet riippuvat suoraan *SbP*-solmusta, ks. tulostuskuvia s. 369–370. Seuraavaksi käyn lyhyesti läpi eri tasoille luokitellut määritteet.

4.1. Etuattribuutit

Etuattribuutit olen perinteiseen tapaan jakanut kahteen ryhmään: tarkenteisiin ja kuvailuattribuutteihin. Tarkenteet ovat kakkostason määritteitä ja ne määrittävät pääsanana referentin spesiestä (notiivista tai kvantitatiivista, Siro 1957, 1960. Vrt. myös Itkonen 1980 ja Vilkuna 1980). Kuvailuattribuutit taas kuuluvat ykköstasolle ja määrittävät pääsanana referentin ominaisuuksia, sitä millainen se on. Jaon on tehnyt jo Bloomfield (1933), suomen kielen tutkimukseen sen on tuonut Penttilä (1967).

Tarkenteiksi olen luokitellut demonstratiivit, omistusgenetiivin ja kvanttorit. Penttilällä (1967) myös prominaaliset adjektiivit (*semmoinen, tällainen*) kuuluvat tarkenteisiin, mutta itse pidän niitä kuvailuattribuutteina, koska ne määrittävät pääsanana referentin laatua pikemminkin kuin tarkentavat sen määräyksilöksi. Lisäksi ne toimivat proadjektiivien tapaan ja ovat muodol-

taan lähempänä adjektiiveja kuin pronomineja. Myös Tuomikoski (1969) pitää niitä pronominipesyeeseen kuuluvina yhdysadjektiiveina.

Tarkenteet voidaan jakaa karkeasti determinantteihin ja kvantto-reihin sen mukaan, kumpaa spesiestä ne määrittävät. Determinantit määrittävät pääsanana referentin definiittistä tai indefiniittistä notiivista spesiestä (mikä yksilö, mistä joukosta osa), kvanttorit sen definiittistä tai indefiniittistä määrää tai paljoutta. Jako ei tietenkään ole selvärajainen: osa kvantto-reista toimii myös determinantteina ja päinvastoin (ns. determinanttikvanttorit). Determinanteiksi tulkittavat määritteet esiintyvät lausekkeessa kuitenkin aina ensimmäisenä. Samanlaisen jaon determinanttien ja kvanttorien välille tekee myös Lyons (1977: 454–455).

Selviä determinantteja ovat demonstratiivipronominit. Jackendoff (1977) pitää determinantteina myös persoonapronomineja silloin, kun ne toimivat substantiivilausekkeen määritteenä: *minä vaivainen syntinen ihminen, te molemmat pojat*⁵. Myös (omistusta ilmaiseva) genetiiviattribuutti kuuluu determinantteihin, vaikka se käyttäytyykin muista determinanteista poikkeavasti. Sen kuvaukseen palaan 5. luvussa. Determinanttisiksi kvantto-reiksi olen luokitellut kvanttorit, jotka esiintyvät SbP:ssä aina ensimmäisenä (esim. *kumpikin, jokainen* ja indefiniittinen *eräs*) tai joiden paikka voi vaihdella demonstratiivien kanssa (totalitiivit *kaikki* ja *koko*). Ymmärrettävistä semanttisista syistä lausekkeen luvulla tosin on merkitystä siinä, voiko demonstratiivi esiintyä kvanttorin kanssa vai ei: **tämä kumpikin poika*, mutta *nämä kummatkin pojat*. Tarkenteiden jaosta ja jakoperusteluista enemmän Jokinen (1986: 46–64).

Kvanttorien kuvauksessa on otettava huomioon myös kvanttorikombinaation mahdollisuus. Kvanttorikombinaatiolla tarkoitan sellaisia sanaliittoja kuin *muutama muu, joku harva* ja *eräs toinen*, joissa kvanttorien järjestys on kiinteä, ja joissa jälkimmäisenä jäsenenä voi esiintyä vain *toinen, muu* sekä joissakin tapauksissa *harva*. Kombinaatiot ovat lähellä leksikaalistuneita fraaseja. Mahdollisten kombinaatioiden sijaan käytetäänkin usein toista samanmerkityksistä rakennetta: *?usea toinen* → *moni muu*.

Ykköstason etumääritteitä ovat kuvailuattribuutit. Ne eivät muodosta homogeenista määritejoukkoa vaan nekin on jaettava kahteen ryhmään, tarkentaviin ja toisaalta tavallisiin kuvailuattribuutteihin. Tarkentavilla kuvailuattribuuteilla on sekä tarkenteiden että kuvailuattribuuttien ominai-

⁵ Lausekkeiden pääsanana on kuitenkin tulkinnanvarainen. Myös persoonapronominia voidaan pitää pääsanana, koska tällöisen lausekkeen toimiessa subjektina taipuu lauseen finiittiverbi sen mukaan: *minä vaivainen syntinen ihminen olen surkea maan matonen*.

suuksia, ja niiden keskinäinen järjestys sekä esiintyminen tarkenteiden kanssa määräytyy hienosyisten semanttisten rajoitusten kautta. Ykköstason määrite on myös kuvailugenetiivi, jota tarkastelen lähemmin luvussa 5.

Tarkentavia kuvailuattributteja ovat prominaaliset adjektiivit, järjestyslukusanat ja superlatiivit. Ne poikkeavat tavallisista kokoon, väriin, muotoon ym. viittaavista adjektiiveista ja esiintyvät aina näitä ennen: *ensimmäinen/suurin suomenkielinen sanomalehti* vs. *?suomenkielinen ensimmäinen/suurin sanomalehti*. Ne määrittävät jäljessään tulevaa lauseketta kokonaisuudessaan eivätkä siten ole semanttisesti tasa-arvoisia muiden kuvailuattribuuttien kanssa (vrt. rekursiivinen puukuvain (6a) s. 358). Järjestyslukusanat ja superlatiivit voivat esiintyä myös tarkenteiden keskellä: *tuo ensimmäinen/suurin Aimon auto*. Näiden ominaisuuksien perusteella kyseiset attributit voisi luokitella myös tarkenteiksi eli kakkostason määritteiksi. Olen kuitenkin kuvannut ne ykköstason määritteinä, sillä ne eivät esiinny ennen kvanttoreita eivätkä demonstratiiveja (**ensimmäiset/*suurimmat useat/monet/kaikki/nuo autot*), ja paikanvaihtelu näyttää koskevan vain genetiiviä, jonka kuvaus on muutenkin ongelmallinen. Usein tarkenneasemainen kuvailuattribuutti sitä paitsi korvataan omalla lausekkeella, jolla on elatiivisijainen jälkiattribuutti (*viides Aimon autoista, suurin Helmin kissoista*)⁶.

Olen karkeasti luokitellut myös partisiippiattribuutit kuvailuattribuuteiksi, koska ne esiintyvät aina determinanttien jäljessä. Määritteettömien partiippien kohdalla tämä on ongelmatonta, mutta monimutkaisemmilla partiippikonstruktiolla on myös tarkenteiden ominaisuuksia, ne tekevät lausekkeesta definiittisen: *Aimon leipoma kakku, viime vuonna raa'asti murhattu pääministeri*. Toisaalta niillä on myös kuvaileva luonne ja ne toimivat mainiosti kuvailuattribuuttien jäljessä: *tavallinen Aimon leipoma kakku, toinen viime vuonna raa'asti murhattu pääministeri*. Partisiippiattribuuttien lähempi tarkastelu jää kuvaukseni ulkopuolelle, sillä se vaatii jälleen lausemaisuuksien eri asteiden sekä kantaverbin syntaktisten ja semanttisten piirteiden tutkimista. Partisiippistamisen yleisiä edellytyksiä on tarkastellut Koivisto (1986).

Tavalliset kuvailuattribuutit, ns. intersektiiviset adjektiivit määrittävät pääsanaa tasa-arvoisesti ja niiden keskinäinen järjestys on suhteellisen vapaa: *pieni pyöreä näkkileipä* vs. *pyöreä pieni näkkileipä*. Adjektiiviattribuuttien järjestystä ja ominaisuuksia ovat lähemmin tutkineet Holma (1966), Penttilä (1967) ja Larsen (1969); ruotsin kielen osalta adjektiiveja on tutkinut mm. Loman (1956, 1964).

⁶ Ilmiö on sukua elatiiviattribuutillisille tarkennelausekkeille (*viisi/kaikki/nuo näistä esimerkeistä*), joissa elatiiviattribuutti ilmoittaa joukon, josta tarkenteen spesifioimat oliot poimitaan.

4.2. Jälkiattribuutit

Substantiivilausekkeen jälkiattribuutteihin kuuluvat lausemääritteet, appositiot, paikallissijaiset jälkiattribuutit ja alakategorioidut määritteet. Ykköstason jälkiattribuutteja, alakategorioituja määritteitä, ovat mm. rektiomääritteet (toivo *paremmasta maailmasta*), *että*-lauseet (ajatus, *että maa on pyöreä*) ja infinitiivikonstruktiot (halu *lähteä mukaan*). Niitä voi olla kerrallaan vain yksi substantiivin määritteenä: *kimpale *graniittia maata*. Kakkostason jälkimääritteitä voi olla useampia (maljakko *Kreikasta, jonka Helmi osti Istanbulista*), mutta niillä on tietty järjestys: appositiot ja paikallissijaiset määritteet tulevat ennen relatiivilauseetta (*maljakko, *jonka Helmi osti Istanbulista, Kreikasta*). Paikallissijaiset jälkiattribuutit olen jakanut ykkös- ja kakkostasolle siten, että rektiota-pauksissa (laki *korkeakoulujen kehittämisestä*) ja nominaalisteissa (usko *tulevaisuuteen*) ne ovat abstrakti- ja deverbaalisubstantiivien alakategorioimia määritteitä, joiden sija määräytyy pääsanana mukaan, muuten ne kuuluvat kakkostasolle (maljakko *Kreikasta*).

