

Nuorten aikuisten kielenhallinta ja meneillään oleva kielenmuutos

Mikko Laasanen¹, Anneli Pajunen¹ & Tuomo Häikiö²

¹ Tampereen yliopisto

² Turun yliopisto

Tämä on hyväksytty käsikirjoitusversio artikkelista Laasanen, M., Pajunen, A., & Häikiö, T. (2022).

Nuorten aikuisten kielenhallinta ja meneillään oleva kielenmuutos. *Virittäjä*, 126(3).

<https://doi.org/10.23982/vir.83371>

Abstrakti

Artikkeli käsittelee kielen natiivihallintaa ja sen huomioimista sosiolingvistisessä meneillään olevan kielenmuutoksen tutkimuksessa. Artikkelissa tarkastellaan iän ja koulutuksen suhteen vertaistettujen nuorten aikuisten kielenhallintaa ja keskitytään derivaatiomorfologian hallintaan. Aineistona on yliopiston pääsykoe, jossa testattiin kielen eri osa-alueita. Pääsykokeeseen kuului derivaatiotesti, jossa piti tunnistaa mahdottomia, mahdollisia ja esiintyviä johdoksia. Kielenhallinnan arviointi perustuu pääsykokeen suoriutumistuloksiin ja ylioppilastutkinnon painotettuihin todistuspistemääriin. Tutkimuksen mukaan ylioppilaskokeen pistemäärä ennusti kokelaan suoriutumista mahdollisten johdosten tunnistuksessa tilastollisesti merkitsevästi. Laadulliseen tarkasteluun otettiin kausatiivi-, passiivi- ja teonnimijohdokset. Näiden hallinnassa havaittiin selkeää vaihtelua ja erilaisia tunnistusstrategioita. Tulosten perusteella nuorten aikuisten kielenhallinta vaihtelee, ja tällä voi olla vaikutuksia kielenmuutoksen tutkimustuloksiin. Mikäli kielenhallintaa ei tutkimuksissa kontrolloida, saatetaan yhteisölliseksi muutokseksi laskea tekijöitä, jotka todellisuudessa heijastavat kehityksellisiä tai hallinnallisia tekijöitä.

Avainsanat: sosiolingvistiikka, meneillään oleva kielenmuutos, kielen natiivihallinta

1. Johdanto

Kielenhallinta vaihtelee sekä kehityksellisesti (ks. Berman 2004; Nippold 2007) että kunkin saavutettavissa olevien kielellisten tietojen ja taitojen osalta (*ultimate attainment*, ks. Kidd, Donnelly & Christiansen 2018; Andringa & Dąbrowska 2019).¹ Kehityksellisestä näkökulmasta katsottuna kielen oppiminen on elinikäinen prosessi eivätkä kaikki puhujat opi kieltä samalla nopeudella. Yksilön näkökulmasta katsottuna kaikki eivät välttämättä saavuta samaa kielenhallinnan tasoa koskaan, vaan hallinnassa on suuria eroja (ks. Hulstijn 2011; Dąbrowska 2012). Täten voidaan sanoa, että puhujat ovat heterogeenisiä kielenoppimisen ja kielenhallinnan sekä sen kehittymisen suhteen. Selkeimmin erot näkyvät perussanaston ja -rakenteiden eli kielen ytimen ulkopuolella kuten erilaisissa kompleksisissa sana- ja lauserakenteissa. Erot tulevat esiin niin lapsia, nuoria ja nuoria aikuisia (ks. Pajunen & Honko toim. 2021) kuin varttuneita aikuisia tutkittaessa (ks. Pajunen & Itkonen 2022). Tässä artikkelissa analysoimme, mitä nuorten aikuisten sanojen tunnistamistesti paljastaa intuitiivisesta sanatiedosta, ja pohdimme, mitä tulokset merkitsevät sosiolingvistisessä kielenmuutoksen tutkimuksessa, jossa hallinta arvioidaan ikäryhmäkohtaisesti samanlaiseksi eikä vaihtelevaksi.

Sosiolingvistisessä meneillään olevan kielenmuutoksen tutkimuksessa muutos luetaan vertaamalla juuri ikäryhmien käyttäytymistä toisiinsa. Kriittinen ikäryhmä muutoksen suunnan ja nopeuden tunnistamisen kannalta on ollut nuorten ryhmä. Tyypillinen nuorten ryhmä on koostunut eri-ikäisistä puhujista eikä puhujien kielenhallintaa ole kontrolloitu. Täten voidaan olettaa, että tyypillisen nuorten ryhmän puhujat sosiolingvistisessä tutkimuksessa vaihtelevat kielenkehityksen ja -hallinnan suhteen. Lisäksi tutkimukset eroavat sen suhteen, minkä ikäisiä puhujia ryhmä pitää sisällään, ts. minkä ikäiset puhujat lasketaan nuoriksi puhujiksi. Koska puhujien kielenhallintaa ei ole kontrolloitu, pidetään puhujaryhmiä implisiittisesti homogeenisinä kielenkehityksen ja -hallinnan suhteen, mitä ne eivät kuitenkaan ole. Tästä voi seurata se, että yhteisölliseksi muutokseksi luetaan tekijöitä, jotka todellisuudessa edustavat kehityksellisiä tai hallinnallisia tekijöitä.

Tämän tutkimuksen tarkoituksena on osoittaa, että jo iän ja koulutuksen suhteen vertaistettujen nuorten puhujien välillä voidaan havaita hallinnallisia eroja. Tämän perusteella oletamme, että vertaistamattomien puhujien välillä eroja voidaan havaita vielä enemmän. Natiivihallinnan testinä on rakenteeltaan kompleksisten sanojen tunnistamistesti eli derivaatiotesti, koska suomen tapaisessa

¹ Artikkelin on muokattu versio Mikko Laasanen väitöskirjan (ks. Laasanen 2021) painetussa versiossa ilmestyneestä käsikirjoituksesta. Kiitämme artikkelin anonyymeja arvioijia rakentavista kommentteista. Mikko Laasanen vastaa artikkelin teoreettisesta osuudesta ja kokonaisargumentaatiosta, Anneli Pajunen kielenhallinnan testistä ja derivaatiotestin laadullisesta tarkastelusta ja Tuomo Häikiö derivaatiotestin tilastollisesta analyysistä.

morfologisesti rikkaassa kielessä derivaation hallinta kehittyy hitaasti (Vainio, Pajunen & Häikiö 2019) ja koska johdosten tunnistaminen edellyttää tietoa kielen eri tasoilta, jotta muodostettu sana noudattaa kielen fono- ja morfotaksia ja on semanttisesti ja tyyllisesti mielekäs (Koivisto 2013: 244–251). Aloitamme esittelemällä lyhyesti kielen natiivihallinnan variaation tutkimusta, minkä jälkeen tarkastelemme sitä, miten kielen natiivihallinnan variaatio on huomioitu sosiolingvistiikassa ja sosiolingvistisessä kielenmuutoksen tutkimuksessa.

2. Kielen natiivihallinnan variaatio

Kielenhallinta perustuu kielen normatiivisuuteen (ks. esim. Itkonen 1978, 2003), mikä tarkoittaa yksinkertaistetusti muotoiltuna sitä, että jotkin kielenkäyttötavat ovat oikein ja jotkin väärin. Kielen normatiivisuus ei ole sidottu preskriptiivisesti normitettuun yleiskieleen vaan kattaa myös arkikielen, erilaiset aluepuhekielet ja murteet. Variaatio natiivien kielenhallinnassa puolestaan perustuu todennäköisesti useisiin eri tekijöihin, mm. kielelliseen älykkyyteen eli kykyyn ymmärtää ja käyttää kieltä monipuolisesti ja vivahteikkaasti² (ks. Andringa & Dąbrowska 2019: 6). Vaikka kielen natiivihallinnan variaation tutkimus onkin viime aikoina lisääntynyt merkittävästi (ks. Chipere 2003, Hulstijn 2011, 2015, 2019; Birdsong 2012; Dąbrowska 2012, 2015, 2019; Kidd, Donnelly & Christiansen 2018; Andringa & Dąbrowska 2019; Pajunen & Honko toim. 2021), ei kyse ole kuitenkaan uudesta huomiosta. Kuten Hadley ja Rispoli huomauttavat (2012: 269), jo 40 vuotta sitten Gleitman ja Gleitman (1970: 137–141) osoittivat, että puhujien kielenhallinta vaihtelee, mikä ei ole linjassa homogeenisin kompetenssin kanssa. Samanlaiseen lopputulokseen päätyy Coppieters (1987), jonka mukaan kieli ei määrää puhujan kielioppia suoraan, vaan kielen natiivipuhujat varioivat kielenhallinnan suhteen.

Ajassa voidaan kuitenkin mennä vielä enemmän taaksepäin. Whitney (1971 [1867]: 12) kirjoitti kielenoppimisen olevan elinikäinen prosessi puhujan kielitaidon kehittyessä ja laajentuessa kykyjen mukaan. Whitneyyn mukaan kyse ei ole joko–tai -tilanteesta, jossa kieli opitaan joko kokonaisuudessaan tai ei ollenkaan, vaan jokainen oppii kykyjensä mukaan. Toisaalla (1979 [1875]: 25) Whitney kirjoittaa lapsen oppivan kielen ytimen, joka mahdollistaa kielellisen kommunikaation muiden puhujien kanssa, mutta samalla korostaa, että tieto kielestä lisääntyy ja kasvaa iän myötä. Osuvasti Whitney (mp.) toteaa, että kielenoppimisen elinikäisyys on helpompi havaita toisen kielen kohdalla kuin ensimmäisen kielen eli äidinkielen kohdalla.

² Tarkkaan ottaen kielellisellä älykkyydellä tarkoitetaan yleisiin kognitiivisiin kykyihin luettavia kykyjä, joita mitataan WAIS-testien (ks. Wechsler Adult Intelligence Scale IV) tapaan kielellisillä testeillä (samankaltaisuudet, sanasto, kielellinen päättely ja ymmärtäminen, kielellinen käsitteen muodostus jne.). Testi on normeerattu myös suomeksi.

Whitneyä yksityiskohtaisemmin natiivihallinnan variaatiosta kirjoittaa Bloomfield (1970 [1927]) menomini-intiaaneja käsittelevässä kirjoituksessaan. Bloomfield (mts. 152) toteaa, että menominien kieltä ei kirjoiteta lainkaan, mutta tästä huolimatta puhujat pitävät toisten puhetta hyvänä ja toisten huonona.³ Esimerkkinä huonosta tai puutteellisesta kielestä Bloomfield (mts. 154) antaa 40-vuotiaan miehen, jonka sanasto on pieni, taivutus barbaarista ja joka muodostaa lauseita ainoastaan muutaman mallin perusteella. Toisaalta Bloomfield (mts. 153) kirjoittaa 60-vuotiaasta naisesta, jonka menomini on kaunista ja idiomaattista. Hymesin (1974 [1972]): 273–274) mukaan kyse on *differentiaalisesta kompetenssista*, eli puhujat osaavat kielensä eritasoisesti. Sittemmin samasta aiheesta on käytetty termiä *vaihteleva kompetenssi* (ks. Ellis 2012).

Dąbrowska (2012) toteaa oletuksen yhtenäisestä kompetenssista olevan myytti ja sanoo natiivipuhujien välillä olevan huomattavia eroja sen suhteen, miten hyvin he kykenevät tuottamaan ja ymmärtämään erityyppisiä kielellisiä rakenteita. Yksi syy eroille on Dąbrowskan (mts. 242) mukaan koulutus. Dąbrowska (2015) arvioi, että on riittävästi teoreettisia perusteita olettaa, että puhujien kognitiivisten kykyjen ja kielenhallinnan variaation välillä on kausaalinen linkki. Dąbrowska (2019) tarkastelee myös eri tekijöiden vaikutusta yksilöiden välisiin eroihin ja nostaa älykkyyssosamäärän tärkeäksi eroja selittäväksi tekijäksi.

Hulstijn (2011) jakaa kielellisen pätevyyden (*linguistic proficiency*) perustason kielelliseen kognitioon (*basic language cognition*) ja ylemmän tason kielelliseen kognitioon (*higher language cognition*). Perustason ilmiöt kuuluvat Hulstijnin mukaan kaikkien puhujien kielelliseen pätevyyteen (vrt. Whitney yllä). Näihin kuuluvat keskeinen leksikko ja yleiset kieliopilliset rakenteet. Ylemmän tason ilmiöt ovat perustason ilmiöiden jatke ja siihen kuuluvat harvinaisempi leksikko ja harvinaisemmin käytössä olevat kieliopilliset rakenteet. Hulstijnin (mts. 230–231) mukaan ylemmän tason ilmaukset ovat leksikaalisesti ja kieliopillisesti monimutkaisempia kuin perustason ilmaukset, ja juuri näiden kohdalla natiivipuhujien hallinnassa on variaatiota. Natiivipuhujien välinen variaatio kielenhallinnassa saa tukea myöhemmällä iällä tapahtuvan kielenoppimisen tutkimuksista (ks. Berman 2004 toim., Ravid 2005; Nippold 2007; Nippold, Frantz-Kaspar & Vigeland 2017). Keskeisenä ajatuksena on, että lapsi oppii kielen ytimen noin viidenteen ikävuoteen mennessä, minkä jälkeen lapset alkavat hiljalleen ymmärtää ja tuottaa monimutkaisempia rakenteita, esimerkiksi passiivia (ks. Nippold 2004; Ravid 2004), kehityksen jatkuessa aikuisuuteen saakka. Tolchinsky (2004: 233–234) kuitenkin huomauttaa, että ikävuotta 5 ei tulisi pitää vedenjakajana yksinkertaisten ja kompleksisten rakenteiden välillä, koska

³ Tämä on hyvä osoitus siitä, että kielen normatiivisuus ei ole sidottu kirjoitusjärjestelmään (ks. ja vrt. Harris 1980: 6–7).

kompleksisuudeltaan samantasoisten mutta funktioltaan erilaisten rakenteiden oppimisessa saattaa olla vuosien ero.

