

Kuidas mõõta süntaktilise kompleksuse arengut lastekeeles

Airi Kapanen

Tallinna Ülikool

Ülevaade. Kuna keel koosneb tasanditest, mille kompleksuse aste erineb, peaks kompleksusega arvestama ka lastekeele uurimisel. Artiklis arutletakse selle üle, mis on süntaktiline kompleksus ja missuguse meetodiga võiks seda mõõta esimest keelt omandava lapse spontaanses kõnes. Vaadeldakse kaht süntaktilise kompleksuse mõõdikut, mida senini on enamasti kasutatud: väljendite keskmist pikkust (MLU) ja produktiivse süntaksi indeksit (IPSyn), mille loomise eesmärk oli näidata süntaksi arengut lastekeeles. Arvestades et MLU ja IPSyn'i näol on tegu kvantitatiivsete mõõdikutega, on osa uurijaid kindlad, et kompleksuse ideaalsel kirjeldamisel tuleks kindlasti vaadelda ka keele omandamise aset leidvaid keeletöötlusprotsesse ning keele muutumist ja varieerumist. Mis tahes mõõdiku kasutamisel on aga esikohal uurija teadlikkus nii mõõdiku võimalustest (mida ja mis ulatuses on võimalik teada saada) ja süntaktilise kompleksuse kui keelenähtuse sisust, kuid ka hoidjakeele rollist keeleomandamisel, kuivõrd laps ei asu kohe omandama keele potentsiaalset kompleksust, vaid hoidjakeelega sarnast keelevarianti.

Võtmesõnad: esimese keele omandamine, süntaktiline kompleksus, väljendi keskmine pikkus, produktiivse süntaksi indeks, eesti keel

Sissejuhatus

Komplekssuse mõõdikut on vaja selleks, et hinnata, kuidas lapse keeleline areng kulgeb. Kuigi on selge, et laps alustab lihtsamast ja liigub keerulisema poole, ei pruugi see areng toimuda täiesti sujuvalt ja ühtlases tempos. Kuna keele omandamisel võivad vahelduda kiire arengu perioodid ja nn settimise perioodid, mil omandatu kindlustub, oleks objektiivsuse huvides kasutada lapse keelelise arengu täpsemaks kirjeldamiseks ühtset kompleksuse mõõdikut. Lisaks on teada, et teadmised eesti keele kasutusmuutrite kohta pole piisavad, seepärast pole ei õppimis- ega õpetamisprotsess mitte optimaalne, vaid intuiitiivne (Eslon 2014: 14). Komplekssuse mõõdik võiks näiteks anda teadmisi ka õppijakeele kasutusmuutrite kohta, mille tundmisel saaks keeleõpe olla optimaalne.

Lastekeele uurimises on kompleksus sage termin (Nieminen 2009: 173), kusjuures „komplekssuse kasv“ on sünonüüm „keelelise arenguga“, ehk teisisõnu: mida edasijõudnum on keelekasutaja, seda komplekssem on tema keelekasutus. Mida kompleksus endast täpsemini kujutab, pole lihtne kindlaks teha, kuna üsna sageli peetakse kompleksust nii tavaliseks, et seda mitmetahulist mõistet ei proovitagi täpsemini kirjeldada (Nieminen 2007: 17; vt ka Cejnarova 2005: 1). Kuna seda mitmedimensioonilist mõistet on raske määratleda, on kõneldud ka kompleksuse deemonist (vt Cejnarova 2005: 9–35 arutlust kompleksuse definitsiooni üle; Nieminen 2007: 33), mida on peetud omal moel ohtlikuks ja mille lähedale on parem mitte minna.

Keeleomandamise uuringutes on eeldatud, et kõigepealt omandavad lapsed lihtsad struktuurid, hiljem keerukad (Smith, van Kleeck 1986), seepärast arvatakse, et omandamise järjekord võib ilmutada ka keele enese kompleksust (Nieminen 2009). Praegu eksisteerib keeleomandamise uurimisvaldkonnas ka arvamusi, et

süntaksi omandamine on mitu korda lihtsam protsess, kui seda seni on arvatud olevat: süntaksi omandamine on väikeste teadmistük- kide omandamise käik ja teadmistükkide valikut juhivad pragma- tilised kaalutlused (Ninio 2011: 178).

Eesti keeleteaduses on eri meetoditel uuritud täiskasvanute tekstikeerukust allkeeliti (Kerge 2003) ning 5.–11. kl õpilaste kirja- liku leksika ja süntaksi arengut, sh kasutatud keelendite sagedust ja leksika sõnaliigilist jaotumust võrreldes haritud täiskasvanuga (Kerge, Uusen, Põlda 2014; Kerge, Uusen, Põlda, Puksand 2014), kuid ka pragmaatilisi tegureid ehk kommunikatiivseid tunnuseid mitmekümnele suulisele ja kirjalikule žanrile omase sõnaliigitasa- kaalu taustal (Kerge, Pajupuu 2010). Pille Eslon (2014) on uurinud morfosüntaktilise ja leksikaalse varieerumise piire ning võrrelnud ilukirjandus- ja õppijakeele kasutusmustreid. Lastekeele valdkonnas on eri sõnaliikide esinemus kõnes süntaksi arengu tunnuseid, lek- sikoni sõnaliigilist jaotumist nii lapse enese kui ka tema vanema keelekasutuses on uuritud ka eesti keele näitel (Argus, Kõrge- saar 2014), samuti on vaadeldud lapse keelelist arengut (Schults, Tulviste, Kaljumäe 2013; Vider 1995), vahel ka MLU taustaks sea- tud (Lääts 2013), kuid kompleksuse kasv keeleomandamisel ei ole siiani ühegi kirjutise keskmes olnud.

1. Kompleksus ja komplitseeritus, keeleomandamine ja hoidjakeel, absoluutne ja struktuuriline kompleksus

1.1. Kas kompleksne tähendab komplitseeritud?

Kompleksus, mis tähendab 'läbipõimunud osadest koosnev, mitut ala hõlmav, liit-, põim-' (Vääri jt 2012: 599), pole sünonüüm oma- dussõnaga „keeruline“. Kuigi lastekeele valdkonnas on märgatud, et lapsed omandavad ka väga kompleksse keele väga kiiresti ja näh- tava pingutuseta (Behrens 2006: 3), võidakse ometi tõlgendada, et

raskused keeleloomes ongi tõendus kompleksisusest: see, mis on raske, on kompleksne (Nieminen 2007: 34). Sageli öeldakse „kompleksne“, kuid mõeldakse „komplitseeritud“; „kompleksne“ võib olla teaduses kasutatav nimetus millegi kohta, mis on väga komplitseeritud ja „komplitseeritud“ võib viidata probleemile, mida ei osata lahendada, kuna ei teata, kuskohast alustada (Cejnarova 2005: 9).

Terminid „kompleksne“ ei peeta lastekeele uurimises nii ohtlikuks või provokatiivseks kui lingvistika muudes valdkondades (Nieminen 2009: 175), kuigi kompleksisuse selgituspüüdluste juures on kritiseeritud, et üsna vähesed keeleteadlased arvestavad kompleksisuse arenemisel keele vastuvõtu protsesside eri rolle; näiteks kas keeleüksust on raske märgata-tajuda, omandada, luua ja välja öelda või sellest aru saada ning kuidas need protsessid mõjutavad keeleüksuse üleüldist kompleksisust (Peters 1983, viidatud Nieminen 2007 kaudu). Ometi ollakse kindlad, et kompleksisus on keele selline omadus, mida kindlasti saab mõõta.

Keeleomandamisuuringutes on päris tavaline, et kompleksisust mõõdetakse valitud keeleüksuste hulgana, tehes kindlaks, mitu morfeemi, sõna või lauseüksust on lausungis. Teisisõnu on levinud tava, et kompleksisust mõõdetakse, st tema uurimiseks kasutatakse (enamasti) kvantitatiivset viisi. Üldiselt arvatakse olevat piisav, kui kasutada väljendi keskmise pikkuse arvutamiseks, olgu morfeemides või sõnedes, 50 kuni 100 lausungit (Lee 1974, Miller 1981, Jalilevand, Ebrahimipour 2014, Williamson 2009: 2).

1.2. MLU kui süntaktilise kompleksisuse mõõtmise levinuim vahend

Süntaktilise kompleksisuse mõõtmise levinumaid vahendeid on MLU (väljendi keskmine pikkus, eesti lastekeeleuurimustes ka VKP). Kuigi MLU näitab morfeemide keskmist hulka uuritud tekstimassiivi lausungis, arvatakse, et väljendi keskmine pikkus

1) on seotud mitme grammatilise pädevusega, muu hulgas ka semantiliste suhete väljendamisega; 2) näitab grammatiliste kategooriate rohkust lausungis, st et lausungis on olemas verb, pronoomen või artikkel; 3) on seotud oskusega produktiivselt kasutada grammatikaelemente, nagu muutetunnused; 4) on tõendus üha keerukamate eitavate ja küsilauseingite olemasolu kohta kõnevoorus (Scarborough jt 1991: 24).

Kuna MLU puhul on tegu kvantitatiivse mõõtevahendiga, tekib küsimus, kas väljendi keskmine pikkus ikka mõõdab süntaktilist kompleksust või mitte. Mida rohkem ühikuid lausungis on, seda komplekssem arvatakse lausung olevat, kuna kompleksuse hindamine on taandatud kvantiteedile, keeleüksuste mingile hulgale, pööramata tähelepanu lausungi seesmisele kvaliteedile. Seepärast ei esitatagi küsimusi, kas lausungi pikkuse mõõtmine on tema kompleksuse kindlakstegemisel kohane mõõdik ja missugust kompleksuse aspekti (või aspekte) lausungi pikkuse mõõtmine võib näidata (vt Nieminen 2007: 18). Seda, et MLU on vaid üldine ja kvantitatiivne mõõdik, mis ei anna selget ülevaadet lauseloome eripäradest, on täheldatud veel mitmes uuringus (nt Schauwers jt 2005, viidatud Lääts 2013: 15-16 kaudu).