Jos pääsana on konkreettistarkoitteinen substantiivi, on jälkiattribuutin käytöllä rajoituksia (Yli-Vakkuri 1970). Myös lausekkeen sijamuodolla näyttää olevan merkitystä: *Helmin kirjahyllyssä on maljakko Kreikasta* vs. *?Aina vain Helmi jaksaa puhua maljakosta(an) Kreikasta*. Jälkiattribuutilliset SbP:t esiintyvät mieluusti kielipollisissa sijoissa, ja erityisesti jälkiattribuutin kanssa samansijaiset muodot ovat huonoja. Luonnollisena selityksenä tälle voi pitää SbP:n kongruenssia ja siitä johtuvia tulkintavaikeuksia: kaksi peräkkäistä samansijaista substantiivina tulkitaan yleensä appositiosuhteessa olevaksi rakenteeksi, mutta tässä tapauksessa kyseinen tulkinta ei ole mahdollinen.

Paikallissijaisten jälkiattribuuttien tulkinta saattaa tuottaa vaikeuksia myös koko lausetta koskevalle analyysille. Finiittiverbistä riippuen jälkiattribuutti voidaan hahmottaa joko substantiivin määritteeksi tai verbiin liittyväksi adverbiksi. Esim. lauseella *Katselin etanan jarrutusta kurvissa* on helposti kaksi tulkintaa sen mukaan puhutaanko jarrutuksesta, joka tapahtuu kurvissa, vai katselusta, joka tapahtuu kurvissa, kun taas lauseella *Aimo söi lenkkimakkaraa taikinakuoressa* ei liene kuin yksi järkevä tulkinta. Jälkiattribuutin oikea pääsana selviää siten laajemman lausekontekstin yhteydessä, usein vasta ottamalla huomioon koko lauseen tekstiyhteys.

Appositiot ovat suhteellisen irrallisia substantiivilausekkeen määritteitä. Ne kongruoivat pääsanansa kanssa ja muistuttavat numeraalien kongruoivia jälkimääritteitä (kpl. 6). Jos appositio esiintyy pääsanana edessä (*leipurimestari Kakkunen, aikakauslehti Virittäjälle*), pidän sitä taipumattomana etuattribuuttina, vrt. *koko* ja kylkiäisadjektiivit *aika, eri* jne. (Jokinen 1986: 93–94).

Relatiivilauseet ovat substantiivin määritteitä, joiden kuvaus vaatii jälleen koko lauserakenteen tuntemusta, relatiivipronominien ominaisuuksien ja pitkänmatkan riippuvuuksien huomioonottamista, joten ne jäävät tämän artikkelin ulkopuolelle. Relatiivilauseiden kuvauksesta GPSG-teoriassa ks. Gazdar ym. (1985:153–). – Appositioita ja relatiivilauseita ei ole mukana tietokoneimplementaatiossani.

4.3. Projektiotasojen pohdintaa

Hahmottelemani substantiivilausekkeen rakenne muistuttaa GPSG:n käyttämää rakennekuvausta (Gazdar ym. 1985: 126). Sen sijaan se poikkeaa Jackendoffin (1977:57, 103) ja Chomskyn (1970) esittämistä X'-syntaksin malleista, joissa määritteet on jaettu kolmelle eri tasolle. Kolmen projektiotason postuloiminen tuntuu kuitenkin turhalta ja mielivaltaiseltakin: jos tasoja käytetään syntaktis-semanttisten erojen tekemiseen, voisi niitä olla periaatteessa mielivaltainen määrä, jotta jokainen hieno ero tulisi esille. Monien eri tasojen käyttö kuitenkin vesittää rakennetta koskevat yleistyksset, mutta toisaalta raja relevantin eron ja erityispiirteen välillä on ohut ja sen vetäminen riippuu osittain siitä, mihin tutkija on sattunut kiinnittämään huomionsa. Varsinkin etuattribuuteilla jako kahdentyyppisiin määritteisiin, tarkenteisiin ja kuvailuattribuutteihin, on semanttisesti paremmin perusteltavissa kuin kolmitasoinen määritejaottelu, ja myös jälkiattribuuttien kohdalla näyttää kaksitasoinen jako alakategorioituihin määritteisiin ja muihin jälkimääritteisiin sekä syntaktisesti että semanttisesti riittävältä. Yksityiskohtaisemmat erottelut on sitten luonnollista kuvata piirteiden avulla, joita käyttäen myös määritteiden samankaltaisuus saadaan esille.⁷

Jackendoffin ja Chomskyn esittämiä X'-teorian versioita voidaan kritisoida myös formaalin kieliteorian pohjalta. Mm. Pullum (1985) kiinnittää huomiota siihen, että X'-teorian käyttö on ollut löyhää ja epämääräistä, ja usein jopa sisäisesti ristiriitaista. Lisäksi hän analysoi X'-teorian perusole-

⁷ Toisaalta on huomattava, että nykyisin syntaksin tutkimuksessa suuntaus on kaikkien tasojen hävittämiseen: monitasoiset, lauseen syntaktis-semanttisesta rakenteesta runsaasti tietoa sisältävät puukuvaimet ovat menettäneet merkityksensä, koska sanojen dependenssisuhteet saadaan esille myös muilla keinoin. Esim. kategoriakieliopissa (Ades & Steedman 1982, Steedman 1987) syntaktisen jäsennyksen tuottama rakennepuu on analyysin sivutuote, jolla ei itessään ole suurta merkitystä. Myös Lauri Carlsonin (1987) säännöllinen unifikaatiokielioppi (RUG, Regular Unification Grammar) käyttää sanojen sisältämää leksikaalista informaatiota lausekerakennesääntöjen sijasta: ainoa syntaktinen sääntö sanoo, että lause koostuu sanoista. Toinen suomen kielen automaattinen jäsentäjä, FPARSE (F. Karlsson 1985), puolestaan tuottaa litteitä rakenteita ja keskittyy lauseen funktionaalisten dependenssisuhteiden kuvaamiseen.

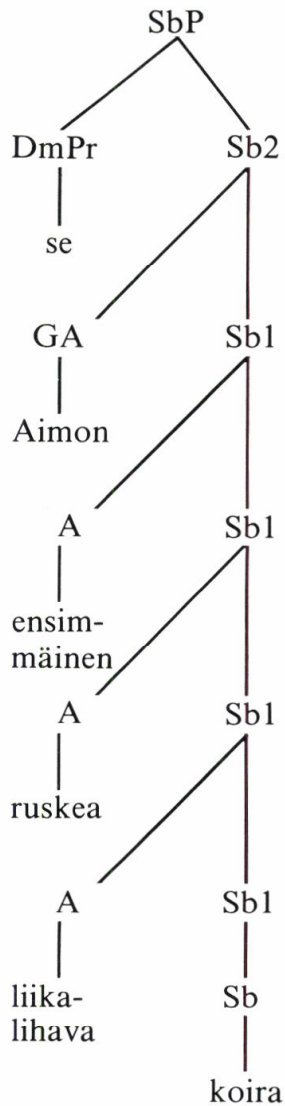
tuksia ja osoittaa, etteivät ne ole välttämättömyksiä. Esim. seuraamisperiaate, jonka mukaan pääsanan projektiotaso on aina yhtä pienempi kuin koko lausekkeen projektiotaso (kaava 5a), voidaan korvata samantasoisuusperiaatteella, jonka mukaan projektiotaso pysyy samana (kaava 5b).

- (5) a. $X^n \rightarrow \dots X^{n-1} \dots$
 b. $X^n \rightarrow \dots X^n \dots$

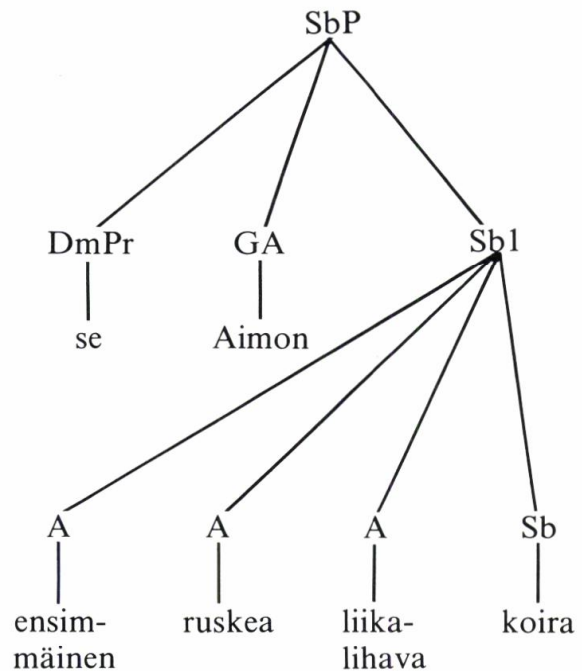
X tarkoittaa lausekkeen pääsanaa ja pisteet mahdollisia etu- ja jälkimääritteitä.