Ehkä selkeimmin myöhemmällä iällä tapahtuva oppiminen näkyy leksikon koossa. Esimerkiksi Nippoldin (2007: 361–362) mukaan 10-vuotias osaa vähintään 20 000 eri sanaa, 15-vuotias 30 000 ja 25-vuotias vähintään 50 000 eri sanaa; luvut tietysti voivat kielittäin vaihdella. Leksikko ei kuitenkaan kasva tyhjiössä vaan rinnakkain syntaktisen kompleksisuuden kanssa, mikä näkyy mm. derivaatiomorfologiassa (Ravid 2004, 2005). Kirjoitettu kieli ja kirjallisuus liittyvät tiiviisti yhteen myöhemmällä iällä tapahtuvan kielenoppimisen kanssa (ks. Ravid & Tolchinsky 2002; Nippold 2004; Tolchinsky 2004, Berman 2016; Salas, Llauradó, Castillo & Taulé 2016). Koska koulutus korostaa kirjoitettua kieltä ja lukutaitoa, on koulutuksella suuri rooli myöhemmällä iällä tapahtuvaan kielenoppimiseen (Nippold 2004: 5). Tästä ei kuitenkaan seuraa, että koulutetumpien kielen hallinta olisi aina parempaa kuin vähemmän koulutettujen, sillä kielellinen tietoisuus ei edellytä koulutusta ja kielenhallinta vaihtelee sekä enemmän että vähemmän koulutetuilla. Koulutus pikemminkin tasaa kuin lisää hallinnan variaatiota (ks. Hadley & Rispoli 2012). Lisäksi on muistettava, että kielen kompleksisuus ei ole sidottu kirjoitettuun kieleen, vaan puhuttu kieli voi olla yhtä kompleksista tai jopa kompleksisempaakin kuin kirjoitettu kieli (ks. Halliday 1985: 76–91).

3. Kielen natiivihallinnan variaatio sosiolingvistisessä meneillään olevan kielenmuutoksen tutkimuksessa ja sosiolingvistiikassa

Tässä luvussa tarkastellaan sitä, miten natiivihallinnan variaatio on huomioitu yhtäältä sosiolingvistisessä meneillään olevan kielenmuutoksen tutkimuksessa ja toisaalta sosiolingvistiikassa yleisemmin. Meneillään olevan kielenmuutoksen tutkimuksen osalta todistusaineisto on kauttaaltaan negatiivista, mikä tarkoittaa sitä, että yhdessäkään tietämässämme meneillään olevan kielenmuutoksen tutkimuksessa puhujien kielenhallintaa ei kontrolloida. Tästä seuraa se, että tutkimuksissa tehdään järjestäen joko tiedostetusti tai tiedostamattomasti eräänlainen homogeenisuusolettama, jonka mukaan kaikki puhujat hallitsevat kielensä yhtä hyvin. Ongelma on siinä, että tämä oletama on ristiriidassa uusimman tutkimustiedon kanssa, jonka mukaan kaikki puhujat eivät hallitse kieltään yhtä hyvin. Sivuuttamalla tämä tosiasia sivuutetaan samalla yksi kielenkäyttöön vaikuttava tekijä, ja seurauksena on, että eri vaihtelutyypit saattavat sekaantua. Ongelma korostuu nuorten ryhmässä, joka on usein kriittinen muutoksen suunnan ja nopeuden tunnistamisessa, sillä nuorilla kielenkehitys saattaa olla eri vaiheessa (ks. Pajunen, Itkonen & Vainio 2015). Nuorten ryhmä voi lisäksi käsittää hyvinkin vaihtelevan ikäisiä puhujia eikä ole olemassa yhtenäisiä kriteerejä sille, minkä ikäiset puhujat tulisi laskea nuoriksi puhujiksi. Tästä

seuraa se, että tyypillinen nuorten ryhmä koostuu kielenkehityksen ja -hallinnan osalta heterogeenisestä joukosta puhujia. Homogeenisuusolettaman mukaan näin ei kuitenkaan olisi. Meneillään olevan kielenmuutoksen tutkimus on yksi sosiolingvistiikan tutkimuskohteista. Voidaan päätellä, että muutoksen tutkimuksessa seurataan enimmäkseen sosiolingvistiikassa vallalla olevia teoreettisia viitekehyksiä ja taustaoletuksia. Koska yksilövariaatio on ollut yksi sosiolingvistiikan keskeisistä tutkimuskohteista jo pidemmän aikaa (ks. esim. Tagliamonte & D’Arcy 2009), voisi olettaa, että myös kielenhallintaa olisi tarkasteltu yhtenä yksilövariaatioon vaikuttavana tekijänä. Näin ei kuitenkaan näytä olevan. On vaikeaa, ellei jopa mahdotonta, löytää kirjoituksia, joissa aihetta käsiteltäisiin yksityiskohtaisesti eli tyyliin puhuja x hallitsee jotkut tietyt kielen rakenteet paremmin kuin puhuja y, ja tämä näkyy tai voi näkyä heidän kielenkäytössään. Toisaalta juuri tämäntyyppisiä tulkintoja tehtiin 1970-luvulla Bernsteinin kooditeoriasta, joka muistuttaa Hulstijnin (2011, 2015) kaksitasoista näkemystä kielellisestä pätevyydestä. Kooditeoriassaan Bernstein (1971: 76–94) tekee eron huolitellun koodin (*elaborated code*) ja rajoitetun koodin (*restricted code*) välille. Huoliteltu koodi on Bernsteinin mukaan eksplisiittinen merkityksen suhteen, vaatii verbaalista suunnittelua eikä välttämättä vaadi puhujalta kontekstisidonnaista tietoa. Rajoitetussa koodissa merkitys on usein implisiittistä ja perustuu yhteisesti jaettuun taustatietoon, jolloin koodi ei vaadi verbaalista suunnittelua, sanasto on rajoitetumpaa ja ei-verbaalinen sisältö tärkeää. Koodien välinen ero liittyy siihen, missä määrin puheakti ja sen konteksti on jaettua tietoa puhujien kesken (ks. Danzig 1995: 150). Bernsteinin rajoitettu koodi liitettiin huonosti koulussa menestyvien puheeseen ja näin Bernstein nähtiin vajateorian⁴ kannattajana (ks. Danzig 1995: 152; Halliday 1995: 133; Kerswill 2007: 58). Bernsteinin kooditeoria ei kuitenkaan ollut alkujaan kompetenssin teoria vaan performanssin (tai *parolen*) (ks. Bernstein 1971: 173–174; Halliday 1995: 131). Kyse oli erilaisesta tavasta käyttää kieltä, ei erilaisesta kyvystä käyttää kieltä. Kuten Labov (1970: 4) toteaa, Bernsteinin näkemys tulkittiin tietystä näkökulmasta, jolloin keskiluokan käyttämä kieli esiintyi parempana kuin alempiluokkaisten kieli. Bernstein (1971: 173) toteaa kuitenkin selkeästi, että ei ole perusteita olettaa, että yksi koodi olisi toista parempi.

Bernsteinin kooditeorian saama vastaanotto osoittaa, että näkemyksiin, joissa puhutaan tai ainakin näytetään puhuttavan kielellisestä puutteesta, on sosiolingvistiikassa saatettu suhtautua hyvinkin kriittisesti. Esimerkiksi Hudsonin sosiolingvistiikan oppikirjassa vuodelta 1980 kirjoitetaan seuraavasti:

The notion of 'linguistic incompetence' concerns the *lack* of the kind of knowledge that is covered by Chomsky's 'competence'. Clearly such a lack is a reality in babies, and in others who for one

⁴ Vajateorian (*deficit theory*) mukaan koulussa huonosti menestyvät oppilaat ovat kasvaneet ympäristössä, joka ei tarjoa riittävästi älyllistä ärsykettä tai vanhemmilta tulevaa oppimista (ks. Hurn 1985 [1978]: 160–162).

reason or another do not speak some particular language: with respect to that language, they are incompetent. Moreover, someone who is only part of the way towards learning some language as a second one (or towards forgetting a language which was his first language) is to that extent incompetent in the language concerned. Controversy enters when it is claimed that some children of school age (or even adults) are incompetent in their first language, compared to with others of the same age. This claim has been made with reference to children from lower-class homes, and is known as DEFICIT THEORY. (Hudson 1980: 214.)

Hudsonin näkemyksestä tulee esiin kaksi seikkaa. Ensin ajatus kielellisestä inkompetenssista nähdään kiistanalaisena ja liitetään vajeteoriaan. Lisäksi inkompetentti puhuja on vauva tai toisen kielen oppija, ei natiivipuhuja. Toiseksi kompetenssi nähdään binaarina käsitteenä: joko puhuja on kompetentti kielen x suhteen tai hän on inkompetentti. Täten kompetentin puhujan kielellinen järjestelmä on täysin kehittynyt, kuten Kerswill (1996: 191) asian ilmaisisi. Tämä tarkoittaa, että puhujat eivät voi olla eriarvoisesti kompetentteja eli mallissa ei ole tilaa differentiaaliselle tai vaihtelevalle kompetenssille (ks. Hymes (1974 [1972]; Ellis 2012). Kielenhallinta ei siis olisi väestötasolla normaali jakautunutta, kuten kognitiiviset kyvyt yleensä ovat (ks. Psykologiliitto). Toisinaan näkee esitettävän, että sosiolingvistiikassa kielenhallintaa on tarkasteltu astetta abstraktimmalla tasolla eli koodien, rekisterien tai tyylin tasolla (ks. esim. Labov 1969: 13–16; Wardhaugh 1998 [1986]: 86–115; Chambers 1995: 4–7; Bell 1997; Tagliamonte 2012: 34–35). Ehkä tunnetuimpana esimerkkinä tästä ovat Labovin varhaiset tutkimukset ei-standardienglannista (esim. 1969, 1970; ks. myös Bernstein 1971). Kuten Maybin (2007: 160) toteaa, pyrkimyksenä on kuitenkin ollut osoittaa, että kyse on erilaisuudesta (*difference*) eli erilaisesta tavasta käyttää kieltä eikä puutteesta (*deficit*) eli puutteellisesta tavasta käyttää kieltä. Vaikka hallinta on mukana jollain tasolla, se on asetettu taka-alalle ja sitä pidetään tietyssä mielessä vähemmän tärkeänä (esim. koodinvaihto puhekielestä kirjakielen vaatii kirjakielen normien hallintaa, ks. Paunonen 1995: 176). Halliday (1995: 140) menee niinkin pitkälle, että hän kutsuu puutteen ja erilaisuuden välistä eroa merkityksettömäksi, sillä mikä tahansa kielellinen ero voidaan hänen mukaansa tulkita kummaksikin. Tämä kuitenkin ohittaa kielen normatiivisuuden eli sen, että jotkin kielenkäyttötavat ovat oikein ja jotkin väärin (ks. Mäkilähde, Leppänen & Itkonen toim. 2019). Kielen natiivihallinnan huomioiminen rekisteritasolla ei myöskään yksinään riitä, sillä muutoksen tutkimuksessa ei tarkastella rekisterien muutosta vaan yksittäisten kielenpiirteiden muutosta. Sosiolingvistisesta kirjallisuudesta on myös vaikea löytää kommentteja siitä, että nuorten kielenkehitys voi olla vielä kesken. Tämä ei ole kovinkaan yllättävää, sillä sosiolingvistiikassa oli pitkään vallalla käsitys, jonka mukaan yksilön kielellinen järjestelmä muotoutuu lapsuudessa ja tämän jälkeen säilyy lähes muuttumattomana, ja vasta reaaliaikainen tutkimus osoitti näkemyksen

vaativan korjaamista tai muokkaamista (ks. Labov 1994: 111–112; Chambers 1995: 184–185; Sankoff, Blondeau & Charity 2001; Kurki 2005; Palander 2005; Sankoff 2005; Sankoff & Blondeau 2007; Mustanoja 2011). Reaaliaikaisen tutkimuksen myötä on alettu puhua elinikäisestä muutoksesta, jonka mukaan osa aikuisista puhujista Kurjen (2013: 210) sanoin ”tempautuu mukaan muutokseen”. Muutoksen vetureina puolestaan toimivat nuoret (ks. esim. Tagliamonte & D’Arcy 2009). Hieman yksinkertaistaen voidaan todeta, että vallitsevan käsityksen mukaan lapsi oppii kielensä, minkä jälkeen seuraa ajanjakso, jolloin lapsen (tai nuoren) kieli alkaa muuttua (*vernacular re-organization*, ks. Labov 2001: 415) ja jolloin hän samalla edistää yhteisössä meneillään olevia kielenmuutoksia (*incrementation*, ks. Labov 2001: 446–465):

[...] it is incrementation, driven by social forces, and actively engaged in by individual speakers **once acquisition is more or less complete**, that is largely responsible for the advancement of innovative linguistic forms [...] (Tagliamonte & D’Arcy 2009: 99; lih. lis.).

Kielenoppimisen (tai omaksumisen) ajatellaan sijoittuvan ensi sijassa lapsuuteen. Nuorilla lapsuuden jälkeisiin kehityksellisiin tekijöihin voidaan kyllä viitata, mutta harvemmin niistä puhutaan yksityiskohtaisemmin (ks. esim. Labov 2001: 415).

Ääntenmuutoksia käsitellessään Labov (1994: 47) on todennut, että 2-vuotiaat eivät anna luotettavaa kuvaa muutoksesta, sillä heidän kielenkehityksensä on vielä kesken. Pari sivua myöhemmin hän toteaa, että tutkimuksissa tulisi ottaa huomioon jopa 8-vuotiaat (mts. 49). Toisin sanoen 8-vuotiaiden kielenkehitys ei Labovin (1994) mukaan ole enää siinä määrin kesken, että se vaikuttaisi tuloksiin. Vastaavasti Kerswill (1996: 191) kirjoittaa useimpien kielen tasojen olevan täysin kehittyneitä ikävuoteen 12 mennessä pois lukien tyylien hallinta ja mahdollisuus pieniin muutoksiin à la King (1969). Luvut ovat ristiriidassa nykykäsityksen kanssa. Niitä voikin pitää vanhentuneina ja esittää, että lapsuuden jälkeinen labiili vaihe kattaa sekä sosiaalisista että kehityksellisistä tekijöistä johtuvia muutoksia, jotka suurimmaksi osaksi päättyvät kielellisen järjestelmän vakiintumisvaiheeseen. Vielä tämäkin on ongelmallista, sillä jos kehityksellisiä tekijöitä ei eroteta sosiaalisista tekijöistä, saattaa tuloksena olla tilanne, jossa yhteisölliseksi muutokseksi luetaan tilanteita, jotka todellisuudessa heijastavat kehityksellisiä tekijöitä.

