

MITEN PRODUKTIIVISEN SYNTAKSIN INDEKSI PALVELEE TOISEN KIELEN TUTKIMUSTA?

Minna Suni
Jyväskylän yliopisto

Lähivertailuja 17.

Toimittaneet Annekatrin Kaivapalu, Külvi Pruuli.

Jyväskylä Studies in Humanities, 53.

Jyväskylä: Jyväskylän yliopisto, 2006, pp. 103–118.

ISBN 951-39-2450-5

ISSN 1459-4323

1 Johdanto

Tässä artikkelissa tarkastellaan sitä, miten produktiivisen syntaksin indeksi eli IPSyn toimii analysoitaessa suomea toisena kielenään oppivien lasten puhetta. Päähuomio on menetelmän käytettävyyden arvioinnissa sekä tutkimusprosessin ja sen tulosten erittelyssä. Produktiivisen syntaksin indeksi on varsinaisesti suunniteltu ensikielen kehityksen seurannan välineeksi, ja sen kehittelytyö ja toimintaperiaate on esitelty toisessa tämän julkaisun artikkelissa (Nieminen s. 69 – 83; ks. myös Nieminen & Torvelainen 2003). Nämä artikkelit muodostavatkin kokonaisuuden, jossa jälkimmäistä ei voisi olla ilman edellistä.

Produktiivisen syntaksin indeksi on alun perin kehitetty englantia ensikielenään käyttävien lasten tuotosten tarkasteluun (Scarborough 1990), ja sen suomenkielinen sovellus (Nieminen & Torvelainen 2004) taas on tarkoitettu suomea ensikielenään puhuvien lasten parissa tehtävään tieteelliseen ja kliiniseen työhön. Siinä missä englanninkielinen IPSyn-mittari on selkeän syntaktinen, on suomenkielinen IPSyn-sovellus luonteeltaan morfosyntaktinen. Syy tähän on selvä: suomen kielen omaksumisessa morfologialla on paljon keskeisempi asema kuin englannissa, ja syntaksin kehitys on siis tiiviissä yhteydessä morfologian kehitykseen.

IPSynin kehittämisen ja vakiinnuttamisen kannalta maahanmuuttajataustaisten lasten ottaminen informanteiksi on paikallaan jo siksi, että englanninkielistä versiota ei ole tiettävästi kokeiltu englantia toisena kielenä oppivien kielenkehityksen seurantaan. Suomenkielisen version käyttö toisen kielen oppimisen kontekstissa on siis uusi askel koko menetelmän käyttöalan kannalta.

Maahanmuuttajataustaiset lapset ovat heterogeeninen, jatkuvasti kasvava ryhmä, jonka kielellisen kehityksen kysymykset koskettavat Suomessa monia ammattiryhmiä: päiväkotien ja neuvoloiden henkilöstöä, perusopetuksessa työskenteleviä opettajia sekä puheterapeutteja. Etenkään alle kouluikäisille suomea toisena kielenä omaksuville lapsille ei ole tarjolla omia kielellisen kehityksen seurantamenetelmiä. Äidinkielenään suomea puhuville laaditut testit taas ovat tehtävänannoiltaan ja sisällöltään usein sellaisia, etteivät ne sovellu suoraan etenkin niille toista kieltä omaksuville lapsille, joiden puheen ymmär-

tämisen taito on sanavaraston suppeuden vuoksi vielä heikko (ks. myös Suni 2004).

Produktiivisen syntaksin indeksin etu tämän kohderyhmän kannalta onkin se, että se voidaan kaikkein luotettavimmin tehdä lapsen spontaanista puheesta, joka kerätään esimerkiksi vapaasta leikki-tilanteesta. Siitä siis puuttuvat monet niistä rajoitteista, joita syntyperäisille suomen kielen puhujille suunnatuissa testeissä vääjäämättä on.

Seuraavan luvun sisältämässä tutkimusprosessin kuvauksessa liikutaan tietoisesti valintojen, kokemusten ja aineiston tasolla, jotta sen pohjalta syntyisi käytännön tuntuma siihen, minkä tyyppistä perehtymistä ja panostusta IPSynmenetelmän käyttö edellyttää alkuvaiheessa. Tällainen reflektiivinen näkökulma tutkimusprosessiin on valitettavan harvoin esillä lingvistisissä tutkimuksissa, vaikka sillä on arvoa etenkin harkittaessa jonkin tutkimusmenetelmän soveltamista uusiin kieliin tai uusille kohderyhmille.

2 Tutkimusprosessi

2.1 Työvaiheet

Tutkimus toteutettiin Jyväskylän yliopiston kielten laitoksessa suomeen toisena kielenä erikoistuvien opiskelijoiden yhteisenä tutkimusharjoituksena.¹ Kukin opiskelija sai ”oman mamulapsen” tuotokset tarkasteltavakseen, mutta käytännössä tehtävät suoritettiin työpareina tai pienryhminä siten, että kaikissa työvaiheissa kokemus, tieto ja vastuu jakautuivat ainakin kahden opiskelijan kesken.

Työvaiheista aineistonkeruu oli tapahtunut etukäteen useista pro gradu -töistä koostuvan hankkeen yhteydessä (ks. lähemmin Stigell 2004, Tuomisto 2004, Pulkkinen 2005)², ja työparien tehtäväksi jäivät siten aineiston transkribointi, analyysin valmistelu ja toteutus sekä tulosten koonti. Reflektoidut kokemukset eri työvaiheista esitellään tuonnempana opiskelijoiden projektiraporteista poimituin anonyymein ottein. Näin tulee kuuluviin tutkimustyön toteuttajien oma ääni eikä ainoastaan ohjaajan tulkinta tapahtumien kulusta.

¹ Tutkimuksen toteuttivat Marina Gävgenen, Päivi Hautalahti, Terhi Hintikka, Silke Jedlicka, Pirita Korpivaara, Ulla Koukkari, Anne Kultha, Evelina Liski, Sari Marin, Annukka Mäenpää, Leena Mäyry, Riikka Nissilä, Maija Oivukkamäki, Sanna Paakala, Henna Puhakka, Saija Taivalmäki, Sanna Tervo, Heidi Toivonen ja Meri Törmä.

