

Aarikka, L., K. Priiki & I. Ivaska (toim.) 2022. Soveltavan kielitieteilijän sormenjälkiä etsimässä: Kielen rakenteet ja niiden käyttäjät Kirsti Siitosen tutkimuksellisinä kiintopisteinä. *In search of the fingerprints of an applied linguist: Linguistic structures and their users as scholarly focal points in Kirsti Siitonen's career. AFinLA-teema / n:o 14. 259–287.*

Hanna Lappalainen<sup>1</sup> ja Katri Priiki<sup>2</sup>

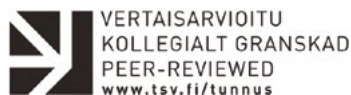
<sup>1</sup>Itä-Suomen yliopisto, <sup>2</sup>Turun yliopisto

## Yksilöllisiä valintoja ja kielitaustan vaikutusta – henkilöviitteiset *hän-*, *he-*, *se-* ja *ne-*pronominit kielielämäkertahaastatteluissa

The article examines variation in the use of the third-person pronouns *hän* and *he* (standard forms for human references) vs. *se* and *ne* (standard forms for non-human references) when referring to people in the interviews collected during the project “A Hundred Finnish Linguistic Life Stories” in 2017–2018. The focus is on two age groups, born in the 1950s and 1990s, and the data includes 32 interviews of people from different social and regional backgrounds. Eight of the interviewees spoke Finnish as their second language. The data is examined both quantitatively (with statistical methods) and qualitatively. The results show that there is considerable variation between individual speakers. Speakers with higher education born in the 1950s were the only relatively homogenous group in the data: they consistently preferred *hän* and *he* in the interviews when referring to human subjects. In the younger age group, being an L2 speaker appeared to be one of the factors affecting pronoun choices.

**Keywords:** quantitative methods, personal pronouns, variation

**Asiasanat:** kvantitatiiviset menetelmät, persoonapronominit, variaatio



## 1 Johdanto

Henkilöviitteisen kolmannen persoonan pronominin vaihtelu on ilmiö, joka erottaa erilaisia suomen kielen varieteetteja toisistaan. Yleiskielessä tehdään selvä ero siinä, että ihmisiin viitataan *hän-* ja *he-*pronomineilla, kun taas muihin referentteihin, kuten eläimiin ja esineisiin, viitataan *se-* ja *ne-*pronomineilla. Erilaisissa normittamattomissa varieteeteissa vaihtelua säätelee kuitenkin monimutkainen verkosto normeja, käytäntöjä ja käsityksiä, ja siksi ilmiötä onkin tarkasteltu eri suunnista. Tutkimuksen kohteena on ollut sekä pronominien käytössä ilmenevä variaatio että käyttöä koskevat käsitykset ja asenteet (esim. Ylikahri 1996; Laitinen 2005; Lappalainen 2010a; Priiki 2017a). Kirsti Siitonen (2008, 2016) on tutkimuksissaan huomionnut molemmat puolet. Tässä artikkelissa tarkastelemme kolmannen persoonan pronominien variaatiota haastatteluaineistossa vertaillen eritaustaisia puhujia<sup>1</sup>.

Pronominien käytössä ilmenevää variaatiota näyttävät aiemman tutkimuksen perusteella selittävän sekä kielenulkoiset että -sisäiset tekijät. Eroja on havaittu niin ikäryhmien, sukupuolten kuin eri tavoin koulutettujen välillä: vanhemmat puhujat käyttävät *hän-* ja *he-*pronomineja enemmän kuin nuoret, naiset ja korkeasti koulutetut enemmän kuin miehet ja matalammin koulutetut (Nuolijärvi 1986: 182–187; Mielikäinen 1988: 107–108; Paunonen 1995 [1982]: 165–177; Ylikahri 1996: 185–186; Lappalainen 2010a: 286–288; Kurki, Siitonen, Väänänen, Ivaska & Ekberg 2011: 95–96). Myös kielitaustalla voi olla merkitystä, sillä ruotsinkielisten ja kaksikielisten suomenpuhujien on havaittu käyttävän enemmän *hän-*pronominia kuin niiden, jotka pitävät itseään yksikielisinä suomenpuhujina; monikossa yhtä selvää eroa ei ole havaittavissa (Paunonen 1990: 202–206; Piiparinen-Rintaluoma 2001: 65–67). Pronominin valintaan saattaa vaikuttaa lisäksi viittauskohde (Lappalainen 2010a: 290–296; Kurki ym. 2011: 97; Priiki 2012, 2016b).

Pronominit ovat myös tilanteisessa vaihtelussa: yleiskieliset variantit ovat tavallisempia haastattelupuheessa ja asiointitilanteissa kuin arkikeskustelussa (Lappalainen 2010a: 296–298; 2010b), ja myös keskustelukumppanilla on vaikutusta: *hän-*pronominia käytetään herkemmin silloin, kun vuorovaikutustilanteessa on läsnä vieraampia osallistujia (Priiki 2016a). Pronominin valinta liittyy kiinteästi myös vuorovaikutustoimintaan, mikä tulee esiin etenkin tarkasteltaessa viittauksia ei-läsnäoleviin osallistujiin (Seppänen 1989, 1998). Edellisten lisäksi alueellinen tausta on relevantti muuttuja. Sekä suomen lounaisissa että kaakkoisissa murteissa on vanhastaan käytetty persoonapronomineja (lounaismurteissa *hän*, *he[e]*; kaakkoismurteissa *hää*, *hyö*) enemmän kuin muissa murteissa, joissa henkilöviittauksissa suositetaan demonstratiiveja *se* ja *ne*. Tämä käy ilmi Siitosen (2008) tutkimuksesta, jossa

1 Kiitämme nimettömiä arvioijia ja AFinLA-teeman toimituskuntaa arvokkaista kommentista. Tutkimus syntyi Lappalaisen aloitteesta ja hän vastasi esiintymien poimimisesta ja koodaamisesta sekä esimerkkien laadullisesta analyysistä. Priiki vastasi tilastollisista analyyseistä ja niiden tulkinna.

hän on laskenut Lauseopin arkiston tallenteista pronominiens esiintymämääriä. Laskelmat tukevat Vilppulan (1989) jo aiemmin esittämää vastaavaa havaintoa.

Kielensisäisistä tekijöistä on syytä mainita yksikön ja monikon ero, logoforiisuus sekä virittyneisyys. Aiemmissä tutkimuksissa on ensinnäkin havaittu, että yleis-kielinen variantti (*hän, he*) on yleisempi yksikössä kuin monikossa (Nuolijärvi 1986: 180–181; Paunonen 1995 [1982]: 171; Lappalainen 2010a: 287; Kurki ym. 2011: 95). Myös logoforiisuus eli käyttö referoinnissa viittaamassa alkuperäiseen puhujaan (esim. N1958b: *ja tota äiti sano et hänel ol aina hirvee paha olla*) on keskeinen *hän*-pronominin esiintymistä selittävä tekijä (Ylikahri 1996; Laitinen 2005; Priiki 2016b, 2017b). Virittyneisyydellä (engl. *priming*) viitataan puolestaan siihen, että puhujat ovat taipuvaisia käyttämään viittauksissaan samaa pronominia kuin juuri edellä on käytetty (Priiki 2016b).

Aiemmassa tutkimuksessa on selvitetty myös kolmannen persoonan pronomineihin liittyviä käsityksiä ja asenteita. Niitä on tutkittu sekä erilaisilla kyselyillä että analysoimalla verkkokeskusteluja. Näissä tutkimuksissa huomion kohteena on ollut ennen muuta käytön suhde ns. elollisuushierarkiaan (elollisuushierarkiasta yleisesti, ks. esim. Sorlin & Gardelle 2018) – erityisesti se, onko toisaalta sopivaa viitata *se*-pronominilla ihmiseen ja toisaalta *hän*-pronominilla eläimeen tai esineeseen (Siitonen 2008, 2016; Priiki 2017a: 74–75; 2021), minkä lisäksi on tarkasteltu sitä, miten tietoisia kielenkäyttäjät ovat omista pronomiivalinnoistaan (Lappalainen 2010a: 310–314). Nämä aiemmat tutkimukset osoittavat, että pronominin valintaan liitetään vahvoja tunteita ja kielenkäyttäjien omia normeja, joilla voi olla vaikutusta myös pronominiens todelliseen käyttöön ja siinä tapahtuviin muutoksiin.

Vaikka suomen kielen kolmannen persoonan pronominit ovat siis saaneet osakseen varsin runsaasti huomiota viime vuosikymmenten tutkimuksessa, niiden vaihtelu on niin monimutkainen ilmiö, että tutkittavaa riittää. Tutkimassamme kieli-elämäkerta-aineistossa erityistä on se, että se sisältää vertailukelpoisia haastatteluja hyvin erilaisilta puhujilta sen sijaan, että olisi keskitytty esimerkiksi tiettyyn alue-murteeseen. Mukana on myös puhujia, jotka ovat oppineet suomea toisena kielenä (S2). S2-puhujien määrä kasvaa jatkuvasti, ja siksi heidät on mielekästä ottaa huomioon variaatiota tarkasteltaessa. Tämänkään piirteen osalta heidän suomeaan ei ole juuri tutkittu (ks. kuitenkin Kuparinen 2001). Niinpä tutkimuksemme täydentää monin tavoin kuvaa kolmannen persoonan pronominiens variaatiosta nykysuomessa. Kiinnitämme huomiota yksilölliseen vaihteluun ja pohdimme kriittisesti perinteistä ryhmäjaottelua esimerkiksi iän, koulutuksen ja sukupuolen mukaan. Ennen tulosten tarkastelua esittelemme aineiston ja analyysimenetelmät.

## 2 Aineisto ja menetelmät

Analyysimme perustuu Sata suomalaista kielellistä elämäkertaa -hankkeen aineistoon. Tässä vuosina 2017–2020 toteutetussa ja Koneen Säätiön rahoittamassa hankkeessa kerättiin kaikkiaan 131 kielielämäkertahaastattelua eri-ikäisiltä suomalaisilta. Suomalaisuus tarkoittaa tässä yhteydessä sitä, että haastatellut asuivat aineistonkeruuhetkellä (v. 2017–2018) Suomessa ja pystyivät osallistumaan haastatteluun suomeksi, ruotsiksi tai suomalaisella tai suomenruotsalaisella viittomakielellä.

Haastattelujen kesto on keskimäärin noin tunti. Suurin osa haastatteluista on opiskelijoiden tekemiä. Koska haastattelijoita on kymmeniä, haastatteluissa on vaihtelua, vaikka kaikki käyttivät samaa haastattelurunkoa. Haastattelut keskittyvät tutkittavien kielellisiin kokemuksiin sekä kielikäsitteisiin ja -asenteisiin. Keskeistä oli käydä läpi tärkeimmät elämänvaiheet ja esimerkiksi perheen merkitys haastateltavan kielenkäytölle, joten haastatteluissa puhutaan varsin paljon ihmisistä ja niissä on suhteellisen runsaasti henkilöviittauksia. Kuten myöhemmin osoitamme, tutkittujen pronominiin määrä kuitenkin vaihtelee. Koska aineistona ovat yksilöhaastattelut eikä haastattelutilanteessa ole yleensä ollut läsnä muita kuin haastattelijaa ja haastateltavaa, kaikki haastatteluissa ilmenevät kolmaspersonaiset pronominit viittaavat ei-läsnäoleviin henkilöihin (vrt. Seppänen 1998). (Ks. tarkemmin hankkeesta ja haastattelujen toteutuksesta Hippi, Lappalainen & Nuolijärvi 2020.)