4. Nuorten aikuisten kielenhallinnan variaatio

Aiemman tutkimuksen perusteella voi todeta, että kielen hallinta vaihtelee kieltä äidinkielenä puhuvilla samoin kuin kieltä toisena tai vieraana kielenä puhuvilla. Hallinnan erot ovat sekä kehityksellisiä että ikäryhmäkohtaisia (suomen kielestä, ks. Pajunen & Honko toim. 2021). Ikäryhmätasolla voidaan erottaa kielenhallinnan keskimääräinen taso ja sen ääripäät, on sitten kyse kouluikäisistä tai aikuisista. Hallinnan erot ovat ehkä suurimmat – tai näkyvät parhaiten –

sanastossa (Keuleers, Stevens, Mandela & Brysbaert 2015) mutta tulevat esiin myös sana-, lause- ja tekstirakenteiden tasolla.

Koska puhujat varioivat kielenhallinnan suhteen, tulee sekä kehityksellinen että ikäryhmäkohtainen kielenhallinta kontrolloida, jotta mahdollinen muutos näyttäytyy luotettavasti. Otamme lähemmin tarkasteltavaksi sosiolingvistisessä tutkimuksessa tärkeän ikäryhmän eli nuoret aikuiset – muutoksen airuet – ja osoitamme, miten koehenkilöryhmän kielenhallinta vaihtelee, ja etenkin, miten paljon se vaihtelee sosiolingvistisen argumentaation kannalta tärkeillä kielen osa-alueilla eli fonologiassa ja morfologiassa.

4.1. Pääsykoe ja kielenhallinta

Tarkastelemme mediaani-ikältään 20-vuotiaita ylioppilastutkinnon suorittaneita nuoria aikuisia. Ylioppilastutkinto rajaa kielenhallinnan variaatiota, koska jo koulutuksen tehtävä on tasata osaamiseroja (Hadley & Rispoli 2012); ylioppilastutkinnon suorittaa ikäluokasta n. 50 % (Kivinen & Hedman 2017). Suomessa kielitaitoa testataan lisäksi monissa kielissä, mikä kasvattaa annettavaa kielikoulutusta ja todennäköisesti myös kielitietoisuutta. Riittävän isossa koehenkilöjoukossa jakoa ei välttämättä tarvitse tehdä sukupuolen mukaan sukupuolten samanlaisuusoletuksen perusteella (*gender similarity hypothesis*), koska kielellinen älykkyys – kielenhallinnan keskeinen taustatekijä – ei ole sukupuoliriippuvaista vaan normaalijakautunutta (Hyde & Linn 1988; Hyde 2005). Normaalijakauman mukaan kustakin ikäluokasta n. 70 % sijoittuisi kielenhallinnassa keskimääräisen osaamisen alueelle ja n. 15 % sen ala- tai yläpuolelle, muutamilla kielenhallinta on ongelmallista tai poikkeuksellisen hyvää (ks. Psykologiliitto). Kielenhallinnan kehitys on kuitenkin elinikäistä ja ylioppilastutkinnonkaan suorittaneet eivät ole kaikki samassa kehitysvaiheessa.⁵ Muutoksen osalta sosiolingvistinen tutkimus osoittaa, että usein nuoret naiset ovat muutoksen kärjessä (ks. Labov 2001: 279–293), joten otamme varmuuden vuoksi myös sukupuolen analyysiin mukaan.

Tarkastelemme lähemmin 225 ylioppilastutkinnon suorittaneen koehenkilön ryhmää ja arvioimme ryhmän kielenhallintaa. Koehenkilöt osallistuivat yliopiston pääsykokeeseen vuonna 2013 ja heistä on käytettävissä monenlaista taustatietoa pääsykokeessa saadun tiedon lisäksi. Koehenkilöt ovat pääasiassa syntyneet 1990-luvun alkuvaiheissa ja 83,3 % heistä oli naisia. Pääsykoe koostui erityyppisistä tehtävistä. Perustamme kielenhallinnan arvioinnin näiden tehtävien suoritustuloksiin ylioppilastutkinnon painotetun todistuspistemäärän lisäksi. Todistuspisteet vaihtelivat välillä 17–56,5 ja pistemedianaani pääsykokeessa oli 35 (ks. Pajunen ym. 2015).

⁵ Pajunen ym. 2015 ja Pajunen & Honko toim. 2021 arvioivat osin samoja nuoria aikuisia ja osoittavat sanaston, merkityksen ja syntaktisen rakenteen hallinnan eroja heidän välillään.

Koehenkilöiden äidinkielen yo-arvosana oli 89 %:lla vähintään magna cum laude approbatur, yleisin arvosana oli eximia (39 %). Äidinkielen arvosanajakauman perusteella voi olettaa, että koehenkilöiden suomen kielen hallinta on hyvä ja melko tasalaatuinen, yltäähän koko ikäluokasta arvosanaan eximia vain alle 20 % (ks. Leino 2002; arviointikriteereistä Eronen 2002).

Pääsykokeen tehtävät jakautuivat kahteen osioon. Ensimmäinen eli kielioppiosio koostui kielen eri osa-alueisiin (morfologia, semantiikka ja syntaksi sekä oikeakielisyys) liittyvistä tehtävistä, joissa oli muodostus-, tunnistus- ja arviointitehtäviä sanan, muodon, lauseen tai pientekstin kontekstissa. Toiseen osioon sisältyi sana-assosiaatiotesti (tarkemmin ks. Pajunen ym. 2015), sanan tunnistamista edellyttävä derivaatiotesti sekä kirjoitelma aiheesta Unelmieni päivä (tarkemmin Pajunen & Honko toim. 2021). Ensimmäisen osion maksimipisteet olivat 55 (mediaani pääsykokeessa 30) ja toisen 90 (mediaani 63,6). Arviointikriteerit kirjattiin etukäteen mallivastauksina tai asteikollisesti mutta kirjoitustehtävä arvioitiin holistisesti.

Pääsykoe on jaoteltu tähän yhteyteen seuraavasti: 1) kielioppi, 2) sanasemantiikka, 3) derivaatiomorfologia ja 4) kirjoitelma. Kieliopin osalta tukeudumme pääsykokeen pisteisiin mutta muiden osalta arviointikriteereitä on täsmennetty (ks. Pajunen & Honko toim. 2021).

Sanasemantiikkaa arvioitiin sana-assosiaatiotestillä, jossa koehenkilön tehtävänä oli valita 60 testisanalle kahdeksasta vaihtoehdosta kunkin sanan lähimmät neljä synonyymia. Derivaatiotestissä piti tunnistaa joukosta *suomen kielen sanaehdokkaita* kielessä oikeasti esiintyvät.⁶ Sanan tunnistaminen (*recognition*) perustuu esimerkiksi käsitteelliseen tuttuuteen mutta koehenkilöt voivat tehdä joko–tai-arvauksia tai nojata apukriteereihin. Merkityksen lisäksi tunnistamiseen vaikuttavia kriteereitä voivat olla sanan yleisyys ja muoto (ks. Balota, Pilotti & Cortese 2001, Schmitt 2010, Wang & Yoselinas 2012). Koehenkilön tehtävänä ei siis ollut arvioida testisanan rakennetta tai derivaatiosuhteita⁷, mutta analyysimme korostaa muotokriteeriä. Kirjoittamistehtävä muistutti kouluainetta ja tulos arvioitiin tekstin kehityksellisten sana- ja rakenneindeksien avulla, jotka mittaavat toisaalta sanastoa, toisaalta sana- ja lauserakenteen kompleksisuutta. Keskitymme jatkossa erityisesti sosiolingvistisessä kontekstissa mielenkiintoisimman eli derivaatiotestin tuloksiin, koska se mittaa riittäväällä aineistolla kielenhallintaa monipuolisesti, so. sekä rakenteen että merkityksen tasolla. Suomi on synteettinen kieli, jossa suurin osa sanoista on sanarakenteeltaan

⁶ Korpusesiintymisen tarkistettiin Helsingin ja Savon Sanomien vuosien 2000 ja 2001 aineistoihin perustuvista korpuksista (Pajunen & Virtanen 2002a ja b). Esiintymiä edellytettiin vähintään yksi kummastakin korpuksesta, jotta kyse ei olisi satunnaismuodosteesta.

⁷ Tarkka tehtävänanto oli: "Ohessa on joukko testisanoja. Tehtäväsi on arvioida, onko kyseessä oikea suomen sana (todella esiintyvä), mahdollinen suomen sana (voisi olla olemassa mutta mielestäsi ei ole) tai suomessa mahdoton sana (ei ole eikä voi olla olemassa). Merkitse vastauksesi kirjoittamalla oikeaan ruutuun X." Testisanat oli listattu ruudukkoon allekkain ja joka sanan jälkeen oli kolme ruutusaraketta vastausvaihtoehdoille; vaihtoehtojen järjestys oli satunnainen. Joka vastauksesta annettiin yksi piste eli vastauksia ei arvotettu, koska tämä olisi helposti vääristänyt tuloksia mahdollisten sanojen osalta.

kompleksisia. Sanasto kasvaa joko sanoja yhdistämällä tai johtamalla. Yhdistämällä sanoja johdetaan nominaalisissa sanaluokissa, mutta johtaminen on mahdollista kaikissa suomen pääsanaluokissa. Kompleksisten sanojen hallinta on siten välttämätöntä uusien sanojen ja luetun ymmärtämiseksi.

Derivaatiotestissä oli 248 substantiivi- tai verbikantaista johdosta, jotka muodostettiin käyttämällä kaikkein yleisimpiä suomen johtimia (esim. kausatiivi-, passiivi- ja teonnimijohtimet) ja toisaalta selvästi harvinaisempia (esim. kollektiivijohdin *-sto* ja *-kko*). Kaikki käytetyt johtimet olivat kuitenkin ainakin jossain määrin produktiivisia. Testissä kantana käytettyihin johtamattomiin sanoihin kuului sekä isoihin että pieniin taivutusperheisiin kuuluvia sanoja; substantiivien ja verbien osuus oli yhtä suuri.⁸ Kaikki kantasanat ja johtimet olivat siis kielessä esiintyviä aineksia mutta niiden yhdistyvyys vaihteli. Mahdottomat johdokset rikkoivat mm. fonotaktisia ja morfofonologisia sääntöjä (esim. *kuiluhtaa*⁹), yhdistyivät vääriin vartalonmuotoihin (esim. sanan perusmuoto sanavartalon sijaan, *käsitön pro kädetön*) tai väärään sanaluokkaan (esim. substantiivi verbin sijaan, vrt. **äitminen* pro (äitinä) *toimiminen* tai *synnyttämisen*). Mukana oli myös johdoksia, joissa kannan ja johtimen merkitykset eivät olleet yhteensopivia (esim. *kourahtaa*). Tämä kriteeri on hankalin, koska se oikeastaan perustuu kuviteltavuuteen,¹⁰ mutta myös mielenkiintoisin, koska kuviteltavuus luo uusia konteksteja – ja uusia sanoja – ja sitä säätelee kyky analogiseen mallintamiseen (ks. esim. Bauer 2001).

Derivaatiotestistä 248 johdoksesta esiintyviä oli 33,1 %, mahdollisia 37,5 % ja mahdottomia 29,4 %. Koehenkilöt onnistuivat tunnistamaan melko odotuksenmukaisesti suomen esiintyvät johdokset ja mahdottomat johdokset (ks. kuvio 1).¹¹ Hajonta oli sen sijaan suurta mahdollisten johdosten tunnistuksessa. Kun esiintyvistä ja mahdottomista johdoksista tunnistettiin keskimäärin n. 70 % oikein, mahdollisista tunnistusprosentti putosi puoleen.¹²

