



MILJA VÄÄNÄNEN

Subjektin ilmaiseminen persoonapäätteellä ja pronomiinilla

Yksikön 1. persoonan muotojen käyttö lounaismurteissa, lounaisissa siirtymämurteissa ja kaakkoismurteissa

1. Johdanto

Tarkastelen tässä artikkelissa¹ verbin persoonapäätteen katoa ja säilymistä yksikön ensimmäisessä persoonassa. Tutkimuksen kohteena on lisäksi subjektipronominin esiintyminen ja sen suhde persoonapäätteeseen. Suomen kielessä subjektin persoonaa voi ilmaista kaksinkertaisesti sekä verbin persoonapäätteellä että subjektipronominilla: *minä mene-n*. Kaksinkertaista persoonanmerkintää on kirjallisuudessa kutsuttu toistoreferenssiksi (Karlsson 1977), jossa jonkin sanan tarkoitteeseen viitataan lauseessa useammin kuin kerran. Persoonakongruenssin tarkoituksena on muun muassa varmistaa finiittiverbin subjektin tunnistaminen: 1. ja 2. persoonassa kongruoiva predikaattiverbi osoittaa aina lauseen subjektin riippumatta siitä, missä kohtaa lausetta subjekti on (Karlsson mas. 383), tai siitä, onko subjektipronominia lainkaan lauseessa ilmipantuna. Puheaktipersoonissahan pronominin ilmipaneminen on ainakin periaatteessa valinnaista toisin kuin kolmannessa persoonassa, jossa on käytettävä ilmisubjektia, jollei tarkoite ole kontekstista pääteltävissä. Normitettu yleiskieli onkin päätynyt käytäntöön, jossa pelkkä verbin persoonapääte on riittävä ja suositeltavakin keino viitata subjektin tarkoitteeseen (normin kehityksestä ja perusteluista ks. Strellman 2005).

Tutkimukset osoittavat kuitenkin, että puhutussa kielessä subjektin ilmaiseminen poikkeaa normitetun yleiskielen käytännöstä. Puhekielessä on nimittäin havaittu olevan yleisintä, että subjektin persoonaa ilmaistaan puheaktipersoonissa sekä persoonapäätteellä että subjektipronominilla. Lappalainen (2004, 81) on tutkinut arkikeskustelua ja havainnut, että subjektipronomini esiintyy 90 prosentissa yksikön ensimmäisen

1 Kiitän artikkeliin saamistani arvokkaista kommentteista kahta *Sananjalan* nimetöntä refereetä. Lisäksi kiitän hyvistä huomioista Marja-Liisa Helasvuota sekä Tomás Leheckaa ja muita Utulingin tilastotyöpajan osallistujia.

persoonan lauseita. Samansuuntaisen tutkimustuloksen on saanut myös Helasvuo (2014a, 68) keskusteluaineistosta tehdyssä tutkimuksessaan: 89 prosentissa lauseita on pronomini ilmipantuna (ks. myös Duvallon 2006, Lappalainen 2006). Toisaalta verbin persoonapäänteen esiintymisessä voi yksikön ensimmäisessä persoonassa olla variaatiota, koska pääte on alttiina puhutun kielen reduktioilmiöille. Sananloppuisen *n:n* kato on tavallinen rennon arkipuheen piirre, jota tavataan monenlaisissa muotoryhmissä (*mentii, semmone, nii, talo etee, mä meni* jne.) Sananloppuisen *n:n* katoa esiintyy yleisesti puheessa, ja yleensä sitä esiintyy sitä enemmän, mitä vapaampi on puhetilanne. Kadon on kuitenkin todettu vanhastaan olevan suomen murteissa yleistä ja varsin systemaattistakin (Rapola 1966, 324–334).

Koska sananloppuisen *n:n* kato on useassa murteessa tavallista, olisi tarpeen selvittää sen yleisyyttä murteita vertaillen. Ilmiötä on kuitenkin kaikkiaan tutkittu vähän, eikä erityisesti persoonapäänteen esiintymistä koskevia tutkimuksia ole. Analysoin tässä artikkelissa persoonapäänteen katoa lounaismurteissa, lounaisissa siirtymämurteissa ja kaakkoismurteissa. Lounaisista siirtymämurteista mukana on kuitenkin vain Porin seudun pitäjiä. Aineistona käytän Lauseopin arkiston murrekorpusta. Mainitut murrealueet ovat valikoituneet mukaan tutkimukseen siitä syystä, että sananloppuisen *n:n* kadon on esitetty olevan niissä erityisen yleistä verrattuna suomen muihin murteisiin (Kettunen 1940, kartat 120–124; Wiik 2006b, kartat 134–156). Toinen valintaperuste on se, että väitöstutkimukseni mukaan (Väänänen tekeillä) subjektin ilmaiseminen pronomiinilla (*mä/mie menen*) on eri murteista yleistä juuri kaakkoisissa ja lounaisissa murteissa, kun taas esimerkiksi savolaismurteet ja pohjoispohjalaiset murteet näyttävät suosivan muita murteita enemmän pronominin poisjättämistä (*menen*).

Kiinnostava kysymys onkin, aiheuttaako persoonapäänteen katotendenssi sen, että subjektipronominia käytetään näissä murteissa enemmän. Ajatus pronominin ja persoonapäänteen esiintymisen korrelaatiosta on ollut fennistiikassa esillä lyhyesti jo Leskisen (1971) kaakkoismurteiden tutkimuksessa, ja myös Lindström ym. (2008) ottavat tämän yhteyden esiin viron murteita koskevassa artikkelissaan. Selvitän tutkimuksessani pronominin ja persoonapäänteen keskinäistä riippuvuutta vielä tarkemmin analysoimalla niiden esiintymistä puhujittain, mikä nähdäkseni paljastaa parhaiten mahdollisen korrelaation. Artikkelini tutkimuskysymykset voi tiivistää seuraavasti:

- 1) Mitkä tekijät vaikuttavat persoonapäänteen säilymiseen tai katoon lounaismurteissa, lounaisissa siirtymämurteissa ja kaakkoismurteissa?
- 2) Onko persoonapäänteen esiintyminen mitenkään yhteydessä subjektipronominin esiintymiseen? Lisääkö persoonapäänteen katoaminen pronominin käyttöä?

Määrittelen aluksi luvussa 2, millaisia redusoitumisilmiöitä sananloppuiseen *n:ään* liittyy. Sen jälkeen esittelen luvussa 3 aineistoni ja ne tekijät, jotka mahdollisesti selittävät päänteen katoa. Analysoin aineistoani tilastollisesti monimuuttujamenetelmän avulla ja käyttämäni satunnaismetsäanalyysin periaatteet kuvaan luvussa 4. Luvussa 5 esittelen

tilastollisen analyysin tulokset ja luvussa 6 tarkastelen vielä persoonapäätteen ja subjekti-pronominin suhdetta puhujittain.

2. Persoonapäätteen kato ja assimilaatio

Sananloppuisen *n:n* heikentymisestä on kirjallisuudessa käytetty monenlaisia nimityksiä. *Redusoituminen* on yhteisnimitys loppuäänteen katoamiselle tai sen assimiloitumiselle seuraavaan äänteeseen (Karlsson 1983, 142). Murrettutkimuksissa on kuitenkin perinteisesti tarkasteltu äänteiden *säilymätapausten* sekä *kadon* tai *heittymisen* variaatiota. Tässä tutkimuksessa käytän nimityksiä *assimilaatio*, *katotapaus* ja *säilymätapaus* niissä yhteyksissä, joissa puhun omista tutkimustuloksistani.

