

PRODUKTIIVISEN (MORFO)SYNTAKSIN INDEKSI LAPSEN KIELEN KOMPLEKSISUUSMITTARINA

Lea Nieminen
Jyväskylän yliopisto

Lähivertailuja 17.

Toimittaneet Annekatrin Kaivapalu, Külvi Pruuli.

Jyväskylä Studies in Humanities, 53.

Jyväskylä: Jyväskylän yliopisto, 2006, pp. 69–83.

ISBN 951-39-2450-5

ISSN 1459-4323

1 Johdanto

Kompleksisuuden voi katsoa olevan yksi lapsen kielen tutkimuksen keskeisiä kiinnostuksenkohteita, sillä kielen kehittymiseen tuntuu implisiittisesti kuuluvan ajatus siitä, että se on samalla jonkinasteista monimutkaistumista ja monipuolistumista, toisin sanoen eri osajärjestelmien kompleksoitumista. Tähän ilmiöön ja sen variaatioon on pyritty pääsemään käsiksi erilaisilla tutkimusmenetelmillä.

Ehkä kaikkein laajimmalle levinnyt ja käytetyin metodi kompleksisuuden selvittämiseksi kielenomaksumistutkimuksissa on ollut MLU (Mean Length of Utterance) eli ilmauksen keskipituus (Brown: 1973). Sen yksinkertaisena peruseriaatteena on laskea sadasta lapsen tuottamasta ilmauksesta morfeemimäärän keskiarvo. Runsaasta kritiikistä huolimatta (ks. Rollins, Snow & Willet 1996: 244–246) tätä metodia on käytetty hyvin paljon. Eri tutkimuksissa MLU:ta luonnehditaan varsin eri tavoin, ja sen on katsottu olevan ainakin kielellisten taitojen (language proficiency) kehityksellinen indeksi (Johnston 2001: 161), rakenteellisen kehityksen yleisindikaattori (Miller 1981: 25), syntaktisen kompleksisuuden ja monipuolisuuden mittari (Rondal, Ghiotto, Bredart & Bachelet 1987: 444) sekä lausekompleksisuuden indeksi (Blake, Quartaro & Onorati 1993: 139). Ei siis ole olemassa varsinaisesti konsensusta siitä, mitä ilmausten keskipituus loppujen lopuksi lapsen tuotoksista kertoo.

Kiistanalaisen MLU:n rinnalle Hollis S. Scarborough (1990) kehitti toisentyypin kompleksisuusmittarin Index of Productive Syntaxin (IPSyn), jossa nimensä mukaisesti analysoidaan lapsen tuotosten syntaktista koostumusta mutta samalla myös näiden rakenteiden produktiivisuutta. Kvantitatiiviseen morfeemien laskemiseen verrattuna tällainen kvalitatiivinen ote on edistysaskel, koska kielen kehitystä ei enää redusoida yhden rakenteellisen yksikön keskimääräiseen esiintymiseen. IPSyniä luonnehditaan avoimesti kompleksisuusmittariksi, mutta tämän(kään) mittarin taustalla olevaa kompleksisuuskäsitystä ei ole pohdittu aiemmin. IPSyn on myös metodina varteenotettava kilpailija MLU:lle. Näistä syistä sen tarkempi tarkastelu on perusteltua.

Seuraavassa käsittelen ensin produktiivisen syntaksin indeksin peruseriaatteita ja sen tulosten antamia näkökulmia kielenainekseen. Otan lyhyesti esil-

le myös muutosprosessin, joka käytiin läpi, kun englanninkielinen mittari muutettiin suomenkieliseksi. Suomalaisversio poikkeaa jonkin verran alkuperäisestä mittarista ja voisi näin toimia esimerkkinä, jos mittari muokattaisiin läheisten sukukieltemme tarpeisiin. Yksityiskohtainen kuvaus muutosprosessista sekä opastus sen käyttöön on luettavissa Niemisen ja Torvelaisen (2003) artikkelista. Lopuksi paneudun tämänkertaiseen pääteemaani eli kompleksisuuteen ja erityisesti siihen, minkälaista kompleksisuuskäsitystä IPSyn edustaa.

2 Produktiivisen syntaksin indeksi

2.1 Peruseriaatteet

Produktiivisen syntaksin indeksi perustuu ennalta valittuihin kielen rakenteisiin, jotka on havaittu kielenomaksumisprosessissa keskeisiksi. Lapsen spontaanista tuotoksesta etsitään näiden rakenteiden reaalistumia erilaisissa konteksteissa. Produktiivisuutta todennetaan juuri tällä reaalistumien erilaisuudella. Jokaisesta mittariin valitusta rakenteesta otetaan huomioon korkeintaan kaksi tarpeeksi erilaista tuotosta ja niistä kummastakin annetaan piste. Jos jotakin rakennetta ei lapsen tuotoksesta löydy lainkaan, se ei tuota pisteitä, yksi rakenteen esiintymä tai useat keskenään samanlaiset esiintymät antavat yhden pisteen ja kaksi tai useampi erilainen toteutuma kaksi pistettä. Analyysin välitön tulos on siis pistemäärä, jota voidaan käyttää tilastollisessa tarkastelussa hyväksi ja jonka avulla yksilöitä tai tutkimusryhmiä voidaan vertailla toisiinsa.