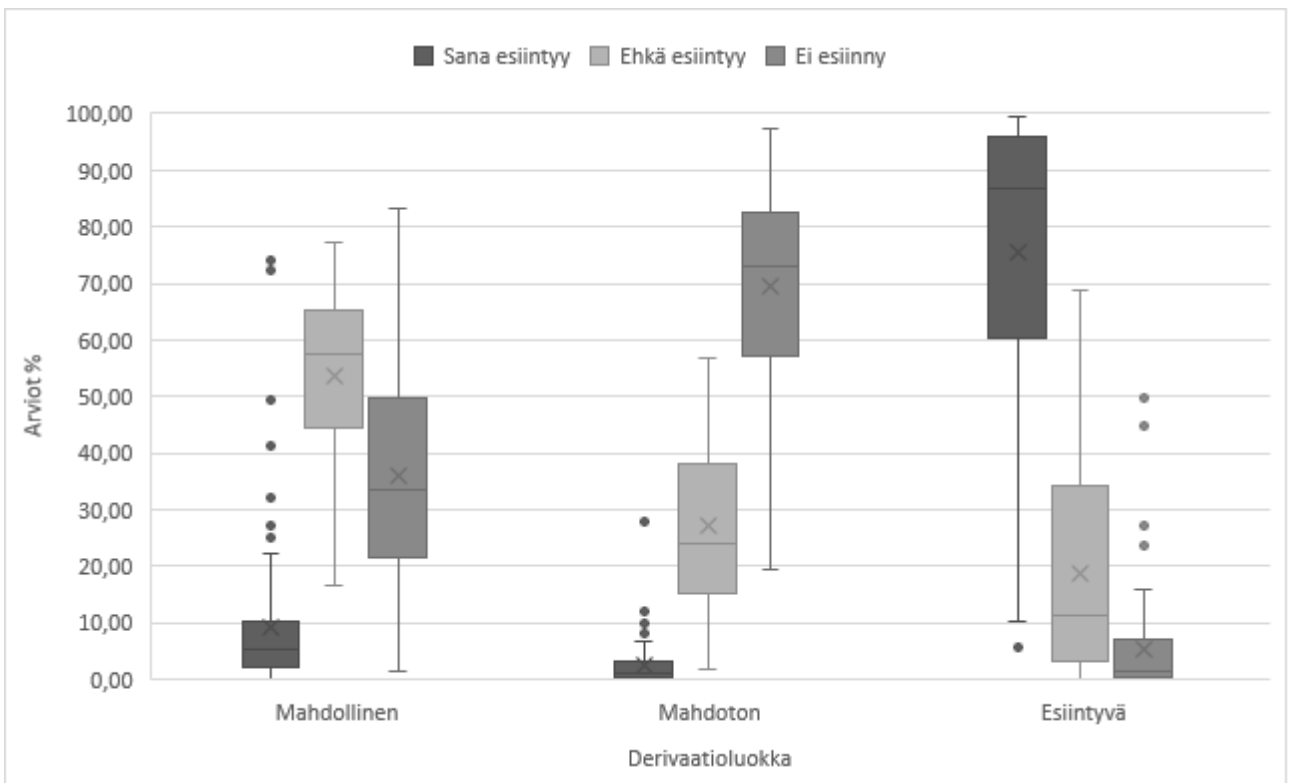
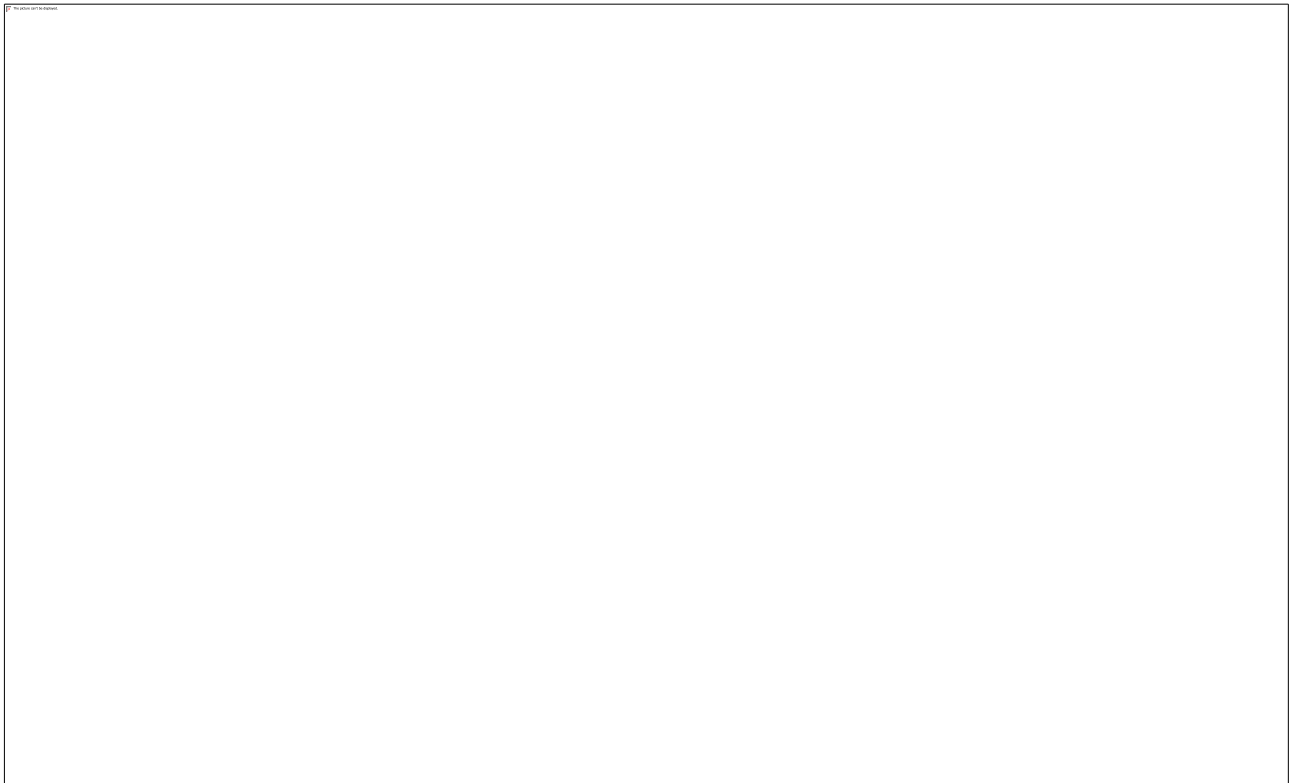
⁸ Johtimien produktiivisuus arvioitiin *Käänteissanakirjan* avulla laskemalla type/token -luku (ks. Bauer 2001), ja kantasanojen ja johtimien frekvenssit ja taivutusperheet ajettiin WordMill-ohjelmalla (Laine & Virtanen 1999).

⁹ Momentaanijohdin liittyy lähinnä *a*-vartaloon (Koivisto 2013).

¹⁰ Onkin mahdollista kuvitella, miltä näyttää käsi, joka hetkeksi kourahtaa, ja missä tilanteessa näin voisi käydä. Johdos ei silti esiinny suomessa.

¹¹ Laatikko-jana-kuviossa 1 aineisto kuvataan jakautuneena kvartiileihin: Laatikko ilmaisee keskeiset 50 % havainnoista ja sen keskellä oleva viiva mediaanin. Laatikosta lähtevät janat ilmaisevat ylä- ja alakvartiilia ja pallot satunnaisia havaintoja (ks. esim. Nummenmaa 2021).

¹² Tarkkaan ottaen keskiarvo oli esiintyvistä johdoksista 75,5 % (keskihajonta 26,2), mahdollisista 53,6 % (keskihajonta 14,6) ja mahdottomista 69,3 % (keskihajonta 17,0). Vastauslomakkeissa oli myös yksittäisiä tyhjiä ruutuja; ne on jätetty kuviosta pois.



Kuvio 1. Suomen kielessä mahdollisten, mahdottomien ja esiintyvien johdosten (n=248) hyväksyttävyyssarviot (n=55800).

Analysoimme pääsykoeaineistoa ensin tilastollisesti juuri suhteessa kykyyn tunnistaa mahdollisia johdoksia (= ns. selitettävä muuttuja eli muuttuja, johon vaikuttavia tekijöitä etsitään), koska pääsykoetulosten perusteella on oletettavaa, että juuri kyky arvioida sanan rakennetta ja

merkityksen mahdollisuutta erottelee nuorten aikuisten kielenhallintaa.¹³ Sen jälkeen analysoimme aineistoa johdostasolla eli laadullisesti.

4.2 Derivaatiotestin tilastollinen analyysi

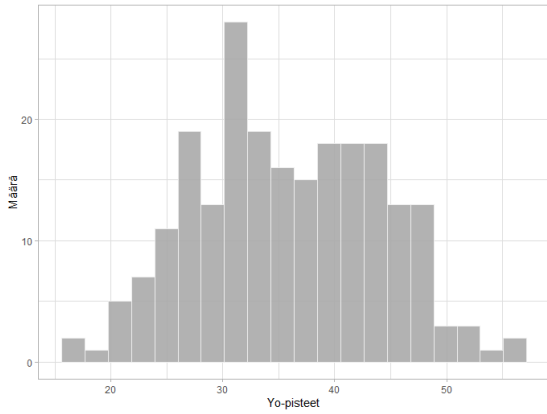
Pääsykoeaineisto analysoitiin R-tilasto-ohjelman versiolla 4.1.1 (R Core Team 2021). Aineistoa mallinnettiin yleistetyllä lineaarisella sekamallilla (glm), jonka avulla pystytään ottamaan huomioon mallissa mukana olevien tekijöiden yksilöllinen vaihtelu. Selitettävä muuttuja oli oikean vastauksen todennäköisyys derivaatiotestissä. Kriittinen selittävä, kategorinen muuttuja oli derivaatioluokka (esiintyvä, mahdollinen, mahdoton), jonka perustasona pidettiin luokkaa mahdollinen. Selittävinä, jatkuvina muuttujina olivat painotetut yo-pisteet, kielioppi (eli pääsykoe 1), sana-assosiaatiotesti (pääsykoe 2), sekä pääsykokeen osion 4 kirjoitelmaindeksit (NTVR, leksikaalinen indeksi, kompleksisuusindeksi sekä johdosprosentti, ks. esim. Wolf Nelson & Van Meter 2007; Pajunen & Honko toim. 2021)¹⁴. Kielelliset muuttujat liittyivät merkitysten hallintaan sekä sanaston ja rakenteiden käytön taitoihin. Lisäksi sukupuoli oli kategorinen, selittävä muuttuja. Kaikkien selittävien muuttujien annettiin olla interaktiossa derivaatioluokan kanssa (ks. taulukko 2). Selittävien, jatkuvien muuttujien kuvailevat tunnusluvut, keskiarvo ja havaintojen vaihtelu, ovat taulukossa 1. Niiden tarkat jakaumat on kuvattu kuvioissa 2–8.

Muuttuja	Keskiarvo	Keskihajonta	Vaihteluväli
Yo-pisteet	35,81	8,32	17,00–56,50
Kielioppi	28,59	13,26	2,00–55,00
Sana-assosiaatio	2,84	0,25	1,20–3,40
NTVR	53,57	6,08	33,80–69,80
Leksikaalinen indeksi	77,52	4,02	61,70–87,90
Kompleksisuusindeksi	1,72	0,32	1,13–3,31
Johdosprosentti	20,95	3,98	7,85–35,20

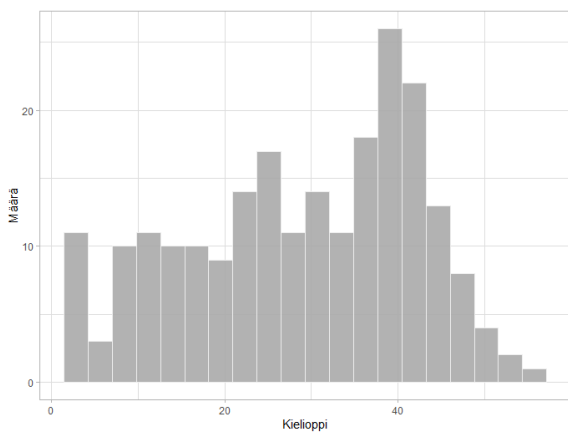
Taulukko 1. Jatkovien muuttujien arvoja kuvailevat tunnusluvut.

¹³ Pääsykokeen pisteistä korkeimmat saatiin näistä osiosta, ja ne olivat ratkaisevia opiskelijavalinnassa. Testiä laadittaessa samaa painotti myös FT, lehtori Kaija Kuiri, jolla oli pitkä kokemus suomen kielen pääsykokeen laadinnasta ja tehtävätyyppien erottelevuudesta. Kiitoksemme ei enää valitettavasti tavoita Kaijaa.

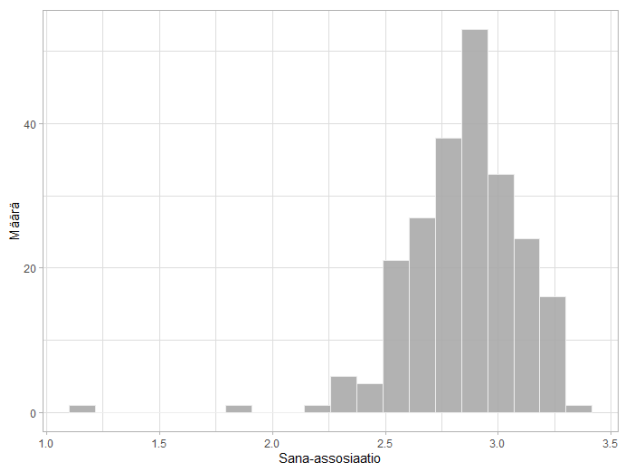
¹⁴ NTVR tarkoittaa tekstissä esiintyvien substantiivien ja verbien suhdetta, ja se ilmaisee tekstin tiiviyyttä. Leksikaalinen indeksi mittaa leksikaalisten sanojen eli substantiivien, verbien, adjektiivien ja tiettyjen adverbiin osuutta tekstin kokonaissanamäärästä ja ilmaisee tekstin informaatiopitoisuutta. Kompleksisuusindeksi kuvaa virkemäisten rakenteiden sivulauseiden ja infiniittisten muotojen määrää ja ilmaisee lauserakenteen kompleksisuutta. Johdosprosentti ilmaisee johdettujen sanojen osuuden kokonaissanamäärästä ja viittaa myös sanatason kompleksisuuteen.



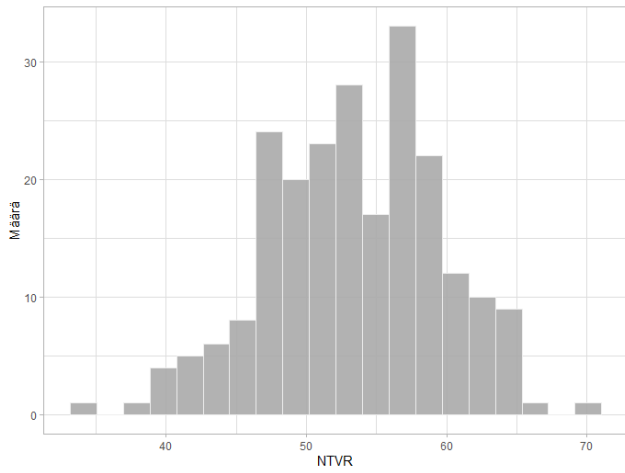
Kuvio 2. Yo-pisteiden jakauma.



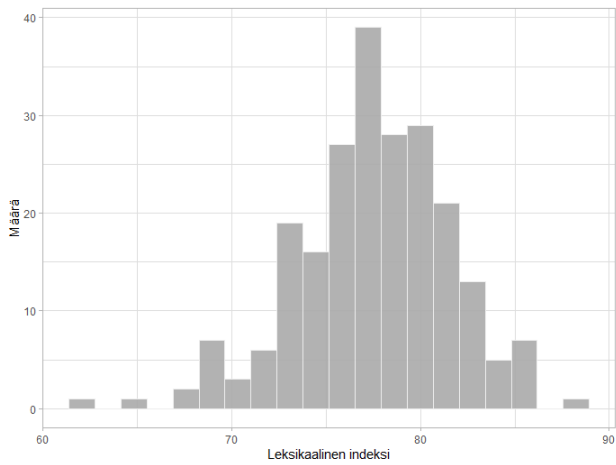
Kuvio 3. Kielioppikokeen (pääsykoe 1) tulosten jakauma.