Loppu-*n* katoaa murteissa monissa muotoryhmissä: yksikön ensimmäisen persoonan verbinpäätteen lisäksi kato on tavallista esimerkiksi illatiivissa, genetiivissä ja passiivissa sekä muun muassa *-hAn-* ja *-kAAn-*päätteissä ja *-inen-*loppuisissa nomineissa. On esitetty, että äänteen asema suhteessa koko morfeemiin vaikuttaa sananloppuisen *n:n* taipumukseen kadota. Kuten Leskinen (1971, 350) toteaa kaakkoismurteiden tutkimuksessaan, katotendenssi on päässyt täydellisimmin toteutumaan niissä muotoryhmissä, joissa *n*:lliset ja *n*:ttömät asut eivät ole oppositiosuhteessa: kato on toisin sanoen täydellisintä niissä morfologisissa kategorioissa, joissa *n* muodostaa vain osan päätekokonaisuutta. Jos sitä vastoin *n* muodostaa morfeemin yksin, säilymätapaukset ja kato vaihtelevat; Leskisen mukaan ne ovat keskenään vapaassa vaihtelussa (mp.). Leskinen on tutkinut persoonapäätteen esiintymistä erikseen kaakkoismurteiden eri alaryhmissä, nimittäin Ylä-Vuoksen, Ala-Vuoksen, Äyräpään, Länsi-Kannaksen, Luoteis-Laatokan ja Raudun. Tutkittavina hänellä ovat olleet kaikki ne muotoryhmät, joissa esiintyy sananloppuinen *n*. Yleisesti Leskisen kvantitatiivisen tarkastelun tuloksista voi todeta sananloppuisen *n:n* katoavan kaikissa murteissa käytännössä kokonaan esimerkiksi illatiivista ja allatiivista; säilymätapaukset ovat niissä lähinnä yksittäisesiintymiä. Toisin on yksikön ensimmäisen persoonan persoonapäätteen laita: vaikka useimmissa murteissa katotapaukset ovat selvästi enemmistönä, on myös murteita, joissa säilymätapauksia on selvä enemmistö. Esimerkiksi Äyräpään murteistossa persoonapäätteen katoja on 70 % ja säilymätapauksia 30 %. Sen sijaan Luoteis-Laatokan murteessa nämä esiintymämäärien suhteet ovat päinvastaiset: katoja on 30 % ja säilymätapauksia 70 %. (Leskinen 1971, 344–349, 356.)

Myös ISK:ssa (§ 39) todetaan, että kato on tyypillisintä niissä muotoryhmissä, joissa *n* kuuluu taivutustunnukseen mutta ei yksin muodosta sitä. Karlsson (1983, 148) kuitenkin otaksuu, että sananloppuisen *n:n* heittymistä koskeva redukoimissääntö koskee kaikkia dentaalinasaleja morfologisesta statuksesta riippumatta. Yksikön ensimmäisen persoonan tapauksessa sananloppuinen *n* kantaa yksin koko morfeemin merkitystehtävää, joten sen täydellisessä kadossa voi hävitä myös subjektin persoonaan viittaava aines. Voisi siten olettaa, että päätteen lyhyys suojelisi persoonapäätettä täydelliseltä kadolta. Koska katoa tavaataan kuitenkin persoonapäätteessäkin (Leskinen 1971; Rapola 1966, esimerkit), loppu-*n:n* esiintyminen ei siis näytä riippuvan pelkästään siitä, muodostaako *n* päätteen yksin.

Verbin persoonapäätte voi siis redusoitua kahdella tapaa: assimiloitumalla seuraavaan äänteeseen tai katoamalla kokonaan. Assimiloituminen on kaikissa murteissa – tai puheessa ylipäättään – hyvin yleistä *k:n*, *m:n* ja *p:n* edellä (Rapola 1966, 330; ISK § 39). Lisäksi sananloppuinen *n* voi puheessa täydellisen katoamisen sijaan muuttua nasaaliseksi vokaaliksi (Wiik 2006, 163). Seuraavissa, Lauseopin arkistosta poimituissa esimerkeissä 1 ja 2 esiintyvät kaikki persoonapäätteen esiintymisen vaihtoehdot, assimilaatio (*saim*), kato (*mie oli*) ja säilyminen (*muistan*). Molemmat esimerkit ovat samalta puhujalta, joka edustaa kaakkoismurteita.

- (1) *leipäunitkii ni *sinne / **saim** panna / **leippi paistaa sekä *ruukkuloita ja *kaikkii / (LaX: Muolaa, 96.)
- (2) siel ku ei ku ei *näet *sillo *Suomes- *olt *mittää *semmosii *viel *puutavara / hommii ni / **mie oli** semmone *pikkune *laps ni / muh... ****muistan** **hyväst ko **issäinkii tul' sielt *rahist *pois ja / (LaX: Muolaa, 103.)

Verbinmuoto voi olla tunnistettavissa yksikön ensimmäiseksi, vaikka loppu-*n* jäisi-kin pois, joten katotapauksissakaan kyse ei välttämättä ole referenssin sekaantumisvaarasta². Esimerkin 3 lopussa oleva *olla*-verbi (*oo*) on kiistatta yksikön ensimmäinen, vaikka siitä sinänsä puuttuukin sananloppuinen *n*. On kuitenkin joitakin verbinmuotoja, joissa katomuodot näyttävät muistuttavan kolmannen persoonan muotoja: esimerkin 4 *käi* on myös kolmannessa persoonassa samannäköinen³. On tosin mahdollista, että ensimmäisessä persoonassa heittyvä loppu-*n* jättää jälkiä sanan foneettiseen hahmoon; viimeinen vokaali voi esimerkiksi olla nasaalistunut, jolloin muoto ei olekaan täysin samanlainen kuin kolmannen persoonan verbinmuoto.

- (3) mä *menen ton *flikan tyyönki *sit mut **hän tule muu *tuaman toho sit ja mä *päasen tänt tuppaa ni sitt- **oo.** / (LaX: Kalanti, 74.)
- (4) mut *kyllä **mie** **nyt ***käi** **aikaihminen jo ni/ **Viipuris aika *pal'jon./ *jonkuu ker ****käi** *jo/ **hevos*kyyvissäkkii mut tuota... (LaX: Antrea, 222.)
- Vrt.: ja/ sit niitä ***käi**/ **Honkase *Antti sielt ***käi** niitä *lehmii **etsimäs ja/ (LaX: Antrea, 108.)

Tässä artikkelissa en kuitenkaan tämän enempää paneudu redusoitumisen foneetikkaan, vaan tyydyn pitämään katoina kaikkia *n:ttömiä* verbinmuotoja. Tutkimukseni

2 Joissakin murteissa on lisäksi yleistä äänteiden heittyminen kolmannessakin persoonassa. Esimerkiksi lounaismurteissa yksikön ensimmäinen saattaa olla redusoituneenakin erilainen kuin yksikön kolmas: *mä oli* mutta *se/hän ol*.

3 Heini Lehtonen (2006) on tarkastellut maahanmuuttajataustaisten helsinkiläisten puheen variaatiota. Hänen tutkimansa puhujat käyttävät paljon *i*-loppuista diftongia yksikön 3. persoonan muodoissa, mikä on puhekielelle hyvin epätyypillistä (*sanoi* vs. *sano*). Lisäksi Lehtosen havainnon mukaan yksikön 1. persoonan sananloppuinen *n* heittyy usein hänen informanttiansa puheessa (*mä sanoi*). Nämä seikat yhdessä voivat hänen mukaansa johtaa maahanmuuttajataustaisten puheessa yksinkertaistettuun verbintaivutukseen, jossa yksikön 1. ja 3. ovat samannäköisiä: *mä sanoi* : *sä sanoit* : *se sanoi*. (Lehtonen 2006, 269–270.)

tavoitteena ei myöskään ole esittää kattavaa kuvausta niistä fonologisista ehdoista, jotka mahdollisesti aiheuttavat sananloppuisen *n:n* redusoitumisen persoonapäätteessä verrattuna muihin muotoryhmiin. Heikki Leskinen on jo esittänyt kaakkoismurteita koskevassa tutkimuksessaan (1971) varsin laajan selvityksen äänneympäristön vaikutuksesta loppu-*n:n* katoon kaikissa kyseeseen tulevilla muotoryhmissä, joten tässä artikkelissa ei ole tarpeen enää toistaa Leskisen tutkimusasetelmia.