² Aineistoa oli kerätty tässä mainittujen pro gradu -töiden aineistoksi; mukana aineistonkeruussa oli myös Maija Oivukkamäki, jonka valmisteilla oleva tutkielma käsittelee suomea toisena kielenä omaksuvien lasten keinoja ilmaista paikkaa.

2.2 Aineisto ja sen transkribointi

Tutkimuksessa oli mukana 15 iältään 4-5-vuotiasta maahanmuuttajataustaista lasta, joiden puhetta oli nauhoitettu ääni- ja videonauhalle helsinkiläisissä päiväkodeissa. Aika, jonka lapset olivat olleet suomenkielisessä päivähoitossa, oli vähintään puoli ja enintään lähes viisi vuotta. Myös kielellisessä taustassa oli runsaasti vaihtelua: informanteista kuusi oli somalinkielisiä, ja muut äidinkielenä olivat arabia, englantia, ibibio, kiina, ranska, thai, urdu, venäjä ja vietnam.

Käytettävissä ollut aineisto ei valitettavasti ollut spontaania, kuten IPSyn-analyysin kannalta olisi ollut parasta. Nauhoitetut tilanteet olivat melko haastattelunomaisia: kukin lapsi suoritti samat strukturoidut tai puolistrukturoidut tehtävät aikuisen johdolla. Eniten ja vapainta puhetta lapsilta oli kirjoittanut kuvasarjan selostustehtävä, joka käsitteli päiväkodin arkea aamusta kotiinläh- töön. Sen yhteydessä juteltiin myös lasten omasta hoitopäivästä.

Vaikka aineistossa oli rajoitteensa, se oli joka tapauksessa koottu yhdenmukaisella tavalla kaikilta informanteilta. Aikuiselta imitoidut ilmaukset suljet- tiin analyysin ulkopuolelle, koska olennaista oli päästä tarkastelemaan lapsen itsenäisesti hallitsemia kielellisiä rakenteita.

Aineiston transkribointi eli litterointi toteutettiin puolikarkeana noudatta- en tietokonepohjaiseen lapsen kielen tutkimukseen kehitetyn CHILDES- ohjelmiston konventioita (ks. McWhinney 2000). Lapsen ja aikuisen puheen- vuorot merkittiin samalla tarkkuustasolla. Sananraja-assimilaatioita ei merkitty, mutta mm. päällekkäispuhunnat kirjattiin tarkasti ylös. Aineisto ei ollut digitaal- lista, joten transkribointi tehtiin C-kaseteilta käyttäen hidastukset ja moninker- taiset toistot mahdollistavia nauhureita.

Transkribointi koettiin kaikkein työläimmäksi työvaiheeksi, mutta myös hyvin hyödylliseksi ja selvästi odotettua antoisammaksi tehtäväksi. Jokainen työpari transkriboi ja tarkisti kahden lapsen tuotokset, eli käytännössä kukin opiskelija transkriboi yhden lapsen osuuden ja tarkisti toisesta tehdyn transkip- tion.

Transkriboinnin hitaus ja vaativuus tulivat monelle yllätyksenä, mutta ko- kemuksen karttuessa työ muuttui helpommaksi. Noin puolituntisen jakson transkribointiin kului neljästä kuuteen tuntia ja lisäksi vielä tarkistamiseen ja korjaamiseen tarvittiin pari tuntia.

Ensimmäisellä kuuntelukerralla jouduin merkitsemään sulkuihin useita epäselviä sa- noja, joita en ollut oikein kuullut. Joskus en kuullut juuri mitään ja moniin sulkuihin tulikin pelkkiä katkoviivoja. Toisella kuuntelukerralla olin jo niin tottunut lapsen pu- hetapaan, että monet sulut saivat väistyä oikeiden sanojen tieltä. Yhteensä litteroin- tiin kului aikaa noin viidestä kuuteen tuntia, mutta valmista tuli.

Usein kävi myös niin, että kun johonkin epäselvään kohtaan myöhemmin palasi uu- delleen, sen kuuli aivan selvästi, vaikkei aikaisemmin ollut saanut siitä mitään selvää. Ensikertalaiselle litteroinnissa oli yllättävää myös se, kuinka hidasta puuhaa sen on. Myös litterointimerkkien kanssa oli aluksi aika pyörryksissä, mutta litteroinnin lop- puvaiheessa ne alkoivat tuntua jo selviltä.

Vaikeinta litteroinnissa olivat siis päällekkäispuhunnat. Myös sananloppuisten n-äänteiden kuuleminen oli usein vaikeaa, ja meillä olikin litterointiparini kanssa useita erimielisyyksiä niistä.

Muutammat kohdat jäivät pysyvästi epäselviksi joko hiljaisen puheäänien, hyvin epätarkan artikulaation tai äänityksen epätasaisen laadun takia. Kaiken kaikkiaan tällaisia ongelmia oli kuitenkin vähän.

S2-lasten [Suomea toisena kielenä oppivien lasten] puheen litteroimiseen liittyy tietysti erilaisia ongelmia kuin syntyperäisen tai aikuisen puhujan puheen litterointiin. Ongelman on esimerkiksi se, että lapset puhuvat usein epäselvästi.

Vaikka tarkistin koko keskustelun vielä kertaalleen loppuksi, joissakin kohdissa lapsen ilmaus jäi kokonaan epäselväksi. Siksi olikin erittäin hyvä, että tarkistimme vielä parityönä toistemme aikaansaannokset.

2.3 Havaintoja informanteista

Jo transkriboinnin aikana tehtiin kvalitatiivisia havaintoja lasten puheesta ja myös niistä tilannetekijöistä, jotka saattoivat vaikuttaa aineiston laatuun ja laajuuteen. Useat työparit saivat tarkasteltavikseen keskenään hyvinkin eritasoisia suomenpuhujia, mikä avasi silmiä sille huomattavalle variaatiolle, jota aineisto kokonaisuudessaankin sisälsi.

Yllätys minulle kylläkin oli se, miten eritasoisia nämä haastatellut lapset olivat. Minun lapseni kun juuri ja juuri osasi muodostaa sanoista lyhyehköjä ja usein vaillinaisia lauseita, niin toinen lapsi pisteli tulemaan vaikeita lauserakenteita ja yhdisteli niitä milloin rinnastus- ja milloin alistuskonjunktioin.