Scarborough (1990) valitsi mittarin alkuperäisversioon mukaan kaikkiaan 56 englannin kielen muotoa ja rakennetta, jotka edustavat neljää eri kategoriata: mukana on 11 nominilauseketta, 16 verbilauseketta, 10 kysymys- ja kieltolauseketta sekä 19 lauserakennetta. Lisäksi jokaisessa kategoriassa on yksi kohta varattu tarkemmin nimeämättömille rakenteille, jotka tutkija kuitenkin arvioi sellaisiksi, että ne tuovat huomionarvoisen lisän lapsen rakennerepertoariin ja siten ansaitsevat pisteen. Näin maksimipistemääräksi muodostuu 120. Kun otetaan huomioon englannin kielen rakenne, on luontevaa, että erityisesti syntaktiset kokonaisuudet korostuvat kaikissa kategorioissa. IPSyn on siis hyvin kieli-kohtainen mittari, joka vaatii huomattavaa muokkaamista, ennen kuin sitä voidaan soveltaa toisessa kielessä.

IPSyn-analyysin aineistona käytetään tavallisesti sataa lapsen tuottamaa ymmärrettävää ilmausta. Näistä ilmauksista kerätään korkeintaan kaksi erilaista kunkin mittarirakenteen toteutumaa, joiden pohjalta pisteet lasketaan. Rakenteiden produktiivisuuden arviointi pohjautuu erilaisuuden periaatteeseen, joka läpäisee koko IPSyn-analyysia. Scarborough (1990) määritteli kolme produktiivisuuskriteeriä, joiden avulla samaa rakennetta edustavien tuotosten erilaisuus määritellään. Leksikaalinen kriteeri (lexical criterion) koskee yksittäisiä sanoja: rakennevalikoimaan mukaan otetuista yksittäisistä sanakategoriasta (esim. erilaiset rinnastuskonjunktiot) täytyy löytyä kaksi eri edustajaa lapsen

tuotoksista kahden pisteen saavuttamiseksi. Kontekstikriteerin (context criterion) avulla arvioidaan suffiksaalisten ainesten produktiivisuutta. Kaksi pistettä saa, jos esimerkiksi monikon tunnus esiintyy kahdessa eri sanassa (esim. *cars, toys*; suomalaisversiossa esim. *autot, leluilla*), joista kumpikaan ei ole vain monikollisena esiintyvä sana (*trousers* 'housut' tai *scissors* 'sakset'). Lausekekriteeri (phrasal criterion) liittyy lauseke- ja lauserakenteiden produktiivisuuden määrittelyyn ja se perustuu erilaisten sanojen määrään. Esimerkiksi genetiivimääritteestä ja nominiedussanasta muodostetun nominilausekkeen *Peter's car* rinnalle toisen pisteen tuovaksi edustumaksi kelpaavat *Peter's bike, John's car* tai *John's bike*, koska niistä jokaisessa on vähintään yksi eri sana kuin *Peter's car* -lausekkeessa. Erilaisten sanojen vaadittu määrä vaihtelee sen mukaan, kuinka monisanaisesta rakenteesta on kyse.

Produktiivisuusperiaatteen lisäksi IPSynille ominaista on kumulatiivisuus. Yksi ainoa lapsen tuottama ilmaus voi antaa pisteitä hyvin moneen mittarissa haettavaan rakenteeseen, ja rakenteet voivat olla keskenään sisäkkäisiä eli yksi rakenne on toisen rakenteen osa. Niinpä *Peter's car* -nominilauseke on esijmerkki genetiivin toteutumasta ja vähintään kahden sanan yhdistelmästä, eli sen voi sijoittaa kahteen kohtaan mittarissa. Mitä monisanaisempia lapsen tuotokset ovat, sitä enemmän niissä todennäköisesti on IPSynissä haettujen rakenteiden sisäkkäisyyttä ja pistepotentiaalia. Kumulatiivisuus antaa tilaa lapsen tuotosten vähittäiselle monimutkaistumiselle ja laajenemiselle, ja jo pienikin edistys lisää pisteitä.

2.2 IPSyn-analyysin antamia näkökulmia lapsen kieleen

Produktiivisen syntaksin indeksi tuottaa oikeastaan neljänlaisia tuloksia, joista tähän mennessä on toistaiseksi käytetty eniten kokonaispistemäärää. Kokonaispistemäärien avulla voidaan tutkia suuriakin koehenkilöryhmiä, käyttää tilastollisia analyyseja ja vertailla koehenkilöryhmiä keskenään. Kokonaispistemäärien laskeminen ja vertailu ei vaikuta kovin paljon edistyneemmältä menetelmältä paljon kritisoituun ilmausten keskipituuden analysointiin verrattuna, onhan molemmissa kyseessä vain numeerinen arvo, joka sellaisenaan ei paljasta lapsen tuotoksista oikeastaan yhtään mitään. Kokonaispisteiden tulkintaa helpottaisi laaja otokseen tuloksiin perustuva normitus, joka määrittelisi muun muassa normaaliin ja poikkeavaan kehitykseen viittaavat pistemäärät. Englannin kielen osalta tällainen pisteiden standardointi on jo tehtykin (Scarborough 1990). Pistenormitukseen liittyvä kvalitatiivinen tieto siitä, minkä rakenteiden käyttöä pistemäärät implikoivat, olisi kuitenkin erittäin tervetullut lisä. Ilman lisäselvitystä kokonaispistemäärät antavat informaatiota itse kielestä yhtä vähän kuin MLU:n keskiarvoluvutkin.