1.3. Mida mõõta, kas absoluutset või struktuurilist kompleksust

Eristatakse kaht tüüpi kompleksust: keele kui süsteemi kompleksus ning lausungite ja väljendite struktuuri kompleksus. Keelekasutaja puutub kokku mõlemaga. Uuriija Lea Nieminen ei räägi süntaktilisest kompleksusest, vaid kasutab terminit „keeleline kompleksus“ ja see tähendab suhet või seost keele ja keelekasutaja vahel (Nieminen 2009: 174): kui keelt kasutatakse, on keelekasutaja kaasatud nii keele vastuvõtmisesse kui ka keele produtseerimisse (Nieminen 2009: 174). Siin selgub kompleksuse üks tähtsaid külgi:

keeleline komplekssus – vahel nimetatud ka süntaktiliseks kompleksuseks (vt nt Culicover 2013: 12) – pole midagi seesugust, mis seisaks keele kasutajast lahus.

Keele kui süsteemi kompleksust on nimetatud ka *absoluutseks kompleksuseks* ning lausungite ja väljendite struktuuri keerukust (struktuurilist kompleksust) *suhteliseks kompleksuseks* (Miestamo 2006, 2008). Osa uurijaid (Dahl 2004, 2008) kasutab samade mõistete tähistamiseks vastavalt termineid *formaalne kompleksus* ja *produtseerimiskompleksus* (vt pikemalt Culicover 2013: 3, 20–24). Kui absoluutne kompleksus on teooriakeskne objektiivne mõiste ja põhineb keelel enesel, siis suhteline kompleksus on subjektiivne ja juba oma loomult dünaamiline ning keskendub keele kasutajale (põhiküsimus „Mis on kompleksne kellele?”; vt Nieminen 2007: 33–35). Teisisõnu, keelelist kompleksust on kaht liiki: absoluutne kompleksus on süsteemi (ehk keele) objektiivne omadus, suhteline kompleksus on „hind”, mida peab maksma keelekasutaja, s.o raskus keelekasutaja jaoks (vt Miestamo 2006, 2008, Culicover 2013: 12). Absoluutne kompleksus tegeleb infohulgaga, mis on vajalik, et süsteemi (siinkohal keelt) kirjeldada: mida rohkem infot see nõuab, seda komplekssem on süsteem (keel). Suhteline kompleksus aga on töötlemise või ka produtseerimise kompleksus ja sel on seos ressursidega, mida on vaja keeleliste representatsioonide, nagu hääliku, sõna jms töötlemiseks (nt mälu).

Osa uurijaid on kindlad, et ideaalne on kirjeldada kompleksust nii, et võetakse arvesse keeletöötlusprotsesse, keele omandamist ning keele muutumist ja varieerumist; nad möönavad, et meil pole ikka veel kompleksuse sõltumatut mõõtjat (Culicover 2013: 11). Laskumata kompleksuse üle pikemasse arutellusse, piisab öelda, et spontaanse keele kompleksuse uurija peab tegema teadliku ja põhjendatud valiku, missugust raamistikku ta uurimiseks kasutab (vt pikemalt selle kohta Nieminen 2007: 38). Kompleksusest kui

keelenähtusest teadlik olla on äärmiselt vajalik, sest paratamatult ei pruugi ühe uuringu raames jõuda analüüsida komplekskssuse kõiki võimalikke mõjureid¹.

1.4. Absoluutne komplekssus ja keeleomandamine

Tõsi on, et nii keele kui süsteemi komplekskusel kui ka struktuuri-
lisel komplekskusel on tihe seos keelekasutajaga ja keele omanda-
misega, ent keele omandamist need raskeks ei muuda. Järgmised
näited illustreerivad seda, kuidas keele omandamisel ei ole takistu-
seks ei keele enese ega ka lausungite ja väljendite struktuur.

Laps on võimeline omandama ka keeruka ehitusega keele. Tugevasti polüsünteetilisi keeli, näiteks navaho ja ketšua, oman-
davad lapsed ei tee morfotaktikavigu – vigu teevad hoopis ketšua
keele vilunud kasutajad (Kelly jt 2014: 61).

Lapsed, kes õpivad hispaania, jaapani ja soome keelt, hakka-
vad mitmesilbilisi sõnu lühendamata ütleva varem kui lapsed,
kes omandavad inglise, saksa või taani keelt (Ota 2006). Osa soome

¹ Näiteks on Soome keeleteadlane Suvi Stolt (2009) uurinud sõnavara varast arengut, seoseid leksikoni ja grammatika vahel väikese sünnikaaluga laste ja ajaliste soome laste seas. Ta leidis, et lapse sool on lapse sõnavara suu-
rusele selge mõju küll ajaliste laste hulgas, kuid mitte väikese sünnikaaluga laste seas. Ema haridus polnud ajaliste laste puhul nende sõnavara kasvuga seotud, kuid väikese sünnikaaluga laste puhul oli seos olemas. Sünnikaalul enesel polnud mingit mõju ajaliste laste sõnavara arengule, kuid sünnikaalu ja aktiivse leksikoni (ingl *expressive lexicon*) vahel oli vanuses 2;0 märkimisväärne mõju. Erisused näitavad, et bioloogilistel ja/või kontekstilistel faktoritel (nt ema haridus) on lapse keelearengule enam mõju just väikese sünnikaaluga laste hulgas. Stolti uurimus rõhutab varase leksikoni tähtsust keeleomandamisel; tema uurimistulemused näitasid, et grammatikat ei omandata eraldi moodulina, vaid et grammatika omandamine on tihedalt seotud lapse ekspressiivse sõnavara kasvuga teise eluaasta lõpuks.

keelt emakeelena omandavaid lapsi on võimelised looma ja ütlema neljasilbilisi sõnu ilma lühendamata alates vanusest 1;7–1;9. See on suur erinevus võrrelduna inglise keelt omandavate lastega, kellest osa jätkab kolmesilbiliste sõnade rõhuta silpide ärajätmist ka pärast vanust 2;10 (Ota 2006: 262).

Ka eesti keelt (põhiolemuselt aglutinatiivne keel, rikas käände-süsteem ja palju tegusõnade pöördevorme, tagasõnad eessõnade asemel) on peetud võrreldes inglise keelega raskesti omandatavaks (vt Ingram 1999: 495), kuid uurimistulemused (Argus 2009, Lipp 1977) näitavad, et eesti keelt pole raske omandada.

Kui vaadelda kompleksuse ja sisendkeele seost, siis on leitud, et lapsed on võimelised omandama üksusespetsiifilise (s.o verbi-) süntaksi² ka siis, kui sisendkeele hulk on väga väike (Ninio 2011: 7). Anat Ninio on leidnud, et 35–50% tegusõnu, mida lapsed kolmes grammatilises põhiseoses kasutasid, olid üllatavalt harvad sisendkeeles, kus neid esines miljoni sõne kohta vähem kui 20 korda (*ibid.*).

Keele kui süsteemi kompleksus loob olenevalt keelest struktuurilise kompleksuse, ometi pole keele omandamine tema süsteemi kompleksuse tõttu üldiselt kunagi olnud raske, sest keelt kui mitmedimensioonilist tervikut ei omandata korraga. Kuna lapsed ja täiskasvanud jagavad samu keelereegleid ja -kategooriaid, pole põhjust oletada, et lapse grammatika on põhiolemuselt teistsugune kui täiskasvanu oma (Pinker 1984, Tomasello 2000: 229, Tomasello 2004: 142–143, Ninio 2011). Seega on keelt omandav laps olukorras, kus ta iga päev puutub kokku nii keele enese ehk süsteemikompleksusega kui ka struktuurilise ehk kõne kompleksusega, ent keele omandamisel juhib teda kogunud keelekasutaja, kelle kõnes esinevate keeleüksuste sagedus (ingl *frequency*), esilduvus (*saliency*) ja jagatud tähelepanu raamistik (*joint attentional frame*)

² Üksusena käsitab Anat Ninio (2011: 7, 161) üksusespetsiifilise raames nimelt ja eelkõige verbi.

minimeerivad struktuurilise kompleksuse (omandamise) võimaliku keerukuse. (Vt nt Tomasello 2003: 22 jm, Argus 2008: 11–15 jm.)

1.5. Hoidjakeel toetab süntaktilise kompleksuse omandamist

Süntaktilise kompleksuse arengu üks mõjutajaid on sisendkeel³, siinses artiklis nimetatakse seda terminiga „hoidjakeel”. Hoidjakeel on dünaamiline kõneregister, mida kasutab täiskasvanu, kõneldes lapsega, selle registri sõnavara ja suhtlusstrateegiad on individuaalsed ning muutuvad lapse kasvades (Kõrgesaar 2014: 193). Eesti hoidjakeele uurijad on üksmeelel, et laste- ja hoidjakeel on teineteisega tihedalt seotud, ning kui üht uurides vaadeldakse paralleelselt ka teise andmeid, siis on võimalik paremaid uurimistulemusi saada (*ibid.*), rääkimata vajadusest uurida keelt just suhtluses: ühelt poolt on hoidjakeel vahend, teiselt poolt aga eesmärk. Laps ei asu kohe omandama keele potentsiaalset kompleksust (eespool ka „absoluutne kompleksus”) – kuna keel omandatakse suhtluses ja kontekstisidusalt, siis kõigepealt omandabki laps igapäevasuhtluseks vajaliku keelevariandi ja siin võib hoidjakeelt võtta kui mõõdikut, sest lapse eesmärgiks on omandada hoidjakeelega sarnane keelevariant⁴. Et aga täiskasvanu paratamatult ei kasuta keelt omandava lapsega eesti keele kõiki potentsiaalseid vorme ja kõiki keeles olemas olevaid kategooriaid, siis kuuluvad hoidjakeele kui mõõdiku hulka just need vormid ja väljendid, mis seal esinevad, kuid ka nende rohkus ja sagedus, mida kõiki saab võrrelda lapse keeles esinevaga.

³ Sisendkeelt on nimetatud erialakirjanduses mitmeti: *motherese, parenthese, child directed speech, input* (vt Argus 2008, Kõrgesaar 2009).

⁴ Ninio järgi esindabki lapse ja täiskasvanu keelevariant kaht arenguperioodi: varast ja hilisemat (Ninio 2011: 47), sest täiskasvanupärast keelekasutust võib pidada keele omandamise eesmärgiks.