Esimerkit 1 ja 2 sisältävät näytteet kahdelta eri informantilta tehtävästä, jossa jutellaan tyypillisestä hoitopäivästä.

Esimerkki 1

H: mm mitäs te teette tuolla ulkona?

L: ulkona tuuleh.

H: mitä te leikitte ulkona?

L: ulkhona (vi-e) u leikhu # eikku ulkkona ulkkona veikku va ois.

H: kenenkäs kanssa sä leikit ulkona?

L: eih ei me tiella uka laatti nalla piäni en nal nalle alftto vöi en tullee al nap auto pois en n tänän on iitä ja sitten hän tänän sivihtä ja sitte tänä tuttu kuis tuostta sitte tättäh.

Esimerkki 2

I 9: ne on+// # ne on # ulkona.

T 2: hmm. # ja mitäs ne siel ulkona # tekee?

I 9: ne leikkii.

T 2: mitä kaikkee ne leikkii?

I 9: ne leikkii omassa puistossa.

T 2: joo. # mitä kaikkee ne tekee siellä?

I 9: öks+// ((huokaus)) yks leikkii ajamista ja yks <leikkii> [>] kottia ja yks leikkii (---) ja yks laskee ja yks mennee keinuuj ja # yks # sanoo että lasketaaj ja yks kattoo # ja kaks kattoo

T 2: <mm> [<]

Esimerkeistä ilmenee, että nämä informantit tuottivat puhetta määrällisesti varsin yhdenmukaisesti, mutta puheen ymmärrettävyydessä ja morfosyntaktisessa laadussa oli sen sijaan selvä ero. Siinä missä esimerkistä 1 on vaikea löytää vaittomasti ymmärrettäviä jaksoja, voisi esimerkin 2 tuotos olla ymmärrettävyytensä ja morfosyntaksinsa kannalta yhtä hyvin suomea äidinkielenään käyttävän vastaavan ikäisen lapsen puhetta.

Osa lapsista tuotti etenkin aineiston alkupuolella puhetta vain niukasti. Tähän vaikuttivat oletettavasti sekä lasten arkuus uudessa tilanteessa että tehtävien haastattelumaisuus.

Alussa lapsi ei ollut kovin puhelias, vastasi vain yksinkertaisesti kysymyksiin, joihin testaaja halusi saada vastaukset. - - Kun lapsi oli vähän tottunut tilanteeseen, hän rupesi puhumaan muistakin asioista.

Informanttini ei ollut keskustelussa kovin aloitteellinen ja omatoiminen; hän ei juurikaan ala itse puhua jostakin asiasta, vaan vastailee enimmäkseen haastattelijan kysymyksiin, usein melko lyhytsanaisesti. Vain muutamassa puheenvuorossa oli kymmenisen sanaa; usein lapsi vastasi vain sanalla tai kahdella.

Lapsen puheenvuorot olivat melko samanlaisia läpi nauhoitustilanteen. Alussa lapsi vastailee hieman lyhytsanaisemmin, mutta keskustelutilanteen edetessä, ehkäpä päästyään alkujännityksestä, lapsi alkaa vastata monisanaisemmin ja keskeyttää enemmän. Pisimmät puheenvuorot sijoittuvat nauhoitustilanteen loppupuolelle.

Niukka tuottaminen ei välttämättä liity kielellisten resurssien niukkuuteen. Vastaavasti runsas tuotos ei aina merkitse sitä, että tuotoksessa olisi myös paljon näyttöä kielen sanaston ja rakenteiden hallinnasta. Informanttien joukossa oli esimerkiksi kohtalaisen puhelias lapsi, joka puoli vuotta aiemmin oli vastaavassa aineistonkeruussa tuottanut vain äidinkieltään. Nyt hän käytti myös suomenkielisiä sanoja, mutta enimmäkseen puhe oli hyvin vaikeatulkintaista ja siksi myös erityisen vaativaa transkriboitavaa. :

Omalle kohdalleni osui litteroitavaksi varsin haasteellinen tapaus. --- Lapsen tuottamat lauseet olivat varsin kirjavia, tässä muutamia esimerkkejä: *Häshäettu lei pushmatutu shn pnt leipä* tai *Kuune setta kuune set tu totaale sittevaattu eisti mite apaattista ei en ikhe mten*. Yritin siis mahdollisimman uskollisesti kirjoittaa litterointiin kaikki äänteet ja ´sanat´joita lapsi käytti.³

Paitsi puheen yleisestä ymmärrettävyydestä, havaintoja kertyi transkriboinnin aikana myös yksittäisistä kielenpiirteistä, kuten konjunktoiden käytöstä tai aikamuotojen vähäisestä vaihtelusta.

2.4 IPSyn-analyysin valmistelu ja toteutus

Produktiivisen syntaksin indeksin alkuperäinen kehittäjä Scarborough (1990) on ottanut analysoitavaksi jokaiselta lapselta sata ymmärrettävää ilmausta. Suomenkielisen version morfosyntaktinen luonne taas on perusteena sille, että

³ Kyseessä saattaa olla sekakielinen tuotos, joten lapsen äidinkieltä hallitseva analysoija olisi luultavasti pystynyt selvittämään paremmin, mitä lapsi kertoo ja kuvailee. IP-Syn-pistemääriin sillä ei olisi kuitenkaan ollut vaikutusta, koska ilmausten ymmärrettävyyden tulee perustua suomenkielisiin ilmauksiin.

analyysi tehdään mieluummin morfeemimäärältään pisimmistä ymmärrettävistä ilmauksista. Nieminen käyttää tekeillä olevassa väitöskirjassaan 80 pisintä ilmausta kultakin lapselta, mutta hänen havaintojensa mukaan jo 20 morfeemimäärältään pisintä ilmausta antavat 80% kokonaispisteistä (Nieminen, suullinen tiedonanto 26.3.2003).

Eri aineistomäärien riittävyttä päädyttiinkin vertaamaan siten, että jokaiselta lapselta valittiin ensin 20 morfeemimäärältään pisintä ymmärrettävää ilmausta, ja analyysi tehtiin tältä pohjalta. Sitten otettiin mukaan kaikki muikin lapsen kielellinen tuotos, ja laskettiin toinen IPSyn-pistemäärä, jota verrattiin ensin saatuun.