Kokonaispistemäärän lisäksi IPSyn-analyysi tuottaa kuitenkin koko joukon kvalitatiivista tietoa, jota tosin on hyödynnetty tähän mennessä raportoiduissa tutkimuksissa vielä varsin vähän. Ensinnäkin pisteet jakautuvat eri osioiden kesken (alkuperäismittarissa nominilausekkeet, verbilausekkeet, kysymys- ja kieltorakenteet sekä lauserakenteet), jolloin ne kertovat osioiden keskinäisistä suhteista ja kielenkäyttäjän potentiaalisista vahvuusalueista. Tähän

tulosten tulkintamahdollisuuteen ovat tarttuneet ainakin Tomblin, Spencer, Flock, Tyler ja Gantz (1999) kokleaimplanti- ja kuulolaitelasten puheentuottamista vertailevassa tutkimuksessa sekä Hadley (1998) selvitellessään kielenkehitykseltään poikkeavien lasten verbilausekkeita. Pitkittäistutkimuksissa osiopisteiden vertailu antaisi viitteitä kulloisistakin muuttumassa olevista osaluista, ja poikittaistutkimuksessa taas olisi mahdollista selvittää, mille osaluelle lasten keskinäiset erot paikallistuvat.

Toinen kvalitatiivinen tulos perustuu lapsen tuotoksista löydettyihin rakenteisiin sekä niiden toteutumisiin. Koska analyysi tehdään käytännössä merkitsemällä IPSyn-lomakkeelle maksimissaan kaksi rakenteen toteutumaa, on mahdollista selvittää yksityiskohtaisesti, mitä mittarin rakenteista lapsi tuottaa ja millaisessa ilmiössä nämä rakenteet lapsen tuotoksissa konkreettisesti toteutuvat. Toki on myönnettävä, että edustumat ovat valikoituneet käytettyjen produktiivisuuskriteerien takia. Lisäksi on muistettava, että tarkasteltavaksi saadaan enintään kaksi tuotosta, joten päätelmiä puhekuviosta tai suosikkirakenteista ei tämän lomakkeelle kirjatun tuloksen pohjalta voi tehdä. Kuitenkin analyysilomakkeen käyttö antaa hyvät mahdollisuudet koehenkilön rakennerepertoarin tarkasteluun muutenkin kuin dikotomisena tuottaa-ei tuotamuuttujana.

Kolmas kvalitatiivinen tulostyyppi on mainittu jo mittarin nimessäkin. Rakenteiden produktiivisuusasteeseen on mahdollista päästä käsiksi lomakemerkintöjen pohjalta. Vaikka kaksi erilaista esiintymää samasta rakenteesta ei vielä takaakaan sen täyttä produktiivisuutta, kyseessä on kuitenkin vihje siitä, että rakenne on juurtunut todennäköisesti vahvemmin kuin vain yhden esiintymän rakenteet. Erityisesti produktiivisuuden määrittelemiseen antaa lisätukea itse analyysivaihe, jossa tutkija on käynyt läpi koehenkilön kaikki tuokset ja punninnut niitä erilaisuusperiaatteiden mukaisesti. Ovatko lomakkeelle kirjatut kaksi tuotosta ainoat koehenkilön koko aineistossa esiintyvät esimerkit tästä rakenteesta? Onko siis aihetta olla vielä varovainen produktiivisuuspäätelmässä? Vai mahtuivatko lomakkeelle vain kaksi ensimmäistä niistä lukuisista erilaisista edustumista, jolloin päätelmät rakenteen produktiivisuudesta ovatkin jo huomattavasti vahvemmalla pohjalla? Onko yhden pisteen antanut rakenne vain yhden toteutuman varassa koehenkilön aineistossa? Vai onko sittenkin kyse puhekuviosta, jonka lukuisat edustumat eivät täytä erilaisuuskriteeriä?

Kvalitatiivisista tulkintamahdollisuuksista suhtautuisin kaikkein skeptisimmin juuri tähän produktiivisuuden arviointiin pelkkien lomakkeelle kirjatujen tulosten avulla, vaikka Scarborough on korostanutkin juuri tätä ominaisuutta nimeämällä mittarin sen mukaan. Kaksi esiintymää ei riitä produktiivisuuden määrittelemiseen. Tomasello (2000a) esittää, että omaksumisen varhaisessa vaiheessa lapset käyttävät kieltä sellaisena kuin ovat aikuisen kuulleet sitä käyttävän. Tietyn taivutusmuodon tai rakenteen tuottaminen ei tarkoita sitä, että sitä osattaisiin soveltaa muihin samantyyppisiin sanoihin. Lapsen kielioppi rakentuisi siis vielä pääosin yksittäisten lekseemien esiintymismuotojen ympärille, ja sääntöpohjaisella produktiivisuudella olisi vasta varsin pieni osuus kielen tuottamisessa. Tällainen kokonaisuuksien omaksumista korostava käyttö-