Hankkeessa haastatelluista olemme tähän tutkimukseen valinneet tutkittavaksi kaksi ikäryhmää, 1950- ja 1990-luvulla syntyneet. Informantit olivat haastatteluhetkellä 20–26- ja 62–67-vuotiaita. Juuri nämä ikäryhmät valittiin, koska niissä haastateltavia on poikkeuksellisen paljon. Mukana on yhteensä 32 informanttia, 16 kummassakin ikäryhmässä – molemmissa yhtä paljon miehiä (8) ja naisia (8). Tutkittavilla on erilaisia alueellisia, ammatillisia ja kielitaustoja. Muutamat ovat kotoisin alueilta, joilla *hän-* tai *hää-* sekä *he-* tai *hyö-*pronominit ovat keskeinen osa murretta, kun taas useimmat ovat kotoisin murrealueilta, joilla *se* ja *ne* ovat valta-asemassa. Kahdeksan informantin äidinkieli on muu kuin suomi, tai heillä on suomen lisäksi toinen lähes yhtä vahva kieli (esim. joruba, unkari, venäjä).<sup>2</sup> Yhtä informanttia (LM1959) lukuun ottamatta kaikkien kielitaitoa voi luonnehtia haastattelun perusteella erinomaiseksi. Useimpien kohdalla ei kielenkäytön perusteella ole helppoa edes kuulla eroa ensikielisiin suomenpuhujiin (ks. Lappalainen, Hippi & Nuolijärvi 2021). 1950-luvulla syntyneet S2-puhujat ovat muuttaneet Suomeen ja oppineet suomen kielen vasta aikuisiässä. Nuoremmat sen sijaan ovat joko asuneet Suomessa koko ikänsä tai

2 Tiedostamme, että käyttämämme ilmaisut, kuten *S2-puhujat*, *äidinkieliset suomenpuhujat* jne., ovat suhteellisia ja ongelmallisiakin, sillä useimpien monikielisten haastateltaviemme suomen kielen taito on erinomainen ja he saattavat kokea suomen olevan vahvempi kieli kuin heidän ensikielensä (vrt. Hippi 2020). Lisäksi ensikieliä voi olla useita eikä ensimmäisenä opittu kieli ole välttämättä äidiltä opittu kieli. Noudatamme kuitenkin aineistonkeruussa käytettyä terminologiaa, jossa äidinkielellä viitataan ensimmäisenä opittuun kieleen.

muuttaneet viimeistään kouluikässä, joten he ovat ehtineet vahvasti sosiaalistua suomen kieleen jo lapsuudessaan tai viimeistään teini-iässä. Lähtökohtamme on, että S2-puhujat ovat kiinteä osa suomalaista kieliyhteisöä ja heitä on siksi mielekästä tarkastella rinnakkain ensikielisten puhujien kanssa (vrt. esim. Lappalainen 2006a; Hippi 2020, *oppija*-termin käsitteestä ja sen monitulkintaisuudesta, ks. myös Suni & Ruuska tässä teoksessa).

Viittaamme informantteihin koodeilla, jotka koostuvat sukupuolesta (M, N), syntymävuodesta ja tarvittaessa kirjaintunnuksesta (a, b, c), jos samana vuonna syntyneitä on useampia. Mikäli koodin alussa on L-kirjain, haastattelu kuuluu hankkeen niin sanottuun laajennusosaan (vrt. sadan haastattelun perusaineisto, ks. tarkemmin Hippi ym. 2020: 16–17). Näin ollen esimerkiksi N1997 tarkoittaa vuonna 1997 syntyneitä naisinformantteja ja LM1994a ja LM1994b vuonna 1994 syntyneitä miesinformantteja, joiden haastattelut kuuluvat hankkeen laajennusosaan.

Tarkastelluissa haastatteluissa tutkittavat tuottavat yhteensä 1605 *hän-*, *he-*, *se-* tai *ne-*pronominia.<sup>3</sup> Esiintymät eivät jakaudu tasaisesti, vaan haastattelua kohti niitä on 10–112. Niinpä vähän pronomineja käyttävien osalta haastattelun antama kuva variaatiosta ei välttämättä ole luotettava, kun taas paljon esiintymiä tuottaneet voivat dominoida aineistossa. Nämä seikat on otettu huomioon tuloksia tulkittaessa, ja käytetyt menetelmät on valittu siten, että ne sopivat epätasaisesti jakautuvan aineiston tarkasteluun. Eroja selittää esimerkiksi se, että osa vastaa haastattelijan kysymyksiin lyhyesti, eikä puheeseen synny viittausetjuja.

Tarkastelemme aineistoa sekä määrällisesti että laadullisesti. Määrällisessä tarkastelussa käytämme tilastollisina monimuuttujamenetelminä satunnaismetsä- ja päätöspuuanalyysseja, jotka perustuvat samaan, aineistoa luokittelevaan algoritmiin. Algoritmi jakaa aineistoa toistuvasti osiin valittujen selittävien muuttujien perusteella siten, että osien välille muodostuu aina tilastollisesti merkitsevä ero. Näin pystytään tarkastelemaan selittävien muuttujien tärkeysjärjestystä sekä niiden keskinäisiä suhteita. Kumpaakin menetelmää on fennistiikassa käytetty jonkin verran viime vuosikymmenen aikana (satunnaismetsäanalyysistä esim. Ivaska 2014: 169–170; Väänänen 2015: 14–15; 2016: 44–46; Priiki 2016b: 117–118; päätöspuumenetelmästä esim. Priiki 2014: 51; Tyysteri 2015: 96; Moshnikov 2021: 98–103 sekä näissä mainitut lähteet). Valitut menetelmät ovat taustaoletuksiltaan sallivia, sopivat pienehkönkin aineiston käsittelyyn eivätkä vaadi esimerkiksi sitä, että tarkasteltavat muuttujat olisivat täysin riippumattomia toisistaan. Tässä tarkasteltavassa aineistossa esimerkiksi iän, koulutustaustan ja äidinkieliyyden välillä voi olla erilaisia keskinäisiä riippuvuuksia. Monimuuttujamenetelmien lisäksi käytämme perinteisempiä tilastomenetelmiä eli ristiintaulukointia ja ruutu- ja janakaavioita. Ristiintaulukoinneilla

3 Muutamat yksittäiset *tämä-* ja *tuo-*pronomien esiintymät olemme rajanneet tarkastelun ulkopuolelle.

osoitamme yhteyksiä yksittäisen muuttujan ja pronomini- vaihtelun välillä, ja ruutu- ja janakaavioilla kuvaamme pronomini- käytön jakautumista aineistossa.

Päätöspuita, satunnaismetsiä ja ristiintaulukointeja varten olemme koodanneet aineiston pronomini- esiintyminä, jolloin puhujat, joilla on paljon pronomini- viittauksia, korostuvat aineistossa. Tarkastelu voidaan toteuttaa myös siten, että hän- tai he-pronominin osuutta kunkin puhujan haastattelussa tarkastellaan määrällisenä muuttujana. Tällöin sekä paljon että vähän pronomineja haastatteluissa käyttävät puhujat vaikuttavat tuloksiin yhtä paljon. Tämä näkökulma aineistoon esitetään ruutu- ja janakaavioilla, jotka havainnollistavat sitä, miten (epä)yhtenäisiä tarkastelut puhujajoukot ovat. Ruutu- ja janakaaviossa ruudun rajat näyttävät, mille alueelle puolet muuttujan arvoista asettuu. Ruudun sisällä oleva viiva osoittaa mediaani- arvoa, ja janat kuvaavat hajontaa. Pisteet ilmaisevat muista poikkeavia arvoja. Ruutu- ja janakaaviot, ristiintaulukoinnit sekä päätöspuuanalyysit tunnuslukuineen on toteutettu SPSS-ohjelmalla (versio 27); satunnaismetsäanalyysi perustuu R-ohjelman party-paketin cforest- ja varimp-funktioihin (R Core Team 2021; Strobl, Malley & Tutz 2009).

Monimuuttuja-analyseissä käytetyt muuttujat ja niiden luokat käyvät ilmi taulukosta 1. Muuttujana on myös informantti, eli tarkastelemme sitä, kuinka keskeistä yksilöllinen vaihtelu on verrattuna perinteisiin sosiaalisiin muuttujiin. Perinteiset muuttujat ovat niin kutsuttuja kiinteän vaikutuksen muuttujia, eli ne voivat saada rajallisen määrän eri arvoja. Yksilö muuttujana poikkeaa tästä, sillä erilainen otos sisältäisi eri puhujat. Tällaista muuttujaa kutsutaan satunnaisvaikutusmuuttujaksi. (Väänänen 2016: 46.) Valituissa menetelmissä on mahdollista sisällyttää analyysiin satunnaisvaikutusmuuttuja perinteisempien kiinteän vaikutuksen muuttujien lisäksi (Tagliamonte & Baayen 2012). Yksilöllisten erojen tarkastelua perustelee se, että aiemmat tutkimukset osoittavat niiden olevan pronomini- vaihtelussa suuria (Nuolijärvi 1986: 187; Paunonen 1995 [1982]: 172; Lappalainen 2010a; Priiki 2016b).

TAULUKKO 1. Tarkasteltavat muuttujat ja niiden luokat.

Muuttuja	Luokat
Informantti	32 puhujaa kukin erikseen
Ikäryhmä	1950-luvulla syntyneet; 1990-luvulla syntyneet
Koulutus	Korkeakoulutus (amk tai yliopisto); muu
Sukupuoli	Nainen; mies
Kielitausta	Suomi äidinkieli; S2
Viittauksen kohde	Perheenjäsen; muu spesifi; geneerinen

Aiemman tutkimuksen perusteella tiedetään, että sekä lingvistisillä että kielenulkoi-  
silla tekijöillä on vaikutusta kolmannen persoonan pronomien variaatioon. Kuten  
taulukosta 1 käy ilmi, kiinnitämme huomiota puhujan ikään, sukupuoleen, koulu-  
tukseen ja äidinkieleen (pelkkä suomi vs. muu kieli [+suomi]) sekä pronomien viit-  
tauskohteen eli referentin laatuun. Koulutuksen osalta tutkittavat on jaettu kahteen  
ryhmään: toiseen kuuluvat ne, joilla on ammattikorkeakoulu- tai yliopistotasoinen  
loppututkinto tai jotka parhaillaan opiskelevat korkeakoulussa, ja toiseen kaikki  
muut. Alueellista taustaa emme ole tarkastelleet tilastollisesti, sillä tutkittavista vain  
kolme asui haastatteluhetkellä *hän*- ja *he*-pronomineja (tai niiden murrevariantteja)  
suosivalla alueella. Luokittelua hankaloitti myös pääkaupunkiseudulla asuvien sijoit-  
telu. Kommentoimme kuitenkin alueellista taustaa yksilöllistä vaihtelua tarkastelles-  
samme. Lisäksi olemme tehneet huomioita puhetilanteen muiden osallistujien eli  
haastattelijan vaikutuksesta, mutta tätäkään emme ole tarkastelleet määrällisesti.

Tarkastelemme myös pronominalinnan yhteyttä siihen, millaisesta henkilö-  
viittauksesta on kyse. Olemme jaotelleet puheena olevat kohteet kolmeen luok-  
kaan: perheenjäseniin, geneerisiin referentteihin ja muihin. Perheenjäseniksi on  
luokiteltu puhujan vanhemmat, isovanhemmat, puoliso, sisarukset, lapset ja lap-  
senlapset. Kaukaisemmat sukulaiset, kuten appivanhemmat, tädit, sedät ja serkut,  
on luokiteltu muihin. Sukulaisia koskeva jaottelu on käytännön syistä mekaaninen,  
sillä haastattelut eivät anna riittävästi tietoa sen arvioimiseen, miten läheisiä kysei-  
set sukulaisuussuhteet ovat. Geneerisiksi on puolestaan luokiteltu viittaukset, jotka  
kohdistuvat yleistäen kehen tahansa tietyn ryhmän edustajaan (ks. VISK § 1407).

Logoforisuus eli käyttö referoidussa katkelmassa viittaamassa alkuperäiseen  
puhujaan on aiemmissa tutkimuksissa todettu keskeiseksi *hän*-pronomien esiinty-  
mistä selittäväksi tekijäksi (Priiki 2016b, 2017b). Tarkastelemassamme haastattelu-  
aineistossa logoforisia tapauksia on vain muutama, ja olemme jättäneet ne tarkas-  
telun ulkopuolelle ja pois laskelmista. Esittämämme havainnot kuvaavat näin ollen  
ei-logoforista käyttöä. Virittyneisyydestä eli siitä, että viittausketjuissa suositaan  
samaa pronomia (Priiki 2016b), esitämme vain joitakin huomioita laadullisen  
analyysin yhteydessä.