Oli melko vaativaa laskea morfeemit sellaisesta lapsen tuotoksesta, joka oli sekä ääntämyksen että rakenteiden osalta usein kaukana natiivien – ei vain aikuisten, vaan myös lasten – tuotoksesta. Toisaalta tehtävää ehkä hieman helpotti se, että informanttien käyttämät morfologiset rakenteet eivät olleet vielä erityisen kompleksisia.

Morfeemien parissa askartelu oli hauskaa ja haastavaa; olo oli kuin pikku työllä, joka on keskittynyt leikeissään äärimmäistä tarkkuutta vaativaan tehtävään. Hauskuus vaihtui pian huolestuneisuudeksi, sillä huomasin, että informantin ilmaukset olivat keskimäärin melko lyhyitä, voittopuolisesti viiden morfeemin pituisia. En sentään ollut aivan tuuliajolla, sillä ilmauksista löytyi myös pidempiä, kymmentä ja sitä suurempia morfeemimääriä sisältäviä ilmauksia.

Lapseni pisimmässä ilmauksessa oli 22 morfeemia. Se oli kuitenkin poikkeus. Kahdenkymmenen pisimmän ilmauksen joukkoon tuli myös ilmauksia, joissa oli vain kuusi morfeemia.

Aineiston haastattelumaisuus vaikutti varmasti siihen, että huomattava osa lasten puheenvuoroista oli vain sanan tai parin pituisia, ja että morfeemien määrä jäi niissä viiden tuntumaan tai sen allekin. Vain harvalla näin lyhyitä ilmauksia kuitenkaan mahtui 20:n ensisijaisesti analysoitavan ilmauksen joukkoon. Toisaalta ilmaus on käsitteenä melko haastava ja häilyväkin tällaisessa empiirisessä tarkastelussa. Elliptisiä ja kesken jääneitä ilmauksia on keskustelussa taajaan. Kiistatta useiksi eri ilmauksiksi jaettavissa olleita puheenvuoroja ei lopulta ollut koko aineistossa kovin paljoa.

2.5 Tuotosten analyysi

Suomenkielisessä IPSyn-versiossa kunkin lapsen tuotoksesta selvitetään 49 eri rakenteen esiintyvyys; rakenteet jakautuvat nomini- ja verbilausekkeisiin sekä lauserakenteisiin. Jos saman rakenteen erilaisia esiintymiä on yksi, pisteitä tulee yksi, ja jos kaksi tai enemmän, pisteitä tulee kaksi. Enimmäispistemäärä on siis 98.

Löydettyjen rakenteiden produktiivisuus pyritään varmistamaan siten, että hyväksyttävyydessä on joitakin rajoituksia. Esimerkiksi pikapuhemuodot *mun* ja *sun* hyväksytään vain kertaalleen genetiivin esiintymäksi; toisen pisteen saaminen edellyttää erilaista genetiivin edustajaa. Nieminen ja Torvelainen

(2003) ovat esitelleet ja perustelleet nämä analyysissa noudatettavat kriteerit selkeästi, joten IPSyn ei vaadi eikä myöskään salli käyttäjiltään mitään omia periaateratkaisuja.

Lomakkeet koettiin selkeiksi ja niiden käyttö kohtalaisen helpoksi. Kun kaikki pisteitä tuovat ilmaukset kootaan lomakkeeseen omille riveilleen, kokonaiskuva lapsen käyttämistä rakenteista hahmottuu melko hyvin jo analyysin edetessä. Jotta analyysi olisi johdonmukaista ja luotettavaa, keskeiset kieliopilliset rakenteet on tietenkin tunnettava ja annettuja kriteerejä on seurattava huolellisesti.

Aluksi ajattelin, ettei lapsen vähistä ja yksinkertaisista ilmauksista saa taulukon täyteen juuri mitään, mutta melkoisen hyvin rivit täyttyivät vähästä materiaalistani huolimatta. Tyhjäksi rivin jouduin jättämään kyllä lähes aina, kun kyse oli vähänkin pidemmästä lauserakenteesta.

Nominilausekkeiden poimiminen oli vielä yllättävän vaivatonta ja nopeaa, mutta etenkin monimutkaisempien verbi- ja lauserakenteiden poimiminen oli hankalaa. Vaikeinta olivat monimutkaisimpien lauserakenteiden etsiminen (esimerkiksi verbi + objekti + 2 adverbialia tai infinitiivilause objektina).

Yllätyksekseni lapsen tuotoksessa oli kuitenkin aika hyvin verbilausekkeita, esimerkiksi mm. III infinitiivi *lukemaan* tai kahden verbin yhdistelmä verbi + verbi I infinitiivissä *ei saa lyöä*. Mitään aikamuotojen vaihtelua lapsen tuotoksesta en kuitenkaan löytänyt.

Tulkintavaikeuksia kuitenkin esiintyi jonkin verran; lasten käyttämien ilmausten merkitys ja funktio eivät suinkaan ole aina kiistattomia.

Erityisesti lauserakenteiden kanssa ongelmia aiheutti oma epävarmuus ja kokemattomuus. - - Jouduin usein miettimään, oliko kyseessä varmasti esim. objekti ja ennen kaikkea oliko lapsi tarkoittanut sen olevan. Juuri siinä olikin tämän projektin suurin haaste - ymmärtää mitä lapsi oli tarkoittanut.

Aineiston haastattelumainen rakenne rajoitti tiettyjen analyysin kattamien rakenteiden esiintymistä. Toisaalta lapsi ei vapaassa leikkitilanteessakaan välttämättä tule käyttäneeksi hallitsemaansa rakennevalikoimaa läheskään kattavasti. Esimerkiksi menneen ajan muodot jäävät puuttumaan, jos toiminta ja sen kuvaus keskittyvät vain tässä ja nyt tapahtuviin asioihin.

Eräs asia kuitenkin vaivasi. Jotkut kysytyt muodot riippuivat jonkin verran myös siitä, mistä asioista puhuttiin, ja haastattelijan tekemistä kysymyksistä. Esimerkiksi toista persoonaa en löytänyt haastateltavan puheesta, vaikka hän sen varmasti osaisikin, sillä haastattelun puheenaiheissa ei tarvinnut sitä muotoa käyttää.