Hoidjakeel toetab süntaktilise kompleksuse järkjärgulist omandamist. Näiteks kuigi mõne sõnaliigi esinemissagedus ei pruugi hoidjakeeles teatud aja jooksul märkimisväärselt muutuda, võib hoidjakeeles eneses samal ajal aset leida hoopis kvalitatiivne muutus: lapsele suunatud kõnes on järjest keerukama ehitusega keeleüksused, nagu tuletised, liitsõnad, ühend- ja väljendverbid (Argus, Kõrgesaar 2014), see aga annab märku varasemaga võrreldes komplekssema tasandi morfoloogiast, mida laps kuuleb. Samas ei kopeeri lapsed pimesi seda, mida vanemad on öelnud (Ninio 2011: 123)⁵: võrreldes samade vormidega neile suunatud kõnes, lihtsustavad lapsed valikuliselt seda, mida lausuvad (samas: 98).

Laps puutub hoidjakeele kaudu kokku nii fonoloogia, morfoloogia, süntaksi kui ka semantikaga. Peale seoste nende keeletasandite vahel on keeleomandamisel oluline ka iga tasandi enese võimalik struktuuriline keerukus. Nii näiteks on morfoloogia omandamisel peetud oluliseks, kui keerukas või lihtne on nimi- või tegusõnavormide moodustus, sest emma-kumma lihtsus-keerukus võib omandamise käiku mõjutada (Xanthos jt 2011, Argus, Kõrgesaar 2014: 39). Süntaksi tasandil ennustab nt hoidjakeeles esinevate liitlausete hulk kindlalt süntaktilist kompleksust ka lapsekeeles (Behrens 2009: 394, Abbot-Shmidt, Behrens 2006, Huttenlocher jt 2002: 349, L. Gleitman, Newport, H. Gleitman 1984: 69). Samas

⁵ Tugiverbiühendeid on lapsele suunatud kõnes suurel määral enam kui lastel endal (laste moodustatud verbi ja objekti konstruktsioonides on neid vaid 8%). Kui Ninio (2011) oma uurimuses vaatles lauseid, milles esines sage tugiverb *give* 'andma', siis selgus, et lapsed ei kasuta seda tugiverbi nii vähe mitte sellepärast, et nad poleks võimelised kasutama sedasama vanematele omast konstruktsiooni, vaid seepärast, et enamik *andma*-verbiga konstruktsioone olid vanema kõnes seotud objektiga „musi“, „kalli“ ja teiste emotsioonisõnadega (Ninio 2011: 123, Nelson 1995). Kuigi *musitama*, *kallistama* jms tegevused on konkreetsed, võivad emotsioonisõnad ise olla lapsele keeleomandamise varasel astmel veel liiga abstraktsed, et neid oma aktiivsesse sõnavarasse võtta.

on leitud, et ka varases vanuses ei erine laste süntaktiline süsteem vanemate omast kvalitatiivselt (Ninio 2011: 62), vaid sarnaneb sellega mitme põhinäitaja poolest (Ninio 2011: 83). Lapsed ei väldi ühe ja sama verbi arvukat kasutamist (kvantitatiivne mõõde) eri tähenduses (kvalitatiivne mõõde), vanemate enda verbivara on samal ajal pea poole (40%) suurem kui lastel (Ninio 2011: 128).

Hoidja- ja lastekeele vahel on märgata ebasümmeetrilist põimumist mitmel keeletasandil: ühelt poolt toimub n-ö alt-üles-protsess: laps omandab keelt, säilitab mälus nii sõnu kui ka konstruktsioone, toimub känkimine, mis tähendab, et meeldejäetavad ühikud pakitakse kokku suuremateks mõtestatud ühikuteks (Allik jt 2002: 112). Mõtestatud ühikute moodustamine leiab aset keele kõigil tasanditel. Teiselt poolt toimub samal ajal n-ö ülalt-alla-protsess, sisendkeele kuul(a)mine. Kuigi vanemad mingil määral lihtsustavad oma keelekasutust⁶, ei kuule laps nende kõne kaudu mitte üksnes lühikesi ja lihtsaid vorme, vaid ka pikki ja keerukaid. Nii tekivad lapse leksikasse keeleomased fonoloogilised, morfoloogilised ja süntaktilised mallid, mille keerukuse aste on ühelt poolt seotud hoidjakeelega, teiselt poolt aga lapse enese kui keeleomandamise aktiivse subjekti kognitiivse arenguga.

⁶ Keeleuurija Elissa Newport ja tema kolleegid (Newport 1976, Newport jt 1975, 1977) uurisid 15 keskklassi ema-tütre suhtlust (lapsed olid vanuses 12–27 kuud). Võrreldi täiskasvanu-lapse ning täiskasvanu-täiskasvanu kõnet. Tulemused näitasid, et lapsele suunatud kõne on enam kompleksne kui kõne, mille täiskasvanu suunas teisele täiskasvanule (Bowermann 1978: 163). Lapsele suunatud kõnet võib seega pidada lihtsaks vaid tinglikult.

1.6. Süntaktiline kompleksus lapse ja täiskasvanu kui kuulaja ja kõneleja vaatenurgast

Sisendkõnest korjavad lapsed üles just sellise vormi, mis sobib nende kommunikatiivse kavatsusega kõige paremini. Üsna loomupäraselt paneb keelt alles omandava inimese lihtsamast grammatikast arusaamine rohkem vastutust pragmaatikale ja diskursuskontekstile (Jackendoff, Wittenberg 2014). Selleks, et mõista väheste morfoloogiliste vormide ja lihtsama ehitusega keelekasutust, nagu see keelt alles omandaval lapsel tavaliselt on, peab kuulaja öeldu tähenduse leidmiseks tavapärasest rohkem pingutama. Väidet sobib ilmestama kahe sõnalause *ema sokk / mommy sock*, mis võib lapsel kanda tähendust 'sokk, mis on ema oma' või 'ema paneb mulle sokki jalga' (Bloom 1970) või ka meenutada mingit soki ja emaga seotud olukorda kõnekohas jne. Lausungi tähendus selgub vaid (kitsamast või laiemast) kontekstist ja selle mõistmine nõuab kuulajalt pingutust või isegi lisateadmisi lapse argipäeva kohta. Pingutust võib vaja olla ka ühe ja sama häälikujärgendi mõistmiseks; näiteks võib sõna „põmma“ ühel ja samal lapsel tähendada mitut asja, nagu 'pall' (objekt) ja 'põrgatama' (tegevus) jms.

Lapse vaatenurgast tähendab keele omandamine seda, et laps kuuleb endale suunatuna sellist kõnet, mis suurel määral koosneb mitmesõnalistest lausungitest. See sunnib keelt omandavat last eristama kõneajas sõnu, et talletada neid edasise äratundmise tarvis oma mällu. (Johnson jt 2013: 3) Kuulaja ja kõneleja puutuvad aga struktuurilise kompleksusega kokku eri vaatenurkadest. Kuulaja, sealhulgas keelt omandav laps kui vähe kogemusega keelekasutaja on sunnitud aktsepteerima keelelised valikud, mida on teinud kõnelev täiskasvanu. Östen Dahl näiteks räägibki valiku struktuurist ja väljundi struktuurist. Kõneleja – ema, isa vm vilunud keelekasutaja – peab pingutama, et mõista keelt omandava lapse lihtsamat

süntaksit. Kuigi jagatakse sama keelesüsteemi, sama grammatikat, on kognitiivsete oskuste areng ning keele süntaktiliste seoste valdamine ja mõistmine lapse ja täiskasvanu puhul märkimisväärselt eri astmel. (Dahl 2004)

1.7. Süntakiline kompleksus ja pragmaatika

Jackendoffi ja Wittenbergi meelest pole süntaks üldine lingvistiline süsteem, vaid keel koosnebki kompleksuse eri astmeist (Jackendoff, Wittenberg 2014: 2). Viimasel 20 aastal on üha tulnud ilmsiks, et lapse nn telegrammistilis⁷ kahe- ja kolmesõnalauseingid varjavad seni arvatust kaugeltki rikkalikumat ja abstraktsemat grammatilist süsteemi (Hyams 2011: 13). Keeleomandamisuurimuste üldine raskuskese on nihkunud sellelt, mida lapsed produtseerivad, sellele, mida nad näivad teadvat, kuid ei oska süstemaatiliselt väljendada.

Lapsekeele pragmaatika kontekstis tuleb näiteks arvestada sellega, et võrreldes täiskasvanuga on lapsel lühem mälu ja tähelepanu kestus, vähem oskust plaanida; ta on lausungi loomisel rohkem piiratud kui keele juba omandanud inimene. Seetõttu nõuab grammatika keelt alles omandavalt lapselt eneselt rohkem ressursi, et teha oma mõttetervikud eksplitsiitseks ja kuulaja pingutus väiksemaks. Võib öelda, et täiskasvanu kõne süntakiline kompleksus muudab semantika ja pragmaatika piirid kuulajale lihtsamini tajutavaks (Jackendoff, Wittenberg 2014: 2).

Pikemad sõnad sobivad omakorda näiteks selle kohta, kuidas lapselt võib rohkem pingutust nõuda oma mõtte n-ö nähtavaks tegemine, mõttele sõna fonoloogilise kesta andmine ja selle sidumine tähendusega. Pikk sõna nõuab kõnelejalt enam ressursi kui lühike:

⁷ Mõni näide, mis muudab lapse varase keele telegrammilaadseks: ärajäetud pronoomenid, abiverbid, muutelõpud ja determinatiivid (*determiners*) ning kogu funktsionaalne ülesehitus, mis neid elemente toetab (Hyams 2011: 14).

seada peab kauem aega hääldama, see võtab töömälus rohkem ruumi (Baddeley 1999) ning võrreldes lühikestega on pikki sõnu pikaajalisest mälust raskem kätte saada (Bock 1982). Seevastu lühemaid sõnu on kergem n-ö kätte saada ja produtseerida (Bock 1982, Keren-Portnoy jt 2010)⁸. Eelnimetatud tuleks arvestada lapse üldise kognitiivse arengu, mälu arenemise kontekstis, unustamata pöörata tähelepanu kindla hoidjate ringi keelele.