lähtöinen malli (usage-based model) mahdollistaa samaa muotokategoriaa tai rakennetta edustavat erilaiset tuotokset ilman, että kyseessä on vielä muodon tai rakenteen produktiivinen tuottaminen abstrahoidun säännön avulla. Myös Langacker (1988) tuo esille, että ihminen käyttää samanaikaisesti sekä kokonaisuuksina opittuja että sääntöjen avulla rakennettuja muotoja, jopa saman muotoryhmän sisällä. Langackerin ajatuksia soveltaen Tomasello (2000a) kuvaa lapsen kielentuottamista "leikkaa ja liimaa" -strategiaksi, jossa uusi ilmaus rakennetaan kierrättämällä aikaisemmista ilmauksista löytyviä sopivia palasia – yksittäisiä sanoja, abstrahoituja kategorioita, osittain abstrahoituja ilmausskeemoja ja kaikkea näiden väliltä – uudessa kontekstissa. Tämä mutkistaa entisestään produktiivisuuden arviointia. Analyysivaiheessa tutkijalle kerääntyneet havainnot rakenteiden esiintymisestä muualla kuin lomakkeelle valikoituneissa esimerkeissä ovat siis ensiarvoisen tärkeitä tietoja produktiivisuuspäätelmien kannalta.

2.3 Suomalaisversio: produktiivisen (morfo)syntaksin indeksi

Kuten jo aikaisemmin mainitsin, IPSyn on kielikohtainen, ja vain senn periaate voidaan siirtää sellaisenaan muiden kielten tutkimukseen. Niinpä IPSynin soveltaminen suomalaisaineiston analysointiin edellytti koko rakennevalikoiman muokkausta suomen kielen ja suomen omaksumisen kannalta relevantiksi. Jo alkujaan oli selvää, että mittari ei voi olla pelkästään tai pääasiallisesti syntaktinen vaan suomen monipuolisen taivutusmorfologian on tultava mittarissa edustavasti esille, samoin syntaktisiin rakenteisiin liittyvä morfologinen komponentti. Näin ollen suomalaisesta IPSynistä lähdettiin tietoisesti rakentamaan morfosyntaktista analysointivälinettä.

Lähtökohtana oli tehdä suomalais-IPSynistä 2;0–3;0-vuotiaiden lasten puhetuotosten analysointiin sopiva väline. Rakennevalikoiman tekemisessä apuna käytettiin olemassa olevia tutkimustuloksia suomen kielen omaksumisesta. Niistä saatuja näkemyksiä täydennettiin analysoimalla 2;0- sekä 2;6-vuotiaiden lasten spontaaneja puhetuotoksia erityisesti morfosyntaksin kannalta. Tähän analyysiin valittiin kummastakin ikäluokasta MLU:lla mitattuna ääripäiden edustajat, jotta kehityksen variaatio saataisiin kartoitettua mahdollisimman edustavasti. Näin mittari erottelisi lapsia ja reagoisi herkästi rakennerepertoaarin muutoksiin. (IPSynin rakennevalikoiman muokkauksesta tarkemmin Nieminen & Torvelainen 2003.)

IPSynin suomalaisversiossa on 46 erilaista erikseen nimettyä suomen kielen ja suomen omaksumisen kannalta keskeistä muotoa ja rakennetta, jotka jakaantuvat kolmeen eri osioon (ks. Liite: IPSyn-rakenteet esimerkkiteututuminen). Nominirakenteita on 9 (nominien sijataivutus, liitepartikkeleiden käyttö nominien yhteydessä, nominilausekkeet), verbirakenteita 16 (verbien taivutus persoonissa, tempuksissa ja moduksissa, liitepartikkeleiden käyttö verbeissä, verbiketjut, kieltoverbirakenteet) ja lauserakenteita 21 (esim. S + V, V + O, S + V + O, sivulause, rinnasteinen lause, adpositiolauseke). Jokaisessa osiossa on lisäksi yksi kohta ennalta määrittelemättömiä rakenteita varten. Maksimipistemääräksi muodostuu näin 98. Scarborough'n (1990) alkuperäisversioon verrat-

tuna erona on se, että kysymys- ja kieltolauseille ei ole varattu omaa kategoriaa. Kieltoilmaukset on sijoitettu verbirakenteisiin, mikä onkin luonnollista, koska kieltoverbi on niin keskeinen osa suomen negaatiota. Kysymyslauseitakaan ei ole jätetty analyysin ulkopuolelle, mutta omaa osiota tai edes erillisiä kysymysrakenteita niiden ei katsottu tarvitsevan tässä mittarissa, koska kysymyksen esittäminen riippuu kovin paljon tilanteesta. Kysymysten esittämisen ei haluttu lisäävän tai vähentävän pisteitä. Niinpä päädyttiin siihen, että kysymykset analysoidaan väitelauseiden tavoin. Lisäksi polaariset *-kO*-kysymykset huomioidaan sekä nomini- että verbiosiossa liitepartikkeleiden käytön yhteydessä. (Rakennevalikoimasta tarkemmin Nieminen & Torvelainen 2003.)

Suomenkielisen version käytössä keskeisenä periaatteena on puhekielisyiden huomioon ottaminen, jolloin itse rakenteen tuottamista korostetaan eikä kirjakielistä tai äärimmäisen huolellista artikulaatiota edellytetä. Toinen rakenteen korostamiseen tähtäävä käytäntö on se, että esimerkiksi taivutuspoikkeamista ei rankaista: *audolla* on esimerkki adessiivista kieliopin vastaisesta morfonologisesta vaihtelusta huolimatta ja *tää ottaa pallo* luetaan S + V + O -rakenteen toteutumaksi, vaikka objektia ei olekaan taivutettu akkusatiivissa. Mahdollisesti puutteellisen ilmiön, kieltä omaksuvan lapsen puhekielisen tuotoksen taustalta halutaan siis nähdä onnistunut rakenne. IPSynin suomalaisversion käyttöön annetaan yksityiskohtaiset ohjeet Niemisen ja Torvelaisen (2003) artikkelissa.