### 3 Kolmannen persoonan pronomien variaation tarkastelua

Tässä luvussa tarkastelemme tekijöitä, jotka vaikuttavat pronomien *hän*, *se*, *he* ja *ne*  
vaihteluun aineistossamme. Tarkastelemme erikseen kahta vertailtavaa ikäryhmää  
sekä yksiköllisiä ja monikollisia pronomineja. Kuten edellä mainitsimme, aiemmissa  
tutkimuksissa on havaittu, että yksiköllinen *hän* on yleisempi kuin monikollinen  
*he*. Tämä pitää paikkansa myös tässä aineistossa: *hän*-pronomineja on yksikkö-

muodoista 47,6 %, *he*-pronominia puolestaan monikkumuodoista vain 27,2 %.<sup>4</sup> Lisäksi yksiköllisten ja monikollisten muotojen variaatio on aineistossa erilaista, ja kuten seuraavissa alaluvuissa osoitamme, niiden vaihteluun vaikuttavat erilaiset tekijät.

Ensimmäisessä alaluvussa (3.1) esittelemme tarkasteltavien pronominien jakauman ikäryhmittäin. Toinen alaluku (3.2) sisältää monimuuttuja-analyysin vaihteluun vaikuttavien tekijöiden tärkeysjärjestyksestä. Koska tässä analyysissä yksilökohtainen vaihtelu osoittautuu tärkeimmäksi tekijäksi, jatkamme tarkastelua jakamalla puhujat päätöspuunen menetelmällä toisistaan tilastollisesti merkitsevästi poikkeaviin ryhmiin (3.3). Tämän jälkeen tarkennamme määrällistä analyysiä kuvaamalla, miten muut siinä huomion arvoiksi osoittautuneet muuttujat eli koulutus- (3.4) ja kielitausta (3.5) sekä viittauksen kohde (3.6) ovat kytköksissä vaihteluun. Näillä muuttujilla on ristiintaulukointien mukaan koko tarkasteltavan aineiston tasolla tilastollisesti merkitsevä yhteys pronominien vaihteluun<sup>5</sup>, kun taas sukupuolella selaista ei tässä aineistossa ole havaittavissa.

### 3.1 Ikäryhmien väliset ja sisäiset erot

Taulukko 2 kuvaa pronominien käyttöä ikäryhmittäin. Ristiintaulukoinnissa merkitsevyys on testattu Pearsonin khiin neliö -testin avulla. Esitämme taulukossa myös sovitetut standardoidut jäännökset, jotka kuvaavat odotuksenmukaisen ja toteutuneen frekvenssin erotusta. Erotusta pidetään tilastollisesti merkitseväenä, jos jäännöksen itseisarvo on suurempi kuin 2.

Kuten on aiemman tutkimuksen perusteella odotuksenmukaista, 1950-luvulla syntyneet käyttävät sekä *hän*- että *he*-pronominia 1990-luvulla syntyneitä enemmän: kaikkien taulukossa esitettyjen lukujen välillä on tilastollisesti merkitsevä ero. *Hän*-pronominin yhteys ikäryhmään on vahvempi kuin *he*-pronominin, mikä ilmenee siitä, että jäännösten itseisarvo on suurempi.

4 Pearsonin khiin neliö -testin mukaan ero on tilastollisesti merkitsevä:  $p < 0,001$ ;  $\chi^2 = 71$ ,  $df = 1$ .

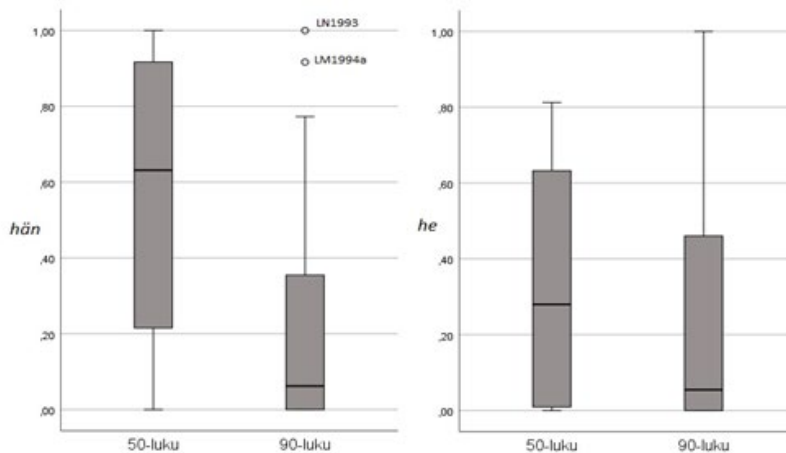
5 Pearsonin khiin neliö -testin mukaan koulutustausta  $p < 0,001$ ;  $\chi^2 = 120$ ,  $df = 1$ ; kielitausta  $p < 0,001$ ;  $\chi^2 = 22$ ,  $df = 1$ ; viittauksen kohde  $p < 0,001$ ;  $\chi^2 = 50$ ,  $df = 2$ .



TAULUKKO 2. Pronominien *hän*, *se*, *he* ja *ne* yhteys puhujan ikäryhmään, ristiintaulukointi sovitettuja standardoituja jäännöksiä käyttäen ( $p < 0,001$ ;  $\chi^2 = 136$ ,  $df = 3$ ).

	<i>hän</i>	<i>he</i>	<i>se</i>	<i>ne</i>	Yhteensä
1950-luvulla syntyneet	287 (9,8)	138 (3,2)	156 (-7,8)	274 (-3,9)	855
1990-luvulla syntyneet	95 (-9,8)	80 (-3,2)	265 (7,8)	310 (3,9)	750
Yhteensä	382	218	421	584	1605

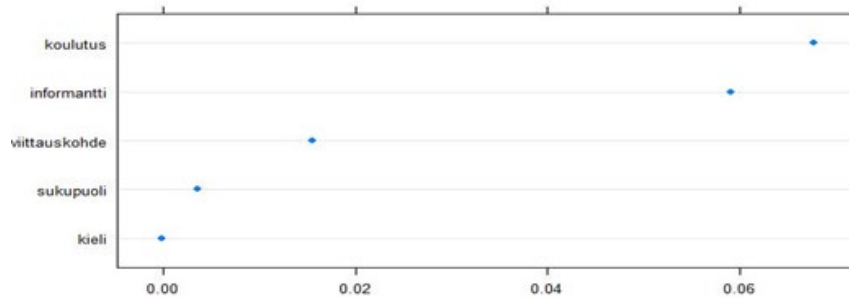
Ruutu- ja janakaaviot (kuvio 1) puolestaan kuvaavat puhujien pronominikäytön hajontaa tarkasteltavissa ryhmissä. *Hän*-pronominin osuutta esittävässä kuviossa huomiota kiinnittää, että janat ulottuvat asteikon ääripäihin: aineistossa on 9 informanttia, jotka käyttävät pelkästään *se*-pronominia (7/9 miehiä) ja 2 informanttia, jotka käyttävät pelkästään *hän*-pronominia (molemmat naisia). 1990-luvulla syntyneistä 7/16 käyttää pelkästään *se*-pronominia, 6/16 pelkästään *ne*-pronominia. Aineiston pienuuden ja suuren hajonnan vuoksi ryhmien mediaanilukujen vertailu ei tuota tilastollisesti merkitseviä tuloksia.<sup>6</sup>

KUVIO 1. Ruutu- ja janakaavio eri vuosikymmenillä syntyneiden haastateltavien *hän*- ja *he*-pronominien osuudesta.

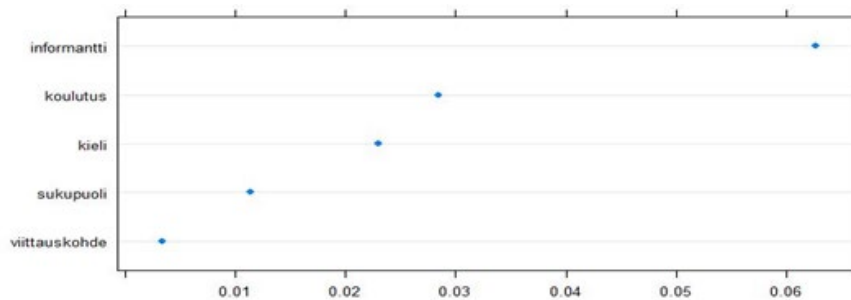
6 *Hän*-pronominin kohdalla ero ei jää kauas merkitsevältä (Yatesin korjattu  $p = 0,077$ ;  $\chi^2 = 3,13$ ;  $df = 1$ ), *he*-pronominin kohdalla jää ( $p = 0,72$ ;  $\chi^2 = 0,13$ ;  $df = 1$ ).

### 3.2 Tärkeimmät vaihtelua selittävät tekijät

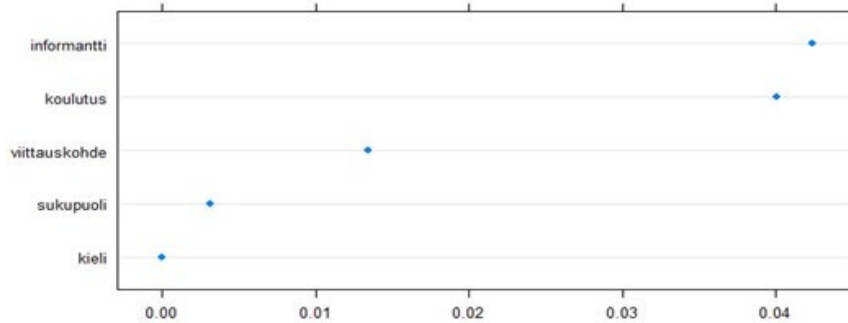
Kuviot 2a ja 2b kuvaavat *hän*- ja *se*-pronominien vaihteluun vaikuttavien tekijöiden tärkeyttä tässä aineistossa ikäryhmittäin; kuviot 3a ja 3b esittävät vastaavan analyysin monikollisista pronomineista. Kuvioissa näkyvät luvut ilmaisevat satunnaismetsä-analyysin tuottamaa muuttujan selittävyys -arvoa (*variable importance*, suomenkielinen termi ks. Väänänen 2015). Numeerisia arvoja olennaisempaa kuvioiden tulkinnassa on selittävien tekijöiden järjestys ja sijainti toisiinsa nähden.



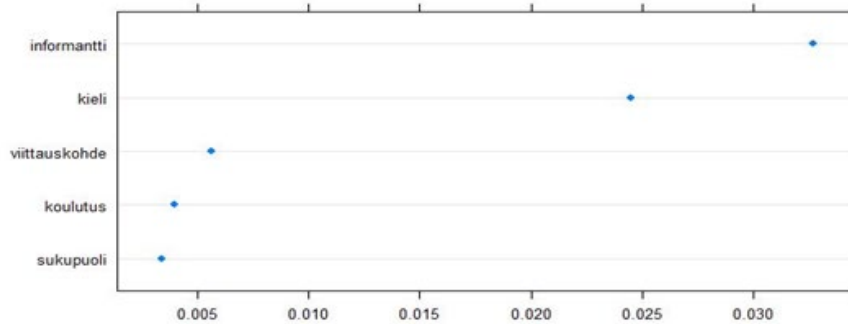
KUVIO 2a. *Hän*- ja *se*-pronominien vaihtelua aineistossa selittävien tekijöiden tärkeys: 1950-luvulla syntyneiden puhujien haastattelut.



KUVIO 2b. *Hän*- ja *se*-pronominien vaihtelua aineistossa selittävien tekijöiden tärkeys: 1990-luvulla syntyneiden puhujien haastattelut.



KUVIO 3a. *He-* ja *ne-*pronominien vaihtelua aineistossa selittävien tekijöiden tärkeys: 1950-luvulla syntyneiden puhujien haastattelut.



KUVIO 3b. *He-* ja *ne-*pronominien vaihtelua aineistossa selittävien tekijöiden tärkeys: 1990-luvulla syntyneiden puhujien haastattelut.

Kuvio 2a näyttää aineiston tilanteen 1950-luvulla syntyneiden puhujien osalta. Kun sitä verrataan kuvioon 2b eli 1990-luvulla syntyneiden puhujien kielenkäyttöön, nähdään eroja. Vanhemmassa ikäryhmässä koulutus ja puhujakohtainen vaihtelu ovat selvästi tärkeimmät tekijät sekä yksiköllisissä että moniköllisissä viittauksissa. Myös viittauksen kohteella näyttää olevan jonkin verran merkitystä nimenomaan vanhemman ikäryhmän pronominiin käytössä, jossa geneeriseen referenttiin viitataan suhteessa useammin *ne*-muodoilla kuin perheenjäsenen; yksikössä erottuvat pelkästään perheenjäseniin kohdistuvat viittaukset. Ero nuorempaan ikäryhmään selittänee se, että 1990-luvulla syntyneissä on suhteessa enemmän niitä, jotka käyttävät vain yhtä varianttia viittauskohteen luonteesta riippumatta, sekä se, että nuorten S2-puhujien variaatio on referentin kohteen vaikutuksen osalta poikkeuksellista. (Ks. tarkemmin alalukua 3.6.) Kielitausta selittää vanhemman ikäryhmän pronomini-vaihtelua tarkastelluista muuttujista heikoimmin. Tätä tulosta tarkasteltaessa on syytä huomata, että 1950-luvulla syntyneiden ryhmässä S2-puhujia on vain kaksi ja he ovat taustaltaan hyvin erilaisia.