Se, käytetäänkö samantasoisuus- vai seuraamisperiaatetta, ei vaikuta kuvauksen laajuuteen. Substantiivilausekkeet, jotka saavat kuvauksen toisessa mallissa, saavat kuvauksen myös toisessa. Sen sijaan lausekkeiden rakennepuut eroavat huomattavasti toisistaan. Jos seuraamisperiaatteesta luovutaan,

(6a)



(6b)



saadaan rekursiivisia rakenteita, joissa jokainen attribuutti yksinään muodostaa pääsanan kanssa lausekkeen. Esim. lauseke *se Aimon ensimmäinen ruskea liikalihava koira* saa kuvan (6a) mukaisen »syvän» rakennepuun. Jos taas seuraamisperiaatteesta pidetään kiinni, vältetään rekursiivisuudelta ja kaikki saman tason määritteet ovat keskenään sisaruksia kuten (6b):ssä.

Rakenteen (6a) puolesta voidaan argumentoida semanttisin perustein: samantasoisillakin määritteillä on järjestys, joka vaikuttaa lausekkeen tulkintaan. (6a):ssa tulkinta saadaan havainnollisesti »paketoitumalla» avulla: attribuutit liitetään yksi kerrallaan määrittämään aikaisemmin muodostettua kokonaisuutta, jolloin muodostuu uusi lauseke, jota taas seuraava attribuutti määrittää jne. Attribuuttien määrittämissuhteilla on näin tietty ala, esim. (6a):ssa adjektiivi *ruskea* määrittää lauseketta *liikalihava koira*, ja järjestyslukusana *ensimmäinen* lauseketta *ruskea liikalihava koira*. Rakennepuu (6b), jossa kaikki saman tason määritteet ovat keskenään sisaruksia ja määrittämissuhteiltaan tasa-arvoisia, ei anna tätä tietoa suoraan, vaan lisäksi on oltava ylimääräinen sääntö, joka määrää tulkinnan muodostettavaksi lineaarisen järjestyksen perusteella.

Samantasoisuus-periaatetta käytetään hyväksi GPSG:ssä, jossa projektio-tasoa osoittava piirre on luokiteltu pääsanalta koko lausekkeelle periytyväksi, ns. HEAD-piirteeksi. Päinvastoin kuin X'-teorian muissa versioissa on GPSG:ssä siis oletusarvona projektiotason muuttumattomuus; projektio-tason muuttuminen on poikkeuksellista ja se täytyy eksplisiittisesti ilmaista säännöissä.

Oman kuvaukseni rakennepuut ovat (6b):n mukaisia implementaatioesityksiä. On kuitenkin muistettava, ettei puun muoto itsessään ole tärkeä vaan se, mitä puukuvaimen ja tasojen avulla pyritään ilmaisemaan: miten hyvin ja laajasti rakenne saadaan niiden avulla ylipäätään kuvattua.

5. Genetiiviattribuutti

Genetiiville on ominaista liikkuvuus: se voi esiintyä vapaasti eri etuattribuuttien joukossa, ei tosin kvanttorikombinaation välissä (7c).

- (7) a. *kaupungin* nuo monet muut aktiiviset asukkaat
 b. nuo *kaupungin* monet muut aktiiviset asukkaat
 c. *nuo monet *kaupungin* muut aktiiviset asukkaat
 d. nuo monet muut *kaupungin* aktiiviset asukkaat
 e. nuo monet muut aktiiviset *kaupungin* asukkaat

Jotta genetiivin paikan vaihtelu voitaisiin kuvata kaksitasoisessa SbP:n rakennemallissa, on kummallekin tasolle postuloitava oma genetiivi. Genetiivi-

viattribuutit jakautuvat näin ollen tarkennegenetiiveihin (GA) ja kuvailugenetiiveihin (DGen) sen mukaan, kumman projektiotason määritteestä on kyse. Tarkennegenetiivit (7a–d) voivat vaihdella muiden tarkenteiden joukossa vapaasti, kuvailugenetiivi (7e) sen sijaan esiintyy aina pääsanansa vieressä muiden mahdollisten kuvailuattribuuttien jälkeen. Tarkennegenetiivin paikkaan liittyvät semanttiset vaikutusalan muutokset, fokusointi- ja painotuserot on helppo kuvata paketoitumalla (6a) avulla.

Genetiivien jako ei ole pelkkä kuvausformalismin aikaansaama tekninen välttämättömyys, vaan sille on olemassa myös selviä semanttisia ja syntaktisia perusteluja. Genetiivien ominaisuudet korreloivat teorian antamien ennustusten kanssa: tarkennegenetiivit eroavat kuvailugenetiiveistä kuten kaksoastason määritteet ykköastason määritteistä yleensäkin (Jokinen 1987a). Lisäksi on huomattava, että samaa substantiiviva voi samanaikaisesti määrittää kaksi eri genetiiviä (sekä tarkenne- että kuvailugenetiivi, (8a–b)), mutta useammat genetiivit muodostavat epäkieliopillisen lausekkeen (8c).⁸

- (8) a. Tästä *eri alueiden* erilaisesta *muutoksen* nopeudesta on yhteiskunnalle haitallisia seuraamuksia.
 b. *Aimon* vanhat *silmälasiensa* sangat menivät poikki.
 c. *Nuo *Aimon* molemmat *isoäidin* vihreät *tuolin* jalat.

Tarkenne- ja kuvailugenetiivin erottaa parhaiten paikan avulla: kuvailuattribuutin jäljessä esiintyvä genetiivi on kuvailugenetiivi. Genetiivin ja adjektiivin keskinäistä paikkaa luonnehtii semanttinen alanvaihtelu, joka liittyy definiittisyyden ja geneerisyyden erottamiseen: *vanha Ladan talvirengas* vs. *Ladan vanha talvirengas*. Edellisessä lausekkeessa on DGen, jälkimmäisessä taas GA, joka tekee lausekkeesta definiittisen.

Tarkenteisiin kuuluva tarkennegenetiivi siis spesifioi pääsanansa referenttiä ja saa yleensä possessiivisen tulkinnan. Se on genetiivimuotoinen substantiivilauseke, jolla ei saa olla jälkimääritteitä⁹. Toisaalta tarkennegenetiivillä voi olla kaikki Sbp:n etuattribuutit määritteinään, erityisesti myös tarkennegenetiivi. Näin rekursiiviset, periaatteessa mielivaltaisen pitkät genetiiviketjut ovat mahdollisia: *Pekan äidin serkun vaimon auton talvirengas* tai *aivokuoren otsalohkon kolmannen aivopoimun tyven (Brocan alueen) vaurio*. Re-

⁸ Koska prototyyppiset substantiivit hyväksyvät kaksi genetiiviä, on ymmärrettävää, että myös nominaaliteissa voi esiintyä kaksi genetiiviä, subjekti- ja objektigenetiivi: *Helmin jokaviikkoinen tenniksen pelaaminen*.

⁹ Jälkimääritteettömyys on yleinen kaikkia etuattribuutteja (adjektiivilausekkeita, postpositiolausekkeen komplementteja jne.) koskeva rajoitus, ja se voidaan helposti kuvata piirteiden avulla. Jos lausekkeessa on jälkiattribuutti, annetaan koko lausekkeelle tätä kuvaava piirre [+POSTM]. Piirre estää lausekkeen esiintymisen etuattribuuttina, sillä etuattribuutilta vaaditaan piirrettä [–POSTM].

kursiivisuus mahdollistaa siten keskiupotuksen substantiivilausekkeen sisällä, mikä matemaattisen kielitieteen ja automaattisen jäsennyksen kannalta on ongelmallista. Genetiivin rekursiivisuutta olen tarkastellut lähemmin muualla (Jokinen 1987b).

Kuvailugenetiivi taas esiintyy yleensä heti substantiivin vieressä ja muodostuu joko yksittäisestä sanasta tai sanaliitosta. Se ei spesifioi pääsanansa referenttiä kuten tarkennegenetiivi, vaan kuvailuattribuuttien tapaan kuvailee sen laatua. Kuvailugenetiiville on siten tunnusomaista geneerisyys: sillä ei ole määrättyä yksittäistä referenttiä, vaan se viittaa substantiivin denotoimaan luokkaan yleensä. Geneerisyys tulee ilmi siinä, ettei tässä asemassa voi esiintyä määräinen genetiivi sen paremmin kuin possessiivinen genetiivikään, jotka molemmat ovat aina referentiaalisia: **sininen minun autoni*, **uusi Laineen auto*. Mahdollinen vastaesimerkki *se teräväkielinen Mikkosen akka*, jossa genetiiviä voitaisiin ajatella possessiivisena, menettää teränsä, kun huomataan, että *Mikkosen* ei ole omistusgenetiivi vaan *akkaa* identifioiva, kuvaileva määrite.

Kuvailugenetiivin suhde pääsanaan on läheinen, yhdyssanamainen, ja usein se kirjoitetaankin tämän kanssa yhteen. Adjektiiviattribuutti sen ja pääsanavan välissä kuulostaa oudolta: *innokas fysiikan opiskelija* vs. *?fysiikan innokas opiskelija*. Kuvailugenetiivi liittyy näin läheisesti sanamuodostukseen, mutta sen postuloiminen itsenäisenä lausekkeena on myös oikeutettua: sillä voi olla omia määritteitä (*innokas Punaisen Ristin jäsen, kohtalokkaat Pariisin-matkan seuraukset*), ja painotuksen avulla se voidaan erottaa irralliseksi määritteeksi (*vihreä tuolin jalka, ei pöydän*).