3. Aineisto ja analyysissa käytetyt muuttujat

Tutkimuksen aineisto on peräisin Lauseopin arkiston murrekorpuksesta. Lauseopin arkiston murreaineisto sisältää yksilöhaastatteluja, jotka on tehty enimmäkseen 1960-luvun lopulla. Haastateltaviksi on valittu mahdollisimman vanhakantaista murretta puhuvia kertojia, jotka ovat yleensä asuneet koko ikänsä samassa pitäjässä tai lähiseudulla. Useimmat ovat syntyneet 1870–1890-luvuilla. Arkiston aineisto soveltuu laajuutensa vuoksi hyvin ilmiön selvittämiseen, koska haastattelut ovat riittävän pitkiä ja niitä on tehty useista samaan murreryhmään kuuluvista pitäjistä. Myös litteroinnit ja annotoinnit on tehty huolellisesti monen tutkijan voimin. Olen siitä huolimatta itse tarkistanut kuuntelemalla jokaisen korpuksesta poimitun finiittiverbin ja jättänyt pois ne tapaukset, joissa päätettä ei voi esimerkiksi taustamelun tai epäselvän puheen takia kunnolla kuulla. Sananloppuisen *n:n* kato on ilmiö, joka voi olla vaikeasti havaittavissa, jos ei siihen kiinnitä erityistä huomiota, mistä syystä eri litteroijat saattavat määritellä kato- ja säilymätapaukset eri tavoin. Itse kuuntelemalla voin siten varmistaa senkin, että tapaukset tulevat yhdenmukaisesti määritellyiksi.

Aineistoa on kaikkiaan 1549 lausetta. Lause tarkoittaa tässä yhteydessä finiittiverbiä täydennyksineen ja määritteineen. Aineisto koostuu kuitenkin pelkästään myöntölauseista. Murrehaastatteluissa esiintyy nimittäin varsin paljon tiettyyppisiä kieltolauseita, joissa on usein käänteinen sanajärjestys ja assimiloitunut persoonapäätte: *em mie tiijä*, *em mää muista* tms. Näiden ilmausten suuri määrä saattaisi vääristää tilastanalyysien tuloksia, joten kieltolauseet on käytännön syistä jätetty tämän tutkimuksen ulkopuolelle. Aineistoa on poimittu Lauseopin arkistosta kolmesta murreryhmästä: lounaimurteista (Kalanti, Lappi, Luvia, Pyhämaa, Rauma), lounaisista siirtymämurteista (Ahlainen, Merikarvia, Noormarkku, Pori) ja kaakkoismurteista (Antrea, Koivisto, Lappee, Luumäki, Muolaa, Nuijamaa, Rautu, Räisälä). Aineistoa on pyritty poimimaan tasapuolisesti sekä länsimurteista (774 lausetta) että itämurteista (775 lausetta). Jokaista pitäjää edustaa Lauseopin arkistossa yksi puhuja.

Tutkimus on luonteeltaan kvantitatiivinen, ja aineistoa analysoidaan monimuuttujamenetelmän avulla. Selitettävänä muuttujana tässä tutkimuksessa on persoonapäätteen esiintyminen. Muuttuja on kaksiluokkainen: PP = verbissä on persoonapäätte, EP = verbissä ei ole persoonapäätettä. Persoonapäätteen assimilaatiotapaukset on analyysin helpottamiseksi yhdistetty säilymätapauksiin, koska tarkasteltavissa murreryhmissä persoonapäätte ei assimiloitu kaikissa äänneympäristöissä. Tavallisimmin assimilaatiota

esiintyy koko aineistossa *k:n* edellä (57 % assimilaatiotapauksista) sekä jonkin verran *m:n* (16 %) ja *p:n* (12 %) edellä. Tutkimissani murreryhmissä ei sen sijaan esiinny assimilaatiota esimerkiksi vokaalin edellä niin kuin monissa muissa murteissa (*minä meni^{i?} illalla*). Tilastollisen analyysin kannalta saattaa olla ongelmallista, jos variaatiota ei ole kaikissa ympäristöissä.

Persoonapäätteen esiintymistä mahdollisesti selittäviksi tekijöiksi olen valinnut kaikkiaan seitsemän kielellistä tai kielenulkoista ilmiötä. Monet tekijöistä on valittu aiemman tutkimuskirjallisuuden perusteella, ja perustelen näiden tekijöiden valintaa jäljempänä. Selittävät tekijät on koottu alla olevaan taulukkoon 1. Selittävät tekijät on listattu taulukossa vasemmalle, kun taas jokaisen tekijän luokat eli eri vaihtoehdot ovat oikealla.

Taulukko 1. Persoonapäätteen esiintymistä selittävät tekijät.

Selittävä tekijä	Tekijän luokat
Murre	lounaismurteet, lounaiset siirtymämurteet, kaakkoismurteet
Pitäjä (PITÄJÄ_PUHUJA)	Ahlainen, Antrea, Kalanti, Koivisto, Lappi, Lappee, Luvia, Luumäki, Muolaa, Merikarvia, Noormarkku, Nuijamaa, Pori, Pyhämaa, Rauma, Rautu, Räisälä
Äänneympäristö (YMPÄRISTÖ)	likvida (<i>l, r</i>), nasaali (<i>m, n</i>), <i>h, j, k, p, s, t, v</i> , vokaali, tauko, absoluuttinen loppu
Lauseen subjekti	subjektipronomini, ei subjektipronominia
Finiittiverbin tyyppi (VTYYPPI)	liikeverbit, muutos-, teko- ja toimintaverbit, puheverbit, kognitiiviset verbit, statiiviset verbit
Referentiaalinen etäisyys (ETÄISYYS)	1–11 lausetta
Tempus	preesens, imperfekti, perfekti, pluskvamperfekti

Äänneympäristö tarkoittaa tässä vain persoonapäätettä seuraavaa äännettä. Koska *r:*ää esiintyy aineistossani melko vähän päätettä seuraavana äänteenä, se on yhdistetty *l:n* kanssa likvidojen luokaksi. Myös eri vokaalit on yhdistetty samaan luokkaan, koska jotkin vokaaleista esiintyvät harvakseltaan. Äänneitä on yhdistelty suuremmiksi ryhmiksi myös sen vuoksi, että vertailu muihin tutkimuksiin olisi helpompaa. Muun muassa Leskinen on yhdistänyt (1971) eri nasaalit yhdeksi ryhmäksi ja samalla tavoin eri vokaalit toiseksi ryhmäksi. Tauko tarkoittaa puheen hetkellistä katkeamista, kun taas absoluuttinen loppu on ilmauksen loppuasema, jossa intonaatio laskee puhunnoksen lopulle tyypillisellä tavalla.

Finiittiverbit on jaettu syntaktis-semanttisin perustein (ks. Pajunen 2001) viiteen luokkaan: liikeverbeihin (mm. *mennä*), muutos-, teko- ja toimintaverbeihin (*tehdä*), puheverbeihin (*sanoa*), kognitiivisiin verbeihin (*muistaa*) ja statiivisiin verbeihin (*olla*). Eri verbilekseemien esiintymisfrekvenssit kuitenkin vaihtelevat. Kahdessa näistä luokista on lekseemi, joka on huomattavasti taajemmin esiintyvä kuin muut luokkansa lekseemit. Tällaisia ovat erityisesti statiiviverbien *olla* ja puheverbien *sanoa*. Statiivisten verbien

luokka muodostuu lähes pelkästään *olla*-verbeistä: niitä on 97 prosenttia luokasta. Tässä yhteydessä en ole erotellut erikseen kopulana toimivia verbejä, vaan kaikki *olla*-verbit on yhdistetty samaan lausetyypistä riippumatta. Puheverbeistä taas *sanoa*-verbillä on 71 prosentin osuus luokastaan. Koska mainitut kaksi lekseemiä ovat niin frekventtejä, voi ajatella statiivisia verbejä ja puheverbejä koskevien tilastollisten havaintojen koskevan ennen kaikkea niitä. Muissa taulukon 1 verbiluokissa on sen sijaan enemmän lekseemien variaatiota.