3 Kompleksisuus ja IPSyn

Vaikka kompleksisuuden kehitys onkin vahvasti mukana lapsen kielen tutkimuksessa joko eksplisiittisesti tai implisiittisesti, käsitteen syvempi pohdinta ja määrittely on valitettavasti usein sivuutettu niin, että vain metodien ja tutkijan esille tuomien tulospainotusten perusteella voi tehdä päätelmiä siitä, miten tutkija kompleksisuuden ymmärtää. Viimeaikaisessa kielitypologiatutkimuksessa kompleksisuuden olemusta on pohdittu useasta näkökulmasta, jotka on syytä ottaa huomioon myös kielenomaksumistutkimuksen piirissä.

Miestamo (painossa) ja McWhorter (2001) erottavat toisistaan absoluuttisen ja suhteellisen lähestymistavan kompleksisuuteen. Absoluuttisella kompleksisuudella tarkoitetaan yksinkertaista systeemin osien määrää. Informaatioteoreettisesta näkökulmasta kompleksisuus on suoraan verrannollinen ilmiön lyhimmän mahdollisen kuvauksen pituuteen (Dahl 2004; Gell-Mann 1996: 81–82): mitä pidempi kuvaus tarvitaan sitä kompleksimmasta ilmiöstä on kyse. Suhteellisella lähestymistavalla kielen kompleksisuuteen taas tarkoitetaan kielenkäyttäjän kokemusta: mitä hän pitää vaikeana, mitä taas helppona. Tällainen kokemus on tietenkin hyvin subjektiivinen ja vaihtelee sen mukaan, millaisesta kielenkäyttäjistä on kyse. Natiivisuomalainen pitää vaikeana todennäköisesti erilaisia asioita kielessä kuin suomea vieraana kielenä oppiva, ja eritaustaisten kielenopiskelijoidenkin kokemukset vaihtelevat keskenään. Vielä oma joukkon-

sa on kieltä äidinkielenään omaksuvat lapset ja heidän vaikeiksi kokemansa asiat kielessä. Suhteellinen kompleksisuus ei myöskään ole välttämättä stabiili ilmiö, koska esimerkiksi kielenoppimis- tai omaksumisprosessin eri vaiheissa erilaiset asiat voivat näyttäytyä hankalina. Oman lisänsä suhteellisen kompleksisuuden määrittelyyn tuo vielä se, otetaanko lähtökohdaksi kielen havaitseminen, tuottaminen tai esimerkiksi semanttisen sisällön hallinta.

Monissa yhteyksissä on havaittu, että absoluuttinen kompleksisuus ei suinkaan tarkoita automaattisesti myös suhteellista kompleksisuutta. Kielenkäyttäjä ei siis välttämättä pidä vaikeana sellaista rakennetta, joka absoluuttisesta näkökulmasta vaikuttaisi hyvinkin kompleksilta. Hän toimii erilaisten yksiköiden varassa kuin kielen kuvaaja, jolle on tärkeää osoittaa rakenteen säännönmukaisuudet. Kielenkäyttäjän resursseihin kuuluu abstraktien sääntöjen ja yksittäisten sanamuotojen lisäksi myös isompia puhekuvioita ja kehyksiä, joita ei tarvitse pilkkoa teoreettisiin alkutekijöihinsä, koska ne toimivat pragmaattisesti laajempina kokonaisuuksina (esim. Laury 2005). Myös lapsen kielen tutkimuksessa on tehty tällaisia havaintoja. Muun muassa Peters (1983), Tomasello (1992, 2000a, 2000b) ja Kauppinen (1998) ovat tuoneet puhekuviot ja opitut kokonaisuudet monipuolisesti esille keskeisenä kielenomaksumisstrategiana. Näiden tutkimusten tulokset ovat vastaesimerkkejä sille oletukselle, että lapsi etenisi omaksumisprosessissaan yksinkertaisemmista monimutkaisempiin rakenteisiin, ainakin jos näitä rakenteita määritellään absoluuttisesta kompleksisuudesta käsin (ks. esim. Kauppinen 1998: 149–150).

Suhteellinen kompleksisuus näyttää siis sangen monimuuttujaiselta ja vaikeasti hallittavalta kokonaisuudelta verrattuna absoluuttiseen kompleksisuuteen. Kuitenkin Kusters (2003: 6) kyseenalaistaa koko absoluuttisen näkökulman ja pitää sitä mahdottomana ajatuksena. Hänen mukaansa kompleksisuus on kielen ja kieltä arvioivan tahon (oppija, omaksuja, tutkija; teoria) välinen suhde eikä autonominen ominaisuus, joka voidaan katsoa kielellä olevan. Näin ollen absoluuttiseksi sanotussa näkökulmassa lopputulos vaihtelee sen mukaan, mitä teoriaa ja miten tutkija soveltaa arvioidessaan kielen kompleksisuutta. Absoluuttisen ja suhteellisen kompleksisuuden sijaan Kusters erottaakin teorialähtöisen ja kielenkäyttäjälähtöisen näkökulman, jotka tarkentuvat kunkin käytetyn teorian sekä kyseessä olevan kielenkäyttäjärühmän mukaan.