2. Kuidas süntaktilist kompleksust mõõta

2.1. Süntaktilise kompleksuse mõõtmine lausungi keskmise pikkuse (MLU) järgi

Süntaktilise kompleksuse mõõtmise kõige lihtsam viis on määrata lausungite keskmine pikkus⁹ (Šuster 2011, vt ka Standish 2005: 2), see on tuntuim töövahend (McWhinney 2000: 42), millega mõõdetakse keeleoskuse taset (*language proficiency*) (Hoff 2012: 275). Lausungi keskmise pikkuse mõõdikut ehk MLU-d on lastekeeleuuringutes siiani üsna palju kasutatud. Lapse vanuse ja tema kasutatud väljendite keskmise pikkuse vahel on leitud tihe seos (Brown 1973, Miller 1981). Samas ollakse veendunud, et MLU iseloomustab keele struktuuri ja selle kompleksust vaid ühemõõtmeliselt, nii et paratamatult kerkib küsimus, kuidas seda tüüpi mõõdikuga mõõta lausungite struktuurilist kompleksust, kui mitmemõõtmelisus on lausungite põhiline omadus (Nieminen 2007: 39). Siiski on leitud, et varasel perioodil peegeldab MLU keele omandamise arengut üsna hästi, kuid hiljem enam mitte (Leiwo 1993). Ollakse üsna üksmeelised, et kui MLU väärtus on 4,000, siis struktuuriline kompleksus

⁸ Näiteks inglise keelt emakeelena rääkima õppivad lapsed väldivad oma kõnes funktsiooniverbe (abiverbe). See on laste viis vähendada öeldu töötlemiseks kuluvat pingutust.

⁹ MLU (*mean length of utterance*) kõrval käibib ka eesti väljendile viitav VKP.

laste lausungites suurel määral enam ei muutu (Nieminen 2009: 181, Huttenlocher jt 2002: 339). Lapsed, kes on ületanud nn Browni V staadiumi (MLU = 4,5) – vt täpsemalt järgmises lõigus –, võivad täiustada lausutu struktuurilist kompleksust, ilma et suurendaksid lausungi pikkust (Jalilevand, Ebrahimipour 2014: 3, Chabon jt 1982).

MLU arvutamiseks sõnastas mõõdiku looja Roger Brown (1973) teatud üheksa kriteeriumit (vt nt ka MacWhinney 2000: 116). Brown oli huvitatud viisist, kuidas morfeemide omandamine võib peegeldada süntaktilist arengut, ja uskus, et MLUm (väljendi keskmine pikkus morfeemides) peegeldab seda täpsemini kui MLUw (väljendi keskmine pikkus sõnedes). Brown kirjeldas lapse keelelist arengut kuue staadiumi kaudu, kus MLU väärtus kasvab vahemikus 1,75–4,5¹⁰ (vt ka Ingram 1999: 50). Üks hilisemaid uuringuid (Klee jt 1989) näitas, et MLU korreleerub vanusega ligikaudu kuni 48 kuud (MacWhinney 2000: 115). Ajal, kui Brown selle mõõdiku lõi, arvutati MLU käsitsi, tänapäeval on seda võimalik teha automaatselt programmiga CLAN.¹¹

Kuigi MLU arvutamiseks leiavad eri autorid piisavat kas 50-st või vähemalt kuni 100-st lausungist (Lee 1974, Miller 1981, Williamson 2009: 2), on osa seisukohal, et mida suurem on analüüsitavate lausungite hulk, seda täpsem on tulemus (Miller, Chapman 1981). Viimast arvamust toetab ka Jodi Tommerdahli ja Christel de Bruijini

¹⁰ Brown (1925–1997) määras oma klassikalises teoses (1973) pikiuuringu põhjal laste – tuntud nimede Adam, Sarah ja Eve all – morfosüntaktilise arengu viis staadiumit: *I staadium*, vanus 15–30 kuud, MLU 1,5–2,9 (keskmine 1,75); *II staadium*, 28–36 kuud, MLU 2,0–2,5 (2,25); *III staadium*, 36–42 kuud, MLU 2,5–3,0; *IV staadium*, 40–47 kuud, MLU = 3,0–3,7; *V staadium*, 42–52+ kuud, MLU 3,7–4,5. Alates 1973. aastast on Browni teost tsiteeritud enam kui 700 korda.

¹¹ Vt nt Brian MacWhinney, Johannes Wagner, Transcribing, searching and data sharing: The CLAN software and the TalkBank data repository. – Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, 11 (2010), 154–173. www.gespraechsforschung-ozs.de (30.08.2014).

uurimus (2008), mille eesmärk oligi teada saada, kuivõrd usaldusväärased on morfosüntaksi analüüsi tulemused tingimustel, kui uurimisaluste lausungite arv on 50, 100, 150 või 200.

MLU-d on arvatud nii morfeemide kui ka sõnede põhjal. On neidki uurimusi, kus MLU arvatati nii morfeemide (MLUm) kui ka sõnede põhjal (MLUw) (nt Marques, Limongi 2011). Morfeemide järgi MLU arvutamise pooldajad ütlevad, et keele põhiüksus on pigem morfeem, mitte sõna, teisalt on keeleuurijaid, kes möönavad, et alati pole lihtne määrata, mis just on morfeem¹² (Williamson 2009: 2). Ka on leitud, et MLUm ja MLUw vahel on tugev sarnasus – MLUm on osutunud vaid õige pisut suuremaks (Dijk, Geert 1999). Siinkohal peab siiski mainima, et erinevus MLUw ja MLUm vahel sõltub väga palju uuritavast keelest. Näiteks eesti keele puhul võivad sõnede ja morfeemide järgi saadud tulemused erineda suurel määral, nagu vorm *kaussidegagi*, mis on üks sõna, kuid sisaldab neli morfeemi (*kaussi-de-ga-gi*).

Enamasti kehtibki eelöeldu inglise keele omandamise kontekstis.¹³ Eesti keel on suurema sünteetilisuse astmega kui inglise keel ja eesti keele omandamise uuringutes arvutatakse MLU sõnedes, sest Eesti lastekeelekorpus on morfoloogiliselt märgendamata (Argus 2008: 9). MLU arvutatakse tänapäeva lastekeeleuurimuses üldiselt sõnedes ka teiste keelte puhul (Behrens 2006: 11), sest tihti tekib probleem, mida määrata morfeemiks¹⁴. Keele omandamise

¹² Vt ka Helen Kõrgesaare artiklit siinsamas kogumikus. *Toim.*

¹³ MLU, aga ka allpool (vt 2.2) jutuks tulev IPSyn ja DSS (*developmental sentence scoring*) olid algselt kavandatud selleks, et analüüsida inglise keelt kõnelevate laste grammatikat; teiste keelte uurimiseks tuleb neid mõõdikuid modifitseerida (Jalilevard, Ebrahimipour 2014: 9).

¹⁴ Williamson (2009: 2) on näiteks arutlenud, kas inglise keele lihtmineviku ebareeglipäraseid vormid (nt *sang, brought, went*) tuleks lugeda üheks morfeemiks, samal ajal kui reeglipärase tegusõnade lihtminevikuvormi arvestada kaheks (*talk-ed, play-ed, push-ed*). Eesti lastekeeleuurijad puutuksid keele ehituse tõttu sedalaadi küsimustega kokku väga sageli.

varasel astmel on morfoloogia sageli omandatud kindlate tervikvormidena ega ole produktiivne ning just vormirikaste keelte puhul ei ole alati selge, mitut morfeemi MLU määramisel arvesse võtta (samas; Hoff 2012: 275).

Lastekeeleuurijad kasutavad keeleandmete transkribeerimiseks enamasti programmi CLAN¹⁵, mis võimaldab keeleandmeid ka analüüsida¹⁶. Kuna CLAN on koostatud inglise keele jaoks, tuleb seda tõika eesti keele materjali analüüsimisel pidevalt arvestada (Argus 2007: 70, Piits jt 2007: 368). Näiteks eesti keele puhul ei esita CLAN sõnasagedustabelis mitte lekseemide, vaid sõnavormide sageduse ning nt *kala* vormid *kala*, *kalaga*, *kalale* loetakse eri sõnadeks (ingl *types*). See omakorda tingib, et statistiliselt on valed ka MLU ja sõnavara variatiivsuse indeks (*type-token ratio*), mille programm arvutab automaatselt (Argus 2007: 70). Eesti lapsekeele andmete analüüsil tuleb uurijal väljendi ja kõnevooru pikkuste täpseks määramiseks teha käsitsitööd, kuna empiirilised tulemused näitavad, et käsitsi arvutatud tulemused ei ühti ei CLAN-i pakutud MLU-ga ega ka MLT-ga (*mean length of turn*, kõnevooru pikkus).

Sõnede hulk lausungis võib osutada keelelisele kompleksusele, kuid ei pruugi seda tingimata teha. Näiteks võib ühe lapse kõnes esineda kordusi (nt *õue õue õue*) ja teise omas mitte (nt *issi läks tööle*), kuid kompleksuse aste olla väga erinev hoolimata sellest, et mõlema näite keskmine pikkus on kolm sõnet.

Vaadakem veel üht näidet lapse kõnest (Argus 2007: 71–72), kus lapse MLU on 3:

*EMA: ei ja, ei ole või?

*LAPS: ei a ee eehe.

¹⁵ Computerized Language Analysis, vt <http://childes.psy.cmu.edu>.

¹⁶ Analüüsi võimaluste kohta vt Hoff 2012: 275, MacWhinney 2000 (võrguväljaanne on uuendatud 24.03.2013).

*EMA: jaa, jah, pane ots külge.

*LAPS: ee ee (Argus 1;7)

Kuigi lapse MLU on toodud näites 3 ehk lapse lausungis on keskmiselt kolm sõnet, ei peegelda selline tulemus adekvaatselt lapse tegelikku keelelist arengutaset. Kolmest sõnest koosnevate lausungite all mõeldakse üldjuhul seda, et lapse lausungis esineb kolm tähenduslikku sõna. (*Ibid.*) Teisedki uurijad (nt Jalilevand, Ebrahimipour 2014: 3) on täheldanud, et kuigi MLU võib olla keelelise arengu üks näitajaid, ei kirjelda see mõõdik grammatikapädevuse struktuurilist kompleksust isegi nende laste puhul, kelle MLU väärtus on ühesugune.