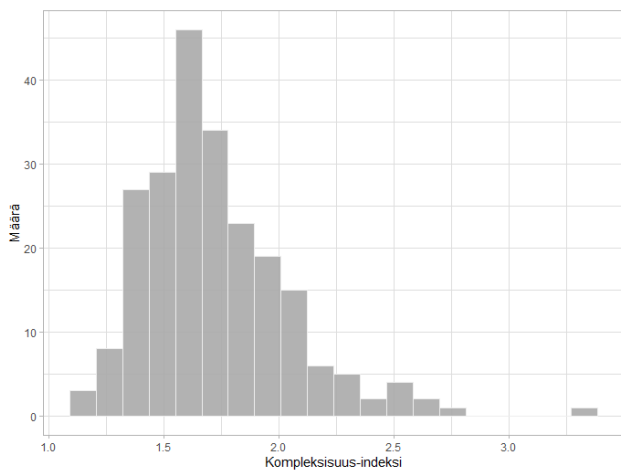
Kuvio 4. Sana-assosiaatiotestin (pääsykoe 2) pisteiden jakauma.



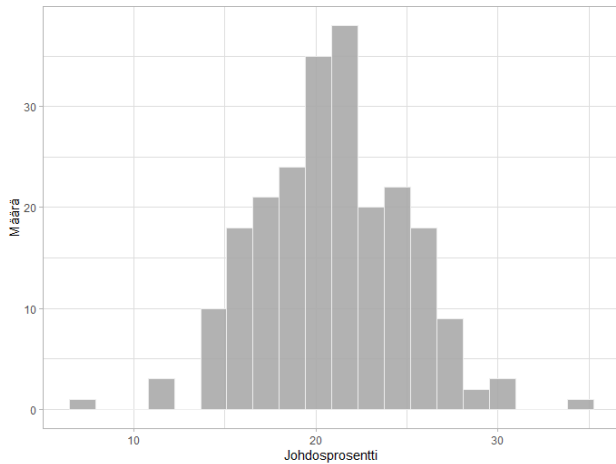
Kuvio 5. NTVR-indeksin jakauma pääsykokeessa 4.



Kuvio 6. Leksikaalisen indeksin jakauma pääsykokeessa 4.



Kuvio 7. Kompleksisuus-indeksin jakauma pääsykokeessa 4.



Kuvio 8. Johdosprosentin jakauma pääsykokeessa 8.

Kaikki malliin tulleet muuttujat muunnettiin z-skaalalle eli uudelleen skaalattiin, jotta ne saatiin vertailukelpoisiksi keskenään. Tällöin kunkin muuttujan keskiarvo on 0 ja keskihajonta 1. Malliin pyrittiin tekemään täysi satunnaisstrukturi (Barr 2013). Singulariteettiongelmiä takia jouduttiin satunnaisstruktuurista kuitenkin poistamaan useita tekijöitä. Lopullinen malli oli muotoa oikein~1+derivaatioluokka*(yopisteet.z+kielioppi.z+assosiaatio.z+ntvr.z+leksikaalisuus_index.z+kompleksisuus_index.z+johdosprosentti.z+sukupuoli)+(derivaatioluokka | kh)+(yopisteet.z + ntvr.z | item).

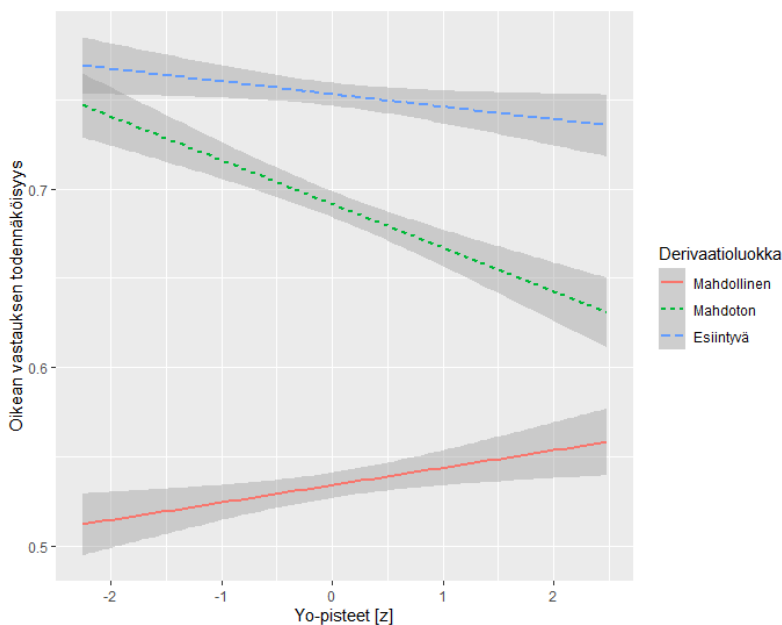
Taulukossa 2 on analyysin tulokset. Kun satunnaisvaikutukset on huomioitu, koko malli selittää aineiston vaihtelusta 54,2 %. Derivaatioluokka selitti tilastollisesti merkitsevästi oikeiden vastausten määrää. Sekä esiintyviin että mahdottomiin sanoihin vastattiin todennäköisemmin oikein kuin mahdollisiin sanoihin. Sukupuolen päävaikutus oli marginaalinen, $p = 0.062$. Miehet suoriutuivat derivaatiotestistä naisia paremmin. Yksikään yhdysvaikutus ei ollut tilastollisesti merkitsevä, mutta koska derivaatioluokan mahdoton ja yo-pisteiden välinen vaikutus oli lähellä merkitsevää, $p = 0,054$, koko yhdysvaikutus on visualisoitu kuviossa 9. Kuvioista nähdään, että mitä paremmat yo-kokeen pisteet kokelaalla oli, sitä paremmin hän suoriutui mahdollisten sanojen luokittelusta. Sen sijaan esiintyvien ja mahdottomien sanojen kohdalla suoriutuminen heikkeni yo-pisteiden funktiona. Myös yhdysvaikutus derivaatioluokan mahdoton ja sukupuolen välillä oli marginaalinen, $p = 0,090$. Se on visualisoitu kuviossa 10. Miehet pärjäsivät paremmin mahdollisten sanojen luokassa mutta heikommin mahdottomien sanojen luokassa. Luottamusvälit ovat kuitenkin vahvasti päällekkäiset, joten tämän vaikutuksen tulkinnassa kannattaa olla varovainen.

Tekijä	Oikean vastauksen todennäköisyys		
	Odds Ratio	CI	p
(Intercept)	1,11	0,80 – 1,55	0,524
derivaatioluokka [mahdoton]	3,19	1,84 – 5,55	<0,001

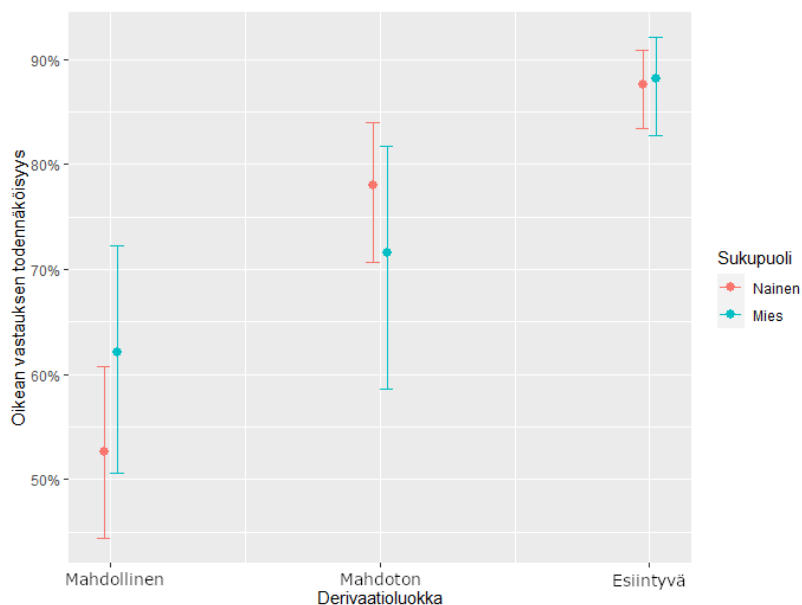
derivaatioluokka [esiintyvä]	6,38	3,99 – 10,19	<0,001
yopisteet,z	1,08	0,91 – 1,29	0,369
kielioppi,z	0,98	0,82 – 1,16	0,782
assosiaatio,z	1,04	0,88 – 1,22	0,668
ntvr,z	1,04	0,86 – 1,25	0,716
leksikaalisuus_index,z	1,11	0,94 – 1,32	0,225
kompleksisuus_index,z	1,03	0,88 – 1,20	0,731
johdosprosentti,z	0,87	0,74 – 1,03	0,118
sukupuoli [M]	1,47	0,98 – 2,20	0,062
derivaatioluokka [mahdoton] *	0,70	0,49 – 1,01	0,054
yopisteet,z			
derivaatioluokka [esiintyvä] *	0,84	0,68 – 1,04	0,102
yopisteet,z			
derivaatioluokka [mahdoton] *	1,34	0,94 – 1,92	0,102
kielioppi,z			
derivaatioluokka [esiintyvä] *	1,04	0,85 – 1,27	0,724
kielioppi,z			
derivaatioluokka [mahdoton] *	0,83	0,59 – 1,17	0,296
assosiaatio,z			
derivaatioluokka [esiintyvä] *	1,02	0,84 – 1,24	0,862
assosiaatio,z			
derivaatioluokka [mahdoton] *	0,93	0,63 – 1,38	0,727
ntvr,z			
derivaatioluokka [esiintyvä] *	0,95	0,76 – 1,19	0,660
ntvr,z			
derivaatioluokka [mahdoton] *	0,87	0,60 – 1,25	0,443
leksikaalisuus_index,z			
derivaatioluokka [esiintyvä] *	0,95	0,77 – 1,17	0,622
leksikaalisuus_index,z			
derivaatioluokka [mahdoton] *	0,92	0,66 – 1,27	0,603
kompleksisuus_index,z			
derivaatioluokka [esiintyvä] *	1,06	0,87 – 1,28	0,578
kompleksisuus_index,z			
derivaatioluokka [mahdoton] *	1,30	0,91 – 1,84	0,148
johdosprosentti,z			
derivaatioluokka [esiintyvä] *	0,97	0,79 – 1,19	0,765
johdosprosentti,z			
derivaatioluokka [mahdoton] * sukupuoli	0,48	0,21 – 1,12	0,090
[M]			
derivaatioluokka [esiintyvä] *	0,72	0,44 – 1,17	0,182
sukupuoli [M]			
Satunnaisvaikutukset			
σ^2	3,29		
τ_{00} item	2,03		
τ_{00} kh	1,13		
τ_{11} item,yopisteet,z	0,01		
τ_{11} item,ntvr,z	0,00		

τ_{11} kh,derivaatioluokkamahdoton	5,06
τ_{11} kh,derivaatioluokkaesiintyvä	1,55
ρ_{01}	-0,51
	-0,31
	-0,91
	-0,73
ICC	0,50
N_{kh}	210
N_{item}	248
Observaatioiden määrä	52079
Marginal R^2 / Conditional R^2	0,089 / 0,542

Taulukko 2. Yleistetyn lineaarisen sekamallin tulokset. Selitettävänä muuttujana on oikean vastauksen todennäköisyys derivaatiotehtävässä. Kaikki jatkuvat muuttujat on muunnettu z-skaalalle. Derivaatioluokan perustasona on mahdolliset sanat.



Kuvio 9. Oikean vastauksen todennäköisyys derivaatiotehtävässä yo-pistemäärän ja derivaatioluokan funktiona. Varjostetut alueet ovat 95% luottamusvälit.



Kuvio 10. Oikean vastauksen todennäköisyys derivaatiotehtävässä sukupuolen ja derivaatioluokan funktiona. Virhejanat ovat 95% luottamusvälit.

4.3 Derivaatiotestin kolmen yleisimmän johdinryhmän hallinta

Tilastollinen analyysi vahvisti näkemystä siitä, että juuri mahdollisten johdosten tunnistus on hankalinta. Se oli myös riippuvaista ylioppilastutkinnon pistemääristä eli todennäköisesti ainakin osaksi kognitiivisista kyvyistä. Sukupuolieroa ei tunnistuksesta löytynyt, mikä on odotuksenmukaista. Sen sijaan ei ollut täysin odotuksenmukaista, että pääsykokeen kielioppiosuus ei vaikuttanut tunnistukseen, mutta tämä voi johtua kielioppiosion tehtävien heterogeenisyydestä. Tarkastelemme aineistoa vielä laadullisesti kolmen aineistossa yleisen johdosryhmän eli kausatiivi-, passiivi- ja teonnimijohdosten avulla. Tavoitteemme on löytää lisäperusteita hyväksyttävyyssarviointien eroille. Derivaatioluokka oli mukana tilastollisessa analyysissä mutta eivät itse johtimet ja johdokset eivätkä niiden arviointierot.

Typologiselta kannalta sekä kausatiivi- että passiivijohtimet ovat kaikkein yleisimmät johdintyyppit, joilla argumenttirakenteita kielissä muutetaan; teonnimijohdin on maailman kielissä rajatumminkin esiintyvä mutta yleinen (ks. Štekauer, Valera & Körtvélyessy 2012). Suomessa kausatiivijohdoksia on verbijohdoksista eniten ja passiivijohdoksia toiseksi eniten. Käänteissanakirjan teonnimijohdoksista *minen*-johtimella muodostettuja on erittäin paljon mutta muut teonnimijohtimet ovat käytöltään rajatumpia (ks. Koivisto 2013: 245). Joka johdosryhmästä oli mukana useita johtimia ja voi hyvin olettaa, että sanojen johtaminen ainakin hyvin produktiivisilla ja funktioltaan läpinäkyvillä johtimilla hallittaisiin aikuisryhmässä hyvin (vrt. Soveri, Lehtonen & Laine 2007). Testisanoista valitut johdokset kattavat yli puolet.