Nuoremmassa ikäryhmässä puhujan koulutustausta on tärkeä selittäjä yksikölisten *hän-* ja *se-*pronominien vaihtelussa (kuvio 2b), mutta monikollisten muotojen (kuvio 3b) variaation selittämisessä koulutuksella ei ole merkitystä. Toisin kuin vanhempien puhujien kohdalla, nuoremmassa ikäryhmässä kielitausta nousee tärkeäksi selittäjäksi. Tämä näkyy etenkin monikkomuodoissa. Vanhempien ryhmässä kielitaustan tarkastelua vaikeuttaa S2-puhujien pieni määrä, kun taas nuorempien ryhmässä S2-puhujien osuus on huomattava (6/16). Ero äidinkieliisiin puhujiin on kiinnostava siinä valossa, että kaikkineen näiden 1990-luvulla syntyneiden nuorten aikuisten puhekielen variaatio ei suuresti poikkea samanikäisistä suomenkielisistä (Lappalainen ym. 2021). On mahdollista, että piirre, jossa kyse on toiseen ihmiseen viittaamisesta ja johon voidaan liittää myös evaluoiva merkitys, on eri tavalla herkkä esimerkiksi oman kielen vaikutukselle tai yleiskielen normien korostamiselle kuin useimmat muut piirteet.

Yksilöllisen vaihtelun tärkeys ei yllätä, sillä, kuten edellä mainitsimme, puhujakohtaiset erot on pronominien käytössä todettu suuriksi jo aiemmissa tutkimuksissa. Lisäksi yksilöllinen vaihtelu on osoittautunut tärkeäksi selittäjäksi monien muiden kielenpiirteiden kohdalla, joita on tarkasteltu tilastollisin menetelmin (esim. Baayen 2008; Tagliamonte & Baayen 2012: 24; Klavan, Pilvik & Uiboed 2015: 213). Koska yksilöllinen vaihtelu osoittautuu keskeiseksi tekijäksi tässäkin aineistossa, seuraava alaluku 3.3 keskittyy tarkastelemaan puhujakohtaisen vaihtelun suhdetta perinteisempiin muuttujiin.

### 3.3 Yksilöllinen vaihtelu

Olemme edellä osoittaneet, että yksilökohtainen vaihtelu ryhmien sisällä on erittäin suurta. Voidaan jopa kyseenalaistaa perinteinen jaottelu iän, koulutustason ja sukupuolen mukaan. Seuraavaksi tarkastelemme idiolektien eroja päätöspuulla. Päätöspuu jakaa aineiston puhujat ryhmiin, joiden välillä on tilastollisesti merkitseviä eroja pronominien käytössä. Samaan ryhmään kuuluvat puhujat siis käyttävät pronomineja samassa suhteessa. Seuraavaksi esitettävä ryhmäjako perustuu SPSS-ohjelman CHAID-algoritmilla<sup>7</sup> tehdyn päätöspuun antamaan jaotteluun. Tilasto-ohjelman tuottaman kuvion sijaan havainnollistamme syntyviä ryhmiä vapaa-muotoisemmilla taulukoilla. Tämä ratkaisu mahdollistaa sen, että voimme esittää samalla, millaisia yhteyksiä iällä, äidinkielellä, koulutuksella ja sukupuolella on päätöspuun tuottamaan ryhmäjakoon. Aineiston puhujien muodostamat, toisistaan *hän-*pronominin käytön osalta merkitsevästi eroavat ryhmät esitetään taulukossa 3; taulukossa 4 on vastaava jaottelu *he-*pronominin käytön osalta. Taulukoissa miesten informanttitunnukset ovat sinisellä, naisten punaisella. Puhujan tunnuksen perään

7 CHAID on lyhenne sanoista *chi-squared automatic interaction detection*. Tällä algoritmilla tuotetaan siis sellainen päätöspuun versio, jossa tilastollinen merkitsevyys lasketaan khiin neliö -testillä.

on merkitty, mikäli tällä on korkeakoulututkinto tai -opintoja (amk, yliopisto) sekä mikäli hänellä on jokin toinen äidinkieli suomen kielen sijaan tai lisäksi (S2).

Hän-pronomeineja haastattelupuheessa				
Ei lainkaan	Vähän (15,3 %)	Jonkin verran (36,7 %)	Paljon (83,2 %)	Erittäin paljon (95,1 %)
LM1955 LM1959 (S2)	N1952 M1954b N1957a	M1952b LN1956 N1957b	M1952a M1954a (yliopisto) N1958a (amk)	LM1950 (S2, yliopisto) M1952c (yliopisto) N1956 (yliopisto) LN1957 (yliopisto) N1958b
M1991 M1993 (amk) LN1995 (yliopisto) LM1995b (yliopisto) M1996b (S2) N1997 M1997	N1991 (S2, yliopisto) N1994 (amk) N1995 (S2, amk)	N1992 (amk) LM1994b (yliopisto) M1996a (S2, amk)	N1996 (S2)	LN1993 LM1994a (S2, amk)

TAULUKKO 3. Aineiston puhujat jaoteltuna *hän*-pronominin käytön osalta ryhmiin, jotka poikkeavat toisistaan tilastollisesti merkitsevästi päätöspuuanalysin<sup>8</sup> mukaan.

He-pronomeineja haastattelupuheessa					
Ei lainkaan	Vähän (4,7 %)	Melko vähän (26,9 %)	Jonkin verran (45,5 %)	Melko paljon (64,4 %)	Paljon (80,8 %)
N1952 N1957a N1957b LM1959 (S2)	M1952b M1954b LM1955	M1954a (yliopisto) LN1956	M1952a N1956 (yliopisto)	M1952c (yliopisto) LN1957 (yliopisto) N1958a (amk) N1958b	LM1950 (S2, yliopisto)
M1991 LN1995 (yliopisto) LM1995b (yliopisto) M1996a (S2, amk) N1997 M1997	N1991 (S2, yliopisto) M1993 (amk) M1996b (S2)	N1992 (amk) N1994 (amk)	LM1994b (yliopisto) N1995 (S2, amk)	N1996 (S2)	LN1993 LM1994a (S2, amk)

TAULUKKO 4. Aineiston puhujat jaoteltuna *he*-pronominin käytön osalta ryhmiin, jotka poikkeavat toisistaan tilastollisesti merkitsevästi päätöspuuanalysin<sup>9</sup> mukaan.

8  $p < 0,001; \chi^2 = 434; df = 4.$

9  $p < 0,001; \chi^2 = 314; df = 5.$

Taulukoista näkyy muun muassa se, että melko paljon, paljon tai erittäin paljon *hän-* ja *he-*pronomineja käyttävien ryhmissä on pääosin 1950-luvulla syntyneitä, mutta mukana ovat myös LM1994a, LN1993 ja N1996. 1950-luvulla syntyneistä kaikki muut ovat korkeakoulutettuja (esim. LM1950 ja M1952c ovat pappeja) paitsi N1958b, joka asuu Etelä-Karjalassa ja käyttää usein murteen mukaisia *hää-*, joskus myös *hyö-*variantteja. Esimerkistä 1 käy ilmi, miten hän viittaa isäänsä *hää-* ja *hänelle-*muodoilla. Hänen osaltaan siis alueellinen tausta lienee olennainen selittävä tekijä ikäryhmän lisäksi.

(1)<sup>10</sup>

Haastattelija: mmm, et se oli niiku arvostettuu tavallaa se muiden kielten osaaminej ja,

N1958b: se oli siis äidin puolelt arvostettuu isähän ei taas antanu niinku sillee arvostanu kieliä millää lailla että,

Haastattelija: joo.

N1958b: **hänelle** oli sit tärkeempiä, tärkeempää esimerkiks matematiikka.

Haastattelija: mm.

N1958b jota **hää** itte oli opiskellu.

Nuorista erottuvat LM1994a, LN1993 ja N1996. LM1994a on kasvanut unkarinkielisessä kodissa ja N1996:n kotikieli on ollut venäjä, mikä voi selittää pronomien käyttöä, kun taas LN1993 on epäodotuksenmukainen *hän-* ja *he-*pronomien suosija: hän on asunut perinteisesti *se-*pronominia suosivilla murrealueilla ja valmistunut nuoriso-ohjaajaksi. Hänellä on taustallaan hieman yliopisto-opintojakin, mutta olennaisempaa on ehkä se, että haastattelija ja LN1993 olivat toisilleen aivan vieraita ja molemmat tuntuvat suhtautuvan tallennustilanteeseen hyvin muodollisesti. Varsinkin haastattelun alkupuolella kumpikin puhuu kaikkineen varsin yleiskielisesti (esim. 2). Tämän voi havaita esimerkiksi monikon 3. persoonan kongruoivista verbi-muodoista (*tekevät, asuvat*), loppuvokaalien säilymisestä translatiivimuodoissa (*lisäksi, maisteriksi*), jälkitavun painottoman *i-*loppuisen diftongin säilymisestä (*valmistuikin*) ja *ole-*muodon käyttämisestä puhekielisemmän *oo-*muodon sijasta.

(2)

Haastattelija: no missäs he, asustelevat tällä hetkellä, sun sisarukset ja öö tota, mitä tekevät työkseen tai opiskelevat tai, sellasta.

LN1993: joo tää pikkusisko tosiaan opiskelee täällä ((nimi poistettu)) yliopistossa, juurikin

10 Esimerkit on litteroitu melko karkeasti. Taukoa on merkitty pilkulla; piste ilmaisee selvästi laskevaa ja kysymysmerkki selvästi nousevaa intonaatiota. Lisäksi litteraattiin on merkitty hakasulkeella päällekkäispuhunnan alku, alleviivauksella selvää painotusta ja kaksoissulkeisiin litteroijan kommentteja. Sulkeissa olevat viivat (–) ilmaisevat, että puheessa on epäselvä kohta.

((kieli poistettu)) ja kirjallisuutta, ja tota, tää mun isosisko-, puoli, **hän** on kotiäitinä, äitinä ja tota, ei oo, mun käsittääkseni mitään koulutusta, peruskoulul lisäksi, ja tota, tää mun veli sitten niin, **hän** on tosiaan ylioppilas ja sitten, opiskeli Helsingissä, kauppatieteitä, ja valmistuikin sieltä maisteriksi, muutama vuosi sitten ja, on nyt sitten, sitten töissä. tekee jotain ekonomin hommia kaiketi, ehkä. ja **hänellä** on myös, myös sitten, nais-, ystävä ja asuvat yhdessä ja näin, mutta ei ole lapsia.

Pelkästään *se*-pronomineja käyttävien ryhmä taas koostuu enimmäkseen 1990-luvulla syntyneistä, mutta mukana ovat myös LM1955 ja LM1959. Kumpikin on matalasti koulutettu mies ja asuu alueella, jolla *se* on valtavaraintti. LM1955 on äidinkielen suomenpuhujia Lapista ja LM1959:n äidinkieli on joruba. Kaikkineen suurin osa (7/9) on miehiä. Pelkkää *ne*-pronominia käyttävien väliltä on vaikeampi löytää mitään yhteistä muuttujaa, mutta suhteellisen harvalla (3/10) on taustallaan korkeakouluopintoja.

### 3.4 Koulutuksen yhteys pronomien käyttöön

Edellä satunnaismetsäanalyysissä (kuviot 2 ja 3) etenkin 1950-luvulla syntyneiden ryhmässä koulutustaso osoittautui pronomien vaihtelua selittäväksi muuttujaksi. Ristiintaulukointi vahvistaa sen satunnaismetsäanalyysin tuloksen, että myös nuorten ryhmässä korkeakoulutuksella on yhteys hieman runsaampaan *hän*-pronomien käyttöön;<sup>11</sup> *he*-pronomien käytössä merkitsevää eroa ei ole.