Seega tuleb lastekeeleuurijal juba enne transkribeeringute analüüsi kaaluda, missugust keelematerjali MLU arvutamiseks kasutada, missugused keeleühikud välja jätta. Näiteks võib kaaluda, kas üneemid (pausitäitehäälikud) märgendada kuidagi nii, et arvutuste jaoks oleks võimalik nad analüüsitava keelematerjali hulgast välja arvata, ning kordused esitada juba transkribeeringutes sel viisil, et neid ei peaks arvestama eraldi sõnedena, vaid ühena. Üheks lekseemiks oleks aga põhjust arvestada kõik semantiliselt samased ja häälikuliselt sarnased sõned, mida transkribeeringus esineb kahel, kolmel, neljal jne korral. Nimelt sisaldab lastekeel, eriti just selle varases staadiumis, palju ühe sõnavormi kordusi (vt Argus 2007: 72). Läbimõtlemata transkriptsioon hakkab analüüsitulemusi oluliselt mõjutama (samas: 73).

Kui aga uurijal on vaja võrrelda eri uurimuste MLU-sid, tuleks teha kindlaks, mille järgi on MLU igas võrdlusaluses uurimuses arvutatud.

Vajadust kasutada MLU kõrval mitut muud mõõtevahendit on tunnistanud laialt (Jalilevand, Ebrahimipour 2014, Šuster 2011, Nieminen 2007 ja 2009, Gillis, Ravid 2009). Neist tulebki allpool juttu.

2.2. Süntaktilise kompleksuse mõõtmine produktiivse süntaksi indeksi järgi

Sage kvantitatiivne vahend, millega süntaktilist kompleksust on mõõdetud, on produktiivse süntaksi indeks IPSyn (*index of productive syntax*) (vt Scarborough 1990, MacWhinney 2000, Sagae, Lavie, MacWhinney 2005, Nieminen 2009). IPSyn põhineb morfoloogilistel vormidel ja süntaktilistel struktuuridel, mis on leitud lapsekeeles olevat põhilised (Nieminen 2009: 177). Indeks on loodud selleks, et lastekeeles saaks uurida individuaalseid erisusi; see on mõeldud parandama MLU-I ilmnunud nõrkusi. IPSyn'i eesmärk on näidata süntaksi arengut lapse keeles (Sagae jt 2005).

Produktiivse süntaksi indeksi töötas 1990. aastal välja Hollis S. Scarborough¹⁷ (vt ka Scarborough jt 1991: 42–43). IPSyn'i loomise peamine mõte oli analüüsida, missugused põhilised süntaktilised struktuurid ja morfoloogilised vormid on esindatud lapse 100-s mõistetavas lausungis (Scarborough 1990, Nieminen 2007: 53). IPSyn mõõdab nelja suurt süntaktilist kategooriat:

- 1) nimisõnafraas. Kokku on siin 12 jaotist, N1–N12, nt omadussõnad, mitmuslikud nimisõnad ja kahe-kolmesõnalised nimisõnafraasid;
- 2) verbifraas. Kokku on siin 17 jaotist (V1–V17), näiteks erinevad verbivormid, adverbid (*now, tomorrow, too, hardly, already, yet*), partikkel või prepositsioon (*up, put back, this on*) jt;
- 3) küsimused ja eitused. Kokku on siin 11 jaotist (Q1–Q11), näiteks intonatsiooniga markeeritud küsimused (*Mine? Want some?*), lihtne eitus ja keelud (*No talk. Not yours. Can't*

¹⁷ IPSyn'i loomiseks uuris Scarborough (1990) pikiuuringuga 15 lapse kõne näidiseid, lapsed olid vanuses 24, 30, 36, 42 ja 48 kuud, kokku sai H. Scarborough 75 kõnenäidist.

find. Don't know), eitust väljendav morfeem subjekti ja verbi vahel (*I not going. You can't have it*), inversiooniga küsimused (*Where was it? What can I do?*), kinnitusküsimused (ingl *tag questions: I'll be the mom, okay? This is fun, isn't it?*) jt;

- 4) lausestruktuurid. Kokku on siin 20 jaotist (S1–S20), näiteks kahesõnalausungid (*Lookit Mom. Here hammer*), subjekti ja verbi järgnevus (*Monkey fall. It goes over there*), relatiivlausungi kasutamine, (*Find the one that fits here*), laused kolme ja enama verbifraasiga jt. (Scarborough jt 1991: 42–43)

Kokku mõõdab IPSyn 60 grammatilist kategooriat: 12 nimisõna, 17 verbi, 11 osa jaotist küsimusi ja eitusi ning 20 tüüpi lausungeid (Scarborough 1990, Scarborough jt 1991, Hassanali jt 2014).

IPSyn mõõdab niisiis pigem morfoloogiliste ja süntaktiliste kategooriate koosesinemust kõnenäites kui sõnede lihtsat hulka selle väljendites. Iga vormi eest võib saada 0 punkti (ei esine kunagi), 1 punkti (esineb kõnenäites ühe korra) või 2 punkti (esineb kaks või rohkem korda). Kuna IPSyn on loodud nii, et see mõõdab, millal teatud grammatilised vormid lapse kõnesse ilmuvad, siis peetakse mingi kindla konstruktsiooni kaht eri esinemisjuhtu piisavaks.

IPSyn'i rakendamiseks on vaja korpust saja transkribeeritud lausungiga (järgnevad lausungid; kordused välja arvatud), tuleb identifitseerida 60 keeleüksust (nn IPSyn'i ühikut), arvutada esinemisjuhud (0 punkti, 1 punkt, 2 punkti) ja lugeda punktid kokku (Sagae jt 2005).

Kuigi ka produktiivse süntaksi indeks IPSyn on kvantitatiivne mõõtevahend, on see analüütilisem kui MLU. Kui MLU hindab tege-likku kõneproduksiooni (*actual speech production*) morfeemides või sõnedes, siis IPSyn keskendub morfosüntaktilise repertuaari suu- rusele ja kvaliteedile (Nieminen 2009: 177, Nieminen 2007: 58–59), mis automaatselt tähendab tegelemist suhtega keele elementide vahel.

Lausungi struktuuri tulebki käsitada kui mitmemõõtmelist ter-
vikut, mis koosneb isesugustest süntaktilistest komponentidest, mis
võivad olla loodud mitmel detailirikkal morfoloogilisel ja süntakti-
lisel tasandil. Lausungi komplekssus ei näita mitte üksi tema osiste
hulk, mida mõõdetakse MLU-ga, vaid ka lausungi enese sisemine
struktuur, mida IPSyn hierarhiliselt näitab.

Ka need mõõdikud ei ole ometi ideaalsed. Komplekssus võib
kasvada vähemalt kahel moel, kas lisades olemasolevale keele-
ainesele uusi osiseid või lisades olemasolevale osisele uusi kihte.
Samas eeldavad MLU ja IPSyn Nieminen (2007: 177) väitel, et kee-
leline komplekssus kasvab vaid ühel moel, kuigi lausungi struktuur
pakub kompleksssuse kasvuks mitut võimalust. Sedasama väidavad
ka Jackendoff ja Wittenberg (2014): süntaks pole sugugi ainus üldine
keelesüsteem, vaid keel koosneb tasanditest, millel on erinev komp-
lekssuse aste.

2.3. Kokkuvõtteks MLU ja IPSyn'i kohta

Lastekeele uurimisel (ka kompleksssuse mõõtmisel) peaks arves-
tama kompleksssuse kui nähtusega, mis hõlmab keele eri tasandeid.
Kõikidel olemasolevatel süntaktilise kompleksssuse mõõdikutel, ka
eelkirjeldatud MLU-l ja IPSyn'il, on vajakajäämisi. Kuigi nii MLU
kui ka IPSyn on struktuurilise kompleksssuse mõõtmiseks eba-
adekvaatsed ega näita lapse keelelise arengu kõiki aspekte, aitavad
need mõõdikud uurijal siiski näha üldisi arengujooni, kui mõlemat
mõõdikut kasutatakse koos (Nieminen 2007, Nieminen 2009: 185).
Peale MLU ja IPSyn'i rakendamise on saadud tulemusi vaja põhja-
likult interpreteerida (vt Nieminen 2007: 86–100). Kuna mõlemad
mõõdikud on keelesõltlikud, tuleb need kohandada kindla keele
tarbeks (Gillis, Ravid 2009: 227).

Kuigi MLU või IPSyn'i kasutamisel võib keelevormide struktuurilise kompleksuse (või selle puudumise) kohta tõenäoliselt saada rohkem teavet kui neid rakendamata, jääb nende mõõdikute kasutusel kõrvale see, mida võib näidata kvalitatiivne analüüs: see on nende keelevormide töötlemine, mis on haaratud kordustesse, ja keeleressursi jagamine pädevama vestluskaaslasega (nt hoidjakeele mõju). Selliseid aspekte kompleksuse mõõdikud ei arvesta, kuna lausungite või sõnade imitatsioonid jäetakse kõrvale ja sama kehtib korratud keelematerjali kohta. Tekib küsimus, kuidas analüüsida kompleksust nii, et seda nähtust ennast ei pisendataks, ning kuidas käsitleda arenevat kompleksust nii, et oleks arvestatud ka keele töötlemise sotsiaalsete aspektidega (Suni, Nieminen 2011: 227–228).

Selge on, et MLU ja IPSyn mõõdavad absoluutset kompleksust (vt Miestamo 2006, Dahl 2004, Nieminen 2007: 56–57, 2009), mis näitab keele kui süsteemi keerukust, kuid need mõõdikud jätvavad kõrvale keelekasutaja ja tema kogemuse keelega (subjektiivse kompleksuse). Keelt absoluutse kompleksuse vaatenurgast analüüsides on uurija see, kes otsustab, missuguseid keeleüksusi analüüsil kaasa arvata, millised mitte (vt Suni, Nieminen 2011: 222). Mõõdikuid tuleb aga kasutamisel kohandada kindla keele jaoks, kuna MLU ja IPSyn on loodud inglise keele omandamise mõõtmiseks. Kui kohandada mõõdikut IPSyn eesti keele tarbeks, tuleks põhimisena arvestada neid kategooriaid, mida kasutatakse just lapsega rääkides; st ei tohiks argumenteerimata hõlmata hoidjakeeles harva esinevaid kategooriaid (nt kaudset kõnet).