On varmasti vain huonoa tuuria, ettei lapsen 20 ilmauksesta löydy yhtään nominin monikkoa, koska lapsen puheessa esiintyi monimutkaisempiakin rakenteita kuin monikko. Myös nominin inessiivin/adessiivin ja nominin illatiivin/allatiivin puute on mielestäni yllättävää.

On sinänsä paikallaan, että nämä aineiston luonteesta ja menetelmän mekaanisuudesta aiheutuvat "epäreiluuden" kokemukset kumoavat perusteettomat odotukset aivan aukottomasta rakenneinventaarista. Vastapainoksi on myös todettava, että erityisen myönteisenä IPSynin käytössä pidettiin sitä, että tämä menetelmä todella kohdistaa huomion siihen, mitä informantti jo osaa. Virheet

ja osaamattomuus eivät ole etualalla. Prosessin edetessä moni havahtuikin aidosti iloitsemaan ”omansa” onnistumisesta ja osaamisesta:

Pitää sanoa, että tämä osuus oli kaikista mielenkiintoisin. Halusimme aina löytää ”omalle” lapselle lisää pisteitä. Morfeemien laskeminen oli kuin kilpailu – mitä enemmän kaikenlaista löydetään, sitä paremmat tulokset saadaan...”

Oli todella HAUSKAA ja yleensä aika helppoa löytää sopivia rakenteita. Olimme aina ”onnellisia” kun olimme saaneet lapselle lisää pisteitä. Kommentit, kuten ”hyvä lapsi” tai ”loistavaa” eivät olleet harvinaisia.

Vaara sympatiapisteiden antamisesta on IPSynissä kuitenkin vähäinen, koska pisteityskriteerit ovat selvät. Jos analyysia voi tehdä yhdessä toisen kanssa, analyysin validius ja reliaabelius on toki turvatumpaa kuin yksin työskenneltäessä.

2.6 Tulosten koonti

Analyysien numeeriset tulokset koottiin taulukkolaskentaohjelmaan, jonka avulla ne järjestettiin ja muunnettiin erilaisiksi kaavioiksi. Visualisoituina tulokset kertoivat nopeasti sen, miten kunkin informantin IPSyn-pistemäärä suhteutuu muiden jokseenkin yhtä kauan suomenkielisessä päivähoitossa olleiden pistemääriin, ja missä suhteessa nomini- ja verbilausekkeista sekä lauserakenteista saadut pisteet ovat toisiinsa ja enimmäispisteisiin.

Tulosten koonti oli luonnollisesti mielenkiintoisin vaihe koko projektissa, sillä oli tietenkin kiinnostavaa nähdä, miten oma lapsi sijoittui muihin verrattuna.

Informanttilapsi sijoittui kaavioon odotusteni mukaisesti: hän ei sijoittunut vuonna 2001 suomenkielisessä päiväkodissa aloittaneiden lasten ryhmässä indeksiluvun suuruuden perusteella kärkipäähän, mutta ei myöskään erottautunut poikkeuksellisen heikon kielitaidon omaavana informanttina.

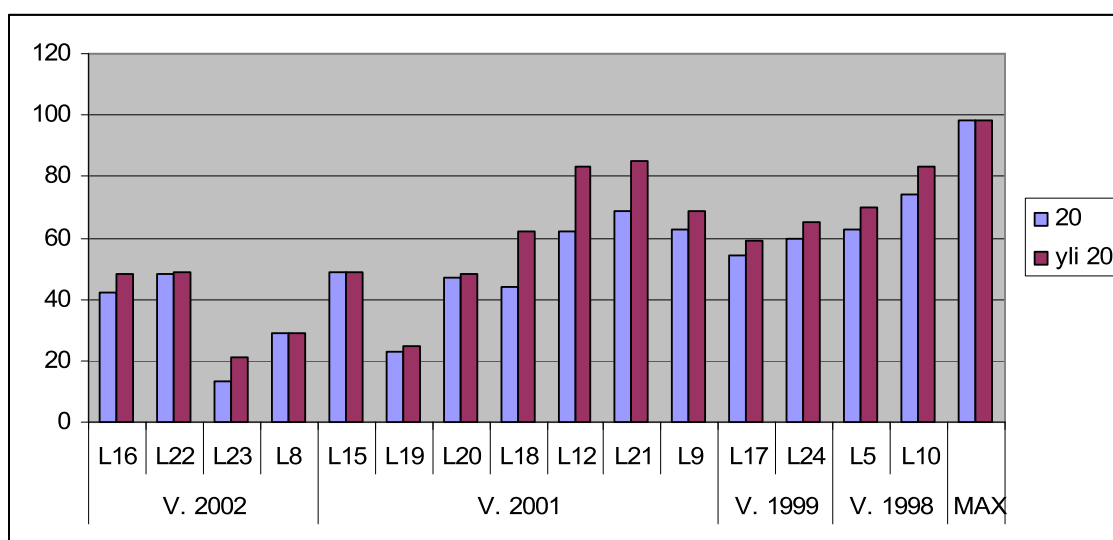
Informanttini puheenvuorot olivat mielestäni aika lyhyitä, ja siksi minulle oli yllätys se, että lapsi sai IPSyn-analyysissä niin hyvän kokonaispistemäärän.

Jos tulokset yllättivät, ne yllättivät poikkeuksetta myönteisesti. Ilman IPSyn-analyysia melkoinen osa kielen rakenteiden hallinnasta olisi siis jäänyt havaitsematta, ja vertailu muihin taas auttoi hahmottamaan sitä, millaista yhdenmu-kaisuutta ja vaihtelua lasten välillä voi odottaa esiintyvän.

3 Tulokset

3.1 Riittävätkö 20 pisintä ilmausta?

IPSyn-analyysin tulokset koottiin kahdessa muodossa, jotta aineiston riittävyttä olisi mahdollista arvioida. Tarkasteltavina olivat ensin vain 20 morfeemimäärältään pisimmästä ymmärrettävästä ilmauksesta lasketut IPSyn-pistemäärät, ja niitä verrattiin sitten pistemääriin, jotka saatiin laskemalla mukaan kaikki mahdolliset pisteitä antavat ymmärrettävät, imitoimattomat ilmaukset. Kaaviosta 1 ilmenee, minkä verran pistemäärät nousivat tarkasteltavaa aineistoa laajentamalla. Äärimmäisenä oikealla on IPSynin enimmäispistemäärä (98).