Lapsen varhaisten verbaalisten tuotosten tutkimuksessa absoluuttinen tai tutkijanlähtöinen kielen kompleksisuuden arviointi tuntuu loogisemmalta lähestymistavalta kuin suhteellinen tai käyttäjälähtöinen suuntaus, koska pienen lapsen omaa kokemusta kielen helppoudesta tai vaikeudesta on hankala selvittää. Välillisesti siihen on kuitenkin pyritty pääsemään käsiksi, sillä omaksumisjärjestyksen on usein ajateltu implikoivan myös suhteellista kompleksisuutta niin, että helpoksi koetut asia opitaan ensin ja vasta sitten vaikeammat. Produktiivisen syntaksin indeksi sijoittuu tutkijalähtöisen tai Miestamon (painossa) termein absoluuttisen kompleksisuuden piiriin. IPSyn rakentaa kuvausta lapsen rakenteellisesta repertoaarista. Mitä suurempi repertoaari on, sen laajemman kuvauksen se vaatii, ja tätä kuvauksen laajuutta ilmentää lapsen saama pistemäärä. Puhtaan informaatioteoreettista lähestymistapaa sekoittaa hiukan

mukaan otettu produktiivisuustekijä, joka tosin voidaan eliminoida laskemalla vain pisteitä antaneiden rakenteiden määrä. IPSyn keskittyy puhtaasti rakenteisiin ottamatta kantaa siihen, miten komplekseja niiden ilmaisemat asiat kognitiivisesti ovat ja mikä osuus tällä voisi olla rakenteen omaksumisessa (ks. Bowerman 1996).

Dahl (2004: 41–42) luonnehtii kieltä systeemiksi, johon sisältyy sekä resursseja että sääntöjä. Resurssit ja säännöt muistuttavat jakoa leksikkoon ja kielioppiin, vaikka täydellisestä vastaavuudesta ei olekaan kysymys. Resurssit ja säännöt ovat kiinteässä suhteessa toisiinsa, osittain häilyvärajaisia ja jopa limittejä käsitteitä. Kun leksikosta valitaan käyttöön esimerkiksi sana *kukka*, sanaan liittyvä leksikaalinen tieto osoittaa juuri tälle sanalla soveltuvat taivutussäännöt kaikkien nominitaivutussääntöjen joukosta. Toisin sanoen, resursseihin sisältyy sääntöjä. Toisaalta myös kielioppi voidaan nähdä resurssina, kun sääntöjen vaikutuksen sijaan painotetaan sitä, kuinka paljon kielenkäyttäjällä on sääntöjä käytettävissään. Resurssit liittyvät kielen ilmaisuvoimaan. Tämän voi todeta niin kielenomaksujan kuin -oppijankin kielen pohjalta. Mitä suurempi sanasto tai kieliopillinen säännöstö kielenkäyttäjällä on hallussaan sitä enemmän ja useammalla tavalla hän pystyy asioita ilmaisemaan, ja puhtaasti resurssien näkökulmasta kieltä voidaan sanoa silloin kompleksimmaksi. IPSynissä tarkasteltavana oleva repertoari on ikkuna kielenkäyttäjän rakenneresursseihin. IPSyn on siis systeemiin kuuluvien resurssien mittari, ja pisteiden lisääntymisen katsotaan implikoivan kompleksimpaa kielisysteemiä.

Mielestäni tärkeä aspekti kielen kompleksisuudessa on se, mitä käytettävissä olevilla resursseilla tehdään: miten ”palikoita” käytetään rakennelmien tekemiseen ja millaiseksi rakennelmat muodostuvat. Millaisia lapsen tuottamat ilmaukset siis ovat rakenteelliselta kompleksisuudeltaan? Samoista palikoista voi rakentaa hyvin erilaisia kokonaisuuksia – tai sitten jättää rakentamispotentiaalin käytön vasta myöhempään vaiheeseen. Resurssit antavat mahdollisuuden ilmaisuvoimaan, mutta mahdollisuuksia ei ole pakko käyttää monipuolisesti tai niiden kaikkia käyttömahdollisuuksia ei vielä ole opittu. Juuri tässä käytön laajuudessa lapset, joilla on samanlaisiksi arvioidut resurssit, saattavat erota hyvinkin paljon toisistaan. IPSynin kumulatiivisuus antaa hiukan tilaa myös tälle tarkastelukulmalle, mutta ilmausten kokonaiskompleksisuus jää kuitenkin yksittäisten rakenteiden korostamisen varjoon. Mittarin fokus on siis kielisysteemissä eikä kielenkäytössä, kuten resurssien yhteydessä jo totesin.

4 Lopuksi

IPSyn on selkeä edistysaskel kvalitatiivisempaan, informatiivisempaan ja tarkempaan suuntaan omaksujan kielen arvioinnissa. Se korostaa osaamista virheiden tai puutteiden sijaan ja ottaa huomioon lapsen todellisen äidinkielen, puhekielen, piirteet. Se antaa esimerkiksi MLU:hun verrattuna yksityiskohtaisempaa tietoa lapsen todellisuudessa käyttämästä kielestä ja sen rakenteellisista

ominaisuuksista pureutumalla lapsen resursseihin. Mittarin hyötyaste paranee heti, kun tuloksissa otetaan huomioon muutakin kuin vain yhteispistemäärä, joka sellaisenaan ei kerro oikeastaan sen enempää kuin MLU:n antama keskiarvokaan.