Kausatiivijohtimista mukana olivat produktiivisin eli *ttA*-johdin sekä *ntA*-, *stA*-, *OidA*- ja *UttA*-johtimet. Näistä *ttA*-johtimella voidaan muodostaa verbijohdoksia sekä verbi- että substantiivikannoista (*hake+ttaa*), muista pääasiassa nominikannoista. Teonnimijohtimista valittiin *minen*, *Us:Ukse* ja *nti*, jotka kaikki liittyvät vain verbikantaan. Suomen kielen eräs typologinen erityispiirre onkin, että kannalla yleensä on sanaluokka (ks. Koivisto 2013). *minen*-johdin yhdistyy lähes mihin tahansa verbikantaan mutta muita teonnimijohtimia käytetään rajatumminkin. *U*-johdin taas on suomen vanhimpia johdinaineita (Räisänen 1988). Keskityimme merkitykseltään passiivisiin ja otamme tarkasteluun kompleksiset *itU*-, *ntU*-, *stU*-, *UtU*- ja *ttU*-johtimet, jotka osaksi ovat vapaassa vaihtelussa, osaksi valikoituneet kannan muodon tai sanaluokan perusteella. Valitut johdintyypit eroavat toisistaan myös siten, että kausatiivi- ja teonnimijohtimet ovat merkitykseltään aika kapeita ja tuottavat samantapaisten merkityksen. *U*-johdin voi painottua kantasanan merkityksen mukaan ja siitä on myös kehittymässä englannin tapaisen promootiopassiivin eli objektin subjektiksi nostavan passiivityypin leksikaalinen vastine (erityisesti *ttU*-johdin, vrt. *x ohittaa asian* vs. *asia ohittuu*). Passiivin osalta pääsykoeaineistossa painotettiin juuri muutoksenalaisuuden arviointia, joten suurin osa *U*-johdoksista kuuluu mahdollisten luokkaan. Ennako-oletuksemme on, että hyväksyttävyyssarvioissa erot näkyisivät siten, että vähiten rajattu kausatiivijohtaminen toisi parhaat tulokset ja että teonnimijohtamisessa produktiivisin ja vähiten rajattu *minen* eroaisi muista teonnimijohtimista. Kolmanneksi oletamme, että passiivijohtamisessa näkyisi toisaalta johtimien produktiivisuusero, ehkä myös *ttU*-johtimen käytön muutoksenalaisuus. Kausatiivijohdoksia oli aineistossa 37 kpl. Näistä 18 oli myös oikeasti esiintyviä sanoja, joiden hyväksyttävyyssprosentti vaihteli yli 50 %-yksikköä (keskiarvo 74,8). Jos *jätättää*, *jujuttaa* ja *astuttaa* tunnistettiin oikeiksi sanoiksi, *hakettaa*-johdoksen tunnisti esiintyväksi enää n. puolet koehenkilöistä, *hierottaa*-johdoksen vielä harvemmat. Mahdolliset johdokset tunnistettiin useimmiten mahdollisiksi (keskiarvo 63,6 %) ja yleisemmin mahdottomiksi kuin esiintyviksi (poikkeuksena *hiiskututtaa* ja *pyykittää*, joita viidennes piti esiintyvinä johdoksina). Mahdottomat tunnistettiin oikein hieman harvemmin kuin mahdolliset (keskiarvo 60,1 %). Koehenkilöistä viidenneksestä lähes puoleen tunnisti mahdollisiksi johdoksiksi sellaisetkin mahdottomat *ttA*-johdokset, joissa kantana oli perusmuoto sanavartalon sijaan (*mennättää* vs. *lampittaa*). Teonnimijohdoksia oli aineistossa yhteensä 76. Näistä *minen* on käytöltään rajaamattomin, ja johdos voidaan muodostaa lähes kaikista verbikannoista mutta ei koskaan suoraan substantiivikannasta. Teonnimen ja sen tuloksen johdin *Us:Ukse* (esimerkiksi *kuvaus*, vrt. *kuvaaminen*) on käytöltään rajatumpi, koska se on produktiivinen vain supistumaverbeistä ja monitavuisista *A*-vartaloista. Rajatuin käytöltään on *nti*-johdin, joka on johdettavissa vain yksitavuisista vartaloista sekä monitavuisista *oida*-vartaloista (esimerkiksi *juonti*, *haravointi*, ks.

Koivisto 2013: 245–246). Testisanoista 42:llä on korpusesiintymiä, 6 on mahdollista ja 28 mahdotonta. Viimeinen ryhmä rikkoo seuraavia muodostussääntöjä: kannan sanaluokka on väärä (etenkin *minen*-johdokset), kannan tavuluku ja loppuvokaali on väärä, vartalo on väärässä astevaihtelumuodossa tai kantana on perusmuoto sanavartalon sijaan. Esiintyvistä teonnimijohdoksista oli arvioitu oikein 76 %, mahdollisista 46 % ja mahdottomista 64 %. Esiintyvät teonnimijohdokset tunnistettiin pääosin hyvin mutta vaihteluväli oli valtava, esimerkiksi vain joka kymmenes tunnisti sanan *peittaus* esiintyväksi sanaksi, puolet sanan *jäävyys*. Mahdollisia teonnimijohdoksia oli vain muutama ja niistä *aarretus* oli ainoa, jota kukaan ei pitänyt mahdollisena, muissa arvioita oli laidasta laitaan. Mahdottomat oli arvioitu mahdollisiksi lähes puolessa tapauksista, jos kyse oli substantiivikantaisesta *minen*-johdoksesta kuten *saariminen* ja *aaltominen*. *Äitminen* ja *pentuminen* olivat kolmanneksen mielestä jopa esiintyviä sanoja, vaikka verbikantaa (esimerkiksi **pennuta*, *pentu-*) ei ole. Johdokset *kansaminen*, *koneminen*, *kouluminen* ja *luolaminen* tunnistettiin useimmiten mahdottomiksi sanoiksi.

Mahdottomat teonnimijohdokset oli pääosin muodostettu substantiivikannoista eli kannan sanaluokka oli väärä, mikä tuo ongelmia myös merkityskriteeriin. Sanaluokkarajaus sivuutettiin *minen*-johdosten yhteydessä lähes aina mutta yllättäen myös *nti*-johdoksissa. Aineistossa oli viisi substantiivikantaista *nti*-johdosta, jotka oli lisäksi muodostettu kaksitavuisesta kannasta: niitä piti mahdollisina n. kolmannes. Johdos **kuvanti* oli 10 %:n mielestä jopa esiintyvä johdos (ehkäpä *Avanti*-ohjelman nimen analogiana). Verbikantaisista *nti*-johdoksista 12 % arvioista oli hyväksyviä tavurajauksista ja perusmuodosta huolimatta (esimerkiksi **tietäänti*). Verbikantaisista mahdottomista *Us:Ukse*-johdoksista tunnistettiin yli 80 % oikein, loput arvioista pitivät niitä kuitenkin mahdollisina tai jopa esiintyvinä huolimatta perusmuodosta (**tuodaus*) tai muista rajausten rikkomuksista.

Kompleksiset *U*-johdokset koostuivat yhteensä 30 mahdollisesta ja mahdottomasta esimerkistä ja parista esiintyvistä *U*-johdoksesta. *U*-johtimet liittyvät sekä substantiivi- että verbikantaan, osassa vaikuttaisi semanttinen tai toisen johtimen blokki estävän johdoksen syntymistä. Aineistossa oli mukana useita *U*-johtimia, joiden produktiivisuus ja ajallinen suosituimmuus hieman vaihtelee. *UtU*-johdin on ollut hyvin produktiivinen, mutta 2000-luvun alussa (*itU*-johdin oli kompleksisista *U*-johtimista produktiivisin, (*istU* toinen ja *ttU*-johdin kolmas.¹⁵

Esiintyvät passiivijohdokset olivat harvinaisia sanoja ja ne tunnettiin huonosti. Mahdollisten osuus painottui passiivijohdoksissa *ttU*-johtimen muutoksenalaisuuden vuoksi. Kirjallisuudessa on satunnaisia löydöksiä mm. uutuustyypeistä *kirjoittua* ja *salakuljettua*. Näiden mallin mukaan testiin

¹⁵ Produktiivisuus on arvioitu laskemalla 24 milj. sanamuodon korpuksesta johdetun hapax legomenon -lekseemin suhde samalla johtimella muodostettujen lekseemien esiintymämäärään (ks. Itkonen & Pajunen 2010; Bauer 2001).

muodostettiin *ttU*-johdokset *talottua*, *?sakottua*, *?syyttyä*, *?käydätyä* ja *?poliisittua*. *Talottua* ja *sakottua* saivat yli 10 % hyväksynnän esiintyvänä sanana ja mahdollisina sanoina niitä piti pitkälti yli 60 % koehenkilöistä. *Poliisittua*-johdosta piti mahdollisena yli puolet koehenkilöistä, vaikka vanhastaan *U*-johdosta ei ole muodostettu inhimilliseen tai elolliseen viittaavista kannoista (ks. Räisänen 1988). Muissa *ttU*-johdoksissa esiintyvä-arvio oli satunnaisempi. *Käydätyä* ja *syyttyä* olivat mahdollisia yli 30 %:n mielestä.

Osa mahdollisista on tuttuja toisen *U*-johtimen yhteydestä (*kehästyä* pro *kehäytyä*, *tyrmiytyä* pro *tyrmiytyä*). Näistä (*i*)*stu*-johtimella muodostettuja hyväksyttiin esiintyviksi aika harvoin (keskimäärin 6 % arvioista) mutta mahdollisina niitä piti 40–50 % koehenkilöistä. Korkeimmat hyväksyttävyyssprosentit sai vanhastaan produktiivinen *UtU* (*varsiutua*, *lumiutua*, *toivoutua*): esiintyvä-arvioita oli kymmenisen prosenttia ja mahdollinen-arviot vaihtelivat 50–70 %:iin. Lähinnä morfofonologisesti selvästi mahdottomat tai väärään kantaan ympätyt rinnakkaistapaukset on tunnistettu mahdottomiksi (esim. **pintaitua*, **tappaattua*) mutta väärästä momentaanivartalosta muodostetun *heruhtua*-johdoksen olisi hyväksynyt mahdolliseksi yli 40 % koehenkilöistä.

Derivaatioluokka/Arviot	Arvio oikein	Keskihajonta	Vaihteluväli
Mahdollinen			
Kausatiivi	63,6	6,3	52–98
Passiivi	56,9	12,7	39–73
Teonnimi	46,1	15,8	22–99
Mahdoton			
Kausatiivi	60,1	10,6	40–74
Passiivi	78,2	16,7	40–74
Teonnimi	64,8	19,9	20–98
Esiintyvä			
Kausatiivi	74,8	22	35–98
Passiivi	-	-	-
Teonnimi	76,1	27,3	10–100

Taulukko 3. Kausatiivi-, passiivi- ja teonnimijohdosten hyväksyttävyyssarvioiden oikeellisuusprosentit derivaatioluokittain.

Johdinryhmien väliset arviointierot eivät ole esiintyvien johdosten osalta suuret (ks. taulukko 3)¹⁶ mutta selvästi suhteessa sanan yleisyyteen. Sanasto onkin nuorilla aikuisilla vielä

¹⁶ Esiintyviä passiivijohdoksia aineistossa oli vain kaksi, joten ne on jätetty taulukosta 3 huomiotta.

kehittymisvaiheessa, ja erityisesti harvinaiset sanat, joilla miljoonassa sanassa on esiintymiä korkeintaan yksi, tunnetaan vaihtelevasti (ks. Pajunen & Honko toim. 2021). Mahdollisissa johdoksissa tilanne oli toisenlainen, erityisesti teonnimien hyväksyttävyyсарviot poikkesivat muista, muiden mahdollisten arviot taas olivat paremmin linjassa aineiston kokonaiskuvaan (vrt. kuvio 1). Mahdottomista passiiveissa arviot olivat johdonmukaisimmin oikeat. Vaihteluvälit olivat kaikissa suuret mutta selittyvät pitkälti arviointistrategioista.

Jokaisessa johdinryhmässä oli mukana yleisiä, produktiivisia ja funktioltaan aika läpinäkyviä johtimia, jotka joko tiivistävät tekstiä tai muuttavat argumenttirakennetta. Ennako-oletus oli, että näissä johtaminen olisi sääntöperäistä (ks. esim. Bertram, Laine & Virkkala 2000; Leminen, Smolka, Duñabeitia & Pliatsikas 2019) ja hyvin aikuisten hallinnassa. Johtimien *ttA*, *minen* ja *UtU* korkea produktiivisuus näkyi arvioinneissa hyvin mutta enemmän hyväksymissuopeutena kuin sääntöjen johdonmukaisena soveltamisena, erityisesti kannan sanaluokkakriteeri oli helppo sivuuttaa. Ilmeisesti johtimen riittävä yleisyys tai yleistymisen ja rajoitusten löystyminen ovat edellyttämysuhteessa: hyvin produktiivinen johdin ikään kuin raivaa väyliä itselleen. Kyse on tavallaan rationaalisesta käyttäytymisestä, jolla tarvittavia sanoja voidaan tuottaa tai tiettyä muotia seurata. Oletimme myös, että *ttU*-johdosten arvioinneista näkyisi johtimen alan laajenemistendensi. Tämäkin oletus piti paikkansa. Myös oletus johtimeen liittyvien rajoitusten vaikutuksesta arviointeihin piti paikkansa.

Mahdottomat sanat muodostettiin soveltamalla kolmea periaatetta. Rikottiin johtimen sanaluokkakriteeriä, muodostettiin johdos sanan perusmuodosta sanavartalon (eli vokaalivartalon) sijaan ja kolmanneksi rikottiin morfofonologiaa ja fonotaktisia sananmuodostussääntöjä. Kuten jo todettiin, väärä kannan sanaluokka jäi koehenkilöiltä hyvin usein tiedostamatta, perusmuoto usein. Tavulukurajaukset jäivät huomaamatta. Sen sijaan morfofonologiset virheet ja eniten niiden määrä toimivat koehenkilöille hälytysmerkkeinä. Ne olivat se virhetyyppi, johon koehenkilöt useimmiten reagoivat. Tämä morfofonologinen tietoisuus osoittaa luonnollisesti, että sanojen tunnistamiseen laadittu derivaatiotestimme mittasi myös niiden kielen ilmiöiden hallintaa, jotka ovat sosiolingvistiksessä tutkimuksessa olleet keskiössä, eli fonologiaa ja morfologiaa.