KAAVIO 1 Informanttien produktiivisen syntaksin indeksi laskettuna heidän 20 pisimmästä ymmärrettävästä ilmauksesta (=20) ja kaikista ymmärrettävistä ilmauksista (= yli 20)

IPSyn pystyi saamaan erot yksilöiden välillä selvästi näkyviin jo 20 pisimmän ymmärrettävän ilmauksen analyysin varassa. Kuitenkin useimpien informanttien pistemäärä nousi, kun mukaan otettiin kaikki aineistoon sisältyneet ymmärrettävät ilmaukset. Aineiston kasvaessa pistemäärät nousivat kaikilla niillä, joilla oli jo 20 pisimmän ilmauksen pohjalta selvästi yli puolet enimmäispisteistä. Informanteilla 18, 12 ja 21 pistemäärien erotus oli erityisen suuri; heistä vain informantti 18 oli saanut alle puolet enimmäispisteistä ilman aineiston laajentamista.

Niillä, joiden IPSyn-pistemäärä jäi 20 pisimmän ilmauksen perusteella alle puoleen enimmäistasosta, tulos ei enää ratkaisevasti muuttunut aineistoa lisäämälläkään. He kaikki olivat olleet suomenkielisessä päivähoitossa alle kaksi vuotta. Tämä merkitsee, että suppeampi näyte saattaa aivan toisen kielen oppimisen alkuvaiheessa sinänsä riittää johdonmukaisten tulosten saamiseen. Aineiston rajaamisessa on kuitenkin riskinsä: informantti 18 on hyvä muistutus

siitä, että ennalta ei voi päätellä, keiden kohdalla aineistonäytettä voi vielä rajata näin pieneksi vaarantamatta tulosten luotettavuutta.

3.2 Suomenkielisen päivähoitoajan yhteys IPSyn-pistemäärään

Suomen kielen rakenteiden hallinnan voidaan odottaa kohenevan ajan myötä, ja siksi voitiin perustellusti odottaa, että IPSyn-pistemäärät olisivat sitä korkeammat, mitä pidempään lapset olivat olleet alttiina suomenkieliselle syötökselle. Suomenkielinen päivähoito on maahanmuuttajataustaisilla lapsilla sekä ajankäytöllisesti että sosiaalisesti keskeisin suomenkielinen ympäristö, ja siksi sen kesto voidaan pitää kielen kehittämisessä keskeisenä muuttujana.

Informantit on järjestetty edellä kaavioon 1 sen mukaan, kuinka pitkään he ovat olleet suomenkielisessä päivähoidossa: vasemmalla ovat vuonna 2002 suomenkielisessä hoidossa aloittaneet ja näiden jälkeen v. 2001 aloittaneet jne. Myös saman vuoden eri kuukausina tulleet on asetettu aikajärjestykseen.⁴

Kaaviosta on nähtävissä selvä nouseva perussuuntaus: mitä pidempään lapsi on ollut suomenkielisessä hoidossa, sitä korkeamman pistemäärän hän todennäköisesti saa. Koska tulos on selkeä ja odotuksenmukainen, IPSynin voidaan pitää luotettavana työkaluna siinä, mihin se on tarkoitettu: se kykenee erottelemaan informantit sen mukaan, miten heidän morfosyntaktisten rakenteiden hallintansa on kehittynyt. Se, että suomi on informanttien toinen kieli, ei tuo näihin tuloksiin mitään poikkeavaa tai erityistä, vaan menetelmä näyttää toimivan toisen kielen kontekstissa aivan yhtä johdonmukaisesti kuin ensikielen kontekstissa.

Kaikki informantit saivat vähintään 20% maksimipistemäärästä, ja korkeimmat pistemäärät jäivät noin 15 pisteen päähän maksimitasosta. IPSyn:in avulla olisi ollut siis mahdollista analysoida morfosyntaktiselta kompleksisuudeltaan sekä nyt analysoituja vahvempia että heikompia tuotoksia. Tuloksista näkyy kuitenkin myös paljon yksilöllistä vaihtelua; yhtä pitkään hoidossa olleet eivät saaneet välttämättä keskenään samantasoisia tuloksia. Samana vuonna suomenkielisessä päivähoidossa aloittaneiden pistemäärät vaihtelevat n. 20:sta n. 50:een (vuoden ajan suomenkielisessä hoidossa olleet eli vuosi 2002) ja n. 25:stä yli 80:aan (kaksi vuotta suomenkielisessä hoidossa olleet eli vuosi 2001). Vaihteluvälit ovat siis huomattavia.

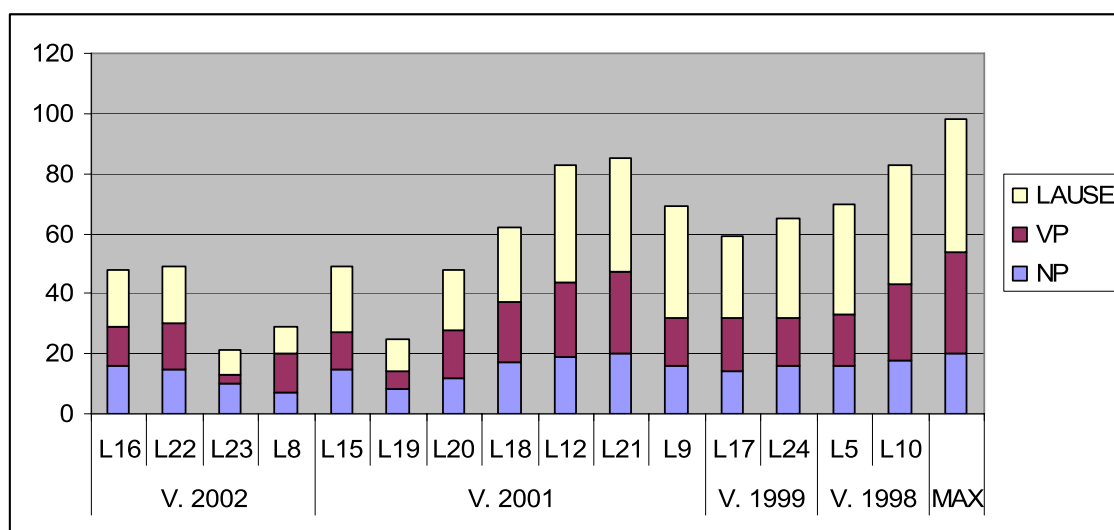
Informantit L23 ja L19 ja jossain määrin myös L8 poikkeavat erityisen selvästi yleisestä tasosta. Osalla heistä oli päiväkodista saadun suullisen tiedon mukaan epäily kielellisen kehityksen erityisvaikeudesta, mutta ei vahvistettua diagnoosia. Ilman tätä taustatietoakin IPSyn-analyysin poikkeava tulos ohjaisi kysymään, mistä muita alempi pistemäärä voisi johtua. Selittävä tekijä voi tuki löytyä myös esimerkiksi tilanteen vieraudesta, lapsen väsymyksestä tai keskittymisen vaikeudesta.