Kuvailugenetiivin määritteet ovat kuitenkin aina kiinteästi yhteydessä itse genetiiviin: *Punainen Risti* ja *Pariisin-matka* muodostavat kokonaisuutena käsitteen, leksikaalistuneen fraasin, joka vastaa yksittäistä leksikaalista entryä. Leksikaalinen entry tarkoittaa leksikossa lueteltua sanamuotoa fonologisine, morfosyntaktisine ja semanttisine piirteineen. Se voi olla itsenäinen hakusana, fraasi tai leksikaalisten sääntöjen avulla saatu johdos, mutta syntaksin näkökulmasta se on aina jakamaton kokonaisuus.

Kuvailugenetiivi ei siten ole substantiivilauseke kuten tarkennegenetiivi, vaan leksikaalinen entry. SbP:n rakenteessa kuvailugenetiivi on siis samantapainen yksikkö kuin demonstratiivipronomini tai kvanttori: se ei muodosta omaa lauseketta vaan se noudetaan substantiivin määritteeksi suoraan leksikosta.¹⁰

¹⁰ Tämä vastaa myös käsitystä siitä, miten leksikaalistuneet fraasit ja sanaliitot on tallennettu mentaaliseen leksikkoon: niitä ei analysoida syntaktisesti vaan ne tajutaan yksinkertaisten käsitteiden tapaan huolimatta siitä, että niillä on oma sisäinen rakenne.

Jos heti pääsanana vieressä esiintyy vain yksi genetiivi *Laineen auto, tuolin jalka*), on ilman semanttista analyysia vaikea ratkaista, kummasta genetiivistä on kyse. Tietokonetoteutuksessani kaikkiin erisnimiin on liitetty piirre [+referentiaalinen], mikä estää erisnimien esiintymisen kuvailugenetiivinä, sillä kuvailugenetiiviltä vaaditaan piirrettä [−referentiaalinen]. Lauseke *vanha Ladan talvirengas* ei siten tule hyväksytyksi, ja lauseke *Ladan talvirengas* saa vain tarkennegenetiivitulkinnan. Toisaalta kyseisen piirteen käyttö tulee yhdistää koko lausekkeen semanttiseen analyysiin, eikä vain yksittäisiin sanoihin, sillä erisnimet voivat tunnetusti esiintyä myös kuvailugenetiivinä (*Mikkosen akka*), eli niitä voidaan käyttää ei-referentiaalisesti.

6. Numeraaleja sisältävät lausekkeet

Numeraaleja sisältävissä lausekkeissa pääsana-määrite -suhde on epäselvä. Determinanttien suhde substantiiviin on yleensäkin tulkinnanvaraista, ks. esim. Lyons 1977. Mm. Hoeksema (1984) on argumentoinut kategoriakieliopin puitteissa sen puolesta, että pääsanana toimii determinantti eikä substantiivi, HPSG:ssä (Head-Driven Phrase Structure Grammar, Pollard & Sag 1987) taas determinantti on substantiivin alakategorioima määrite. X'-teorian perusoletusten mukaan determinantti ei voi olla pääsana, koska tällöin koko lauseke olisi determinanttilauseke, ei substantiivilauseke, mikä johtaisi kummallisuuksiin muun analyysin kohdalla.

Yleinen käsitys on, että numeraali toimii pääsanana yksikön nominatiivissa, jossa substantiivi on partitiivimuotoinen, ja substantiivi toimii pääsanana muissa sijoissa, joissa numeraali ja substantiivi kongruoivat. Leiwo (1977) tosin tulkitsee lukusanan kaikissa tapauksissa substantiivin määritteeksi. Olen kuitenkin kuvannut kaikki numeraalin sisältävät lausekkeet SbP:tä muistuttavina kaksitasoisina numeraalilausekkeina (*NumP*), joissa numeraali on pääsana. Ratkaisuni perustuu X'-teorian pääsanakäsitteeseen ja numeraaleja sisältävien lausekkeiden yhdenmukaiseen kuvaukseen.

Yksikön nominatiivi -tapauksessa numeraali on selvästi pääsana, ks. esim. Hakulinen & Karlsson (1979:146). Muissa sijoissa numeraalin pääsanamaisuutta voi perustella sillä, että se määrää ympärillään olevien sanojen luvun: numeraalin edessä esiintyvät, kongruoivat määritteet (demonstratiivit ja adjektiivit) ovat monikollisia, mutta lukusana itse ja sen jälkeen tulevat attributit yksiköllisiä: *näillä koulunsa päättäneillä kahdella iloisella pojalla, ne vaikeat kaksikymmentä vuottamme*. Numeraalin edessä olevat määritteet tosin voivat olla yksiköllisiäkin, mutta tällöin lausekkeen merkitys muuttuu: kyse on kollektiivisesta luennasta, ei distributiivisesta. Kollektiivinen luenta riippuu myös numeraalia seuraavasta substantiivista, ja on

yleensä mahdollinen vain määrää ja aikaa denotoivilla substantiiveilla, jotka voidaan tulkita paitsi erillisiin olioihin, myös yhteen kokonaisuuteen viittavina sanoina: *Niistä/Siitä kolmesta viikosta jäi mukava muisto* mutta *Niistä/*Siitä kolmesta lauseesta ei osattu piirtää puukuvaimia*. Asiaan lienee ensimmäisenä kiinnittänyt tarkempaa huomiota Göran Karlsson (1971). Myös Lauri Karttunen (1984) on tarkastellut sitä.

On kuitenkin huomattava, että obliikvisijoissa numeraali ja substantiivi näyttävät muodostavan tasa-arvoisen sanaliiton, jossa kumpikaan ei suoraanaisesti hallitse toista. Varsinkin jos numeraalin edessä on määritteitä, tuntuu niiden tulkinta yksinomaan numeraalin tai substantiivin attribuuteiksi harhaanjohtavalta, ja sen sijaan tekee mieli yhtyä Setälän (1973:83) näkemykseen määritteistä lukusanan ja substantiivin yhteisinä määritteinä. X'-teorian mukaan lukusanan edessä esiintyvät attributit voivat kuitenkin kuulua vain lausekkeen pääsanalle, numeraalille. Määritteellä ei voi olla useita eri pääsanoja, sillä lausekkeessa on vain yksi pääsana, jonka määritteitä kaikki muut lausekkeen jäsenet ovat. Toisaalta demonstratiivi ja genetiivi voidaan tulkita numeraalin ja substantiivin yhteisiksi määritteiksi, ks. s. 364.

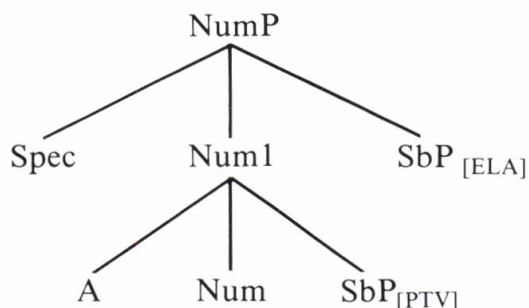
Numeraalin jälkeen tulevat määritteet muodostavat yhdessä substantiivilausekkeen, joka toimii numeraalin jälkiattribuuttina (*viisi Helmin tarkkaan mietittyä esimerkkiä*). Yksikön nominatiivissa tämä on partitiivisijainen, mutta muissa sijoissa se kongruoi pääsanana kanssa. Kongruoiva jälkiattribuutti tosin on epätavallinen määrite, koska jälkiattribuutit koostuvat etupäässä alakategorioiduista määritteistä, joiden sijan määrää pääsana, sekä lauseattribuuteista. Ainoa kongruoivaksi jälkiattribuutiksi luokiteltava määrite, appositio, taas esiintyy pääsanansa suhteen melko itsenäisenä, täydentävänä lisänä, eikä niinkään tämän hallitsemana määritteenä: *Helsingissä, Suomen pääkaupungissa tai ystävälleni Esmeraldalle*. Toisaalta, kuten edellä jo tuli esille, muodostavat numeraali ja substantiivi obliikvisijoissa sanaliittomaisen kokonaisuuden, jossa kumpikin esiintyy suhteellisen itsenäisenä jäsenenä. Ajatus numeraalilausekkeen kongruoivista jälkimääritteistä ei siten ole täysin outo.

Numeraalin jälkiattribuuttina voi olla myös elatiivisijainen, monikollinen substantiivilauseke: *viisi koirista, kolmelle noista vaikeista esimerkkilauseista*. Jälkiattribuuttien välillä on merkitysero: partitiivisijainen ja kongruoiva SbP viittaa oliolajiin ja ilmoittaa ne oliot, joiden lukumäärästä on kyse, elatiiviattribuutti taas viittaa tiettyyn oliojoukkoon, josta numeraalin ilmoittama lukumäärä oliota poimitaan. Jälkiattribuutit eivät sulje toisiaan pois, mutta substantiivien oma semantiikka aiheuttaa rajoituksia yhdessä esiintyvälle määritteille: *neljä poikaa tytöistä* ei saa järkevää tulkintaa,

sen sijaan *neljä poikaa entisistä oppilaistani* on käypä lauseke. Rajoitukset on helppo ymmärtää attribuuttien erilaisten semanttisten funktioiden pohjalta. Myös jälkiattribuuttien järjestys (partitiivisijainen attribuutti ennen elatiivisijaista) on tässä mielessä luonnollinen: ensin ilmoitetaan mitä poimitaan, sitten mistä nämä oliot poimitaan.