4.4. Derivaatiomorfologian hallinnan variaatio ja vaikutus muutoksen tunnistamiseen

Yleensä oletetaan, että sananjohto olisi keskeisiltä osiltaan hallinnassa jo kouluikäisillä (Carlisle 2000; Larsen & Nippold, 2007: 65; Nippold & Sun 2008). Derivaatioltaan rikkaissa kielissä hallinnan kehitys jatkuu kuitenkin pitkään (ks. hepreasta Ravid & Avidor 1998). Vainio ym. (2019) testasivat suomesta derivaatiomorfologian hallintaa priming-tutkimuksessa, jossa koehenkilöt tekivät leksikaalisen päätöksen kohdesanaan vaikuttavan johdetun prime-sanan ja johtamattoman kohdesanan suhteesta siten, että johdetut sanat olivat joko kielessä esiintyviä sanoja, pseudosanoja

tai johtamattoman sanan kanssa yhteen kuulumattomia (esimerkiksi *kouraista*/**kourahtaa/kohdata* > *koura*). Koehenkilöitä oli yhteensä 300 ja ryhmä koostui 12- ja 15-vuotiaista koululaisista sekä korkeakouluopiskelijoista. Testi mittasi sekä kehityksellistä muutosta että ikäryhmätasoa ja osoitti, että derivaatiomorfologian hallinta kehittyy myös suomessa hitaasti: koululaiset eivät ole vielä aikuistasolla. Derivaatiotestiamme vaikeutettiin aikuiskontekstiin ja tavoitteena oli osoittaa ikäkausivaihtelu myös koulutuksen suhteen kontrolloidussa nuorten aikuisten ryhmässä. Derivaatiotesti osoitti, että hallinta aikuisryhmässä vaihteli ja että koehenkilöillä oli johdosten tunnistamisessa erilaisia strategioita. Esiintyvissä sanoissa se oli sanan tuttuus, mahdollisissa sanoissa painottui joko muodon moitteettomuus tai merkityksen kuviteltavuus mutta virheellisten sanojen tunnistus perustui morfofonologisiin virhekesaumiin. Osallistujista vain muutama pärjäsi erinomaisesti joka derivaatioluokan arvioinnissa, vastaavasti heikonlaisia oli muutama. Muilla oli vaihtelevia arviointistrategioita: yksi ryhmä keskittyi arvioimaan oikeita sanoja ja erottamaan ne virheellisistä kiinnittämättä mahdollisiin huomiota, toinen selvisi hyvin esiintyvistä ja mahdollisista johdoksista mutta ei tunnistanut mahdottomia. Mielenkiintoisin on ehkä ryhmä, joka arvioi mahdolliset johdot joko mahdollisiksi tai esiintyviksi, ja ryhmä, joka piti niitä mahdollisina tai mahdottomina. Edellisellä ryhmällä on kielenmukaisesta sanarakenteesta intuitio, mihin viittasi tilastoanalyysikin, jälkimmäisellä ei ole.

Derivaatiotesti osoitti selvästi, että iältään ja koulutukseltaan kontrolloidussa koehenkilöryhmässä morfologinen hallinta vaihtelee ja vaihtelua selittää kielellisten tekijöiden lisäksi ylioppilastutkinnon pistemäärä. Taustatekijöinä on luonnollisesti kognitiiviset kyvyt sekä monenlaiset kieli-intuitioon pohjaavat arviointistrategiat. Tilastoanalyysissä sanasto- ja rakennemuuttajat todennäköisesti peittyivät näiden strategioiden moninaisuuteen.

Hallinnan tarkastelun lopputulema on siis se, että jos ylioppilastutkinnon suorittaneelta nuorelta aikuiselta, joka hallitsee derivaatiota heikosti, testattaisiin vaikkapa sananjohtamiseen tai yleensä sanarakenteeseen liittyviä muutoksia, hän antaisi tukensa yleisten sanojen osalta oikeille ilmiöille mutta muuten saattaisi viitata sanaluokka- ja fonotaksirajoitusten sekä sanavartaloiden roolin heikentymiseen. Mikäli henkilö taas kuuluisi derivaatiota hyvin hallitseviin, ei hän näin todennäköisesti tekisi. Täten muutoksen tutkimuksen tulokset riippuvat siitä, millaisia puhujia ikäryhmä pitää sisällään. Heikosti osaavista koostuva ikäryhmä saattaisi viitata meneillään olevaan muutokseen, mutta hyvin osaavista koostuva ikäryhmä ei. Sekalainen ryhmä taas ennustaisi epäselvää tilannetta, jossa muutos saattaisi olla aluillaan tai jossa kyse voisi olla stabiilista vaihtelusta.

Derivaatiotaidot kehittyvät hitaasti ja edellyttävät monimutkaisen sääntöverkoston hallintaa mutta myös laajenemismahdollisuuksien ymmärtämistä eli analogisen ajattelukyvyn kehittyneisyyttä.

Tämä prosessi ei rajoitu lapsuuteen tai nuoruuteen vaan jatkuu vielä aikuisuudessa. Produktiivisten johtimien osalta käyttö laajenee ilmeisen analogisesti erilaisia malleja hyväksi käyttäen (esim. *taloja rakennetaan pellolle > pelto rakentuu täyteen taloja > ? pelto talottuu*). Kyse on analogisesta kyvystä hahmottaa ja hyödyntää kielen tarjoamia mahdollisuuksia (analogiasta ks. Itkonen 2005). Täten jotkut puhujat kykenevät soveltamaan oppimiaan malleja ja tuottamaan harvinaisempaa mutta silti korrektia sanastoa, kun taas toisilla nämä yritykset saattavat johtaa harhaan. Yleisellä tasolla tilanne ei juurikaan eroa niistä tilanteista, joissa puhujan tavoitteena on muodostaa kirjakielen mallin mukaisia sanoja, mutta kirjakielen normiston heikko hallinta johtaa virheisiin (ks. esim. Leino 1992).

5. Lopuksi

Tutkimuksemme osoittaa, että iän ja koulutuksen suhteen vertaistettujen nuorten puhujien välillä voidaan havaita hallinnallisia eroja. Tämän perusteella teemme sen oletuksen, että sosiolingvistikseen meneillään olevan kielenmuutoksen tutkimuksissa käytetyt vertaistamattomat nuorten ryhmät sisältävät vaihtelua todennäköisesti vielä enemmän. Suurin osa sosiolingvistikseen meneillään olevan kielenmuutoksen tutkimuksesta on keskittynyt kielen äänne- ja muotopiirteisiin, minkä perusteella natiivihallinnan variaation merkitys voidaan pyrkiä sivuuttamaan vetoamalla esimerkiksi johdosten epätyypillisyyteen tutkimuskohteena. Kyse ei kuitenkaan ole siitä, onko jokin kielenpiirre tyypillinen vai epätyypillinen kielenmuutoksen tutkimuskohde, vaan yksinkertaisesti siitä, vaikuttavatko kielen hallinnalliset ja kehitykselliset tekijät tarkasteltavien piirteiden käytön yleisyyteen vai eivät. Lisäksi, kuten tutkimuksemme osoittaa, sananjohtamisen hallinta vaatii tietoa kielen fonologiasta, fonotaksista ja morfologiasta eli juuri niistä kielenilmiöistä, joihin sosiolingvistinen meneillään olevan kielenmuutoksen tutkimus on keskittynyt. On kuitenkin todennäköistä, että kielenhallinnan variaation vaikutus ei ole vakio vaan sen vaikutus voi olla suurempi, pienempi tai jopa olematon riippuen tutkittavasta kielenpiirteestä. Missä kulkee raja ja milloin kielenhallinta voidaan jättää kontrolloimatta? Tähän kysymykseen voidaan vastata ainoastaan määrittelemällä kaikille normaalin kielenkehityksen läpikäyneille natiivipuhujille yhteinen kielen alue, jossa hallinnallisia eroja ei ole tai ne ovat minimissään, eli ns. kielen ydin (ks. Whitney 1979 [1875]: 25).

Lähdeluettelo

ANDRINGA, SIBLE – DĄBROWSKA, EWA 2019: Individual differences in first and second language ultimate attainment and their causes. – *Language Learning* 69 (1) s. 5–12.

<https://doi.org/10.1111/lang.12328>

- BALOTA, D. A. – PILOTTI, M. – CORTESE, M. J. 2001:. Subjective frequency estimates for 2,938 monosyllabic words. – *Memory & Cognition*, 29(4) s. 639–647.
<https://doi.org/10.3758/BF03200465>
- BARR, DALE J. 2013: Random effects structure for testing interactions in linear mixed-effects models. – *Frontiers in Psychology* 4 s. 328. <https://doi.org/10.3389/fpsyg.2013.00328>
- BAUER, LAURIE 2001: *Morphological productivity*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486210>
- BELL, ALLAN 1997: Language style as audience design. – Nikolaus Coupland & Adam Jaworski (toim.), *Sociolinguistics. A reader and coursebook* s. 240–250. Basingstoke: Macmillan.
<https://doi.org/10.1017/S004740450001037X>
- BERMAN, RUTH toim. 2004: *Language development across childhood and adolescence*. Amsterdam: Benjamins. <https://doi.org/10.1075/tilar.3>
- BERMAN, RUTH 2004: Between emergence and mastery. The long developmental route of language acquisition. – Ruth A. Berman (toim.), *Language development across childhood and adolescence* s. 9–34. Amsterdam: Benjamins. <https://doi.org/10.1075/tilar.3.05ber>
- BERMAN, RUTH 2016: Linguistic literacy and later language development. – J. Petera, Melina Aparici, Elisa Rosado & Naymé Salas (toim.), *Written and spoken language development across the lifespan. Essays in honour of Liliana Tolchinsky* s. 181–200. Cham: Springer.
https://doi.org/10.1007/978-3-319-21136-7_12
- BERNSTEIN, BASIL 1971: *Class, codes and control. Volume 1. Theoretical studies towards a sociology of language*. London: Routledge & Kegan Paul.
- BERTRAM, RAYMOND– LAINE, MATTI – VIRKKALA, MINNA 2000: The role of derivational morphology in vocabulary acquisition. Get by with a little help from my morpheme friends. – *Scandinavian Journal of Psychology* 41 s. 287–296. <https://doi.org/10.1111/1467-9450.00201>
- BIRDSONG, DAVID 2012: Three perspectives on non-uniform linguistic attainment. – *Linguistic Approaches to Bilingualism* 2 (3) s. 255–259. <https://doi.org/10.1075/lab.2.3.02bir>
- BLOOMFIELD, LEONARD 1970 [1927]: Literate and illiterate speech. – Charles F. Hockett (toim.), *A Leonard Bloomfield anthology* s. 147–156. Bloomington: Indiana University Press.
<https://doi.org/10.2307/451863>
- CARLISLE, JOANNE. F. 2000: Awareness of the structure and meaning of morphologically complex words. Impact on reading. – *Reading and Writing* 12 s. 169–190.
<https://doi.org/10.1023/A:1008131926604>
- CHAMBERS, J. K. 1995: *Sociolinguistic theory. Linguistic variation and its social significance*. Oxford: Blackwell.

- CHIPERE, NGONI 2003. Understanding complex sentences. Native speaker variation in syntactic competence. Basingstoke: Palgrave Macmillan. <https://doi.org/10.1057/9780230005884>
- COPPIETERS, RENÉ 1987: Competence differences between native and near-native speakers. – *Language* 63 (3) s. 544–573. <https://doi.org/10.2307/415005>
- DĄBROWSKA, EWA 2012: Different speakers, different grammars. Individual differences in grammatical abilities. – *Linguistic Approaches to Bilingualism* 2 s. 219–253. <https://doi.org/10.1075/lab.2.3.01dab>
- DĄBROWSKA, EWA 2015: Individual differences in grammatical knowledge. – Ewa Dąbrowska & Dagmar Divjak (toim.), *Handbook of cognitive linguistics* s. 649–667. Berlin: de Gruyter. <https://doi.org/10.1515/9783110292022-033>
- DĄBROWSKA, EWA 2019: Experience, aptitude, and individual differences in linguistic attainment. A comparison of native and nonnative speakers. – *Language Learning* 69 (1) s. 72–100. <https://doi.org/10.1111/lang.12323>
- DANZIG, ARNOLD 1995: Applications and distortions of Basil Bernstein’s code theory. – Alan R. Sadovnik (toim.), *Knowledge and pedagogy. The sociology of Basil Bernstein* s. 145–170. Norwood, N. J.: Ablex Pub. Corp.
- ELLIS, NICK 2012: Variable competence. – *Linguistic Approaches to Bilingualism* 2 (3) s. 264–268. <https://doi.org/10.1075/lab.2.3.04ell>
- ERONEN, RIITTA 2002: Ylioppilasaineen ja sen arvostelun avainsanoja. – *Kielikello* 2/2002.
- GLEITMAN, LILA R. – GLEITMAN, HENRY 1970: *Phrase and paraphrase. Some innovative uses of language*. New York: W. W. Norton & Company Inc.
- HADLEY, PAMELA A. – RISPOLI, MATTHEW 2012: A lifespan perspective on individual differences in grammatical abilities. – *Linguistic Approaches to Bilingualism* 2 (3) s. 269–272. <https://doi.org/10.1075/lab.2.3.05had>
- HALLIDAY, M. A. K. 1985: *Spoken and written language*. Victoria: Deakin University.
- HALLIDAY, M. A. K. 1995: Language and the theory of codes. – Alan R. Sadovnik (toim.), *Knowledge and pedagogy. The sociology of Basil Bernstein* s. 127–143. Norwood, N. J.: Ablex Pub. Corp.
- HARRIS, ROY 1980: *The language-Makers*. Ithaca, New York: Cornell University Press.
- HUDSON, R. A. 1980: *Sociolinguistics*. Cambridge: Cambridge University Press.
- HULSTIJN, JAN H. 2011: Language proficiency in native and nonnative speakers. An agenda for research and suggestions for second-language assessment. – *Language Assessment Quarterly* 8 (3) s. 229–249. <https://doi.org/10.1080/15434303.2011.565844>