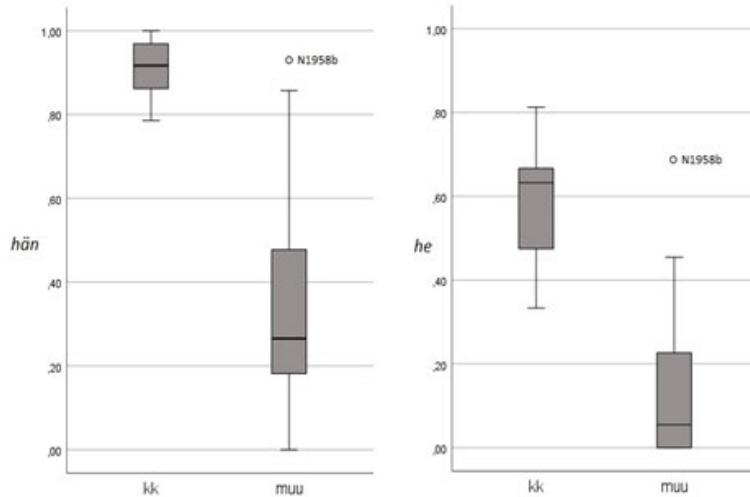
Kuvioiden 4a ja 4b ruutu- ja janakaaviot (seur. sivulla) havainnollistavat tutkittavien pronomien käytön jakautumista heidän koulutustaustansa mukaan. Vaikka enemmistö nuoremman ikäryhmän puhujista suosii *se*- ja *ne*-pronomineja (kuvio 4b, mediaania ilmaiseva viiva sijoittuu alas), hajonta on suurta sekä korkeakoulutettujen että muiden keskuudessa, eivätkä erot ryhmien mediaaniluvuissa ole tilastollisesti merkitseviä. Kolme korkeakoulutettua nuorta ei haastattelussa käytä *hän*-pronomia lainkaan, kun taas matalammin koulutettu LN1993 käyttää sitä yksinomaisesti (taulukko 3).

Vanhemmassa ikäryhmässä (kuvio 4a) ilmenee tilastollisesti merkitsevä ero koulutuksen suhteen sekä *hän*- että *he*-pronomien osuutta kuvaavien mediaanilukujen välillä.<sup>12</sup> Kuvion 4a vasen ruutu- ja janakaavio havainnollistaa sitä, että aineiston korkeasti koulutetut 1950-luvulla syntyneet puhujat suosivat etenkin yksiköllistä *hän*-pronomia varsin yksimielisesti: ruutu on pieni ja janat lyhyet. Matalammin koulutettujen joukossa muista poikkeaa jo edellä esillä ollut eteläkarjalainen N1958b (ks. esim. 1), joka käyttää paljon sekä *hän*- että *he*-pronomineja ja niiden murrevariantteja.

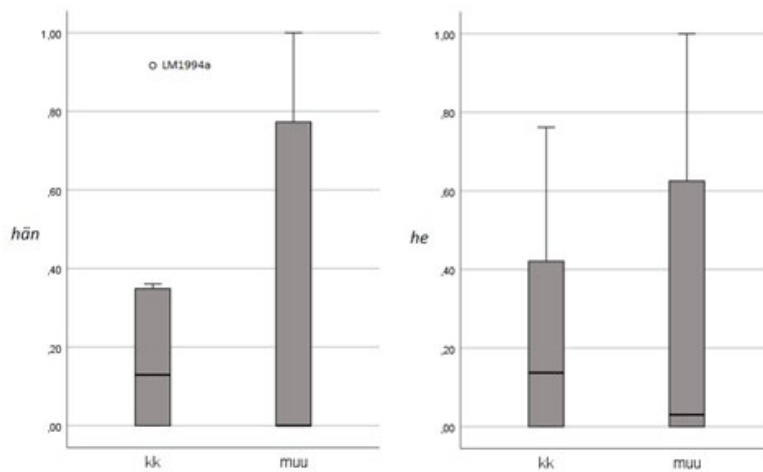
11  $p = 0,032; \chi^2 = 5, df = 1$ .

12 Fisherin tarkan testin mukaan kummassakin  $p = 0,026, df = 1$ .

274 YKSILÖLLISIÄ VALINTOJA JA KIELITAUSTAN VAIKUTUSTA – HENKILÖVIITTEISET HÄN-, HE-, SE- JA NE-PRONOMINIT KIELIELÄMÄKERTAHAASTATTELUISSA



KUVIO 4a. Ruutu- ja janakaavio 1950-luvulla syntyneiden eri koulutustaustaisten haastateltavien *hän*- ja *he*-pronominien osuudesta.



KUVIO 4b. Ruutu- ja janakaavio 1990-luvulla syntyneiden eri koulutustaustaisten haastateltavien *hän*- ja *he*-pronominien osuudesta.

Seuraavat katkelmat havainnollistavat, miten akateemistaustaiset käyttävät tyypillisesti yleiskielisiksi miellettyjä *hän*- ja *he*-pronomeja ja vähemmän koulutetut puhujat puolestaan puhekielisinä pidettyjä *se*- ja *ne*-variantteja. LN1957 on korkeasti koulutettu nainen, joka on kotoisin Pohjois-Pohjanmaalta mutta asunut jo useita vuosikymmeniä Tampereen seudulla. Kummallakaan alueella *hän*- ja *he*-



pronominin käyttö ei ole vanhastaan tyypillistä, mutta LN1957 käyttää niitä haastattelussa yksinomaisesti. Hän toimii tällä hetkellä opetustehtävissä. Esimerkissä 3 hän puhuu useista spesifeistä henkilöistä, joista *hän*-pronominilla hän viittaa veljeensä (perheenjäsen) ja opettajaansa, *he*-pronominilla opettajaan ja tämän mieheen. LN1957:n puhetaapaa ei voi kuitenkaan muutoin luonnehtia kovin yleiskieliseksi, sillä siinä on lukuisia yleisiä ja yleistyviä puhekielen piirteitä (esim. *mä, mun, mulla; tärkeä, muutto*; vrt. Mantila 2004: 325–326), samoin kuin joitakin hänen länsimurteisesta taustastaan kertovia variantteja (esim. *jokka*).

(3)

LN1957: mu isovelji, oli myöskin, sillä tavalla että **hän** oli siellä samassa, ää oppilas-asuntolassa et se totta kai helpotti, öö, jollain tavalla sitä mutta mä koen sillä et mä oon ollut aika itsenäine aina et iham pienestä asti, et sitte ku mä Ouluun muutin, kuustoistavuotiaana ni mä muutin, yksin eikä siellä ollut ke tään, valmiiks tuttua? mulla oli kyllä sit semmone opettaja-, paris- tai siis opettaja ja **hänen** miehensä jokka, vanhemmat ei oikeestaan sitä muutto ei kumpaakaan muutto pystyneet tukemaan millään tavalla, päivävastoi isä oli aika lailla vastaan sitä et mä edes lähen lukioon, ja oliki vastaan ja sitte mulla oliki semmonen, pariskunta, jon- jokka oli mun, entine opettaja matika opettaja ja **hänen, hänem** miehensä, jokka sitte, oli mun hyviä ystäviä vähä pitivät minua niinku ottotyttönä? jollain tavalla niin, **heidän** tuki oli aika tärkeä silloin ku mä muutin Ouluun kuustoistavuotiaana.

N1957a on puolestaan eteläpohjalainen pankkitoimihenkilö, jolla on merkonomien koulutus. Esimerkissä 4 hän puhuu puoliosastaan, johon hän viittaa pronominilla *se* peräti kuudesta. Eteläpohjalaiselle murteelle tyypillisesti pronomini lyhenee sulautuessaan sitä seuraavaan *olla*-verbiin (vrt. esim. Lyytikäinen, Rekunen & Yli-Paavola 2013: 259). N1957a käyttää jonkin verran myös *hän*-pronominia viitatessaan muihin kuin perheenjäseniin, tyypillisesti spesifioimattomiin tietyn ryhmän edustajiin, mitä voi pitää hieman yllättävänäkin (ks. tarkemmin alalukua 3.6). Eteläpohjalaisen murteen mukainen *se*-pronominin käyttö on linjassa N1957a:n kaikkineenkin vahvan murteellisuuden kanssa (ks. tarkemmin Nuolijärvi 2020: 359–361).

(4)

Haastattelija: okei ja sitte, tämä sun aiva- aiva oma, nykyinen perhe että, puoliso, onko, eri ta-, eri puhetaapaa ku sulla

N1957a: aiivan täysin ((nauraa))

Haastattelija: ((nauraa)) onko-,

276 YKSILÖLLISIÄ VALINTOJA JA KIELITAUSTAN VAIKUTUSTA – HENKILÖVIITTEISET  
HÄN-, HE-, SE- JA NE-PRONOMINIT KIELIELÄMÄKERTAHAASTATTELUISSA

N1957a: ((etunimi)) puhuu niin eri lailla ku minä enempi niinku, voisko sanoa, kirjakiältä.  
s\_on aina puhunu eri lailla ku minä.

Haastattelija: joo. ooksä huomannu mitä muutosta. pualisom puhetavassa?

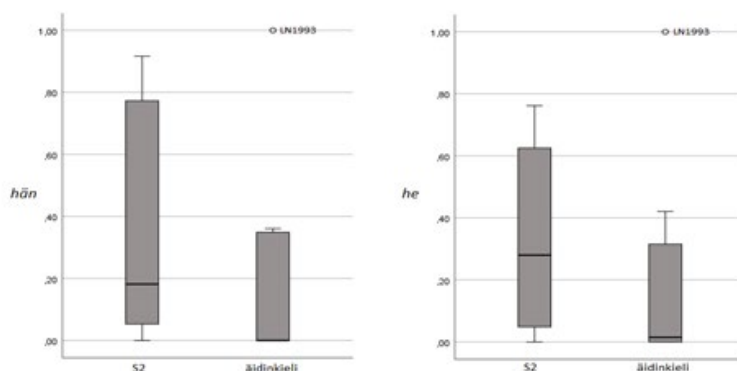
N1957a: e. s\_om pitäny pintsansa. ((nauraa))

Haastattelija: ((naurahdus))

N1957a: s\_on mum miälestä puhunu aina nuan. eikä s\_oo tyyliäsäm muuttanu. että aattelis  
ku se asuu- tullup Pohojammallet takaasi ni ei, se puhuu samalla lailla ku aina  
ennenki. mä en oos saanu sitä oppimahan, ((nauraa))

### 3.5 Äidinkieliset ja S2-puhujat

Edellä alaluvussa 3.2 osoitimme, että puhujan kielitausta on tarkasteltavien pronomien vaihtelun selittämisessä keskeinen tekijä vain nuoremmassa ikäryhmässä. Kielitausta vaikuttaa erityisesti monikollisten muotojen käyttöön. Ruutu- ja jana-kaavio kuviossa 5 osoittaa, etteivät sen paremmin 1990-luvulla syntyneet suomenkieliset kuin S2-puhujatkaan ole pronomien käytössään yhteneväisiä joukkoja. Äidinkielisillä suomenpuhujilla sekä *hän*- että *he*-pronomien osuuden mediaaniarvo asettuu alemmas kuin S2-puhujilla. Ero ei kuitenkaan ole tilastollisesti merkitsevä, sillä otoskoko on tässä pieni: tarkastelluista 16:sta 1990-luvulla syntyneestä kuudella on jokin muu äidinkieli kuin suomi. Muista nuorista äidinkielisistä suomenpuhujista poikkeaa LN1993, joka haastattelutilanteessa käyttää yksinomaan yksinomaan *hän*- ja *he*-pronomineja (ks. tarkemmin alalukua 3.3, esim. 2).



KUVIO 5. Ruutu- ja jana-kaavio 1990-luvulla syntyneiden äidinkielisten suomenpuhujien ja S2-puhujien *hän*- ja *he*-pronomien osuudesta.

Eniten huomiota S2-puhujien *he*-pronomien käytössä kiinnittävät tapaukset, joissa viittauskohde on selvästi geneerinen (ks. tarkemmin alalukua 3.6). Alaluvussa 3.2

esitetyn satunnaismetsäanalyysin mukaan viittauksen kohde ei 1990-luvulla syntyneillä puhujilla ole erityisen tärkeä vaihteluun vaikuttava tekijä mutta ei täysin merkityksetönkään. Satunnaismetsäanalyysi ei siis kerro tarkemmin, millainen yhteys viittauksen kohteella on pronominien vaihteluun. Arvelemme analyysin tuloksen mahdollisesti liittyvän siihen, että tässä aineistossa 1990-luvulla syntyneet S2-puhujat käyttävät *he*-pronominia äidinkieliä suomenpuhujia tavallisemmin etenkin geneerisissä viittauksissa.