3. Missugune peaks olema kompleksuse mõõdik, mida võib pidada heaks

Lise Menni ja Jill Duffieldi järgi (2014) peaks kompleksuse mõõtmisel olema täidetud järgmised tingimused:

1. Kompleksuse mõõtmine peaks suutma ennustada pingutust, mida tehakse lingvistiliste üksuste töötlemiseks (ingl *processing effort*). Mõõtevahend peaks suutma korrektselt prognoosida pingutust, mida nõuab üksikute lausungite või terviklike mitme lausungi üksuste (*passages*) lausumine. Kõnelõigu töötlemise pingutus võib sõltuda kindlast lausungisse valitud leksikaüksusest. Ehkki keeleteadlased analüüsivad üksiklausungit nii, et jagavad selle mitmeks osaks, võib inimmeel tollesama lausungi talletada kui terviku ja töödelda seda kas kui tervikut või ühtaegu nii tervikuna kui ka läbi analüüsituna. Mõõtmisel võib selguda, et ei pruugigi olla erinevust selle vahel, kas mõni keeleüksus säilitatakse mälus kui tervik (nt *võta või jäta*) või on üksuse osadel ennustatavad seosed (nt ühilduv nimisõna fraas *punastesse karpidesse*, kus *-te* või *-de* viitab mitmusele ja *-sse* näitab ruumisuhet, kombineerudes kummagi tüve leksikaalse tähendusega).
2. Kompleksuse mõõdik peab sobituma keele erisuguse kasutuse juhtudega. Ei saa olla vaid üht mõõdikut, sest see, mis nõuab pingutust kõnelejalt, ei ole alati võrreldav selle pingutusega, mida teeb kuulaja (nt kuulaja peab tegema kindlaks, millele kõneleja osutab, tõlgendama mõnd võrdlust jms). See, mida on raske õppida (nt ebareeglipärased vormid), võib olla suhteliselt lihtne kasutada, kui see kord on juba omandatud. Üleüldise kompleksuse mõõtmist peab olema võimalik lahutada alaeasmärkideks, on vaja mitut laias laastus võrreldavat, kuid mõnikord vastandlikku mõõdet (Menn, Duffield 2014). Näiteks

standardses keeletöötlusmudelil ei ole arvestatud järjestikuste seostega, mis sõnade vahel võivad olemas olla, ja nii jäetakse tähelepanuta konstruktsioonid ja idioomid. Selle asemel, et loendada sõnu ja määrata nende liik, võib olla otstarbekas uurida ka seda, missugustes (sagedastes või püsi-) väljendites sõnu on kasutatud.

3. Komplekssuse mõõdikud peavad suutma prognoosida koostoimet, kuidas sagedus ja struktuur vastastikku üksteist mõjutavad. Töötlemise pingutus on sageduse ja struktuuri koostoime funktsioon mitmel tasandil.

Struktuur on kooslus, mis koosneb teatud viisil struktureeritud osade väärtustest, mis üksteisest erinevad, ning on osalt määratletud selle kaudu, millised on struktuuri teised osad (Karlsson 2002: 60). Keele struktuursus avaldub sümbolilisuses, allsüsteemides ning suhetes allsüsteemide sees ja vahel. Igas allsüsteemis on mitut liiki üksusi, üksused realiseeruvad kategooriatena. Süntaksi lausekategooriateks on näiteks väit- ja küsilauseid jne, fraasikategooriateks nimisõna- ja verbifraas jne. Morfoloogia kategooriateks on nt eesti keele käände- ja pöörde lõhud jm sõnaliigi morfoloogiale omane. Leksikoni põhiüksus on sõna ja kategooriateks eri liiki sõnad ja väljendid. Kõigi allsüsteemide elemendid moodustavad süntagmasid, mille elemendid on omavahel süntagmaatilistes seostes. Ka kõrvuti asetsevate sõnade tähendus(ed) on süntagmaatilises seoses. Teisalt on iga keeleelement paradigmaatilistes seostes nende elementidega, millega teda saab asendada. Näiteks on foneemid /p, l, m, s, k/ paradigmaatilises seoses kontekstis „_uu“, samuti nagu sõnad *linn*, *buss*, *üllatus*, *see* kontekstis „___ on suur“ (Karlsson 2002: 33–34).

Näitena töötlemise pingutusest kui sageduse ja struktuuri koostoime funktsioonist mitmel tasandil võib tuua kõnejadast

sõna äratundmise, mis nõuab teadlikkust fonoloogiast (silbid, foneemid jt), sõna tähenduse mõistmist, mida toetavad nii temaga koos kasutatud sõnad ja nende tähendus kui ka suhtlusolukord. Eelöeldule lisandub arusaamine kasutatud sõna(de) morfoloogiliste tunnuste tähendusest.

Näiteid struktuuri ja sageduse koosmõjust pakub kirjandus inglise keele põhjal. Nii kasutatakse tegusõna „shrink” ‘kokku tõmbama’ sagedamini kogeja argumendistruktuuris (nt *The sweater shrank two sizes*) kui mõnes muus (nt *They shrank the sweater two sizes*). Või teine näide: sagedus suhteverbide raamistikus loob ootuse, mis mõjutab kuulaja (või lugeja) töötlusmalle ja arusaamist. Sellega seostub, et teatud afaasia puhul saavad inimesed lausetest paremini aru siis, kui verb on oma kõige sagedamini kasutatud freimis ehk raamistikus (vt Gahl jt 2003), või et normkõnelejad töötlevad mitmuslikku sõna suhteliselt kiiremini, kui seda sõna harilikult kasutataksegi mitmuses (vt Menn 2013: 5). Näide struktuuri ja sageduse vastasmõjust on ka kollokatsioonid, s.o sõnade järgnevused, mis sageli koos eksisteerivad (nt *it’s, gimme, suppose we, ladies and gentlemen, wash the dishes*) ega olene sellest, kas järgnevus moodustab süntaktilise üksuse või kas tal on hästi defineeritud ja tuntud tähendus (Menn 2013: 4).

4. Komplekssuse mõõdikud peavad suutma toime tulla ka võistlemisega (seda võib mõista kui kontekstisobilike või -tüüpiliste keelendite, suhtlustoimingute ja kõnekäitumise joonte suhtelist ennustamatust). Töötlemine on võistlemisest tugeval määral mõjutatud; näiteks võistlevad leksikaalsed üksused (pöördu mine lapse poole, kasutades tema õe-venna nime), võistlevad suhtlusolukorra tõlgendused ning näiteks ka poolikuks, lõpetamata jäänud struktuuri võimalik keeleline ja suhtlustähendus. (Menn, Duffield 2014)

5. Komplekssuse mõõdikud tuleb kohandada pragmaatikale ja kontekstile. Pragmatikal on tohutu mõju keelendite töötlemisele ja nende mõistmisele psühholingvistikas (samas; vt ka Gibson 1998, Lai, Curran, Menn 2009).

Kokkuvõte

Süntaktilise komplekssuse mõõtmise tulemuse määravad paljus valikud, mida teeb keeleomandamise uurija. Seetõttu määrab komplekssuse mõõtmise tulemuse just uurija teadlikkus uuritavast keelest ja selle ehitusest, komplekssusest endast ning uurimismeetoditest ja oskusest neisse kriitiliselt suhtuda.

Ehkki nii MLU kui ka IPSyn on struktuurilise komplekssuse mõõtmiseks ebaadekvaatsed ega näita lapse keelelise arengu kõiki aspekte, aitavad need mõõdikud uurijal siiski näha üldisi arengujooni, kui mõlemat mõõdikut kasutada koos.

Lühidalt võib MLU ja IPSyn'i kohta öelda järgmist.

- 1) Mõlemad mõõdikud on kvantitatiivsed, kuid komplekssuse uurimisel on vajalik ka keelematerjali kvalitatiivne analüüs. Näiteks lausungi komplekssust ei näita mitte üksnes tema osiste hulk, mida mõõdetakse MLU-ga, vaid ka lausungi enese sisemine struktuur.
- 2) Komplekssuse mõõtmisel on väga tähtis uurija teadlikkus ehk valikud ja otsustused, mida ta teeb, nagu nt Nieminen (vt 2007: 38) otsustas, et tema mõõdab absoluutset komplekssust, ja põhjendas, miks.
- 3) Milline tahes mõõdik on juba ette ebatäiuslik, sest keele komplekssus on väga mitmetahuline nähtus. Seetõttu on tähtis uurija teadlikkus nii uuritavast keelest kui ka mõõdikute sisust ja eesmärgist ehk sellest, mida ja miks mõõdetakse.

- 4) Inglise keele (flekteriva keele) tarbeks loodud mõõdikud ei sobi teist tüüpi ehitusega keelele (nt aglutinatiivne keel) ja on üsna loomulikki, et see nii on. Seega ei saa otse üle võtta keeleteaduslikes artiklites muude keelte kohta saadud uurimistulemusi, ja oletada, et samad tulemused võiksid üksi ühele kehtida ka eesti keele kohta.
- 5) MLU on loodud aastal 1974 ja IPSyn 1990. Kui praeguseks ei oleks avastatud, et need mõõdikud on ebatäiuslikud, oleks midagi valesti, seda enam, et mõlema puhul on tegu kvantitatiivsete töövahenditega (see tõik ise ütleb varjamatult, et tegu on ebatäiuslike vahenditega, sest kvantiteet vajab alati kvalitatiivset selgitust). Keeles, nagu ka mujal, on lisaks kvantiteedile olemas ka kvalitatiivne mõõde, kuigi need on tihedalt seotud.

Süntaktilise kompleksuse mõõtmise täiuslikku viisi või vahendit pole olemas, kuna üsna palju näib olevat nii keelesõltlikke (objektiivseid) kui ka keeleväliseid (subjektiivseid) tegureid, mida süntaktilise kompleksuse mõõtmist kavandades tuleks või võiks hõlmata. Laps avastab keeleomandamise käigus mitme allsüsteemiga loomulikku keelt, mille grammatika ei eksisteeri niisama, vaid selle eesmärk on väljendada ja konstrueerida tähendusi (Karlsson 2002: 255) ning tähendused luuakse suhtluses. Seetõttu on kompleksuse uurimisel peale keele kohta sõnastatud normide ja reeglite tundmise vajalik n-ö pehmem, sotsiaalne lähenemine – tõlgendus, mille annab keeleuurija.