Verbiluokituksen merkitys tässä tutkimuksessa liittyy siihen, että muutamilla verbilekseemeillä on erityisiä ja kiteytyneitä tehtäviä diskurssissa. Erityistehtävään vakiintuneet rakenteet voivat olla alttiimpia myös reduktioilmiöille. On tunnettua, että kielenkäytössä usein toistuvat rakenteet kokevat äänteellisiä muutoksia (Bybee 2007, 11). Tutkimuksissa onkin havaittu kielellisen rakenteen esiintymistiheyden olevan yksi merkittävä reduktiota selittävä tekijä (esim. Hollmann & Siewierska 2011). Esimerkiksi englannissa *don't* redusoituu yleisimmin frekventeimmässä rakenteissa (*I don't know* tms.) (Bybee & Scheibman 1999). Myös suomesta on havaittu, että frekventit kognitiiviset verbit muodostavat kiteytyneitä konstruktioita, joihin kuuluu subjektipronomini ja joihin voi liittyä myös äänteellistä kulumista (Helasvuo 2014a, 70–76). Tässä tutkimuksessa kiinnostuksen kohteena on erityisesti *sanoa*-verbi, joka yksikön ensimmäisessä persoonassa toimii referoidun puheen merkitsimenä. Murrehaastattelu on narratiivista kerrontaa, jossa *sanoa*-verbin muodostama kiinteä rakenne (*mie sanoi* tms.) esiintyy usein ja jossa se merkitsee osan ilmausta puhujan itsensä aiemmin sanomaksi:

- (5) **mie sanoi** **no **nyt ko se/ *viel- o **poika/ *nyt *sem *myöp mut ku se *männöö *sinne **mie sanoi** ni **sit et sitä *saa/ **rahalla etkä *millää (LaX: Lappee, 70.)

Olla-verbi puolestaan on Lauseopin arkiston useimmin esiintyviä sanoja, ja se on yleinen verbi yksikön ensimmäisessäkin persoonassa. Narratiivisessa diskurssissa sitä esiintyy esimerkiksi puhujaa itseään luonnehtivissa predikaatiivilauseissa: *ni / mie oli semmone pikkune laps ni / muh... muistan hyväst* (ks. esimerkki 2 aiemmin). Hypoteesini on, että mahdollinen verbityypin merkitys persoonapäätteen katoamisessa liittyy nimenomaan *olla*- ja *sanoa*-verbeihin, koska ne ovat tutkitussa aineistossa taajaan esiintyviä.

Referentiaalinen etäisyys tarkoittaa etäisyyttä saman tarkoitteen edelliseen mainintaan kontekstissa (käsitteestä ks. Givón 1983). Etäisyys saa arvon 1, jos samaan tarkoitteeseen (eli puhujaan itseensä) on viitattu heti edellisessä lauseessa. Jos välissä on yksi lause, jossa on eriviitteinen subjekti, etäisyys saa arvon 2 jne. Etäisyyttä on kuitenkin laskettu edeltävässä kontekstissa vain kymmenen lauseen verran. Kaikki kymmentä lausetta kauempana esiintyvät saavat etäisyyden arvoksi 11.

Vaikka tarkoitteen referentiaalinen etäisyys edelliseen mainintaan onkin diskurssin ilmiö, jota on yleensä tarkasteltu kolmannen persoonan referentiaalisia ketjuja tutkittaessa, sen on todettu selittävän hyvin subjektipronominin esiintymistä myös yksikön ensimmäisessä persoonassa. Lindström ym. (2008) ovat tutkineet viron murreaineiston avulla niitä tekijöitä, jotka vaikuttavat yksikön ensimmäisen persoonan pronominin (*mina/ma*) esiintymiseen tai poisjättämiseen. Tutkimuksen keskeisimpiä havaintoja on,

että pronomini jää useimmin pois, jos referentiaalinen etäisyys on 1, toisin sanoen myös edellisessä lauseessa on yksikön ensimmäiseen persoonaan viittaava muoto. Myös Helasvuon (2014b, 466) suomea koskevassa tutkimuksessa todetaan, että subjektipronomini jää todennäköisimmin lauseesta pois, jos referentiaalinen etäisyys on 1. Persoonapääteellinen verbi voi viitata subjektin tarkoitteeseen myös ilman pronominia. Edellä mainituissa tutkimuksissa on selvitetty etäisyyden vaikutusta pronominin esiintymiseen, mutta tässä tutkitaan nyt sitä, onko etäisyydellä referentin edelliseen mainintaan samalla tavoin merkitystä myös persoonapäätteen esiintymisessä. Jääkö pääte todennäköisemmin pois, jos sama tarkoite on ollut esillä edeltävissä lauseissa?

Aikamuodolla ei alkualettamukseni mukaan ole sinänsä erityistä vaikutusta päätteen esiintymiseen, mutta se on analyysissa mukana siitä syystä, että sanavartalot voivat olla eri tempusmuodoissa erilaisia, mikä päätteen kannalta ilmenee erilaisina äänneympäristöinä. Varsinkin preesensissä ja imperfektissä verbin vartalo voi puheessa olla varsin erilainen eri aikamuodoissa (esim. *mä oon, mä olin; mie lähen, mie läksin*). Liittomuodoissa taas persoonamuotoinen verbi on aina *olla*, mikä voi myös vaikuttaa päätteen säilymiseen tai katoon.

4. Satunnaismetsämenetelmä

Aineistoa on analysoitu tilastollisesti. Menetelmänä on käytetty satunnaismetsäksi (engl. *random forest*) kutsuttua analyysityökalua, joka on kehitetty päätöspuumenetelmien pohjalta. Päätöspuuanalyysit ovat luokittelevia monimuuttujamenetelmiä. Kun tässä tapauksessa selitettävänä on kaksiluokkainen muuttuja (PP = persoonapäätte tai EP = ei persoonapäätettä), päätöspuuanalyysi kertoo, kuinka hyvin annettujen selittävien tekijöiden perusteella voi ennustaa PP:n tai EP:n todennäköisyyttä aineistossa. Analyysi pyrkii toisin sanoen löytämään sen tekijän, jonka avulla aineistoa voi parhaiten luokitella luokkiin PP ja EP. Prosessi käy aineistoa läpi jakaen sitä ryhmiin ja arvioiden selittäviä tekijöitä siltä kannalta, kuinka suuria eroja sen avulla saa aikaan PP- ja EP-ryhmien välille. Analyysin tuloksena syntyy puukuvio, jossa puun jokaisessa haarassa aineisto jakaantuu ryhmiin yhä uudestaan ja uudestaan niin kauan, kuin tilastollisesti merkitseviä eroja saadaan aikaan. (Baayen 2008, 161.)

Yksittäisen päätöspuun ongelmana voi toisinaan olla analyysin hierarkkisuus: prosessissa valikoituu yksi muuttuja tärkeimmäksi eli puun ensimmäiseksi haaraksi, mutta mahdollisesti lähes yhtä hyvät selittäjät eivät sen sijaan pääse analyysissa kovinkaan hyvin esiin. Ongelma on ilmeinen analyyseissa, joissa aineisto on pienehkö mutta selittäviä muuttujia useita, kuten tämän tutkimuksen tapauksessa on laita. Tässä tutkimuksessa pyrin löytämään kaikki merkitykselliset tekijät, jotka selittävät persoonapäätteen esiintymistä. Näistä syistä olen valinnut menetelmäksi satunnaismetsän, joka sietää monia muita tilastomenetelmiä paremmin aineiston pienuutta ja epätasaisuutta ja jossa selittävien tekijöiden mahdollinen kollineaarisuuskaan ei ole tämän menetelmän käytön este (Tagliamonte & Baayen 2012, 24).