Esimerkissä 5 informantti N1995, jonka ensikieli on kurdi, puhuu yleisesti nuorista, ja esimerkissä 6 informantti LM1994a, joka on kasvanut unkarinkielisessä kodissa, viittaa *he*-pronominilla yleistävästi romaneihin. Kummassakaan tapauksessa haastattelijan edeltävä repliikki ei tarjoa mallia pronominivalintaan. Nuorista puhuttaessa *he*-pronomini tuntuu vielä epäodotuksenmukaisemmalta kuin romaneihin viitattaessa, sillä jälkimmäisessä tapauksessa yleiskielistä pronominivalintaa voisi motivoita halu puhua mahdollisimman kunnioittavasti tästä etnisestä vähemmistöstä, johon kohdistuu paljon rasismia. *He*-variantti korostuu toistuessaan esimerkissä 6 peräti neljä kertaa.

(5)

Haastattelija: osaat sä, sanoo jotain sanoja jotka ois vaik tyypillisiä nuorisolle, tai nuorille. joita vanhemmat ihmiset ei käyttäis. oot sä huomannu.

N1995: mm, ((naurahdus)) mm. kyl mun mielestä aika lailla niinku ne sanat mitä nuoret käyttää om myöski, sanotaan vaik et, ne, kolme- nelikymppisten, niinku, ((huokaus))

Haastattelija: joo.

N1995: ky- kyl **heki** käyttää samoi sanoja?

(6)

Haastattelija: ni ootsä ikinä kiinnittäny huomioo et sä huomaisit eroja niinku, erilaisten ihmisryhmien kielenkäytössä.

LM1994a: no tietysti. että kyllähän sen huomaa tosi paljon esimerkiksi ehkä niinku mitä, mä oon ollu aika paljon elämässäni romanien kans tekemisissä ja niinku esim **heillä** huomaa, et niinku vähemmistönä, **heidän** suomen kielen käyttönsäki hyvin tyypillistä. et tavallaan miten **he he** puhuttelee toisiaan tai niinku, jopa se äänen painotus ja kaikki on niinku, hyvin niinku omal-, lal-, lal-, laisensa.

### 3.6 Viittauksen kohde ja puhekuppanien vaikutus toisiinsa

Kuten luvussa 2 kerroimme, olemme jaotelleet puheena olevat henkilöt kolmeen luokkaan: perheenjäseniin, muihin spesifeihin referentteihin ja geneerisiin referentteihin. Geneerisillä tarkoitamme viittauksia, jotka kohdistuvat kehen tahansa tietyn ryhmän edustajaan (ks. Vilkuna 1992: 149–155; VISK § 1407, 1408; ks. myös esim. 5–6 edellä). Kuten Vilkuna (mts. 137) toteaa, suomessa geneerisyys on pääteltävä kontekstista. Yleistäviä viittauksia käytetään esimerkiksi kuvattaessa tapoja ja normeja (VISK § 1407). Haastatteluissa esiintyy niissä käsiteltävistä aiheista johtuen varsin paljon tällaisia epäspesifejä viittauksia.

Seuraavat esimerkit havainnollistavat tyypillisiä geneerisiä viittauksia. Esimerkissä 7 LN1956 viittaa *se*-pronominilla yleistäen paikalliseen ihmiseen, joka voi olla periaatteessa kuka tahansa; esimerkissä 8 M1952b:n *ne*-pronomini viittaa kuvitteellisiin englantilaisiin puhekuppaneihin.

(7)

LN1956: ei, ei paikallinen ihminen täällä joka, joka kulkee tuolla mettässä niin, ei **se**, tietosesti, raiskaa sitä eikä, ei[kä,

Haastattelija: [nii,

LN1956: tee sinne semmos-, tuhoja semmosia että **se**, häirittis sitä,

Haastattelija: m[m,

LN1956: [luontoa siellä että kyllä me ajatelhaan että meän lapsetki pitäs voija täällä, ellää ja, olla että,

(8)

M1952b: koska jos mä puhun vaik englantii nii ehän mä puhu sitä kuitenkaa sataprosenttisesti vaik mä sanon et mä pärjään sillä.

Haastattelija: mm mm.

M1952b: jos (--) mä meen, mennee jonneki, Englantii ja, ni totta kai **ne** huomaa et mä en oo, syntyjään puhuja.

Alaluvussa 3.2 esitetyn satunnaismetsäanalyysin mukaan viittauksen kohde selittää etenkin 1950-luvulla syntyneiden puhujien pronomnivaihtelua. Sitä, millaisia yhteyksiä aineistossa löytyy, selvennämme seuraavaksi ristiintaulukointien avulla. Aineisto on tässä kohdin rajattu vanhempaan puhujaryhmään. Yksikkö- ja monikko-muodot on ristiintaulukoitu erikseen, sillä monikkomuodot esiintyvät tässä aineistossa yksiköllisiä pronomineja useammin geneerisinä. Erillisellä tarkastelulla saadaan kiinnitettyä huomio eroon *hän*- ja *se*-pronominien sekä *he*- ja *ne*-pronominien välillä.

TAULUKKO 5a. Pronominien *hän* ja *se* yhteys viittauksen kohteen tyyppiin 1950-luvulla syntyneillä puhujilla, ristiintaulukointi sovitettuja standardoituja jäännöksiä käyttäen ( $p = 0,002$ ;  $\chi^2 = 12$ ,  $df = 2$ ).

	<i>hän</i>	<i>se</i>	Yhteensä
Perheenjäsen	185 (3,5)	74 (-3,5)	259
Muu spesifi	70 (-2,4)	55 (2,4)	125
Geneerinen	32 (-1,8)	27 (1,8)	59
Yhteensä	287	156	443

TAULUKKO 5b. Pronominien *he* ja *ne* yhteys viittauksen kohteen tyyppiin 1950-luvulla syntyneillä puhujilla, ristiintaulukointi sovitettuja standardoituja jäännöksiä käyttäen ( $p < 0,001$ ;  $\chi^2 = 19$ ,  $df = 2$ ).

	<i>he</i>	<i>ne</i>	Yhteensä
Perheenjäsen	45 (4,0)	43 (-4,0)	88
Muu	52 (0,2)	101 (-0,2)	153
Geneerinen	41 (-3,4)	130 (3,4)	171
Yhteensä	138	274	412

Yksikkömuotojen osalta (taulukko 5a) ero on merkitsevä perheenjäsenen ja muuhun spesifiin henkilöön kohdistuvien viittausten välillä. Monikkomuotojen variaatiossa (taulukko 5b) ero on merkitsevä geneeristen ja perheenjäsenen kohdistuvien viittausten kohdalla (suluissa ilmaistu standardoitu jäännös on itseisarvoltaan suurempi kuin 2). 1950-luvulla syntyneet puhujat siis suosivat perheenjäseniin viitatessaan *hän*- ja *he*-pronominia (jäännös on positiivinen), muihin spesifeihin henkilöihin viitatessaan *se*-pronominia ja geneerisissä viittauksissa *ne*-pronominia. Eroa yksikön ja monikon välillä voi selittää se, että yksikössä geneerisiä viittauksia esiintyy tässä aineistossa yleisesti ottaen melko vähän.

Esimerkki 9 havainnollistaa tyypillistä tapaa viitata perheenjäseniin, tässä tapauksessa äitiin, *hän*-pronominilla. Haastateltavan äiti on jo edesmennyt, mikä

280 YKSILÖLLISIÄ VALINTOJA JA KIELITAUSTAN VAIKUTUSTA – HENKILÖVIITTEISET  
HÄN-, HE-, SE- JA NE-PRONOMINIT KIELIELÄMÄKERTAHAASTATTELUISSA

voi lisätä pyrkimystä viitata tähän arvostavasti eli yleiskielistä varianttia käyttäen. Kyseisessä katkelmassa on kolme peräkkäistä *hän*-viittausta, joten voidaan puhua myös virittyneisyydestä (vrt. Priiki 2016b).

(9)

Haastattelija: puhutaan vielä, lisää sun perheestä, ja lähisukulaisista, alotetaan ensin sun vanhemmista, mistä päin he olikaan kotosin,

M1954a: mun äiti oli kotoisin ((narinaa)) se on tänä päivänä Hyvinkäätä Hyvinkään kaupunkia mut sillon kun **hän** syntyi, kauan sitten niin tota, kymmenluvulla tuhatyheksänsataakymmenluvulla ni se oli, Padasjokea.

Haastattelija: okei.

M1954a: Padasjoen kunnan aluetta ja tota ei kun tota ei missään nimessä Padasjokea kun Hausjärvi.

Haastattelija: okei joo.

M1954a: Hausjärveä mut sitten ne kunnanrajat on muuttunu siitä tuli sitte Hyvinkään maalaiskuntaa ja tänään se on Hyvinkään kaupunkia ja, ja tota, mut ihan agraariseutuhan se ympäristö siinä on edelleenki ja tota, **hän** om maanviljelijän tytär, tänään sitä tilaa pitää **hänen**, yhden, neljästä veljestään, vanhin poika.

Haastattelija: joo.

Monella haastateltavalla viittaustavat kuitenkin vaihtelevat, jopa saman vuoron aikana, kuten tapahtuu esimerkissä 10, jossa helsinkiläinen N1992 puhuu veljestään. Tässä tapauksessa *hän*-pronomini näyttäisi liittyvän koodinvaihtoon: puhuja muuttaa yhtäkkisesti puhetapaansa, artikuloi hetken korostetun huolellisesti (merkitty litteraattiin kursiivilla), ja käyttää yleiskielisiä variantteja (esim. *syntynytkin* vs. *syntynytki*; ks. koodinvaihdosta Lappalainen 2004: luku 6).

(10)

Haastattelija: öö entäs sisaruksii, onko,

N1992: on pikkuveli joka on, kaksikyvyt ((naureskelee)), kaksytkolmevuotias? joo. ööm, no **hän** on meistä se joka on syntynytkin Vantaalla että **hän** on ihan, kunnan vantaalainen ja, nyt **seki** asuu kyllä Helsingissä nykyään, jaa juu.

Pronominivalinta on siinä määrin leksikaalinen ilmiö, että sen voi olettaa olevan herkkä mukautumaan vuorovaikutustilanteessa. Tähän on kiinnitetty huomiota aiemmassa tutkimuksessa (esim. Lappalainen 2006b: 257–258; 2010a: 296–302; 2011: 181–189; Mielikäinen & Palander 2014: 196). Tältä pohjalta voi olettaa, että haastatelijan pronominalinnoilla saattaa olla vaikutusta haastateltavan puheeseen ja päinvastoin. Mitään selvää yhteyttä ei voi kuitenkaan tutkimusaineistossamme havaita.

Suurin osa haastatteliijoista käyttää kysymyksissään *hän-* ja *he-*pronomineja. Vaikka joissakin tapauksissa haastattelijan pronomiivalinta edeltää haastateltavan pronominin käyttöä, suurin osa haastateltavien esiintymistä on kuitenkin spontaaneja niin, että lähikontekstissa ei ole mitään haastattelijan antamaa mallia. Ei ole tietenkään mitään keinoa todistaa, etteikö jossain tilanteessa tapahtuisi mukautumista, mutta aineistossa on runsaasti esimerkkejä siitä, miten haastattelija ja haastateltava käyttävät eri variantteja, kuten voi havaita seuraavista LM1955:n haastattelusta poimituista katkelmista.

(11a)

Haastattelija: joo, onks **hän** ollut pitkään siellä.

LM1955: mm, kyllä **se** varmaaj jo kymmenev vuotta nyt ollu.

(11b)

Haastattelija: no entäs isovanhemmat t- tii- onko **he** ollu, elossa sun, [lapsuuden aikana.

LM1955: [on joo.

Haastattelija: joo.

LM1955: kyllä mum- niinku äitin va-, isovanhemmat, ja,

Haastattelija: ja **he** oli täältä Lapista s[itte.

LM1955: [joo **ne** oli tosta, Särkijärveltä iham Muoniosta ja,

Haastattelija: joo.