⁴ Informantti L17 oli ollut ensin äidinkielisessä ja sitten kaksikielisessä päivähoidossa ennen siirtymistä suomenkieliseen ryhmään. Hänet on sijoitettu kaavioon sen mukaan, milloin hän aloitti kaksikielisessä päivähoidossa.

Joka tapauksessa on myönteistä, että IPSyn näyttää erottavan tapaukset, joissa jo on pidemmän havainnoinnin pohjalta päädytty epäilemään jotakin erityistä. Se merkitsee, että menetelmä voi olla hyvinkin hyödyllinen apuneuvo kliinisessä työssä, kun maahanmuuttajataustaisella lapsella epäillään yksilöllisen kielenkehityksen poikkeavaa kulkua. Menetelmän pitkäikäisy on tällöin paikallaan.

3.3 IPSyn-pistemäärän rakentuminen

Kolmas tutkimuksen näkökulma oli se, miten IPSyn-pistemäärät rakentuvat sekä yksilötasolla että yleisesti ottaen. IPSyn-analyysin kokonaispistemäärä on kunkin informantin nominilausekkeista, verbilausekkeista ja lauserakenteista saamien pistemäärien summa (ks. lähemmin Nieminen s. xx-xx tässä julkaisussa). Kaaviosta 2 ilmenee, miten tutkimuksen eri informanttien produktiivisen syntaksin indeksit muodostuivat näistä osista. Mukaan tähän analyysiin on otettu kaikki informanttien tuottamat ymmärrettävät imitoimattomat ilmaukset.



KAAVIO 2 IPSyn-pistemäärien muodostuminen informanteittain. NP= nominilausekkeista saadut pisteet, VP=verbilausekkeista saadut pisteet, LAUSE= lauserakenteista saadut pisteet.

Kaaviota on selkeintä tarkastella siltä kannalta, missä määrin eri informanttien indeksin rakentuminen muistuttaa maksimipistemäärän rakentumistapaa. Useimpien pylväissä eri osasten suhteet ovat varsin samankaltaiset kuin ne maksimisuorituksessa olisivat – vieläpä niin, että mitä pidempään lapsi on ollut suomenkielisessä päivähoitossa ja mitä korkeamman pistemäärän hän on saanut, sitä samankaltaisempi pylvään rakenne yleensä on suhteessa maksimisuoritukseen.

Selkeimmin indeksiluvun rakentuminen poikkeaa yleisestä linjasta edellä jo mainittujen informanttien L23 ja L8 tapauksessa. Poikkeama painottuu verbilausekkeisiin siten, että toisella niistä saatu pistemäärä on suhteessa kokonai-

suuteen varsin matala ja toisella taas korkea. Näin siis kielenkehityksen mahdollinen erityisyys tulee esiin myös tässä lähemmässä tarkastelussa.

Nominilausekkeet näyttävät olevan suuremmin yhteydessä kokonaispistemäärään kuin muut IPSynin komponentit. Ne myös ennustavat verbilausekkeitä paremmin sitä, mille tasolle lauserakenteiden hallinta kullakin informantilla asettuu. Tämä saattaa selittyä sillä, että lauserakenteista saatavaan pistemäärään vaikuttavat melko paljon sellaiset kompleksiset adverbiaalilausekkeet ja objektirakenteet, jotka edellyttävät myös nominilausekkeiden hallintaa.

Jos tarkastellaan korkeimmat pistemäärät saavuttanutta puolta informanteista – kaaviossa informantit L18:sta alkaen oikealle – havaitaan, että ero suhteessa enimmäispistemäärään syntyy heillä ennen kaikkea verbilausekkeiden hallinnasta. Se merkitsee, että 2-4-vuotisen suomenkielisessä ympäristössä olemisen myötä nämä lapset ovat ehtineet omaksua keskeiset nominirakenteet jo varsin kattavasti, mutta kompleksiset verbirakenteet ovat vasta kehitymässä.

4 Menetelmän arviointia

IPSynin käyttö koettiin yllättävän helpoksi ja nopeasti omaksuttavaksi. Aikaa vievin työvaihe oli tuotosten transkriointi, kun taas vaikeinta oli morfeemimäärien laskeminen välillä hyvinkin epäselvistä tuotoksista. Indeksien laskutapa pidettiin myös hyvin oikeudenmukaisena, koska se tunnustaa jo orastavankin osaamisen eikä rankaise esimerkiksi puhekielisyydestä tai itse keksityistä sanoista. Yleisesti ottaen lapsista oli transkriointivaiheessa syntynyt vähemmän osaava kuva kuin millaiseksi käsitys muotoutui varsinaisen IPSyn-analyysin myötä. IPSyn onnistui siis nostamaan esiin sellaista kielen hallintaa, johon ei vielä tarkallakaan kuuntelulla saatu otetta.

Produktiivisen syntaksin indeksin suomenkielinen versio (Nieminen & Torvelainen 2003) palveli kaiken kaikkiaan hyvin maahanmuuttajataustaisten lasten suomen kielen rakenteiden hallinnan mittarina. Se oli mittarina johdonmukainen; pistemäärät nousivat sitä korkeammiksi, mitä pidempään suomenkielisessä hoidossa olleista lapsista oli kyse. Tämä perussuuntaus oli selvä, vaikka lasten välillä esiintyikin yksilöllistä vaihtelua.