Numeraalin sisältävien lausekkeiden analyysi saadaan näin yhtenäiseksi. Numeraali on leksikaalinen kategoria, joka muodostaa oman lausekkeensa, NumP:n. Se saa kongruoivia, yleensä monikkomuotoisia etumääritteitä sekä kahdentyyppisiä jälkimääritteitä: joko elatiivisijaisia, monikollisia substantiivilausekkeita tai partitiivisijaisia ja kongruoivia, yksiköllisiä substantiivilausekkeita. Numeraalilausekkeen rakennepuu on seuraava:

(9)

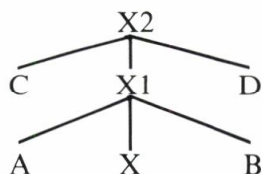


Jälkiattribuuttien järjestys ja erilainen semantiikka on ilmaistu tasojen avulla: lähinnä pääsanana esiintyvät partitiivisijaiset jälkiattribuutit ovat ykköstason määritteitä, elatiivisijaiset jälkiattribuutit kuuluvat kakkostasolle.

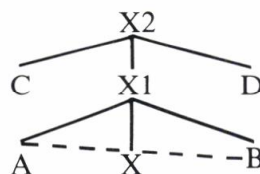
Rakennepuun avulla on helppo nähdä, että etuattribuuttien intuitiivisesti luonnollisin tulkinta numeraalin ja substantiivin yhteisinä määritteinä on mahdollista tarkenteille mutta ei kuvailuattribuuteille. Kuvailuattribuutti ei voi määrittää substantiivia, koska SbP_[PTV] on sen kanssa samantasoinen numeraalin määrite, ja samantasoiset määritteet ovat keskenään tasa-arvoisia¹¹. Toisaalta attribuutit määrittävät pääsanalauseketta kokonaisuudessaan, joten ne voidaan tulkita yhteisiksi määritteiksi kaikille niille lausek-

¹¹ Puukuvaimien avulla havainnollistettuna tämä tarkoittaa sitä, että rakennepuu (10a) on hyvinmuodostunut, kun taas rakennepuuta (10b), jossa A:lla on kaksi pääsanana (X ja B), ei hyväksytä.

(10a)



(10b)



keille, joista pääsanana koostuu. Koska tarkenteiden pääsanana on Num1, joka sisältää numeraalin ja sen partitiivisijaisen (tai kongruoivan) jälkimääritteen, SbP:n, voidaan tarkenteet siis tulkita näiden yhteisiksi määritteiksi. Näin tulkiten tarkenteet kuitenkin myös määrittävät Num1-lausekkeeseen kuuluvia adjektiiviattribuutteja, mikä taas ei ole suotavaa.¹²

Tarkenteiden esiintymiseen NumP:ssa liittyy mielenkiintoisia rajoituksia. Numeraalilausekkeessa ei ole genetiiviattribuuttia, jos sillä on elatiivisijainen jälkiattribuutti (**Helmin viisi noista esimerkeistä*), eikä numeraalin partitiivisijainen jälkiattribuutti yleensä sisällä demonstratiivipronomia (**viisi tuota merkillistä esimerkkiä*), vaikka demonstratiivi elatiivisijaisessa jälkiattribuutissa on tavallista. Tässä kohdin lausekerakennesääntöni hieman yligeneroivat, sillä ne hyväksyvät edelliset lausekkeet kieliopillisiksi.

Ensimmäisessä tapauksessa yligenerointi voidaan estää käyttämällä piirteitä samaan tapaan kuin niitä käytetään estämään jälkiattribuuttien esiintyminen etuattribuuteissa (ks. alaviite 9). Toisessa tapauksessa yligenerointi voidaan välttää myös siten, että lausekerakennesäännöissä partitiivisijainen jälkiattribuutti määritellään Sb1-lausekkeeksi eikä maksimaaliseksi substantiivilausekkeeksi. Demonstratiivipronominithan ovat kakkostason määritteitä, eivätkä ne siten voi esiintyä jälkiattribuutissa, joka muodostuu pelkästään ykköstason lausekkeesta. X'-teoriassa määritteinä saa kuitenkin olla vain joko maksimaalinen projektio tai kieliopillinen formatiivi (joka laajasti määriteltynä tarkoittaa leksikaalista entryä), joten Sb1-lauseke on teorian puitteissa sopimaton jälkiattribuutiksi. Toisaalta on myös huomattava, että demonstratiivipronomini voi esiintyä, jos koko jälkiattribuutti on monikollinen. Esim. kauppatorilla on mahdollista pyytää *Saanko neljä noita hyvännäköisiä haukia*, joka merkitykseltään vastaa suurin piirtein lausetta *Saanko neljä noista hyvännäköisistä hauista*. Numeraalin ja tarkenteiden välinen suhde vaatii kuitenkin tarkempaa tutkimusta kuin mitä tässä on mahdollista esittää, joten olen antanut sääntöjeni yligeneroida.

Numeraalien kuvauksessa on vielä huomattava plurale tantum -sanat, joiden yhteydessä numeraali on aina monikossa. Plurale tantum -sanaa voitaisiin pitää rakenteen pääsanana, sillä lukusanan numerus määräytyy nimen omaan substantiivin kieliopillisen monikollisuuden perusteella. Jos oletamme, että lukusanat muodostavat omia lausekkeitaan, voidaan asia ajatella myös toisinpäin. Lukusana määrää paitsi jälkiattribuuttina esiintyvän substantiivilausekkeen sijan myös numeruksen: monikkomuotoinen lukusana

¹² Ylemmän tason määritteiden pääsanaksi määritelläänkin yleensä alemman projektiotason pääsanana, ja määrittelyä jatketaan, kunnes päädytään leksikaaliseen kategoriaan. Näin ylemmän tason määritteilläkin on vain yksi pääsanana, nim. leksikaalinen kategoria.

vaatii kongruoivan ja partitiivisijaisen jälkiattribuuttinsa aina monikkoon, yksiköllinen lukusana puolestaan yksikköön (vrt. kuitenkin em. demonstraatiivin sisältävät SbP:t). Lauseke *kahdet koirat* on siis kieliopillinen, ja sen outous johtuu vain siitä, että sille on vaikea löytää semanttista tulkintaa. Lauseketta ei lukusanan täsmällisyyden vuoksi voi tulkita epämääräiseen joukkoon viittaavaksi ilmaukseksi (*Kymmenet/sadat/monet koirat haukkuivat ympärilläni*), eikä se substantiivin semantiikan vuoksi helposti saa plurale tantum -tulkintaakaan (vrt. kuitenkin naparetkeilijä, joka tarvitsee koiravaljakkoonsa *kahdet koirat*, eli kaksi erillistä koiravaljakkoa; Vilkuna 1980). Toisaalta myös lauseke *kaksi saksea* saa tulkinnan, vaikka se tällöin ei viittaakaan kahteen leikkaamisessa käytettävään apuvälineeseen (*Kahdet saksat kuuluvat joka kodin varustukseen*), vaan ravun yksittäiseen saksen puoliskoon. Se, että *kaksi häätä* tai *kolme tikapuuta* eivät saa tulkintaa, kuvataan yksinkertaisesti siten, että kyseisten substantiivien yksikkömuodot eivät esiinny leksikossa vapaina muotoina.

7. Substantiivilausekkeen rakennesäännöt

Substantiivilausekkeen säännöt olen kirjoittanut ID/LP-muodossa, joissa tavallisten PS-sääntöjen sisältämä informaatio on jaettu kahtia. Välittömät dependenssisäännöt (Immediate Dependency Rules) kertovat, mitä tytärkonstituentteja tietty konstituentti voi sisältää; lineaarisen järjestyksen säännöt (Linear Precedence Statements) taas ilmoittavat, missä järjestyksessä sisarusten tulee esiintyä, jos niiden keskinäinen järjestys on määrätty. ID/LP-säännöt antavat mahdollisuuden vapaan konstituenttijärjestyksen kuvaamiseksi, vaikka onkin huomattava, että vapaus koskee vain samaan konstituenttiin kuuluvia sisaruksia, ei eri konstituenttien sisältämien osakonstituenttien järjestystä. ID/LP-säännöistä enemmän Gazdar ym. (1985: 44–50), Gazdar & Pullum (1982:18–22), Vilkuna (1986: 206–211).

Rakennekuvaus on implementoitu tietokoneelle Helsingin yliopiston tietokoneolingvistiikan tutkimusyksikössä. Kuvauksen olen toteuttanut yhdessä Lauri Carlsonin kanssa Lisp-kielisellä HUG (Helsinki Unification Grammar) -jäsenyysformalismilla. Lauri Carlson on tehnyt tähän alun perin Lauri Karttusen kehittämään jäsentäjään (Karttunen 1986) ID/LP-ohjelmopakkauksen, joten säännöt voidaan kirjoittaa ID/LP-formaatissa. ID/LP-HUG on yksi unifikaatiopohjaisista kieliopista (ks. Shieber ym. 1984, Shieber 1986; hyvä suomenkielinen johdatus unifikaatioon on Carlson 1988).