Mittarin työstäminen uuden kielen käyttöön on melko työläs tehtävä, ja siksi olisikin eduksi, jos samaa mittaria voisi soveltaa moneen tarkoitukseen. Mikään ei estä käyttämästä sitä esimerkiksi toisen tai vieraan kielen oppimisen tutkimuksessa, kuten Minna Sunin seuraava artikkeli tässä kokoelmassa osoittaa. Samoin läheiset sukukielet, kuten suomi ja viro, olisivat mielenkiintoinen IPSyn-kokeilun kohde, koska silloin kieltenvälinen vertailu olisi mahdollisuuksien rajoissa toisin kuin rakenteiltaan täysin erilaisten kielten välillä. Tämä edellyttäisi yhteistyötä ja tämänhetkisen suomenkielisen mittarin muokkausta myös esimerkiksi vironkieliseksi versioksi, mikä lienee jo merkittävästi helpompi tehtävä kuin suomalais-IPSynin rakentaminen englanninkielisen mittarin periaatteiden pohjalta.

IPSynin mittaama kompleksisuus on luonteeltaan absoluuttista, kielen rakenteisiin pureutuvaa. Siinä ei arvioida rakenteiden vaikeutta tai helppoutta kielenomaksujan näkökulmasta. Kielisysteemistä IPSyn tarkastelee käytössä olevia resursseja, jotka Dahlin (2004: 42) mukaan edustavat kielen rikkautta ja ilmaisuvoimaa ja ovat siten systeemin kompleksisuuden toinen komponentti. Toisen systeemin kompleksisuuden komponentin muodostavat ne säännöt, joiden avulla resursseja käytetään. Nämä säännöt eivät kuitenkaan ole IPSynin tarkastelun kohteena. Mittarin tulosten validi tulkinta kompleksisuuden suhteen on mahdollista, kun tietää, mihin mittari kielessä todellisuudessa pureutuu.

LÄHTEET

- Blake, J., Quartaro, G. & Onorati, S. 1993. Evaluating quantitative measure of grammatical complexity in spontaneous speech samples. *Journal of Child Language* 20, 139–152.
- Bowerman, M. 1996. The origins of children's spatial semantic categories: cognitive versus linguistic determinants. Teoksessa J. J. Gumperz & S. C. Levinson (toim.) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press, 145–176.
- Brown, R. 1973. *A first language. The early stages*. Harmondsworth: Penguin Education.
- Dahl, Ö. 2004. *The Growth and Maintenance of Linguistic Complexity*. Philadelphia, PA: John Benjamins.
- Gell-Mann, M. 1996. *Kvarkki ja jaguaari. Seikkailuja yksinkertaisessa ja monimutkaisessa*. (Suom. R. & T. Tuomi.) Helsinki: WSOY.
- Hadley, P.A. 1998. Early verbal-related vulnerability among children with specific language impairment. *Journal of Speech, Language and Hearing Research* 41, 1384–1397.
- Johnston, J. R. 2001. An alternate MLU calculation: Magnitude and variability of effects. *Journal of Speech, Language, and Hearing Research* 44, 156–164.
- Kauppinen, A. 1998. *Puhekuviot, tilanteen ja rakenteen liitto. Tutkimus kielen omaksumisesta ja suomen konditionaalista*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kusters, W. 2003. *Linguistic Complexity. The Influence of Social Change on Verbal Inflection*. LOT Dissertation Series 77. LOT: Utrecht.
- Langacker, R. W. 1988. A Usage-Based Model. Teoksessa B. Rudzka-Ostyn (toim.) *Topics in Cognitive Linguistics. Current Issues in Linguistic Theory* 50. Amsterdam: John Benjamins, 127–161.
- Laury, R. 2005. Complexity of spoken languages. Clause combining in interaction. Konferenssiesitelmä. Approaches to Complexity in Language, 24.–26.8.2005, Helsinki.
- McWhorter, J. H. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5, 125–166.
- Miestamo, M. (painossa). On the feasibility of complexity metrics. Ilmestyy teoksessa K. Kerge & M-M Sepper (toim.) *FinEst Linguistics. Proceedings of the annual Finnish and Estonian conference of linguistics. Tallinn, May 6–7, 2004*. Tallinn TPÜ/EKI.
- Miller, J. F. 1981. *Assessing language production in children. Experimental procedures*. London: Edward Arnold.
- Nieminen, L. & Torvelainen, P. 2003. Produktiivisen syntaksin indeksi – suomenkielinen versio. *Puhe ja kieli*, 23, 119–132.
- Peters, A. 1983. *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

- Rollins, P. R., Snow, C. E. & Willet, J. B. 1996. Predictors of MLU: semantic and morphological developments. *First Language* 16, 243–259.
- Rondal, J. A., Ghiotto, M., Bredart, S. & Bachelet, J-F. 1987. Age-relation; reliability and grammatical validity of measures of utterance length. *Journal of Child Language* 14, 433–446.
- Scarborough, H. S. 1990. Index of productive syntax. *Applied Psycholinguistics* 11, 1–22 .
- Tomasello, M. 1992. *First Verbs. A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
- Tomasello, M. 2000a. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11, 61–82.
- Tomasello, M. 2000b. The item-based nature of children’s syntactic development. *Trends in Cognitive Sciences* 4, 156–163.
- Tomblin, J.B., Spencer, L., Flock, S., Tyler, R. & Gantz, B. 1999. A comparison of language achievement in children with cochlear implants and children using hearing aids. *Journal of Speech, Language and Hearing Research* 42, 497–511.