Satunnaismetsäanalyysi perustuu samaan algoritmiin kuin edellä kuvattu binääri-nen päätöspuu, mutta tässä menetelmässä aineistosta otetaan useita satunnaisotoksia. Aineistosta erotetaan osa erilleen (*in-bag observations*), ja tästä osa-aineistosta menetelmä valitsee satunnaisesti analyysiin mukaan eri tekijöitä ja selvittää tällä tavoin, millaisia vaikutuksia on kunkin selittävän tekijän mukaan ottamisella tai poisjättämisellä. Analyysin tarkkuutta lisää se, että otoksista syntyneitä puita testataan arvioimalla, kuinka hyvin niiden perusteella voi ennustaa selitettävän ilmiön esiintymistä jäljelle jääneessä aineistossa (*out-of-bag observations*). (Tagliamonte & Baayen 2012, 22; Ivaska 2014.) Kuten satunnaismetsän käsitteeseen sisältyvä metaforakin kertoo, useista puista muodostuva metsä on selittävydessään enemmän kuin yksittäinen puu.

Satunnaismetsä tekee listauksen selittävien tekijöiden tärkeysjärjestyksestä (R:n funktio *varimp* eli *variable importance*). Jos selittävällä tekijällä näyttää olevan yhteys selitettävään ilmiöön, *varimp*-analyysi varmistaa tuon yhteyden vertaamalla todellista selittävää tekijää satunnaistekijään (*randomly permuted predictor*). Menetelmä toisin sanoen sekoittaa vertailuaineistoa vaihtamalla aineiston havainnot sattumanvaraisesti ja katkaisemalla näin selittävän tekijän alkuperäisen yhteyden selitettävään muuttujaan. Koska sekoitettu vertailuaineisto on luotu sattumanvaraisesti, pitäisi satunnaistekijän olla selittävydessään olennaisesti huonompi kuin oikean selittävän tekijän. Jos selittävyys huononee, alkuperäinen tekijä on siten todettavissa aidosti merkitseväksi. Mitä suurempi ero todellisen tekijän ja satunnaistekijän välillä on, sitä tärkeämpi selittäjä on kyseessä. (Tagliamonte & Baayen 2012, 23.)

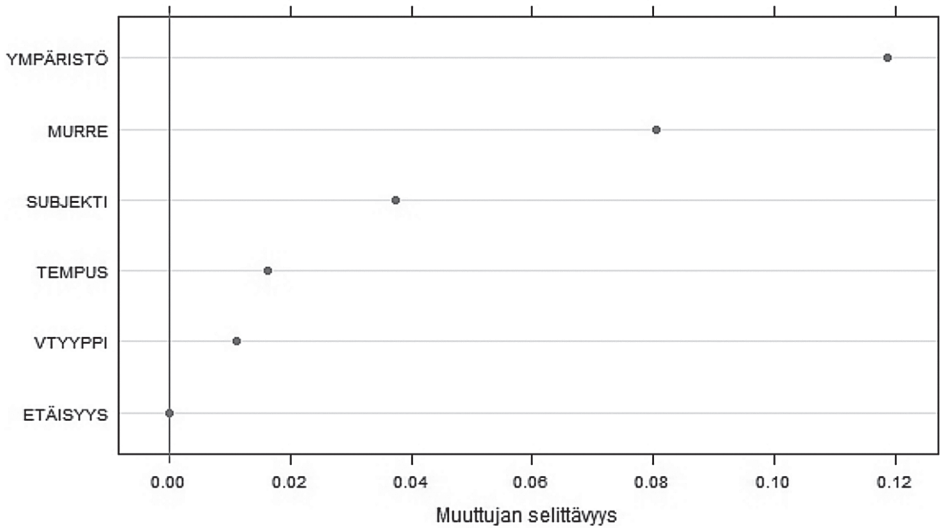
5. Persoonapäätteen esiintymiseen vaikuttavat tekijät

Analysoin satunnaismetsämenetelmän avulla ensin koko aineiston ja sen jälkeen tarkastelen erikseen lounaismurteita, lounaisia siirtymämurteita ja kaakkoismurteita.

Satunnaismetsäanalyysi asettaa muuttujat tärkeysjärjestykseen sen mukaan, kuinka hyvin kukin muuttuja selittää tutkittua ilmiötä. Kuviossa 1 on esitetty muuttujien tärkeysjärjestys koko aineistosta siten, että tärkein muuttuja on ylhäällä ja vähiten tärkein alhaalla. Kuvion vaaka-akseli (*Muuttujan selittävyys* eli *variable importance*) puolestaan kertoo, kuinka paljon muuttujat eroavat toisistaan: mitä enemmän oikealla muuttujan selittävyttä kuvaava piste on, sitä selittävämpi tekijä on kyseessä. Sen sijaan mitä lähemmäksi vasemmalla olevaa nollaa piste kuviossa asettuu, sitä vähemmän tärkeä kyseinen muuttuja on. Tagliamonte ja Baayen (2012, 24) määrittelevät merkityksettömiksi ne tekijät, jotka asettuvat kuviossa nollan tuntumaan.

Satunnaismetsäanalyysi paljastaa, että äänneympäristö on persoonapäätteen katoa tai säilymistä parhaiten selittävä tekijä koko aineistossa. Seuraavaksi parhaita selittäjiä ovat murreryhmä (lounaismurteet, lounaiset siirtymämurteet ja kaakkoismurteet) ja lauseen subjektin esiintyvyys (subjektipronomini tai ei subjektipronominia). Nämä kolme tekijää eroavat muista selvästi. Kuitenkin myös kaksi seuraavaa, *tempus* ja *verbityyppi*, poikkeavat suhteellisen paljon listan viimeisenä olevasta muuttujasta eli referentiaalisesta

Kuvio 1. Satunnaismetsäanalyysi koko aineistosta.



$C = 0.774$

etäisyydestä, joten niilläkin voi katsoa olevan jonkinlaista selitysarvoa. Sitä vastoin referentiaalinen etäisyys ei siis analyysin mukaan selitä persoonapäätteen esiintymistä. Palaan merkityksellisiksi osoittautuneisiin tekijöihin lähemmin jäljempänä. Sitä ennen tarkastelen murteiden välisiä eroja.

Kuvio 1 osoitti murreryhmän olevan toiseksi tärkein selittäjä. Taulukkoon 2 on koottu persoonapäätteellisten (PP) ja -päätteettömien (EP) finiittiverbien lukumäärät kussakin murreryhmässä.

Taulukko 2. Persoonapäätteellisten ja -päätteettömien finiittiverbien määrät murreryhmittäin.

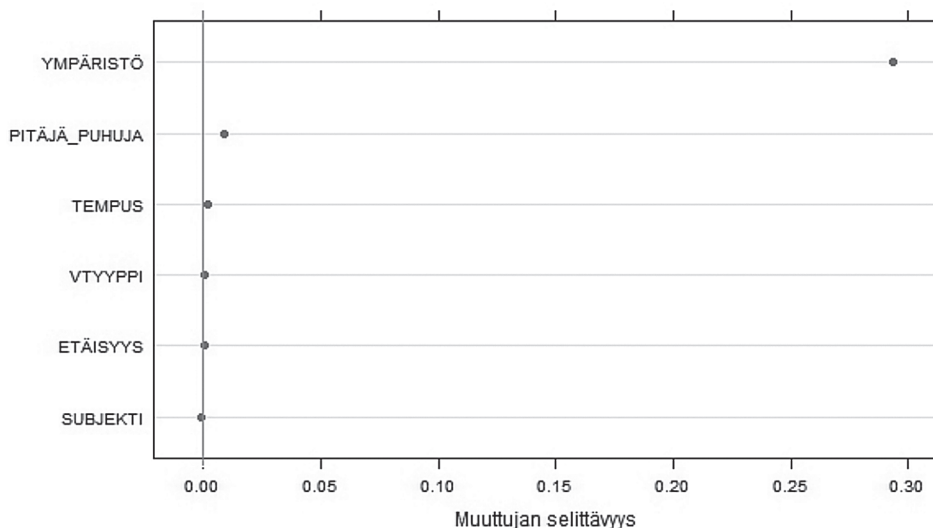
	PP	EP	Yht.
Lounaismurteet	160 (33 %)	332 (67 %)	492 (100 %)
Loun. siirtymämurteet	80 (28 %)	202 (72 %)	282 (100 %)
Kaakkoismurteet	365 (47 %)	410 (53 %)	775 (100 %)
Yht.	944 (39 %)	605 (61 %)	

Murreryhmistä eniten päätteen katotapauksia on lounaisissa siirtymämurteissa: 72 prosenttia finiittiverbeistä on ilman päätettä. Sen sijaan kaakkoismurteissa päte katoaa kokonaan vain 53 prosentista verbejä. Lounaismurteet sijoittuvat näiden kahden väliin: päätteen katotapauksia on 67 prosenttia. Murteiden välinen ero on tilastollisesti merkitsevä (Pearsonin $\chi^2=43,412$; $df=2$; $p<0,001$).