- HULSTIJN, JAN H. 2015: *Language proficiency in native and non-native speakers. Theory and research*. Amsterdam: Benjamins. <https://doi.org/10.1075/llt.41>
- HULSTIJN, JAN H. 2019: An individual-differences framework for comparing nonnative with native speakers. Perspectives from BLC theory. – *Language Learning* 69 (1) s. 157–183. <https://doi.org/10.1111/lang.12317>
- HURN, CHRISTOPHER J. 1985 [1978]: *The limits and possibilities of schooling. An introduction to the sociology of education*. Toinen painos. Boston: Allyn and Bacon, Inc.
- HYDE, JANET 2005: The gender similarities hypothesis. – *American Psychologist* 60 s. 581–592. <https://doi.org/10.1037/0003-066X.60.6.581>
- HYDE, JANET – LINN, MARCIA 1988: Gender differences in verbal ability. A meta-analysis. – *Psychological Bulletin* 104 s. 53–69. <https://doi.org/10.1037/0033-2909.104.1.53>
- HYMES, D. H. 1974 [1972]: On communicative competence. – J. B. Pride & Janet Holmes (toim.), *Sociolinguistics: selected readings* s. 269–293. Harmondsworth: Penguin.
- ITKONEN, ESA 1978: *Grammatical theory and metascience. A critical investigation into the methodological and philosophical foundations of 'autonomous' linguistics*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.5>
- ITKONEN, ESA 2003: *What is language? A study in the philosophy of language*. Yleisen kielitieteen julkaisuja 8. Turku: Turun yliopisto.
- ITKONEN, ESA 2005: *Analogy as structure and process. Approaches in linguistics, cognitive psychology and philosophy of science*. Amsterdam: Benjamins. <https://doi.org/10.1075/hcp.14>
- ITKONEN, ESA – PAJUNEN, ANNELI 2010: *Empiirisen kielitieteen metodologia*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KERSWILL, PAUL 1996: Children, adolescents, and language change. – *Language Variation and Change* 8 s. 177–202. <https://doi.org/10.1017/S0954394500001137>
- KERSWILL, PAUL 2007: Social class. – Carmen Llamas, Louise Mullany & Peter Stockwell (toim.), *The Routledge companion to sociolinguistics* s. 51–61. London: Routledge.
- KEULEERS, EMMANUEL – STEVENS, MICHAËL – MANDERA, PAWEŁ – BRYLSBAERT, MARC 2015: Word knowledge in the crowd. Measuring vocabulary size and word prevalence in a massive online experiment. – *The Quarterly Journal of Experimental Psychology* 68 (8) s. 1665–1692. <https://doi.org/10.1080/17470218.2015.1022560>
- KIDD, EVAN – DONNELLY, SEAMUS – CHRISTIANSEN, MORTEN H. 2018: Individual differences in language acquisition and processing. – *Trends in Cognitive Sciences* 22 (2) s. 154–169. <https://doi.org/10.1016/j.tics.2017.11.006>

- KING, ROBERT D. 1969: *Historical linguistics and generative grammar*. New Jersey: Prentice-Hall Inc.
- KIVINEN, OSMO – HEDMAN, JUHA 2017: Ylioppilastutkinnon suorittaneiden ikäluokkaosuudet laskussa. blogit.utu.fi/utu/2017/02/07 [Luettu 20.8.2020.]
- KOIVISTO, VESA 2013: *Suomen sanojen rakenne*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KURKI, TOMMI 2005: *Yksilön ja ryhmän kielen reaaliaikainen muuttuminen. Kielenmuutosten seuraamisesta ja niiden tarkastelussa käytettävistä menetelmistä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KURKI, TOMMI 2013: Näennäisaikametodi ja muutosmallit. – *Sananjalka* 55 s. 202–212.
- LAASANEN, MIKKO 2021: *Meneillään oleva kielenmuutos ja kielellinen vaihtelu. Variationistisen sosiolingvistiikan teoriaa*. Tampereen yliopiston väitöskirjat 429. Tampere: Tampereen yliopisto.
- LABOV, WILLIAM 1969: *A study of Non-Standard English*. Washington, D. C.: Center for Applied Linguistics.
- LABOV, WILLIAM 1970: The logic of nonstandard English. – James E. Alatis (toim.), *Linguistics and the teaching of English to speakers of other languages or dialects. Georgetown University Round Table on Languages and Linguistics 1969* s. 1–39.
- LABOV, WILLIAM 1994: *Principles of linguistic change. Volume 1: Internal factors*. Oxford: Blackwell.
- LABOV, WILLIAM 2001: *Principles of linguistic change. Volume 2: Social factors*. Oxford: Blackwell.
- LAINEN, MATTI – VIRTANEN, PATRIK 1999: WordMill. Lexical Search Program. Turku: University of Turku.
- LARSEN, J. – NIPPOLD, MARILYN A. 2007: Morphological analysis in school-age children. Dynamic assessment of a word learning strategy. – *Language, Speech and Hearing Services in Schools* 38 s. 201–212. [https://doi.org/10.1044/0161-1461\(2007/021\)](https://doi.org/10.1044/0161-1461(2007/021))
- LEINO, PENTTI 1992: Ylioppilasaineiden kielivirheet. Mistä ne kertovat? – Anita Julin, Kaija Parko, Pirkko Tiuraniemi & Sakari Viinamäki (toim.), *Ylioppilasaineita 1992* s. 170–188. Helsinki: Suomalaisen Kirjallisuuden Seura.
- LEINO, PENTTI 2002: Ylioppilasaine – äidinkielen taitojen mittari. – *Kielikello* 2/2002.
- LEMENEN, ALINA – SMOLKA, EVA – DUÑABEITIA, JON A. – PLIATSIKAS, CHRISTOS 2019: Morphological processing in the brain. The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex* 116 s. 4–44. <https://doi.org/10.1016/j.cortex.2018.08.016>

- MAYBIN, JANET 2007: Language and education. – Carmen Llamas, Louise Mullany & Peter Stockwell (toim.), *The Routledge companion to sociolinguistics* s. 157–163. London: Routledge.
- MUSTANOJA, LIISA 2011: *Idiolekti ja sen muuttuminen. Reaaliaikatutkimus Tampereen puhekielestä*. Tampere: Tampereen yliopisto.
- MÄKILÄHDE, ALEKSI – LEPPÄNEN, VILLE – ITKONEN, ESA (toim.) 2019: *Normativity in Language and Linguistics*. Amsterdam: Benjamins. <https://doi.org/10.1075/slcs.209>
- NIPPOLD, MARILYN A. 2004: Research on later language development. International perspectives. – Ruth A. Berman (toim.), *Language development across childhood and adolescence* s. 1–8. Amsterdam: Benjamins. <https://doi.org/10.1075/tilar.3.04nip>
- NIPPOLD, MARILYN A. 2007: *Later language development. School-age children, adolescents, and young adults*. Kolmas painos. Austin, Tx.: PRO-ED.
- NIPPOLD, MARILYN A. – SUN, LEI. 2008: Knowledge of morphologically complex words. A developmental study of older children and young adolescents. – *Language, Speech, and Hearing Services in Schools* 39 s. 365–373. [https://doi.org/10.1044/0161-1461\(2008/034\)](https://doi.org/10.1044/0161-1461(2008/034))
- NIPPOLD, MARILYN A. – FRANTZ-KASPAR, MEGAN W. – VIGELAND, LAURA M. 2017: Spoken language production in young adults. Examining syntactic complexity. – *Journal of Speech, Language and Hearing Research* 60 s. 1339–1347. https://doi.org/10.1044/2016_JSLHR-L-16-0124
- NUMMENMAA, TAPIO 2021: *Tilastotieteen käsikirja*. Helsinki: Tammi.
- PALANDER, MARJATTA 2005: *Lapsuudesta keski-ikään. Seuruututkimus itäsavolaisen yksilömurteen kehityksestä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- PAJUNEN, ANNELI – VIRTANEN, PATRIK 2002a: Helsingin Sanomat 2000–2001. Morfosyntaktisesti analysoitu tietokanta, 24 miljoonaa sananmuotoa.
- PAJUNEN, ANNELI – VIRTANEN, PATRIK 2002b: Savon Sanomat 2000–2001. Morfosyntaktisesti analysoitu tietokanta, 8 miljoonaa sananmuotoa.
- PAJUNEN, ANNELI – ITKONEN, ESA – VAINIO, SEPPO 2015: Sanamerkityksen hallinta nuorilla aikuisilla. – *Virittäjä* 119 (2) s. 160–187.
- PAJUNEN, ANNELI – ITKONEN, ESA 2022. How we know the meaning of adjectives: by intuition or by context. Esitelmä 19.8.2022. The 8th conference of Scandinavian association for language and cognition. Åbo Akademi University. <https://blogs2.abo.fi/salc8/wp-content/uploads/sites/208/2022/06/Abstract-book.pdf>. Viitattu 5.8.2022.
- PAJUNEN, ANNELI – HONKO, MARI (toim.) 2021: *Suomen kielen hallinta ja kehitys. Peruskoululaiset ja nuoret aikuiset*. Helsinki: Suomalaisen Kirjallisuuden Seura.

PAUNONEN, HEIKKI 1995: Puhesuomen muuttuva omistusmuotojärjestelmä. – *Virittäjä* 99 (4) s. 162–200.

PSYKOLOGILIITTO. Kognitiivisten testien pistemäärien sanalliset kuvaukset.

https://www.psyli.fi/files/836/sanalliset_kuvaukset.pdf. Viitattu 31.7.2022.

RAVID, DORIT 2004: Derivational morphology revisited. Later lexical development in Hebrew. – Ruth A. Berman (toim.), *Language development across childhood and adolescence* s. 53–81.

Amsterdam: Benjamins. <https://doi.org/10.1075/tilar.3.07rav>

RAVID, DORIT 2005: Emergence of linguistic complexity in later language development. Evidence from expository text construction. – Hava, Bat-Zeev Shyldkrot & Dorit Ravid (toim.), *Perspectives on language and language development. Essays in the honor of Ruth A. Berman* s. 337–355. Cham: Springer. https://doi.org/10.1007/1-4020-7911-7_25

RAVID, D. – AVIDOR, AVRAHAM 1998: Acquisition of derived nominal in Hebrew. Developmental and linguistics principles. *Journal of child language* 25 s. 229–266.

<https://doi.org/10.1017/S0305000998003419>

RAVID, DORIT – TOLCHINSKY, LILIANA 2002: Developing linguistic literacy. A comprehensive model. – *Journal of Child Language* 29 s. 417–447. <https://doi.org/10.1017/S0305000902005111>

R CORE TEAM 2021: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

RÄISÄNEN, ALPO 1988: *Suomen kielen u-johtimiset verbit*. Helsinki. Suomalaisen Kirjallisuuden Seura.

SALAS, NAYMÉ – ANNA LLAURADÓ, ANNA – CASTILLO, CHRISTINA – TAULÉ, MARIONA 2016:

Linguistic correlates of text quality from childhood to adulthood. – Joan Perera, Melina Aparici, Elisa Rosado & Naymé Salas (toim.), *Written and spoken language development across the lifespan. Essays in honour of Liliana Tolchinsky*. Cham: Springer. https://doi.org/10.1007/978-3-319-21136-7_18

SANKOFF, GILLIAN 2005: Cross-sectional and longitudinal studies. – Norbert Dittmar, Klaus J.

Mattheier & Ammon Ulrich (toim.), *Sociolinguistics. An international handbook of the science of language and society* s. 1003–1013. Berlin: de Gruyter.

<https://doi.org/10.1515/9783110171488.2.7.1003>

SANKOFF, GILLIAN – BLONDEAU, HÉLÈNE – CHARITY, ANNE 2001: Individual roles in a real-time change. Montreal (r->R) 1947–1995. – Hans Van de Velde & Roeland van Hout (toim.), *'r-atics*.

Sociolinguistic, phonetic and phonological characteristics of /r/ s. 141–157. *Etudes & Travaux* n^o4. Bruxelles: Institut des Langues Vivantes et de Phonétique.

- SANKOFF, GILLIAN – BLONDEAU, HÉLÈNE 2007: Language change across the lifespan. /r/ in Montreal French. – *Language* 83 (3) s. 560–588. <https://doi.org/10.1353/LAN.2007.0106>
- SCHMITT, NORBERT 2010: *Researching vocabulary. A vocabulary research manual*. Hampshire: Palgrave and Macmillan.
- SOVERI, ANNA – LEHTONEN, MINNA – LAINE, MATTI 2007: Word frequency and morphological processing in Finnish revisited. *The Mental Lexicon* 2 (3) s. 359–385.
<https://doi.org/0.1075/ml.2.3.04sov>
- ŠTEKAUER, PAVOL – VALERA, SALVADOR. – KÖRTVÉLYESSY, LÍVIA. 2012: *Word-formation in the world's languages. A typological survey*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511895005>
- TAGLIAMONTE, SALI 2012: *Variationist sociolinguistics. Change, observation, interpretation*. Oxford: Wiley-Blackwell.
- TAGLIAMONTE, SALI – ALEXANDRA, D'ARCY 2009: Peaks beyond phonology. Adolescence, incrementation and language change. – *Language* 85 (1) s. 58–108.
<https://doi.org/10.1353/lan.0.0084>
- TOLCHINSKY, LILIANA 2004: The nature and scope of later language development. – Ruth A. Berman (toim.), *Language development across childhood and adolescence* s. 233–247. Amsterdam: Benjamins. <https://doi.org/10.1075/tilar.3.15tol>
- VAINIO, SEPO – PAJUNEN, ANNELI – HÄIKIÖ, TUOMO 2019: Acquisition of Finnish derivational morphology. School-age children and young adults. – *First Language* 39 (2) s. 139–157.
<https://doi.org/10.1177/0142723718805185>
- WANG, WEI-CHUN & YONELINAS, ANDREW, P. 2012: Familiarity is related to conceptual implicit memory: An examination of individual differences. – *Psychonomic Bulletin & Review* 19 (6) s. 1154–64.
- WARDHAUGH, RONALD 1998 [1986]: *An introduction to sociolinguistics*. Kolmas painos. Oxford: Blackwell.
- WECHSLER ADULT INTELLIGENCE SCALE IV. Helsinki: Hogrefe. Psykologien kustannus.
- WHITNEY, WILLIAM DWIGHT 1971 [1867]: *Language and the study of language. Twelve lectures on the principles of linguistic science*. – Michael Silverstein (toim.), *Whitney on language: selected writings on William Dwight Whitney* s. 7–110. Cambridge, Mass.: The MIT Press.
- WHITNEY, WILLIAM DWIGHT 1979 [1875]: *The life and growth of language. An outline of linguistic science*. New York: Dover Publications, Inc.
- WOLF NELSON, NICHOLA – VAN METER, ADELIA M. 2007: Measuring written language ability in narrative samples. – *Reading and Writing Quarterly* 23 (3) s. 287–309.

<https://doi.org/10.1080/10573560701277807>