LIITTEET

LIITE 1 IPSyn-rakenteet esimerkkitoteutumineen

Tässä liitteessä esitellään IPSyniin valitut rakenteet sekä annetaan niistä jokaisesta esimerkkejä, jotka on poimittu 2;6-vuotaiden lasten ilmauksista. IPSynin käytön seikkaperäiset ohjeet on luettavissa Niemisen ja Torvelaisen artikkelista (2003).

Nominirakenteet

N1 Monikko

Esimerkit

nä ujohti nää peitot

minä eikij jejuiija

N2 Inessiivi/adessiivi

äiti nukkuu täällä sängyssä

om mujja toinen poniki

N3 Elatiivi/ablatiivi

pois kyyistä

tää maittuu yvältäki

N4 Illatiivi/allatiivi

se tuli istuun minum pöytään

se laittaa tam pöyvälle

N5 Partitiivi

minä lainaan tätä ponia

N6 Genetiivi

peitom päälle tev vois panna

se tuli istuun minum pöytään

N7 Liitepartikkeli nominissa/adverbissa

lusikallaki

vieläkö

N8 2 sanan nominilauseke

tässä tuuvaan niitä heiniä

se om minum paikka

N9 3 sanan nominilauseke

tuo ol lammät tuo oikeem pikkune

N10 muu nominirakenne

ottaa omaan kyytiinsä

siellä ei isommat jalatkaa

Verbirakenteet

V1 Yksikön/monikon 1. persoona

Esimerkit

minäkin tulen kyytiä

mitäs sittem me kassotaan

V2 Yksikön/monikon 2. persoona

sinä ajat tällä

te teette kuakaa

V3 Passiivi

jääkettä annetaan

ei mennä alle

V4 Imperfekti

mä soitin sille

V5 Perfekti

on tommosta tetty

en ook keittäny

V6 Imperatiivi

tuu akeep ponit täältä pois

V7 Konditionaali

toss+ois sem paikka

V8 III infinitiivi

sänkyyn nukkumaa

tää yks tyttö on ajamassa

V9 Liitepartikkeli verbissä

ei tää mahukkaa

V10 Verbi + adverbiaali

kuka jillä voi ajjaa

V11 Verbi + 2 adverbiaalia

sittäm mää meel leissuu

V12 Verbi + I infinitiivi

voisit+sää ajaat tällä

V13 Verbi + III infinitiivi

nääkim menee äitiv viereen nukkumaa

V14 Verbi + I inf. + III inf.

tää yksi tyttö on ajamassa

V15 Kieltoverbi + verbi

mina aluum mennäl leikkimää

V16 Kieltoverbi + verbi + I inf.

tää ei anna ennää lääkkeitä

V17 Muu rakenne

mä en osaas sitä tehdä

ei halua mennä nukkumaan

en saa avattua

Lauserakenteet

S1 Mikä tahansa vähintään

2 sanan yhdistelmä

Esimerkit

ne kattois näitä

tossa auto

S2 S + V

S3 [Määre + S] + V

S4 Predikatiivilause

S5 V + O

S6 V + [määre + O]

S7 S + V + O

S8 V + O + Advli

S9 V + O + 2 Advli

S10 S + V + Advli

S11 S + V + 2 Advli

S12 S + V + [Määre + Advli]

S13 S + V + O + Advli

S14 S + V + O + 2 Advlia

S15 Infinitiivilauseke objektina

S16 Pre/postpositioliouseke

S17 Rinnastuskonjunktio

S18 Alistuskonjunktio

S19 Rinn.konj. yhdistää sanoja

S20 Rinn.konj. yhdistää lauseita

S21 Sivulause

S22 Muu rakenne

*tää hyppäsi sem päähä**paan tille lääketä**tän alka meni kippee**s+om pieni**jääkettä annetaa**kuka on ostanup ponit**minä tien tätä juokaa**kerta tä äitin tukka on harjattu**tetät tuo niitä**nyt te korjaa enkaat**tää laitetaan kyytiin**laita sinä tolle valmiiks ruokaa laita**laitetaan se istumaan tuohon**hiellä kaupassa oli laivoja**nyt mää heitän sen tohol lälvee**nää menee lenskarin kyytii**mää teem ponillel luokaa**mä kaaja äkkiä noi maaha**mää haluan ostaa auton**tää vauva nukkuu äitiv vieressä**rakkoli on tässä ja sakset**s+ei osaak ku s+om vie pieni**kaikki pöytää ja istumaa**nyt o yö ja kaikki nukkuu**mietittäis missä on niitä tankkeja**s+ei ossaa ajjaat tätä lahkalia*

Maksimipistemäärät:

Nominirakenteet 20 pistettä

Verbirakenteet 34 pistettä

Lauserakenteet 44 pistettä

Yhteensä 98 pistettä