Seuraavassa esitän satunnaismetsäanalyysit vielä erikseen jokaisesta murreryhmästä. Murreryhmittäinen tarkastelu nimittäin osoittaa, että päätteen katoamisen tai esiintymisen ehdot vaihtelevat jonkin verran eri ryhmissä. Kuviossa 2 murreryhmän tilalla on nyt yhtenä selittäväenä tekijänä pitäjä (kuviossa PITÄJÄ_PUHUJA). Ei tosin ole varmaa, että tarkasteltavina ovat nimenomaan pitäjien väliset erot. Koska yhtä pitäjää edustaa Lauseopin arkistossa yleensä vain yksi puhuja⁴, voi yhtä lailla todeta tarkastelevansa tässä myös puhujien välisiä eroja.

Kuvio 2. Satunnaismetsäanalyysit murreryhmittäin.

2a. Lounaismurteet.

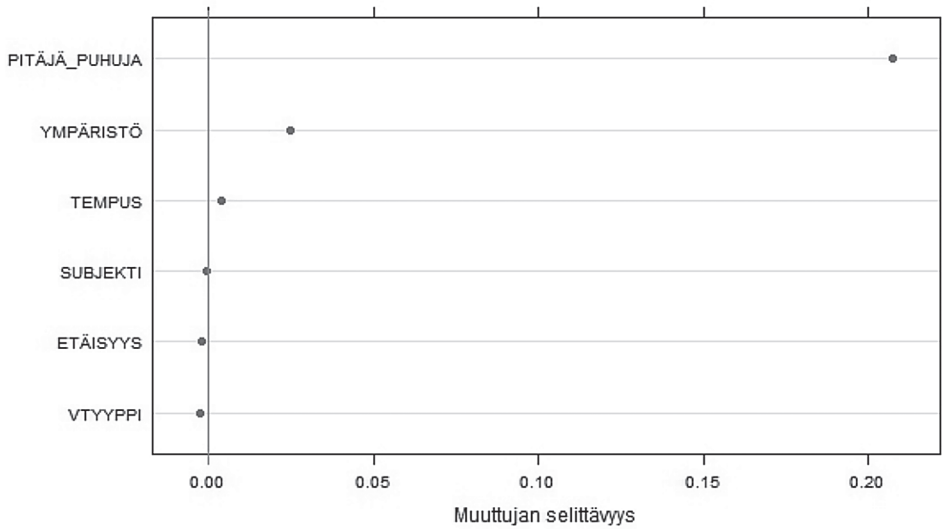


C=0.940

Lounaismurteissa (kuvio 2a) äänneympäristö on hyvin selvästi tärkein persoonapäätteen esiintymistä määräävä tekijä. Pitäjä/puhuja-tekijäkin poikkeaa kuviossa nolasta, mutta sen merkitsevyys jää kuitenkin kauas äänneympäristöstä. Äänneympäristö on merkittävä selittäjä myös lounaisissa siirtymämurteissa (2b), mutta sitä tärkeämmäksi tässä murreryhmässä nousee kuitenkin pitäjä/puhuja. Molemmissa länsimurteissa kahden tärkeimmän tekijän jälkeen on esitetty tempus, joka kuvioissa 2a ja 2b kyllä poikkeaa nolalinjasta mutta kuitenkin varsin vähän. Muilla analyyseissa mukana olevilla tekijöillä (verbityyppi, referentiaalinen etäisyys ja subjektipronominin esiintyminen) ei näissä länsimurteissa näytä olevan merkitystä persoonapäätteen kannalta. Subjektipronominin vaikutukset päätteen käyttöön tulevat kyllä toisella tavalla esiin, kun ilmiötä tarkastellaan puhujittain (ks. luku 6).

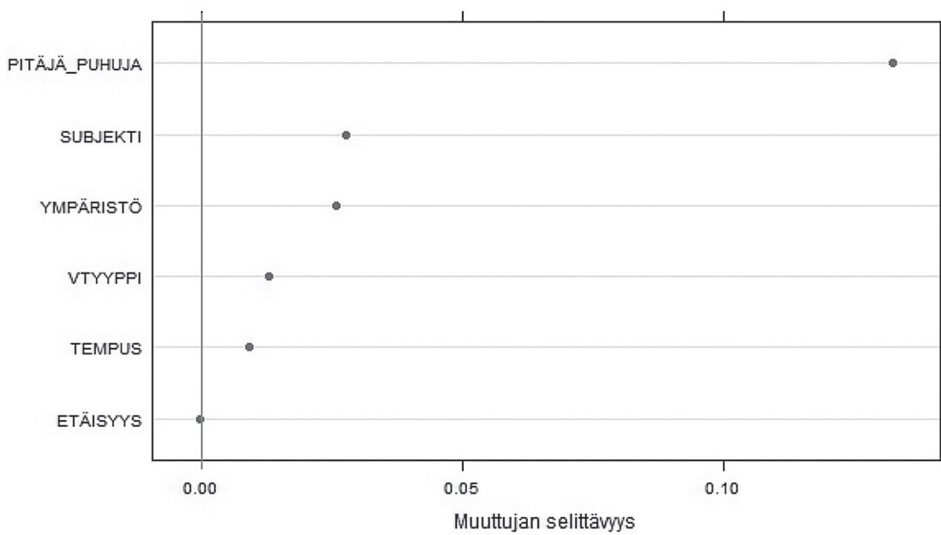
4 Tähän tutkimukseen valituista paikkakunnista Merikarvian aineisto koostuu kuitenkin avioparin yhteishaastattelusta, jossa haastateltavana on kaksi informanttia.

2b. Lounaiset siirtymämurteet.



C=0.940

2c. Kaakkoismurteet.



C = 0.855

Kaakkoismurteissa useallakin tekijällä on kuvion 2c mukaan arvoa persoonapäätteen esiintymisen selittäjinä. Pitäjä/puhuja poikkeaa tässäkin ryhmässä selvästi muista tekijöistä merkittävytydessään. Subjektipronominin esiintyminen lauseessa on selvästi merkityksellinen toisin kuin tutkituissa länsimurteissa – sen selittävyys ohittaa äänneympäristönkin. Koko aineiston analyysissä (kuvio 1) todettiin lauseen subjektin olevan kolmanneksi tärkein selittäjä, mutta sen merkitykseen vaikuttavat siis lopulta vain kaakkoismurteet. Myös verbityyppi ja tempus poikkeavat selvästi nollasta.

Kaikissa murreryhmissä pitäjä/puhuja ja äänneympäristö osoittautuivat siis selvästi tärkeimmiksi tekijöiksi persoonapäätteen esiintymisessä. Yhteistä kaikille murreryhmille on myös se, että referentiaalisella etäisyydellä ei ole merkitystä. Länsimurteissa useimpia muita tekijöitä pitäjää/puhujaa ja äänneympäristöä lukuun ottamatta ei voi pitää merkittävinä, sen sijaan kaakkoismurteissa useakin tekijä näyttää selittävän päätteen esiintymistä. Tarkastelen persoonapäätteen esiintymistä puhujittain tarkemmin jäljempänä omassa alaluvussa. Seuraavassa luon ensin katsauksen äänneympäristön merkitykseen kaikissa murteissa ja sen jälkeen tarkastelen lyhyesti verbityypin ja aikamuodon merkitystä kaakkoismurteissa.