Olin odottanut tutkimuksen tuloksia innostuksella ja kiinnostuksella. Oli erittäin mielenkiintoista vertailla lapsia ja nähdä, että suomenkielinen hoito vaikuttaa aika paljonkin kielen taitoon. Toisaalta oli jännää, että myös eroja löytyi lasten väliltä, vaikka he olivat olleet yhtä pitkään hoidossa.

IPSyn-asteikon rajat eivät tulleet tässä tutkimuksessa mukana olleiden informanttien kohdalla vielä vastaan, vaikka mukana oli taidoiltaan varsin eritasoisia lapsia. Menetelmä erotteli lapset hyvin, ja se näytti myös tunnistavan tapaukset, joissa oli mahdollisesti taustalla kielellisen kehityksen erityisvaikeuksia.

Selkeän rakennekeskeisenä ja mekaanisenakin menetelmänä IPSyn tietenkin häivyttää suuren osan yksilön kommunikatiivisesta kielitaidosta. Tässä tutkimuksessa myös aineiston haastattelumaisuus vaikutti samaan suuntaan ja heikensi sikäli lasten mahdollisuuksia antaa ”hyvä näyttö” osaamisestaan.

Oma ennakkotuntemukseni ennen aineiston analyysia oli, että hän olisi parempi ja osaisi ilmaista enemmän asioita, joita IPSyn mittaa. Hän oli oma-aloitteinen (kyseli kuvasta asioita ja kommentoi sitä ilman, että testaja olisi kysynyt kohdasta). Vaikutti myös siltä, että informantti olisi halunnut sanoa enemmän kuin pystyi. (Aloitti jutun, mutta ei sitten saanutkaan kerrottua sitä: *"Mä ku mä mä oli mun kotissa sitte minä ei mä oli mä oli mun koti"*.)

Kuitenkin on muistettava, että IPSyniä ei ole tarkoitettukaan toiminnallisen tai kommunikatiivisen kielitaidon mittariksi, vaan nimenomaan työkaluksi, jolla voidaan tarkastella rakenteiden produktiivistumista ja kompleksisuuden kehittymistä. Analyysin tuloksena on sekä monipuolinen inventaario kunkin informantin hallitsemasta rakennevalikoimasta että suoraan tilastolliseenkin käsitteelyyn soveltuvaa numeerista tietoa niin yksilö- kuin ryhmätasollakin.

Kokonaisuutena IPSyn oli positiivinen kokemus. - - IPSynin hyviä puolia on sen yksinkertaisuuden lisäksi myös tulosten selkeys. Niitä on helppo vertailla keskenään ja vaikka tehdä samalle informantille pitkittäistutkimus.

5 Lopuksi

Produktiivisen syntaksin indeksi soveltuu hyvin suomea toisena kielenä oppivien parissa käytettäväksi. Nyt informantit olivat lapsia, kuten IPSynin perusajatuksena on alusta alkaen ollut, mutta todennäköisesti menetelmä toimisi myös aikuisten maahanmuuttajien tuotosten analyysissa ainakin suomen kielen oppimisen alkuvaiheessa. Sitä voisi tutkimustyössä käyttää niin lapsilla kuin aikuisillakin esimerkiksi rakenteiden oppimisjärjestyksen seurantaan, kunhan informantteja olisi vielä enemmän ja tutkimus olisi mahdollista toteuttaa pitkittäisesti.

Myös maahanmuuttajalasten kielenkehityksen erityiskysymyksiä selvittäessä IPSyn voi osoittautua arvokkaaksi avuksi, koska se ei edellytä erityisen testitilanteen järjestämistä eikä vaadi lapselta sen enempää kulttuuristen konventioiden hallintaa kuin tehtävänantojen ymmärtämistäkään.

Seuraava askel olisi kuitenkin verrata suomea äidinkielenään ja toisena kielenään omaksuvien lasten IPSyn-tuloksia asettaen rinnakkain esimerkiksi 3-vuotiaat natiivit ja 3 vuotta suomenkielisessä päivähoitossa olleet maahanmuuttajataustaiset lapset. Myös keskenään samaa äidinkieltä puhuvien ja suomea toisena kielenään omaksuvien lasten tarkastelu omana ryhmänään toisi uusia kiinnostavia näkökulmia, kunhan informanttijoukko olisi riittävän suuri.

Erityisen toivottavaa olisi, että IPSynistä saataisiin käyttöön myös esimerkiksi vironkielinen sovellus, jolloin vertailujen tekeminen lähisukukielten kesken avaisi jälleen uusia tutkimuksellisia mahdollisuuksia sekä ensikielen että toisen kielen saralla.

LÄHTEET

- McWhinney, B. 2000. *The CHILDES Project: Tools for analyzing talk*. Mahwah, Nj.: Erlbaum.
- Nieminen, L. & Torvelainen, P. 2003. Produktiivisen syntaksin indeksi - suomenkielinen versio. *Puhe ja kieli* 23:3, 119-132.
- Pulkkinen, R. 2005. Ajan ilmaisut 4-5-vuotiaiden maahanmuuttajalasten suomen kielessä. Suomen kielen pro gradu -työ, Jyväskylän yliopiston kielten laitos.
- Scarborough, H. S. 1990. Index of Productive Syntax. *Applied Psycholinguistics* 11, 1-22.
- Stigell, L. 2004. Toisto 4-5-vuotiaiden S2-lasten puheessa. Suomen kielen pro gradu -työ, Jyväskylän yliopiston kielten laitos.
- Suni, M. 2004. Toinen kieli tutuksi. *Puheterapeutti* 3, 4-6.
- Torvelainen, P. 2005. 2;0-vuotiaiden lasten fonologisen kehityksen variaatio: puheen ymmärrettävyyden, sananmuotojen tavoittelun ja tuottamisen sekä sananmuotojen yksikonsonanttisuuden ja -vokaalisuuden tarkastelu. Suomen kielen lisensiaatintyö, Jyväskylän yliopiston kielten laitos.
- Tuomisto, S. 2004. Päiväkoti-ikäisten maahanmuuttajalasten kuvailukeinot: ihan niinku talo tulee talo tästä. Suomen kielen pro gradu -työ, Jyväskylän yliopiston kielten laitos.