Üks lahendusi, kuidas luua süntaktilise kompleksuse täpsem mõõdik, näib olevat teha selle eesmärgi huvides koostööd erialade vahel. Keeleüksuse loomisele kuuluvat pingutust saab mõõta ilmselt neurolingvist, kes uurib aju aktiivsust (vt Lise Menni viidatud jm töid). Vajalikke lisaandmeid saaksid aga keeleomandamist uurivale keeleteadlasele anda näiteks ka muud neuroteadlased, arvutimudelite loojad, arengupsühholoogid.

Kirjandus

- Abbot-Smith, Kirsten; Heike Behrens 2006. How known construction influence the acquisition of other constructions: The German periphrastic passive and future constructions. – *Cognitive Science*, 30, 995–1026.
- Allik, Jüri; Riina Häidkind, Jaanus Harro, Mart Viikmaa, Kairi Kreegipuu, Marika Rauk, Eve Kikas, Tiia Tulviste, Peeter Tulviste, Aavo Luuk, Anu Realo, Kenn Konstabel, Helle Pullmann, Monika Schmidt, Maaja Vadi, Maie Kreegipuu 2002. *Psühholoogia gümnaasiumile*. Tartu: Tartu Ülikooli kirjastus.
- Argus, Reili 2007. Eesti lastekeelekorpusse morfoloogilisest märgendamisest. Tallinna Ülikooli eesti filoloogia osakonna toimetised, 9, 65–86. http://evkk.tlu.ee/wwwdata/kogumik2007/korpusekogumik_argus.pdf (27.07.2014).
- Argus, Reili 2008. Eesti keele muutemorfoloogia omandamine. (= Tallinna Ülikooli humanitaarteaduste dissertatsioonid 19.) Tallinn: Tallinna Ülikooli kirjastus.
- Argus, Reili 2009. Acquisition of Estonian: some typologically relevant features. – Helle Metslang (ed.), *Estonian in Typological Perspective. Language Typology and Universals / Sprachtypologie und Universalienforschung*, 62 (1/2), 91–108.
- Baddeley, Alan 1999. *Essentials of Human Memory*. Hove: Psychology Press.
- Behrens, Heike 2006. The input-output relationship in first language acquisition. – *Language and Cognitive Processes*, 21 (1–3), 2–24.
- Behrens, Heike 2009. Usage-based and emergentist approaches to language acquisition. – *Linguistics*, 47 (2), 383–411. doi 10.1515/LING.2009.014
- Bloom, Louis 1970. *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: M.I.T. Press.
- Bloom, Paul 2004. Myths of word learning. – Geoffrey Hall, Sandra R. Waxman (eds.). *Weaving a Lexicon*. Cambridge, MA: M.I.T. Press.
- Bock, Kathryn J. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. – *Psychological Review*, 89, 1–47.
- Bowerman, Melissa 1978. Semantic and syntactic development: A review of what, when, and how in language acquisition. – Richard Louis

- Schiefelbusch (ed.), *Bases of Language Intervention*. Baltimore: University Park Press, 97–189.
- Brown, Roger 1973. *A First Language: The Early Stages*. London: George Allen & Unwin.
- Cejnarova, Andrea 2005. *Complexity of the Big and Small*. Stellenbosch: University of Stellenbosch.
- Chabon, Shelly S.; Louise Kent-Udolf, Donald B. Egolf 1982. The temporal reliability of Brown's Mean Length of Utterance (MLU-m) measure with post-stage V children. – *Journal of Speech, Language, and Hearing Research*, 25, 124–128.
- Culicover, Peter W. 2013. *Grammar & Complexity: Language at the Intersection of Competence and Performance*. Oxford: Oxford University Press.
- Dahl, Östen 2004. *The Growth and Maintenance of Linguistic Complexity*. Philadelphia: John Benjamins.
- Dahl, Östen 2008. Grammatical resources and linguistic complexity: Sirionó as a language without NP coordination. – Matti Miestamo, Kaius Sinnemäki, Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 153–163. doi: 10.1075/slcs.94
- Dijk, Marijn van; Paul van Geert 1999. *Short-term Variability in Child Language*. (Ms.) Netherlands: University of Groningen.
- Eslon, Pille 2014. Morfosüntaktilise ja leksikaalse varieerumise piiridest: ilukirjandus- ja õppijakeele kasutusmuutrite võrdlus / Constraints on morphosyntactic and lexical variability. – *Eesti Rakenduslingvistika Ühingu aastaraamat / Estonian Papers in Applied Linguistics*, 10, 55–71. doi: <http://dx.doi.org/10.5128/erya.v0i10.254>
- Gahl, Susanne; Lise Menn, Gail Ramsberger, Daniel S. Jurafsky, Beth Elder, Molly Rewega, Audre L. Holland 2003. Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. – *Brain and Cognition*, 53, 223–228.
- Gibson, Edward 1998. Linguistic complexity: Locality of syntactic dependencies. – *Cognition*, 68, 1–76.
- Gillis, Steven; Dorit Ravid 2009. *Language acquisition*. – Sandra Dominiek, Jan-Ola Östman, Jef Verschueren (eds.), *Cognition and Pragmatics*.

- Handbook of Pragmatics Highlights 3. Amsterdam, Philadelphia: John Benjamins, 201–249. doi: 10.1075/hoph.3.12gil
- Gleitman, Lila; Elissa L. Newport, Henry Gleitman 1984. The current status of the motherese hypothesis. – *Journal of Child Language*, 11, 43–79.
- Hassanali, Khairun-nisa; Yang Liu, Aquiles Iglesias, Tamar Solorio, Christine Dollaghan 2014. Automatic generation of the index of productive syntax for child language transcripts. – *Behaviour Research Methods*, 46 (1), 254–262. doi: 10.3758/s13428-013-0354-x
- Hoff, Erika 2012. *Research Methods in Child Language: A Practical Guide*. First Edition. UK: Blackwell.
- Huttenlocher, Janellen; Marina Vasilyeva, Elina Cymerman, Susan Levine 2002. Language input and child syntax. – *Cognitive Psychology*, 45, 337–374. <http://psychology.uchicago.edu/people/faculty/levine/HuttenlocherVasilyeva2002.pdf> (27.07.2014).
- Hyams, Nina 2011. Missing subject in early child language. – Jill de Villiers, Tom Roeper (eds.), *Handbook of Generative Approaches to Language Acquisition*. The Netherlands: Springer, 13–52.
- Ingram, David 1999. *First Language Acquisition: Method, Description and Explanation*. The 5th edition. Cambridge: CUP.
- Jalilevand, Nahid; Mona Ebrahimipour 2014. Three measures often used in language samples analysis. – *Journal of Language Acquisition and Development (JCLAD)*, 2 (1), 1–12.
- Johnson, Elisabeth K.; Mybeth Lahey, Mirjam Ernestus, Anne Cutler 2013. A multimodal corpus of speech to infant and adult listeners. – *Journal of the Acoustical Society of America*, 134, EL534-EL540. doi: 10.1121/1.4828977
- Jackendoff, Ray; Eva Wittenberg 2014. What you can say without syntax: a hierarchy of grammatical complexity. – Frederick J. Newmeyer, Laurel Preston (eds.), *Measuring Grammatical Complexity*. Oxford: OUP, 65–82.
- Karlsson, Fred 2002. *Üldkeeleteadus*. Tallinn: EKSA.
- Kelly, Barbara; Gillian Wigglesworth, Rachel Nordlinger, Joseph Blythe 2014. The acquisition of polysynthetic languages. – *Language and Linguistics Compass*, 8 (2), 51–64.

- Keren-Portnoy, Tamar; Marilyn M. Vihman, Rory A. DePaolis, Chris J. Whitaker, Nicola M. Williams 2010. The role of vocal practice in constructing phonological working memory. – *Journal of Speech, Language, and Hearing Research*, 53, 1280–1293. doi:10.1044/1092-4388(2009/09-0003)
- Kerge, Krista 2003. Keele variatiivsus ja *mine*-tuletus allkeelte süntaktilise keerukuse tegurina / Dynamics of the Syntactic Complicacy of the Journalistic and Fiction Texts through the 20th Century. (= TPÜ humanitaarteaduste dissertatsioonid 10.) Tallinn: TPÜ kirjastus.
- Kerge, Krista, Anne Uusen, Halliki Põlda 2014. Teismeeva loovkirjutiste sõnavara ja selle hindamine / Teenage vocabulary and its assessment in creative writing. – *Eesti Rakenduslingvistika Ühingu aastaraamat / Estonian Papers in Applied Linguistics*, 10, 157–175.
- Kerge, Krista; Anne Uusen, Halliki Põlda, Helin Puksand 2014. Loovkirjutiste süntaksimuutujate areng teismeevas / Development of syntactic parameters of teenage creative writing. – *Emakeele Seltsi aastaraamat*, 59 (2013), 46–76.
- Klee, Thomas; Mary Schaffer, Susan May, Irene Membrino, Karen Moughey 1989. A comparison of the age-MLU relation in normal and specifically language-impaired preschool children. – *Journal of Speech, Language, and Hearing Research*, 54, 226–233. doi:10.1044/jshd.5402.226
- Kõrgesaar, Helen 2009. Hoidjakeelele omastest joontest. – *Oma Keel*, 2, 28–37.
- Kõrgesaar, Helen 2014. Küsimused eesti lapsele suunatud kõnes leedu ja vene keele taustal / Interrogatives in Estonian child-directed speech. – *Eesti Rakenduslingvistika Ühingu aastaraamat / Estonian Papers in Applied Linguistics*, 10, 193–207. doi:10.5128/ERYa10.12
- Lai, Vicky Tzuyin; Tim Curran, Lise Menn 2009. Comprehending conventional & novel metaphors: An ERP study. – *Brain Research*, 1284, 145–155.
- Lee, Laura Louise 1974. *Developmental Sentence Analysis*. Evanston: Northwestern University Press.
- Leiwo, Matti 1993. *Lapse keeleline areng*. Tallinn: TPÜ kirjastus.
- Lipp, Ellen 1977. The acquisition of Estonian inflections. – *Journal of Child Language*, 4 (2), 313–319. <http://dx.doi.org/10.1017/S030500090001689>