Äänneympäristö osoittautui siis satunnaismetsäanalyyseissa erityisen merkittäväksi tekijäksi lounaismurteissa mutta vähemmän merkitykselliseksi muissa ryhmissä. Taulukkoon 3 on vertailun helpottamiseksi koottu kaikki analysoidut äänneympäristöt. Taulukossa tarkastellaan jokaista äännettä tai äänneryhmää sen mukaan, kuinka suuri osuus

Taulukko 3. Persoonapäätteen esiintyminen eri äänneympäristöissä.

Ympäristö	Lounaismurteet				Loun. siirtymämurteet				Kaakkoismurteet				Yht.
	EP		PP		EP		PP		EP		PP		
	N	%	N	%	N	%	N	%	N	%	N	%	
likvida	8	100	0	0	6	100	0	0	17	57	13	43	44
nasaali	57	84	11	16	36	67	18	33	61	43	80	57	263
<i>h</i>	3	100	0	0	8	89	1	11	24	80	6	20	42
<i>j</i>	31	97	1	3	15	100	0	0	45	63	27	37	119
<i>k</i>	1	1	77	99	18	58	13	42	44	46	52	54	205
<i>p</i>	1	8	11	92	4	50	4	50	10	45	12	55	42
<i>s</i>	65	88	9	12	33	75	11	25	67	66	34	34	219
<i>t</i>	2	7	27	93	15	63	9	37	24	44	31	56	108
<i>v</i>	20	91	2	9	10	83	2	17	20	51	19	49	73
vokaali	84	84	16	16	40	73	15	27	77	68	37	32	269
tauko	30	97	1	3	12	86	2	14	18	36	32	64	95
abs. loppu	30	86	5	14	6	60	4	40	3	12	22	88	70

kyseisen äänteen tai äänneryhmän edellä esiintyvistä persoonapäätteistä on katotapauksia (EP) ja kuinka suuri osuus säilymätapauksia (PP). On kuitenkin syytä huomata, että muutamien äänteiden, kuten *h:n*, *p:n* ja likvidojen, kokonaisfrekvenssi jää varsinkin lounaisissa siirtymämurteissa melko pieneksi, joten niiden osalta tulokset jäävät epävarmoiksi.

Lähempi äänneympäristön tarkastelu osoittaa, että lounaismurteissa katotapauksia esiintyy suhteellisen selvästi vain tiettyjen äänteiden edellä ja vastaavasti säilymätapauksia toisten äänteiden edellä. Lounaismurteissa likvidojen (*l*, *r*) ja *h:n* edellä esiintyy vain katotapauksia. Päätte on enimmäkseen kadonnut silloin, kun päätettä seuraava äänne on *j* (97 % tapauksista on katoja), *v* (91 %), *s* (88 %), vokaali (84 %) tai nasaali (84 %). Lisäksi kato on hyvin yleistä myös tauon edellä (97 %) ja absoluuttisessa lopussa (86 %). Nämä havainnot käyvät täydellisesti yksiin Wiikin (2006a) näkemyksen kanssa: hänen mukaansa sananloppuinen *n* häviää lounaismurteissa absoluuttisessa lopussa, vokaalin edeltä, *h:n*, *s:n*, vokaalin, likvidan ja nasaalin edeltä (mts. 203). Wiik on kuitenkin tarkastellut sananloppuista *n*:ää kaikissa muotoryhmissä, ei pelkästään yksikön ensimmäisen persoonan persoonapäätteessä.

Niitä ympäristöjä, joissa päätte lounaismurteissa säilyy, on huomattavasti vähemmän, mutta säilyminen vaikuttaa toisaalta olevan niissä hyvin systemaattista. Päätte näyttää säilyvän erityisesti klusiilien edellä: *k:n* edellä säilymätapauksia on 99 %, *p:n* edellä 92 % ja *t:n* edellä 93 %. *K*- ja *p*-ympäristöistä suuri osa on mitä ilmeisimmin assimilaatioita. Vaikka katotapaukset ovatkin lounaismurteissa enemmistönä, ovat *k* ja *p* äänneympäristöjä, joissa päätte kuitenkin säilyy assimiloituneena jokseenkin säännöllisesti. Mainitut äänneympäristöt ovatkin kaikissa murteissa *n:n* assimiloituneen muodon vakiintuneita ympäristöjä (Rapola 1966, 330). Päätteen säilyminen *t:n* edellä taas johtuu Rapolan (1966, 329) mukaan ”suojaisesta lausefoneettisesta asemasta”, loppu-*n:n* säilyminen tässä asemassa onkin hänen mukaansa ominaista kaikille murteille joitakin kaakkoismurteita lukuun ottamatta.

Kolmesta tutkitusta murrealueesta katotapauksia on suhteellisesti eniten lounaisissa siirtymämurteissa (ks. taulukko 2). Lähinnä vain katotapauksia esiintyy lounaismurteiden lailla likvidojen edellä ja lisäksi *j:n* edellä. Myös frekvenssiltään vähälukuisen *h:n* edellä esiintyy lähinnä katoja. Nekin ympäristöt, joissa esiintyy molempia mutta katotapaukset ovat enemmistönä, ovat siirtymämurteissa samoja kuin lounaismurteissa. Päätte katoaa todennäköisimmin, jos seuraava äänne on *v* (83 %), *s* (75 %) tai vokaali (72 %) tai jos päätettä seuraa tauko (86 %). Huomionarvoista kuitenkin on, että katotapauksia ei näissä äänneympäristöissä ole suhteellisesti niin paljon kuin lounaismurteissa. Sen sijaan säilymätapaukset eivät siirtymämurteissa ole enemmistönä missään äänneympäristössä. Niissäkin ympäristöissä, joissa päätte lounaismurteissa säilyy tai assimiloituu lähes aina, kato näyttää siirtymämurteissa olevan suhteellisen yleistä: *k:n* edellä katoja on 57 %, *p:n* edellä 50 % ja *t:n* edellä 63 %. Lyytikäinen ym. (2013, 80) mainitsevat, että sanojen loppu-*n* on erityisesti Porin seudulla lähes aina kadonnut. Taulukossa 3 esittelemäni havainnot eivät osoita loppu-*n:n* katoavan ainakaan persoonapäätteessä aina, mutta kato näyttää kyllä olevan näissä murteissa laajaa ja koskevan useimpia äänneympäristöjä.

Satunnaismetsäanalyysi osoitti, että äänneympäristö on kaakkoismurteissa vasta kolmanneksi tärkein persoonapäätteen esiintymistä selittävä tekijä. Jakaantuminen kato- ja säilymätapauksiin eri äänneympäristöissä onkin kaakkoismurteissa tasaisempaa kuin lounaismurteissa ja lounaisissa siirtymämurteissa. Melko selviä päätteen katoympäristöjä ovat tässäkin murreryhmässä *h* (80 % katotapauksia), vokaali (68 %), *s* (66 %) ja *j* (63 %). Tosin prosenttiosuudet ovat selvästi pienempiä kuin länsimurteiden vastaavat osuudet. Monissa äänneympäristöissä katotapauksia on lähes yhtä paljon kuin säilymätapauksia. Tällaisia äänneympäristöjä ovat likvidat (57 % katotapauksia), *v* (51 %), *k* (46 %), *t* (44 %), *p* (45 %) ja nasaalit (43 %). Säilymätapauksetkin ovat siis joissakin äänneryhmissä niukkana enemmistönä, ja lisäksi ne ovat selvänä enemmistönä myös absoluuttisessa lopussa (88 % säilymätapauksia) ja tauon edellä (64 %).