Ensin haastattelija viittaa haastateltavan tyttären *hän-*pronominilla, mutta LM1955 käyttää kysymystä välittömästi seuraavassa vastauksessaan tästä *se-*muotoa (esim. 11a). Hieman myöhemmin haastattelija viittaa kysymyksissään kahdesti haastateltavan isovanhempiin *he-*variantilla, mutta haastateltava käyttää jälleen eri varianttia eli *ne-*pronominia (esim. 11b).

## 4 Päätelmät

Olemme tarkastelleet tässä artikkelissa kolmannen persoonan pronomien käyttöä 2010-luvulla kerätyssä haastatteluaineistossa. Olemme tutkineet variaation yhteyttä ennen muuta kielenulkoisiin muuttujiin – eli puhujien ikään, sukupuoleen, koulutukseen ja kielitaustaan – mutta myös lingvistisiin tekijöihin eli yksikön ja monikon eroon ja referentin luonteeseen. Logoforiset tapaukset jäivät niiden vähäisyyden vuoksi tarkastelun ulkopuolelle, eivätkä aineiston epätasainen jakauma ja pie- nuus mahdollistaneet myöskään alueellisen taustan systemaattista huomioimista. Olemme keskittyneet vertailemaan kahden ikäryhmän, 1950-luvulla ja 1990-luvulla

syntyneiden kielenkäyttöä toisiinsa sekä yksiköllisten ja monikollisten muotojen käyttöä.

Olemme huomioineet tarkastelussa myös yksilöllisen vaihtelun, sillä aiemmissa tutkimuksissa puhujien väliset erot on todettu keskeiseksi *hän-* ja *se-*pronominien vaihteluun vaikuttavaksi tekijäksi. Puhujakohtainen vaihtelu on tässäkin aineistossa suurta, ja olemme pyrkineet tarkastelussamme suhteuttamaan yksilöllisiä eroja perinteisiin ryhmäjakoisiin sekä pohtimaan, miksi tietyt puhujat poikkeavat esimerkiksi omasta ikäryhmästään tai samantasoisesti koulutetuista.

Tarkastelemamme vanhempi ikäryhmä käyttää *hän-* ja *he-*pronomeineja odoituksen mukaisesti enemmän kuin nuorempi ikäryhmä. Ikäryhmien ero on erityisen suuri *hän-*pronominin frekvenssissä: 1950-luvulla käyttävät suhteessa selvästi enemmän yleiskieliseksi miellettyä *hän-*varianttia kuin 1990-luvulla syntyneet, jotka suosivat puolestaan enemmän puhekielistä *se-*varianttia. Osoitamme, että frekvenssieron lisäksi ryhmiä erottaa se, että vaihteluun vaikuttavat hieman erilaiset tekijät. Myös yksikkö- ja monikkomuotojen väliltä löytyy eroja paitsi frekvenssissä myös siinä, mitkä muuttajat ovat vaihtelun kannalta tärkeimpiä.

1950-luvulla syntyneiden ryhmässä suurempaa *hän-* ja *he-*pronominien frekvenssiä selittää yksilöllisten erojen lisäksi etenkin koulutustausta. Vanhemman ikäryhmän korkeakoulutetut ovatkin ainoa pronominien käytöltään varsin yhtenäinen ryhmä aineistossa. 1990-luvulla syntyneillä koulutus selittää jossain määrin vain yksikkömuotojen vaihtelua, ja sen sijaan keskeiseksi osoittautui kielitausta. Tätä tulosta selittää S2-puhujien painottuminen nuorempaan ikäryhmään. Aiemmissa tutkimuksissa nuorten on todettu suosivan henkilöviittauksissa *se-* ja *ne-*pronomeineja (Mielikäinen 1988: 107–108; Paunonen 1995 [1982]: 165–177; Ylikahri 1996: 185–186; Lappalainen 2010a: 286–288; Kurki ym. 2011: 95–96). Tässä tarkastellussa haastatteluaineistossa myös 1990-luvulla syntyneissä on kuitenkin *hän-* ja *he-*pronominien suosijoita, eikä vaihtelu nuoren ryhmän sisällä selity yksiselitteisesti esimerkiksi kielitaustalla tai alueellisella taustalla. Viittauksen kohteella on merkitystä etenkin vanhemmassa ikäryhmässä, mikä voi liittyä tarkastelussa käyttämäämme referenttien melko mekaaniseen jaotteluun, jossa ydinperhe korostuu. Perheenjäseniin viitataan yleisemmin persoonapronomeineilla, geneerisiin referentteihin demonstratiivipronomeineilla. Aineistosta löytyy kuitenkin viitteitä myös siitä, että monikollisten geneeristen referenttien kohdalla pronominien käytössä voisi olla eroja myös nuoremmilla puhujilla niin, että S2-puhujat eroavat perinteisessä mielessä suomenkielisistä.

Ikäryhmittäiset erot herättävät kysymyksen siitä, onko meneillään sukupolvittainen muutos, jonka seurauksena *hän-*pronominin käyttö hiljalleen hiipuu. Tällaiseen päätelmään on syytä liittää vahvoja varauksia, sillä reaaliaikaiset tutkimukset ovat antaneet viitteitä siitä, että *hän* on herkkä ikäkausivaihtelulle ja sen käyttö saattaa lisääntyä huomattavastikin iän myötä. Näin on tapahtunut esimerkiksi 1950-luvulla syntyneille helsinkiläispuhujille (Lappalainen, Mustanoja & O'Dell 2019: 576) ja ylei-



semminkin Helsingin puhekielessä (Kuparinen, Mustanoja, Peltonen, Santaharju & Leino 2019: 42). Tämän piirteen muutosherkkyys liittyy sen vahvaan normittamiseen ja erityisluonteeseen toiseen henkilöön viittaavana muotona. Näin ollen siihen voi yhdistyä tulkinta puhujan asennoitumisesta referenttiä kohtaan (ks. Lappalainen 2010a: 315; Siitonen 2008; 2016: 88; Kurki ym. 2011: 95; Nenonen 2020: 50–57). Esimerkiksi työelämään liittyvät paineet saattavat johtaa siihen, että puhuja tietoisesti tai tiedostamattaan muuttaa puhetapaansa (ks. esim. Mustanoja 2011: 373). Tätä kautta ikä kytkeytyy myös koulutukseen, jota vanhemmilla informanteilla on useimmiten enemmän kuin nuoremmilla.

Aiemmassa tutkimuksessa on keskitytty äidinkieleltään suomenkielisiin, ruotsinkielisiin ja suomen ja ruotsin suhteen kaksikielisiin puhujiin. Sen sijaan tietoa sellaisten puhujien kolmannen persoonan pronomien käytöstä, jotka käyttävät suomen rinnalla jotain muuta kieltä kuin ruotsia, ei toistaiseksi juuri ole (ks. kuitenkin Kuparinen 2001). Siitosen (2016: 390) kansanlingvistikissa kyselytutkimuksessa vastaajat yhdistävät *hän*-pronominin kaksikielisyyden ohella myös oppijankieleen. Omassa tutkimuksessamme S2-puhujia on vain kahdeksan, mutta tulokset antavat viitteitä siitä, että S2-tausta voi olla yksi tekijä, joka selittää runsaampaa *hän*- ja *he*-pronominien käyttöä, mikä tukee Kuparisen aiempia havaintoja (mas. 65 - 67). Ero voi näkyä jossakin rajatussa kontekstissa, kuten geneerisiin referentteihin viittaamisessa. Tämän puhujajoukon heterogeenisuus kuitenkin viittaa siihen, ettei S2-tausta sinänsä riitä selittämään yksilöiden välisiä eroja. S2-puhujien pronominvaihtelua kannattaisi kuitenkin tutkia lisää laajemmalla aineistolla. Kun pronomien valintaan liittyy normittamisen myötä kunnioituksen osoittamisen aspekti, on mahdollista, että joitakin S2-puhujia heidän kieli- ja kulttuuritaustansa ohjaa yleiskielisempiin valintoihin, mistä kertoo se, että monet S2-puhujat korostavat saman hankkeen aineistoon perustuvassa tutkimuksessa teittelyn tärkeyttä enemmän kuin samanikäiset suomenkieliset (Lappalainen 2020: 280). Tältä pohjalta voi myös miettiä, millaisia asioita pitäisi korostaa, kun S2-opetuksessa tarkastellaan kolmannen persoonan pronomineja ja annetaan neuvoja niiden käytöstä. Tämä kytkeytyy laajempaan pohdintaan normittamattoman puhekielen asemasta S2-opetuksesta, mistä on keskusteltu jo pitkään (ks. Lauranto 1995).

Yksilötasolla voidaan todeta, että erilaiset tekijät vaikuttavat eri suuntiin. Tässä tarkasteltujen muuttujien lisäksi ilmiöön vaikuttaa muita tekijöitä, jotka voivat saada ryhmäkeskiarvojen perusteella epätodennäköisen henkilön suosimaan tai karttaamaan tiettyjä pronomineja. Kaikkiin näihin tekijöihin ei pääse oman aineistomme avulla käsiksi, sillä haastattelussa ei ole esimerkiksi kysytty erikseen kolmannen persoonan pronomineja koskevista asenteista.

Tutkimuksen tuloksia tulkittaessa on muistettava myös aineiston luonne. Kyseessä on erityisesti kielenkäyttöön ja kieliasenteisiin liittyvä tutkimushaastattelu, jossa paitsi haastattelutilanne myös sen teema on voinut ohjata yleiskieliseen puhe-tapaan. Tätä on voinut lisätä se, että useimmiten haastattelija ja haastateltava eivät

ole tunteneet kovin hyvin toisiaan, ja sekin, että useimmat haastattelijat olivat tässä roolissa ensikertalaisia. Syytä on mainita sekin, että suuri osa henkilöviittauksista keskittyy haastattelun alkuun, jossa puhutaan perheestä ja elämänvaiheista. Aiempien tutkimusten mukaan haastattelun alussa puhetapa on usein formaalimpi kuin loppupuolella, jolloin tallentamiseen liittyvä jännitys on jo laantunut (Saanihahti 1990: 76). Tiivistäen voi todeta, että tutkimamme puhujat saattavat hyvinkin esimerkiksi arki-keskustelussa käyttää enemmän puhekielisiä *se-* ja *ne-*pronomineja kuin analysoimissamme haastatteluissa.

Tarkastelemassamme haastatteluaineistossa *hän-* ja *he-*pronominit esiintyvät toisinaan yhdessä muiden yleiskielen piirteiden kanssa. Ainakin yhdellä informantilla niiden käytöllä on kuitenkin selvä kytkös murretaustaan. Alueellisen vaihtelun tavoittaminen vaatisi laajemman aineiston, jossa olisi riittävästi edustajia myös niiltä murrealueilta, joilla *hän/hää-* ja *he/hyö-*varianttien osuus on ollut vanhastaan vahva. Houkuttelevaa olisi tutkia myös muita ikäryhmiä niin tämän hankkeen haastatteluista kuin muistakin aineistoista. Haastatteluaineisto edustaa omaa lajityyppiään esimerkiksi siinä, että logoforisia pronomineja, jotka aiempien tutkimusten (Priiki 2016b, 2017b; ks. myös Nenonen 2020: 47–50) mukaan muodostavat arkisemmassa vuorovaikutuksessa merkittävän osan *hän-* ja *he-*pronomiinien käytöstä, esiintyy haastatteluissa vain vähän. Kattavan kuvan saaminen kolmannen persoonan pronominivariaatiosta Siitosen (2008, 2016) tutkimusten vertailukohdaksi ja täydennykseksi edellyttäisikin erilaisten, tähänastisia tutkimuksia monipuolisempien kielenkäyttäjryhmien ja -käyttötilanteiden tarkastelua.

## Kirjallisuus

- Baayen, R. H. 2008. *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Hippi, K. 2020. Monikielisyteen suhtautuminen kielielämäkertojen valossa. Teoksessa K. Hippi, H. Lappalainen & P. Nuolijärvi (toim.) *Suomalaisten kielellistä elämää. Sata suomalaista kielellistä elämäkertaa -hankkeen satoa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 105–142.
- Hippi, K., H. Lappalainen & P. Nuolijärvi 2020. Sata suomalaista kielellistä elämäkertaa hankkeena ja aineistona. Teoksessa K. Hippi, H. Lappalainen & P. Nuolijärvi (toim.) *Suomalaisten kielellistä elämää. Sata suomalaista kielellistä elämäkertaa -hankkeen satoa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 11–50.
- Ivaska, I. 2014. Edistyneen oppijansuomen avainrakenteita. Korpusnäkökulma kahden kielimuodon tyypillisiin rakenteellisiin eroihin. *Virittäjä*, 118 (2), 161–193.
- Klavan, J., M.-L. Pilvik & K. Uibo 2015. The use of multivariate statistical classification models for predicting constructional choice in spoken, non-standard varieties of Estonian. *SKY Journal of Linguistics*, 28, 187–224.
- Kuparinen, K. 2001. Suomen kielen kaksi koodia oppijan asenteissa ja kielenkäytössä. Teoksessa P.-R. Piiparinen-Rintaluoma (toim.) *Suomi toisena kielenä ja variaatio*. Kakkoskieli 4. Helsinki: Helsingin suomen kielen laitos, 7–86.