- Lääts, Riina 2013. 5-aastaste sisekõrva implantaadiga laste kõnetaju ja -loome nelja juhtumi kirjeldus. Magistritöö. Käsikiri TÜ haridusteaduste instituudis. Tartu: Tartu Ülikool.
- MacWhinney, Brian 2000. The CHILDES project. Tools for analyzing Talk. – Electronic edition. Part 1: The CHAT transcription format, 30.03.2014. <http://childes.psy.cmu.edu/manuals/chat.pdf> (27.07.2014).
- Marques, Suelen Fernanda; Suelly Cecilia Olivan Limongi 2011. Mean length utterance (MLU) as a measure of language development of children with Down syndrome. – *Journal da Sociedade Brasileira de Fonoaudiologia*, 23 (2), 152–157. <http://dx.doi.org/10.1590/S2179-64912011000200012>
- Menn, Lise 2013. Psycholinguistic and neurolinguistic evidence for a rich model of linguistic representation or, why linguistic representations must be redundant. Handouts. (Ms.) <http://linguistics.uoregon.edu/wp-content/uploads/2013/09/Menn-Handout.pdf> (27.07.2014).
- Menn, Lise; Cecily Jill Duffield 2014. Looking for 'Gold Standard' to measure language complexity: What psycholinguists and neurolinguists can (and cannot) offer to formal linguists. – Frederick J. Newmeyer, Laurel Preston (eds.), *Measuring Grammatical Complexity*. Oxford: OUP, 281–302.
- Miestamo, Matti 2006. On the feasibility of complexity metrics. – Krista Kerge, Maria-Maren Sepper (eds.), *FinEst Linguistics. Proceedings of the Annual Finnish and Estonian Conference of Linguistics*. Tallinn, May 6–7, 2004. Publications of the Department of Estonian of Tallinn University, 8, 11–26.
- Miestamo, Matti; Kaius Sinnemäki, Fred Karlsson (eds.) 2008. *Language Complexity, Typology, Contact, Change*. Helsinki: University of Helsinki. doi: 10.1075/slcs.94
- Miller, Jon F. 1981. *Assessing Language Production in Children*. Baltimore: University Park Press.
- Miller, Jon F.; Robin S. Chapman 1981. The relation between age and mean length of utterance in morphemes. – *Journal of Speech, Language, and Hearing Research*, 24, 2, 154–161.
- Nelson, Raymond J. 1975. Behaviourism, finite automata and stimulus-response theory. – *Theory and Decision*, 6, 249–267.

- Newport, Elissa L. 1976. Motherese: The speech of mother's to young children. – N. John Castellan, David B. Pisoni, George R. Potts (eds.), *Cognitive Theory*, Vol. II. Hillsdale: Erlbaum, 177–217.
- Newport, Elissa L.; Henry Gleitman, Lila Gleitman 1977. Mother I'd rather do it myself: Some effects and noneffects of maternal speech style. – Catherine E. Snow, Charles A. Ferguson (eds.), *Talking to Children: Language Input and Acquisition*. Cambridge: CUP.
- Nieminen, Lea 2007. A Complex Case: A Morphosyntactic Approach to Complexity in Early Child Language. (= Jyväskylä Studies in Humanities 72.) Jyväskylä: University of Jyväskylä. <https://jyx.jyu.fi/dspace/bitstream/handle/123456789/13421/9789513928582.pdf?sequence=1> (27.07.2014).
- Nieminen, Lea 2009. MLU and IPSyn measuring absolute complexity / VKP ja IPSyn absoluutse keerukuse mõõtjana. – *Estonian Papers in Applied Linguistics / Eesti Rakenduslingvistika Ühingu aastaraamat*, 5, 173–185. <http://dx.doi.org/10.5128/ERYa5.11>
- Ninio, Anat 2011. *Syntactic Development: Its Input and Output*. New York: Oxford University Press.
- Ota, Michiko 2006. Input frequency and word truncation in child Japanese: Structural and lexical effects. – *Language and Speech*, 49 (2), 261–295.
- Pajupuu, Hille; Krista Kerge 2010. Characteristics and assessment of educated L1 and L2 dialogue. – Rosario Caballero Rodrigues; Jesus Pinar Sanz (eds.), *Modos y formas de la comunicacion humana / Ways and Modes of Human Communication*. Cuenca: Universidad Castilla-La Mancha, 339–348.
- Pinker, Steven 1984. *Language Learnability and Language Development*. Cambridge: Harvard University Press.
- Piits, Liisi; Meelis Mihkla, Tõnis Nurk, Indrek Kiissel 2007. Designing a speech corpus for Estonian unit selection synthesis. – Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek, Mare Koit (eds.), *Proceedings of the 16th Nordic Conference of Computational Linguistics NODALIDA-2007*. Tartu: University of Tartu.
- Peters, Ann 1983. *The Units of Language Acquisition*. Cambridge: CUP.
- Rauh, Gisa 2010. *Syntactic Categories: their Identification and Description in Linguistic Theories*. Oxford: OUP.

- Sagae, Kenji; Alon Lavie, Brian MacWhinney 2005. Automatic measurement of syntactic development in child language. – Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL), 197–204. <http://childes.psy.cmu.edu/grasp/acl05-ipsyn.pdf> (27.08.2014).
- Scarborough, Hollis S. 1990. Index of productive syntax. – *Applied Psycholinguistics*, 11, 1–22.
- Scarborough, Hollis S.; Leslie Rescorla, Helen Tager-Flushberg, Anne E. Fowler, Vicki Sudhalter 1991. The relation of utterance length to grammatical complexity in normal and language-disordered groups. – *Applied Psycholinguistics*, 12, 23–45.
- Schauwers, Karen; Steven Gillis, Paul Govaerts 2005. Language acquisition in children with a cochlear implant. – Paul Fletcher, Jon F. Miller (eds.), *Developmental Theory and Language Disorders*. (= Trends in Language Acquisition Research 4.) Amsterdam: John Benjamins, 324–329.
- Schults, Astra; Tiia Tulviste, Kai Kaljumäe 2013. Eesti laste esimesed sõnad: MacArthuri-Batesi suhtlemise arengu testi tulemused. – *Eesti Arst*, 1, 21–27.
- Smith, Carlota S.; Anne van Kleeck 1986. Linguistic complexity and performance. – *Journal of Child Language*, 13 (2), 389–408. <http://dx.doi.org/10.1017/S0305000900008126>
- Standish, Russell K. 2005. The importance of the observer in science. – *The Two Cultures: Reconsidering the division between the Sciences and Humanities*. UNSW conference proceedings (July 2005) (Ms.). <http://arxiv.org/abs/physics/0508123> (27.08.2014)
- Stolt, Suvi 2009. Language in acquisition. Early lexical development and associations between lexicon and grammar – Findings from Full-Term and Very-Low-Birth-Weight Finnish Children. Helsinki: University of Helsinki. <https://helda.helsinki.fi/bitstream/handle/10138/19892/language.pdf?sequence=2> (27.07.2014).
- Suni, Minna; Lea Nieminen 2011. Complexity and interaction: comparing the development of L1 and L2. – *Eesti ja soome-ugri keeleteaduse ajakiri (ESUKA) / Journal of Estonian and Finno-Ugric Linguistics (JEFUL)*, 2 (2), 215–236.

- Šuster, Simon 2011. Measuring lexical complexity of the language production of the World-of-Warcraft gaming community. http://www.let.rug.nl/suster/unpublished/wow_complexity.pdf. Ms.
- Tomasello, Michael 2000. Do your children have adult syntactic competence? – *Cognition*, 74, 209–253.
- Tomasello, Michael 2003. *Constructing a Language: A Usage-Based Account of Language Acquisition*. Cambridge: Harvard University Press.
- Tommerdahl, Jodi; Christel de Bruijn 2008. Is One Language Sample Enough in Morphosyntactic Analysis? https://www.reading.ac.uk/web/FILES/cls/CLS_Tommerdahl,deBruijn.pdf (27.07.2014).
- Vider, Kadri 1995. 2–3-aastaste eesti laste sõnavara. Diplomitöö. Käsikiri TÜ eesti ja üldkeeleteaduse instituudis. Tartu: Tartu Ülikool.
- Williamson, Graham 2009. Mean Length of Utterance. [www.sltinfo.com http://www.sltinfo.com/wp-content/uploads/2014/01/mean-length-of-utterance.pdf](http://www.sltinfo.com/wp-content/uploads/2014/01/mean-length-of-utterance.pdf) (27.07.2014).
- Xanthos, Aris; Sabine Laaha, Steven Gillis, Ursula Stephany, Ayhan Aksu-Koç, Anastasia Christofidou, Natalia Gagarina, Gordana Hrzica, F. Nihan Ketrez, Marianne Kilani-Schoch, Katharina Korceky-Kröll, Melita Kovačević, Klaus Laalo, Marijan Palmović, Barbara Pfeiler, Maria D. Voeikova, Wolfgang U. Dressler 2011. On the role of morphological richness in the early development of noun and verb inflection. – *First Language*, 31 (4), 461–479. doi: 10.1177/0142723711409976

How to measure the development of syntactic complexity in child language

Airi Kapanen
Tallinn University

As language consists of layers with different degrees of complexity, there is a need to take complexity into consideration in child language research. Here, the central aim is to give a critical overview of syntactic complexity theory and the methods that are used to measure this phenomenon in the spontaneous speech of a child acquiring his/her first language.

Based on multiple viewpoints (see *Kirjandus* above), the two main methods and tools are discussed: measuring syntactic complexity in children's spontaneous speech by MLU (mean length of utterance) and IPSyn (Index of Productive Syntax). Besides critical comparison of the two, the usefulness of these tools—created for a fleective language, such as English—in studying other types of languages is examined; e.g. can they also be used with Estonian, an agglutinative type of language.

Whatever measuring tools one is planning to use, it is the researcher's own conscious knowledge of syntactic complexity and the possibilities of measuring it with certain tools that come first. In addition, the role of motherese, i.e. mother's, father's or any other caretaker's child-directed speech has to be taken into account: at the very beginning, the child does not acquire the potential complexity of the whole language, but only of a register similar to his/her caretaker's speech.

Keywords: L1 acquisition, structural complexity, absolute complexity, input language, motherese, child-directed speech