Säännöt, joiden mukaan pieni SbP-jäsentäjäni analysoi substantiivilausekeita, ovat seuraavat (sääntöjen kirjoitusasu on hieman pelkistetty siitä, miten ne on kirjoitettu IDLP-HUG:iin):

ID-säännöt (välittömät dependenssisäännöt):

- 1: $Sb2 \rightarrow (DmPr), (Sb2_{[GEN, -POSTM]}), (Q), (APr), Sb1, (PostM)$
- 2: $Sb1 \rightarrow (PrA), A2^*_{[-POSTM]}, (DGen_{[GEN]}), Sb, (Sb2_{[PTV]})$
- 3: $A2 \rightarrow A1, (Sb2)$
- 4: $A1 \rightarrow (AdA), A$
- 5: $Num2 \rightarrow (DmPr), (Sb2_{[GEN, -POSTM]}), Num1, (Sb2_{[ELA]})$
- 6: $Num1 \rightarrow (PrA), A2^*_{[-POSTM]}, Num, (Sb2_{[PTV]})$

Sulut merkitsevät optionaalisuutta. Säännöissä 2 ja 6 esiintyvä tähti taas on ns. Kleenen tähti, joka tarkoittaa iteratiivisuutta. Sillä merkitty symboli voi esiintyä säännössä äärettömän monta kertaa; säännöt sallivat siis useampien adjektiivien esiintymisen samassa lausekkeessa. Toisaalta tähdellä merkityn symbolin ei tarvitse esiintyä kertaakaan, joten optionaalisuutta merkitsevät sulut voidaan jättää sen ympäriltä pois.

Säännöissä 1 ja 5 esiintyvä symboli $Sb2_{[GEN, -POSTM]}$ tarkoittaa tarkennegenetiiviä, joka on genetiivisijainen, jälkiattribuutiton substantiivilauseke. Säännössä 1 esiintyvä symboli APr taas viittaa adjektiivisiin pronomineihin, jotka esiintyvät kvanttorikombinaation jälkimmäisenä kvanttorina. Kakkostason jälkiattribuutteja ei ole implementoitu, vaan niiden tilalla säännössä 1 on vain paikanvaraaja $PostM$.

Säännössä 2 on erikseen mainittu pronominaaliset adjektiivit, PrA , mutta muut tarkentavat ykköstason määritteet ovat implementoimatta. Jälkiattribuuteista mukana on vain kvanttorimaisten substantiivien partitiivisijainen jälkimäärite, $Sb2_{[PTV]}$.

Säännöt 3 ja 4 kuvaavat adjektiivilauseketta. Säännössä 3 esiintyvä SbP tarkoittaa adjektiiveihin liittyvää rektiosubstantiivie (kuuluisa runoistaan). Säännöissä 2 ja 6 adjektiivilausekkeelta vaaditaan piirrettä $[-POSTM]$, jonka avulla estetään lausekkeet **kuuluisa runoistaan mies* ja **kuuluisat runoistaan viisi miestä*. Lyhenne AdA säännössä 4 taas viittaa adjektiivin vahvikkeena olevaan adverbiin (*kovin, hyvin* jne).

Säännöt 5 ja 6 kuuluvat numeraalilausekkeelle. SbP toimii elatiivisijaisena jälkimääritteenä $Num2$ -tasolla (sääntö 5), ja partitiivisijaisena tai kongruoivana jälkimääritteenä $Num1$ -tasolla (sääntö 6). Substantiivilausekkeiden sijaero näkyy säännöissä hakasulkuihin merkittynä sijapiirteenä.

LP-säännöt (lineaarista järjestystä koskevat säännöt):

LP-säännöissä kategoriasymbolit ovat tarpeeksi yksikäsitteisiä, joten en ole merkinnyt niihin piirrerajoituksia näkyviin. Säännöissä 5 sijapiirteet ovat kuitenkin välttämättömiä, sillä muuten säännöstöstä tulisi ristiriitainen.

Sääntöön 1 liittyvät LP-säännöt:

$DmPr < Q$

$Q < APr$

$APr < Sb1$

$Sb1 < PostM$

$Sb2 < Sb1$

Sääntö 1 ei siis spesifioi genetiiviattribuutin GA paikkaa muuten kuin toteamalla, että sen on oltava ennen lausekkeen pääsanaa, Sb1:tä (viimeinen LP-sääntö). Koska LP-sääntöjen 'edeltää'-relaatio, $<$, on transitiivinen, ei jokaista 'edeltää'-relaatiota tarvitse eksplisiittisesti merkitä näkyviin. Esim. kaksi ensimmäistä LP-sääntöä yhdessä sanovat, että DmPr on ennen APr:a, koska DmPr on ennen Q:a ja tämä taas on ennen APr:a.

Sääntöön 2 liittyvät LP-säännöt:

$PrA < A2$

$A2 < DGen$

$DGen < Sb$

$Sb < Sb2$

Sääntöön 3 ei tarvitse liittää järjestystä spesifioivaa LP-sääntöä, koska adjektiivi ja sen rektiömääräite voivat vaihdella vapaasti.

Sääntöön 4 liittyvä LP-sääntö:

$AdA < A$

Sääntöön 5 liittyvät LP-säännöt:

$DmPr < Num1$

$Sb2_{[GEN]} < Num1$

$Num1 < Sb2_{[ELA]}$

Sääntöön 6 liittyvät LP-säännöt:

$PrA < A2$

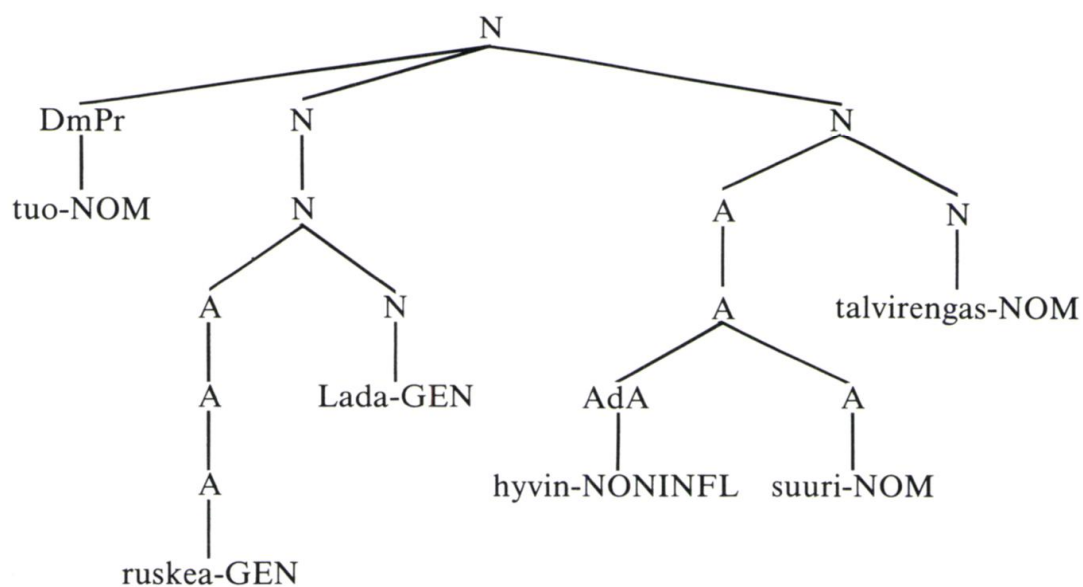
$A2 < Num$

$Num < Sb2$

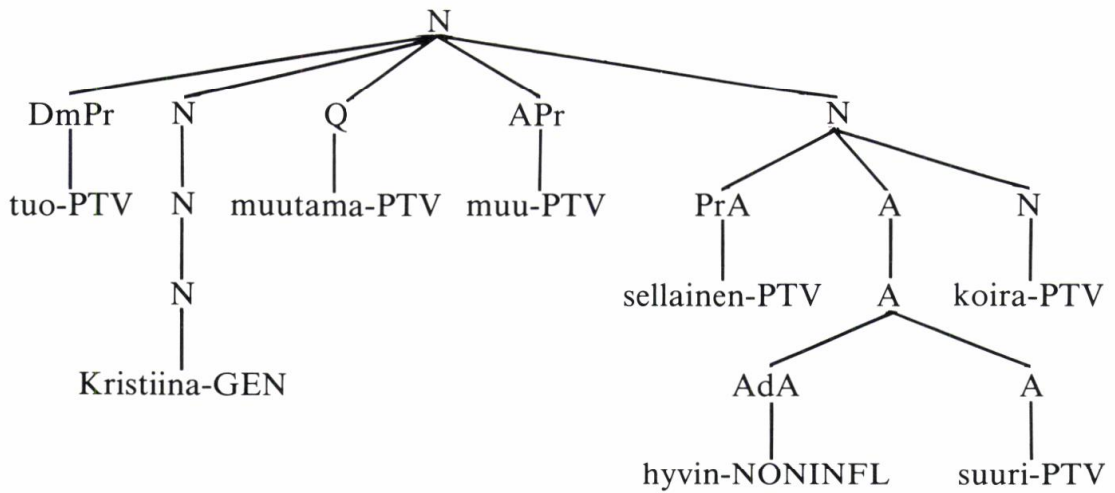
Sääntöjen avulla voidaan jäsentää mm. seuraavat lausekkeet (1)–(18).