- Kuparinen, O., L. Mustanoja, J. Peltonen, J. Santaharju & U.-P. Leino 2019. Muutosmallit kolmen aikapisteen pitkittäisaineiston valossa. *Sananjalka*, 61, 30–56. <https://doi.org/10.30673/sja.80056>.
- Kurki, T., K. Siitonen, M. Väänänen, I. Ivaska & J. Ekberg 2011. Ensi havaintoja Satakuntalaisuus puheessa -hankkeesta. *Sananjalka*, 53, 84–108. <https://doi.org/10.30673/sja.86706>.
- Laitinen, L. 2005. *Hän*, the third speech act pronoun in Finnish. Teoksessa R. Laury (toim.) *Minimal reference. The use of pronouns in Finnish and Estonian discourse*. Helsinki: Suomalaisen Kirjallisuuden Seura, 75–106.
- Lappalainen, H. 2004. *Variaatio ja sen funktiot. Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lappalainen, H. 2006a. Kelan asiointikeskustelut Helsingin puhekielen variaation kuvastajina. Esimerkkinä yksikön 1. ja 2. persoonan persoonapronominien käyttö. Teoksessa K. Juusela & K. Nisula (toim.) *Helsinki kieliyhteisönä*. Helsinki: Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos, 275–298.
- Lappalainen, H. 2006b. *Mie vai mä, sinä vai te?* Virkailijoiden kielelliset valinnat itsen ja vastaanottajaan viitattaessa. Teoksessa M.-L. Sorjonen & L. Raevaara (toim.) *Arjen asiointia. Keskusteluja Kelan tiskin äärellä*. Helsinki: Suomalaisen Kirjallisuuden Seura, 241–284.
- Lappalainen, H. 2010a. *Hän vai se, he vai ne?* Pronominivariaatio ja normien ristiveto. Teoksessa H. Lappalainen, M.-L. Sorjonen & M. Vilkuna (toim.) *Kielellä on merkitystä. Näkökulmia kielipolitiikkaan*. Helsinki: Suomalaisen Kirjallisuuden Seura, 279–324.
- Lappalainen, H. 2010b. *Se ja hän* puhutussa kielessä. *Kielikello*, 4/2010, 4–7.
- Lappalainen, H. 2011. Förändring, variation och variationens funktioner. En fallstudie av variationen i bruket av personliga pronomen på en finsk landsbygdsort. Teoksessa G. Bockgård & J. Nilsson (toim.) *Interaktionell dialektologi*. Uppsala: Institut för språk och folkminnen, 159–197.
- Lappalainen, H. 2020. Millaista kohteliaisuutta suomalaiset arvostavat asiakaspalvelussa? Teoksessa K. Hippi, H. Lappalainen & P. Nuolijärvi (toim.) *Suomalaisten kielellistä elämää. Sata suomalaista kielellistä elämäkertaa -hankkeen satoa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 253–287.
- Lappalainen, H., K. Hippi & P. Nuolijärvi 2021. S2-puhujien variaatio ja variaatiota koskevat havainnot. Esitelmä AFinLan syysseminaarissa 11.–13.11.2021. Jyväskylän yliopisto.
- Lappalainen, H., L. Mustanoja & M. O'Dell. 2019. Miten ja milloin yksilön kieli muuttuu? Helsingiläisidialektien muutos ja muutoksen tutkimuksen menetelmät. *Virittäjä*, 123 (4), 550–581. <https://doi.org/10.23982/vir.67808>.
- Lauranto, Y. 1995. Normi, rekisteri ja S2-opetus. *Virittäjä*, 99 (2), 261–263.
- Lyytikäinen, E., J. Reikonen & J. Yli-Paavola 2013. *Suomen murrekirja*. Helsinki: Gaudeamus.
- Mantila, H. 2004. Murre ja identiteetti. *Virittäjä*, 108, 322–346.
- Mielikäinen, A. 1988. Naiset puhekielen säilyttäjinä ja uudistajina. Teoksessa L. Laitinen (toim.) *Isosuinen nainen. Tutkielmia naisesta ja kielestä*. Helsinki: Yliopistopaino, 82–111.
- Mielikäinen, A. & Palander, M. 2014. *Miten suomalaiset puhuvat murteista. Kansanlingvistinen tutkimus metakielestä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Moshnikov, I. 2021. *NUT-partisiipin variaatio rajakarjalaismurteissa*. Dissertations in Education, Humanities, and Theology, 168. Joensuu: Itä-Suomen yliopisto. <http://urn.fi/URN:ISBN:978-952-61-3744-5>.
- Mustanoja, L. 2011. *Idiolekti ja sen muuttuminen. Reaaliaikatutkimus Tampereen puhekielestä*. Acta Universitatis Tamperensis 1605. Tampere: Tampereen yliopisto. <https://urn.fi/urn:isbn:978-951-44-8417-9>.

- Nuolijärvi, P. 1986. *Kolmannen sukupolven kieli. Helsinkiin muuttaneiden suurten ikäluokkien eteläpohjalaisten ja pohjoissavolaisten kielellinen sopeutuminen*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nuolijärvi, P. 2020. Monipaikkainen kielenkäyttäjä 2010-luvun Suomessa. Teoksessa K. Hippi, H. Lappalainen & P. Nuolijärvi (toim.) *Suomalaisten kielellistä elämää. Sata suomalaista kielellistä elämäkertaa -hankkeen satoa*. Helsinki: Suomalaisen Kirjallisuuden Seura, 332–374.
- Paunonen, H. 1990. "Se on hän suomeksi." Huomioita kaksikielisten helsinkiläisten suomesta. Teoksessa I. Almqvist, P. Cederholm & J. Lainio (toim.) *Från Pohjolas pörten till kognitiv kontakt. Vänskrift till Erling Wande den 9. Maj 1990*. Stockholm Studies in Finnish Language and Literature 6. Stockholm: Stockholm University, 190–214.
- Paunonen, H. 1995 [1982]. *Suomen kieli Helsingissä. Huomioita Helsingin puhekielen historiallisesta taustasta ja nykyvariaatiosta*. Helsinki: Helsingin yliopiston suomen kielen laitos.
- Piiparinen-Rintaluoma, P.-R. 2001. Ruotsinkielisten nuorten puhesuomen variaatiosta. Teoksessa P.-R. Piiparinen-Rintaluoma (toim.) *Suomi toisena kielenä ja variaatio*. Kakkoskieli 4. Helsinki: Helsingin yliopiston suomen kielen laitos, 87–171.
- Priiki, K. 2012. Läsnaolijaan viittaava *hän* Harjavallan, Kokemäen ja Huittisten nykypuhekielessä. *Sananjalka*, 54, 30–49. <https://doi.org/10.30673/sja.86713>.
- Priiki, K. 2014. Kaakkois-Satakunnan henkilöviitteiset *hän, se, tää* ja *toi* subjekteina. *Sananjalka*, 56, 86–107. <https://doi.org/10.30673/sja.86732>.
- Priiki, K. 2016a. Henkilöviitteisten kolmaspersoonaisen pronomien vaihtelu keskustelukumppanin mukaan. *Puhe ja kieli*, 36 (2), 123–144.
- Priiki, K. 2016b. Puhutun suomen kielioppia ja yksilöllistä vaihtelua. Kvantitatiivinen tutkimus *hän*-pronominista Kaakkois-Satakunnan nykypuhekielessä. *Sananjalka*, 58, 112–135. <https://doi.org/10.30673/sja.86748>.
- Priiki, K. 2017a. Hän, se, tää vai toi? Vuorovaikutussosiolingvistinen tutkimus henkilöviittauksista Kaakkois-Satakunnan nykypuhekielessä. *Annales Universitatis Turkuensis C432*. Turku: Turun yliopisto. <https://urn.fi/URN:ISBN:978-951-29-6719-3>.
- Priiki, K. 2017b. The Finnish logophoric pronoun *hän*—a quantitative approach. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 8 (2), 327–349. <https://doi.org/10.12697/jeful.2017.8.2.11>.
- Priiki, K. 2021. Verkkokeskustelijoiden käsitykset *hän*- ja *se*-pronomien norminvastaisesta käytöstä. *Virittäjä*, 125 (3), 320–346. <https://doi.org/10.23982/vir.98383>.
- R Core Team 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Saanilahti, M. 1990. Huomioita puheen murteellisuustason vaihtelusta haastattelutilanteessa. Teoksessa M. Saanilahti, M. & K. Nahkola (toim.) *Kielen vaihtelu ja puhujan valinnat. Tutkimuksia virtolaisten kielenkäytöstä, murteentuntemuksesta ja kielellisistä asenteista*. Virtain tutkimuksia 7. Tampere: Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos, 39–82.
- Seppänen, E.-L. 1989. Henkilöön viittaaminen puhetilanteessa. Teoksessa A. Hakulinen (toim.) *Suomalaisen keskustelun keinoja I*. Kieli 4. Helsinki: Helsingin yliopiston suomen kielen laitos, 195–222.
- Seppänen, E.-L. 1998. *Läsnaolon pronominit*. Tämä, tuo, se, ja hän viittaamassa keskustelun osallistujaan. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Siitonen, K. 2008. Sävypronomini *hän*. *Sananjalka*, 50, 87–109. <https://doi.org/10.30673/sja.86682>.
- Siitonen, K. 2016. Ihmistarkoitteinen *hän* puhutussa suomessa – kansanlingvistinen kyselytutkimus. *Acta Philologica*, 49, 382–323.

- Sorlin, S. & L. Gardelle 2018. Anthropocentrism, egocentrism and the notion of Animacy Hierarchy. *International Journal of Language and Culture*, 5 (2), 133–162. <https://doi.org/10.1075/ijolc.00004.gar>.
- Strobl, C., J. Malley & G. Tutz 2009. An introduction to recursive partitioning. Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14 (4), 323–348.
- Tagliamonte, S. A. & R. H. Baayen 2012. Models, forests and trees of York English. *Language variation and change*, 24 (2), 135–178. <https://doi.org/10.1017/S0954394512000129>.
- Tyysteri, L. 2015. *Aamiaiskahvilasta ötökkätarjontaan. Suomen kirjoitetun yleiskielen morfosyntaktisten yhdyssanarakenteiden produktiivisuus*. Annales Universitatis Turkuensis, C408. Turku: Turun yliopisto. <https://urn.fi/URN:ISBN:978-951-29-6184-9>.
- Vilkuna, M. 1992. *Referenssi ja määräisyys suomenkielisten tekstien tulkinnassa*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Vilppula, M. 1989. Havainnot *hän-* ja *he-*pronomien käytöstä suomen murteissa. *Virittäjä*, 93 (3), 389–400.
- VISK = A. Hakulinen, M. Vilkuna, R. Korhonen, V. Koivisto, T. R. Heinonen & I. Alho 2004. *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. Verkkoersio, viitattu 13.1.2022. Saatavissa: <http://scripta.kotus.fi/visk>.
- Väänänen, M. 2015. Subjektin ilmaiseminen persoonapäätteellä ja pronominilla. Yksikön 1. persoonan muotojen käyttö lounaismurteissa, lounaisissa siirtymämurteissa ja kaakkoismurteissa. *Sananjalka*, 57, 7–28. <https://doi.org/10.30673/sja.86735>.
- Väänänen, M. 2016. *Subjektin ilmaiseminen yksikön ensimmäisessä persoonassa. Tutkimus suomen vanhoista murteista*. Annales Universitatis Turkuensis, C430. Turku: Turun yliopisto. <https://urn.fi/URN:ISBN:978-951-29-6664-6>.
- Ylikahri, K. 1996. Referoinnin laajentumat *hän-*, *he-* ja *se-*, *ne-*pronomien käytössä Siikaisten murteessa. *Virittäjä*, 100, (2), 182–203.