1. tuo ruskea koira
2. sellainen ruskea koira
3. Ladan talvirengas
4. tuo Ladan hyvin suuri talvirengas
5. tuo ruskean Ladan hyvin suuri talvirengas
6. tuo hyvin suuri graniitin kimpale
7. tuon Kristiinan koira
8. noita tuon ruskean koiran suuria serkkuja
9. Kristiinan serkun Ladan vanha talvirengas
10. Kristiinan sellainen graniitin kimpale
11. muutamia sellaisia suuria graniitin kimpaleita
12. noita Kristiinan muutamia muita sellaisia hyvin suuria koiria
13. kimpale Kristiinan graniittia
14. kimpale kimpaleesta graniittia
15. kolme samaa suurta kirjainta
16. kolme Kristiinan ruskeaa koiraa
17. noista kolmesta Kristiinan ruskeasta koirasta
18. noista Kristiinan kolmesta ruskeasta koirasta

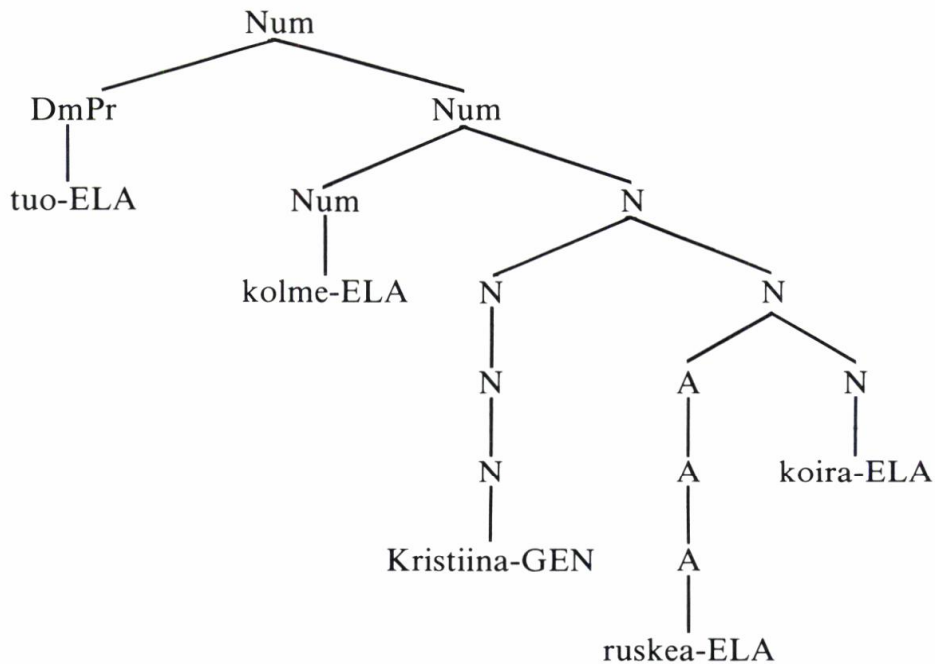
Esim. 1. tuo ruskean Ladan hyvin suuri talvirengas



Esim. 2. noita Kristiinan muutamia muita sellaisia hyvin suuria koiria



Esim. 3. noista kolmesta Kristiinan ruskeasta koirasta



Esimerkkeinä tietokoneanalyysistä ovat lauseiden 5, 12 ja 17 puukuvaimet. Substantiivin tunnuksena on niissä N. Tasoja ei ole merkitty lausekkeiden tunnuksiin, mutta ne tulevat kuitenkin näkyviin puukuvaimessa: leksikaalisen entryn yläpuolella oleva symboli viittaa leksikaaliseen kategoriaan, sitä

seuraava symboli ykköstason lausekkeeseen ja sitä seuraava kakkostason lausekkeeseen. Leksikaaliset entryt on lisäksi varustettu sijaa osoittavalla lyhenteellä.

On huomattava, että kuvaukseni on paljon laajempi kuin sen tämänhetkinen tietokoneimplementaatio.

8. Lopuksi

Suomen substantiivilausekkeen rakenne lienee varsin tunnettu ja tarkkaan tutkittu jo aikaisemmin. Kuten jo alussa mainitsin, ei sitä kuitenkaan ole kuvattu minkään formaalin kielioppiteorian puitteissa, ja kuvaus on siltä osin jäänyt eräällä tavalla puolitiehen: huolellisesti koottu aineisto on olemassa, mutta sen käsittely laajemissa yhteyksissä on jäänyt vähemmälle huomiolle. Substantiivilausekkeen X'-teoreettisella kuvauksella pyrin näin ollen avaamaan uusia näkökulmia vanhaan aiheeseen: tulkitsemaan aineiston uudelleen, jotta sitä voitaisiin tuoreella tavalla käyttää koko kieltä koskevan tutkimuksen eteenpäin viemiseen.

Kyse ei tietenkään ole sattumanvaraisesta tosiasioiden sekoittamisesta vaan kattavan teorian järjestelmällisestä luomisesta. Formaali kuvaus ei ole päämäärä sinänsä; se on tapa saada kuvauksesta eksplisiittinen, loogisesti koherentti ja yksikäsitteinen. Näin voidaan selventää sellaisia lingvistisiä ilmiöitä ja ilmiöiden välisiä suhteita, jotka intuitiivisesti jäävät huomaamatta tai ainakin vaille johdonmukaista kuvausta. X'-teoriaan pohjautuvassa kuvauksessani mm. genetiiviattribuuttien jako kahdeksi erityyppiseksi määritteeksi saa selkeät perustelut ja jälkiattribuuttien erilainen suhde pääsanaan terävöityy. Toisaalta kuvaukseni tuo havainnollisesti esille sen, ettei X'-teorian mukainen lausekerakennesäännöstö ole ongelmaton edes prototyyppisimmän lausekkeen, SbP:n, kohdalla: hajoavat rakenteet ja appositiot ovat hankalia kuvata ja kuvailuattribuuttien ja tarkenteiden väli-maastoon kuuluvat määritteet vaikeita luokittaa. Numeraaleja sisältävät lausekkeet eivät nekään saa tyydyttävää kuvausta, jos pitäydytään tiukasti X'-teorian periaatteisiin.

LÄHTEET

- ADES, A. E. & MARK J. STEEDMAN 1982. On the order of Words. *Linguistics and Philosophy* 4 s. 517–558.
- BARWISE, JON & ROBIN COOPER 1981: Generalized Quantifiers and Natural Language. *Linguistics & Philosophy* 4 s. 159–219.
- BLOOMFIELD, LEONARD 1933: *Language*. Holt, Rinehart and Winston, New York.

- CARLSON, LAURI 1987: RUG: Regular Unification Grammar. Yleisen kielitieteen laitoksen tutkijaseminaarissa keväällä 1987 esitetty moniste.
- 1988: Unifikaatio kielioppiformalismina. Teoksessa Östman, Jan-Ola, Urpo Nikanne, Kristiina Jokinen & Kalevi Wiik (toim.) SKY 1988. Suomen kielitieteellisen yhdistyksen vuosikirja 1988 s. 79–90. Suomen kielitieteellinen yhdistys, Hki.
- CHOMSKY, NOAM 1970: Remarks on Nominalization. Teoksessa Jacobs, Roderick & Peter Rosenbaum (eds.) Readings in English Transformational Grammar s. 184–221. Ginn and co. Waltham, Massachusetts.
- GAZDAR, GERALD 1982: Phrase Structure Grammar. Teoksessa Jacobson, Pauline & Geoffrey K. Pullum (eds.) The Nature of Syntactic Representation s. 131–186. D. Reidel, Dordrecht.
- GAZDAR, GERALD, EWAN KLEIN, GEOFFREY K. PULLUM & IVAN A. SAG 1985: Generalized Phrase Structure Grammar. Harvard University Press, Cambridge, Massachusetts.
- GAZDAR, GERALD & GEOFFREY K. PULLUM 1982: Generalized Phrase Structure Grammar: a Theoretical Synopsis. Indiana University Linguistics Club, Bloomington, Indiana.
- HAKULINEN, AULI & FRED KARLSSON 1979: Nykysuomen lauseoppia. Suomalaisen Kirjallisuuden Seura, Hki.
- HAKULINEN, AULI & MARJA-LEENA SORJONEN 1986: Keskustelun kuvaamisesta. Teoksessa Leino, Pentti & Jyrki Kalliokoski (toim.) Kieli 1 s. 5–38. Helsingin yliopiston suomen kielen laitos, Hki.
- HAWKINS, JOHN A. 1984: Modifier-Head or Function-Argument Relations in Phrase Structure? The evidence of some word order universals. *Lingua* 63 s. 107–138.
- HOEKSEMA, JACOB 1984: Categorical Morphology. Drukkerij van Denderen B.V., Groningen.
- HOLMA, TAPANI 1966: Positiivisesti kongruoivien edussanetta edeltävien attribuuttien keskinäisestä järjestyksestä huomioon otettuna genetiiviattribuutin paikka. *Laudaturtutkielma*. Jyväskylän yliopiston suomen kielen laitos.
- HOPPER, PAUL & SANDRA A. THOMPSON 1984: The Discourse Basis for Lexical Categories in Universal Grammar. *Language* 4 s. 703–752.
- JACKENDOFF, RAY 1977: X syntax: A Study of Phrase Structure. *Linguistic Inquiry Monograph Two*. MIT Press, Cambridge, Massachusetts.
- JOKINEN, KRISTIINA 1986: X'-teoria ja suomen substantiivilausekkeen rakenne. Pro gradu -tutkielma. Helsingin yliopiston yleisen kielitieteen laitos.
- 1987a: Genetiiviattribuutin kuvauksesta. Esitelmä XIV Kielitieteen päivillä Turussa 9.5.1987.
- 1987b: The Structure of NPs and Finite State Grammars. Esitelmä The Tenth Scandinavian Conference of Linguistics -konferenssissa Bergenissä 11.6.1987.
- KANGASMAA-MINN, EEVA 1980: Suomen kielen persoonallisesta passiivista. — *Sananjalka* 22 s. 57–70.
- KARLSSON, FRED 1977: Syntaktisten kongruenssijärjestelmien luonteesta ja funktioista. *Vir.* 81 s. 359–391.
- 1985: Process Grammar. Teoksessa Karlsson, Fred (ed.) *Computational Morphosyntax*. Research Report 1981–85. University of Helsinki, Department of General Linguistics. Publications 13.
- KARLSSON, GÖRAN 1971: Kardinaalilukusanää edeltävän attribuutin numeruksesta. *Sananjalka* 13 s. 52–80.
- KARTTUNEN, LAURI 1984: Agreement with Numerals in Finnish. Unfinished draft.
- 1986: D-PATR: A Development environment for Unification-Based Grammars. CSLI Report 61. Stanford.
- KOIVISTO, HELINÄ 1986: Verbien partisiippistamisen ja partisiippiattribuuttien käytön ongelmia. Teoksessa Nahkola, Kari (toim.) *Folia Fennistica & Linguisti-*

- ca. Esitelmiä ja tutkielmia. s. 27–66. Tampereen yliopiston suomen ja yleisen kielitieteen laitoksen julkaisuja 13, Tampere.
- LARSEN, SINIKKA 1969: Adjektiiviattribuuttien järjestyksestä. *Laudaturtutkielma*. Tampereen yliopiston suomen kielen laitos.
- LEIWO, MATTI 1977: Kielitieteellisiä näkökohtia viivästyneestä kielenkehityksestä. *Studia Philologica Jyväskyläensia* 10, Jyväskylä.
- LOMAN, BENGT 1956: Om relation mellan ordföljd och betydelse hos framförställda attributiva bestämningar till substantiviska huvudord. *Arkiv för nordisk filologi* 73 s. 218–244.
- 1964: Adjektivkombinationer. *Förhandlingar vid sammankomst för att dryfta frågor rörande svenskans beskrivning I. Dupl.* s. 72–81, Stockholm.
- LYONS, JOHN 1977: *Semantics I–II*. Cambridge University Press, Cambridge.
- MUYSKEN, PIETER 1982: Parametrizing the Notion 'Head'. *Journal of Linguistics Research* 2.3 s. 57–75.
- PENTTILÄ, AARNI 1963: Suomen kielioppi. WSOY, Porvoo.
- 1967: Attribuuttien järjestyksestä. *Vir.* 71 s. 259–267.
- POLLARD, CARL & IVAN A. SAG 1987: *Head-Driven Phrase Structure Grammar. Information-based Syntax and Semantics. Volume 1: Fundamentals*. To appear in *CSLI Lecture Notes 12*, Stanford.
- PULLUM, GEOFFREY K. 1985: Assuming Some Version of X-bar Theory. *Chicago Linguistic Society* 21/1 s. 323–353.
- SELLS, PETER 1986: *Lectures on Contemporary Syntactic Theories*. University of Chicago Press, Chicago.
- SETÄLÄ, E. N. 1973: Suomen kielen lauseoppi. 16. painos. Tarkistanut Matti Sadeniemi. Otava, Keuruu.
- SHIEBER, STUART M. 1986: *An Introduction into Unification-Based Approaches to Grammar*. University of Chicago Press, Chicago.
- SHIEBER, STUART M., LAURI KARTTUNEN & FERNANDO C. N. PEREIRA 1984: *Notes from the Unification Underground: a Compilation of Papers on Unification-Based Grammar Formalisms*. Technical Report 327, Artificial Intelligence Center, SRI International, Menlo Park, California.
- SIRO, PAAVO 1957: Suomen kielen subjektista kielen rakenteen osana. *Vir.* 61 s. 181–190.
- 1960. Suomen yksinäislauseen perustavia ongelmia. *Sananjalka* 2 s. 36–51.
- STEEDMAN, MARK 1987: Gapping as Constituent Coordination. Esitelmä ASL/LSA Conference on Language and Logic -konferenssissa Stanfordinissa heinäkuussa 1987.
- TUOMIKOSKI, RISTO 1969: Asemoilmaustemme luokittelusta. *Suomi* 114:4. Suomalaisen Kirjallisuuden Seura, Forssa.
- VENNEMANN, THEO & R. T. HARLOW 1977: *Categorial Grammar and Consistent VX Serialization*. *Theoretical Linguistics* 4 s. 227–254.
- VILKUNA, MARIA 1980: Referenssi ja spesies suomenkielisten tekstien tulkinassa. *Lisensiaatintyö*. Helsingin yliopiston suomen kielen laitos, Hki.
- 1986: Konstituenttirakenteen ja sanajärjestyksen ongelmia: suomen hajoavat infiniittirakenteet ja lauserakennekielioppi. Teoksessa Leino, Pentti & Jyrki Kalliokoski (toim.) *Kieli* 1 s. 183–223. Helsingin yliopiston suomen kielen laitos, Hki.
- YLI-VAKKURI, VALMA 1970: Konkreettisen substantiivin paikallissija-attribuutti suomen paikallissijajärjestelmässä. *Sananjalka* 12 s. 35–71.
- ZWICKY, ARNOLD M. 1985: *Heads*. *Journal of Linguistics* 21 s. 1–29.

The structure of Finnish noun phrases

KRISTIINA JOKINEN

The article studies the structure of the Finnish noun phrase in the framework of a GPSG-style X'-theory. Special attention is paid to genitive attributes and phrases with numerals. Shortcomings and virtues of X'-theoretic description are discussed. Also a short survey is given of an implementation of the description on Xerox Lisp Machines.

A distinction is made between *noun phrase* and *nominal phrase*. The former denotes a syntactic category with a noun as head; the latter is an abbreviation for a class of categories which can function as subjects or objects in a sentence.

Pre- and postmodifiers of a noun are divided into two bar-levels, which correspond roughly to linear order and semantic distinctions among different modifiers. On the first level, premodifiers consist of descriptive modifiers, while postmodifiers consist of subcategorized arguments; on the second level, premodifiers include specifiers, and postmodifiers include adjuncts. Specifiers refer to demonstrative pronouns, quantifiers and (possessive) genitives, and they specify the referent with respect to its definiteness and givenness. Descriptive modifiers, on the other hand, describe different properties or the quality of the referent, and they include adjectives and descriptive genitives. Descriptive modifiers also contain specifier-type attributes like superlatives, ordinal numerals and the so-called pronominal adjectives.

Subcategorized arguments are modifiers whose morphosyntactic properties are governed by the head noun: partitive case postmodifiers of quantifier-type nouns (*pullo olutta* 'a bottle of beer'), and various complements (*ehdotus uudeksi laiksi* 'a suggestion for a new law', *kyky oppia* 'an ability to learn', *ajatus että lähtisimme* 'the idea that we would go'). Second-level postmodifiers consist of other clausal attributes, locative case modifiers which cannot be considered subcategorized arguments

(*maljakko Kreikasta* 'a vase from Greece'), and appositions.

Two kinds of genitives can appear in Finnish noun phrases. The specifier genitive, which mostly has a possessive reading, occurs on the second bar-level and forms a noun phrase with no postmodifiers. It may consist of an arbitrarily long genitive chain (*aivokuoren otsalohkon kolmannen aivopoimun tyven vaurio* 'a lesion of the root of the third gyrus of the frontal lobe of the cerebrum'), and it is a potential source of recursive center-embedded constructions. Though there is an order constraint among specifiers (demonstratives precede quantifiers), the genitive may appear freely among them. However, there are differences in scope and focus of the phrases.

The descriptive genitive, which belongs to the first level modifiers, is peculiar to Finnish. It follows adjectives (if there are any), and appears closest to the head. It does not form a phrase, but is made up of a lexical entry, i.e. a single word or a lexicalized, conceptually coherent totality (*ahkera fysiikan opiskelija* 'an industrious student of physics', *innokas Punaisen Ristin jäsen* 'an eager member of the Red Cross'). Thus it cannot form recursive constructions. It gets a generic reading.

The difference between specifier and descriptive genitives is most clearly exemplified in the definiteness of noun phrases. The specifier genitive makes the referent definite (*tuolin vihreä jalka* 'the green leg of the chair'), while the descriptive one does not (*vihreä tuolin jalka* 'a green leg of a chair, a green chair leg').

The head in phrases with numerals is controversial. It is suggested that numerals form a constituent of their own. Their description thus becomes more coherent, and the X'-theoretic concept of head is maintained.

Numerals have two types of postmodifiers: a first-level modifier which tells what kind of referents are referred to,

and a second-level modifier which specifies the group or set from which referents are picked out. The first postmodifier is in the partitive case if the numeral is in the nominative case (*viisi koiraa* 'five dogs'); otherwise it agrees in case and number with the numeral (*viidelle koiralle* 'to five dogs'). The latter postmodifier is always in the elative case regardless of the case of its head (*viisi koirista* 'five of the dogs'). If both modifiers occur in the same NP, the partitive-and-agreeing modifier precedes the elative one, which is conceivable on the basis of their semantics.

Premodifiers of a numeral appear in the plural, although the numeral and its postmodifier are singular. Premodifiers of a numeral make the whole phrase definite, cf. *viisi ruskeaa koiraa* 'five brown dogs' vs. *ruskeat viisi koiraa* 'the five brown dogs'. If the noun in a phrase denotes time, then premodifiers can appear in the singular, but the interpretation of such a construction differs from the one with premodifiers in the plural.