Nämäkin havainnot ovat yhteneväisiä aiemman tutkimuksen kanssa. Leskisen mukaan (1971) sananloppuisen *n:n* katoa esiintyy kaikissa kaakkoismurteissa eniten *h:n*, *s:n* ja vokaalin edellä sekä jossain määrin myös klusiilin edellä, mutta sen sijaan säilymätapauksia esiintyy hänen mukaansa suhteellisen paljon terminaalisisessä asemassa. Äyrämöismurteita tutkinut Ruoppila (1955, 96) on toisaalta aiemmin todennut, että mitään ratkaisevaa vaikutusta loppu-*n:n* esiintymiseen ei ole sillä, alkaako seuraava sana vokaalilla vai konsonantilla vai onko kyseessä absoluuttinen loppu. Ruoppila esittää sen sijaan (mts. 99), että persoonapäätteen tapauksessa vaikutuksensa olisi painoiseikoilla: jos verbi on painollinen, päätte on säilynyt, mutta jos pronomini puolestaan on painollinen, *n* katoaa.

Verbityyppi osoittautui merkitykselliseksi vain kaakkoismurteissa. Alkuloetusten mukaisesti persoonapäätte jää pois erityisesti puheverbeissä (66 % verbeistä on ilman persoonapäätettä) ja statiivisista verbeistä (65 % on ilman päätettä). Persoonapäätteettömiä on kaakkoismurteiden aineiston verbeistä kaikkiaan 53 prosenttia ja persoonapäätteellisiä 47 prosenttia (ks. taulukko 2). Mainittujen verbityyppien esiintymämäärien erot odotuksenmukaiseen esiintymämäärään ovat merkitseviä (Pearsonin $\chi^2 = 22,8$; $df = 5$; $p < 0,001$). Näissä verbiluokissa kuitenkin lekseemit *sanoa* ja *olla* ovat siis frekvenssinsä vuoksi hallitsevia, joten voi ajatella tulosten koskevan käytännössä lähinnä niitä. Muut verbiluokat eivät vaikuta merkittävästi persoonapäätteen esiintymiseen.

Kaakkoismurteissa jossain määrin merkityksellinen on myös aikamuoto. Olennaista on kuitenkin vain ero preesensin ja imperfektin välillä. Imperfektimuotoisista verbeistä persoonapäätte jää pois 57 prosentissa tapauksista, mutta preesensissä olevista verbeistä harvemmin: vain 40 prosenttia verbeistä on ilman päätettä. Erot odotuksenmukaisiin frekvensseihin ovat merkitseviä ($p < 0,001$; Fisherin tarkka testi).

6. Persoonapäätte ja subjektipronomini

Satunnaismetsäanalyysin mukaan ainoastaan kaakkoismurteissa on merkitystä persoonapäätteen säilymiseen sillä, onko lauseessa ilmipantu subjektipronomini vai ei. Subjektipronominien lähempi tarkastelu osoittaa, että ilman pronominia esiintyvät verbit suosivat

vahvasti persoonapäätettä. Vastaavasti päätte jää todennäköisemmin pois, jos lauseessa on pronomini. Kaakkoismurteiden subjektipronominittomista verbeistä 70 prosenttia on persoonapäätteellisiä. Jos lauseessa sen sijaan on ilmipantu pronomini, persoonapäätte on vain 38 prosentissa verbeistä. Erot odotuksenmukaisiin esiintymämääriin ovat merkitseviä (Pearsonin $\chi^2 = 64,2$; $df = 1$; $p < 0,001$).

Leskinen (1971, 365) on tehnyt samansuuntaisen havainnon tutkimistaan kaakkoismurteista: hänen aineistossaan päätten katoa esiintyy 78 prosentissa tapauksista, jos lauseessa on pronomini, mutta subjektittomissa lauseissa (tarkoittaa pronominittomia) katoa esiintyy hänen mukaansa vain 48 prosentissa tapauksista. Subjektipronominin ja verbin persoonapäätteen välistä yhteyttä on tutkittu myös viron murteista. Lindströmin ym. (2008, 170) havaintojen mukaan subjektipronominin ja päätten välillä on tilastollisesti merkitsevä yhteys: pronomini esiintyy 60,6 prosentissa lauseista, jos verbissä on persoonapäätte *n*, mutta sen osuus lisääntyy 75,9 prosenttiin, jos päätte puuttuu.

Seuraavassa on esimerkki Koivistolta. Katkelman ainoat persoonapäätteelliset verbit (*söin*, *rynttysin*) esiintyvät ilman subjektipronominia. Toisaalta katkelmassa on aiemmin myös kaksi muutakin verbiä, jotka ovat pronominittomia, mutta niissä ei kuitenkaan ole myöskään persoonapäätettä (*oti*, *mäni*).

- (6) **mie *koppasi **esliina **helmahei **oti** sitä **rieskaa siit **uuni *suust ni / ja ja / **mäni** sit **sauna *uunii taas **katsomaa ja / **söin** *rynttysin / täy... ni *täys kylhä **mie** **olikii . / (LaX: Koivisto, 267.)

Kaakkoismurteissakaan on tuskin kyse niin tietoisesta puheen suunnittelusta, että puhujat valitsisivat yksittäiseen ilmaukseen joko subjektipronominin tai persoonapäätteen. Kyse voi olla paremminkin puhujien välisistä eroista: jotkut puhujista käyttävät vähemmän subjektipronominia, ja voi olettaa heidän suosivan silloin vastaavasti persoonapäätettä, jotta referenssin selvyys säilyy. Tarkastelen seuraavassa yksittäisiä puhujia ja heidän käyttämiään persoonapäätteitä ja pronomineja.

Taulukkoon 4 on koottu tässä tutkimuksessa mukana olevat puhujat, joista siis kukin edustaa yhtä pitäjää murrealueellaan. Merikarvialta on kaksi puhujaa, A ja B. Jokaiselle puhujalle on laskettu kaksi suhdelukua: sekä persoonapäätteiden että persoonapronominien osuus kaikista hänen käyttämistään yksikön ensimmäisessä persoonassa olevista finiittiverbeistä (myöntölauseissa). Suhdeluku kertoo toisin sanoen sen, kuinka suuri osuus puhujan käyttämistä muodoista on persoonapäätteellisiä sekä kuinka suuressa osassa verbeistä on ilmipantu pronomini. Suhdeluvun maksimiarvo on 1, joka tarkoittaa sitä, että päätettä tai pronominia käytetään kaikissa mahdollisissa tapauksissa (pronomini-idekseistä ks. Paunonen 2001).

Taulukko 4. Persoonapäätteen ja subjektipronominin suhdeluvut puhujittain/pitäjittäin.

	Pitäjä/puhuja	Persoonapäätte	Subjektipronomini	Y1-verbejä (kpl)
Kaakkoismurteet	Antrea	0,41	0,76	59
	Koivisto	0,19	0,80	193
	Lappee	0,17	0,90	78
	Luumäki	0,92	0,71	53
	Muolaa	0,52	0,85	71
	Nuijamaa	0,39	0,81	54
	Rautu	0,73	0,42	190
	Räisälä	0,60	0,78	77
Lounaismurteet	Kalanti	0,25	0,90	204
	Lappi Tl.	0,33	0,76	46
	Luvia	0,41	0,74	95
	Pyhämaa	0,39	0,84	92
	Rauma	0,36	0,84	55
Lounaiset siirtymämurteet	Ahlainen	0,04	0,98	139
	Merikarvia A	0,65	0,93	57
	Merikarvia B	0,70	0,74	23
	Noormarkku	0,69	0,85	26
	Pori	0,07	0,98	42

Satunnaismetsäanalyysi osoitti puhuja/pitäjä-tekijän olevan merkityksellinen etenkin kaakkoismurteissa ja lounaisissa siirtymämurteissa, mutta ei niinkään lounaismurteissa. Tulos saa yksityiskohtaisemman vahvistuksen oheisesta taulukosta 4, jossa puhujia on tarkasteltu erikseen. Taulukosta ilmenee, että persoonapäätteiden suhdeluvuissa on kaakkoismurteissa varsin suuria eroja: alin suhdeluku on 0,17 ja ylin 0,92. Lähes yhtä suuria eroja puhujien välillä löytyy lounaisista siirtymämurteista, joissa persoonapäätteen alin suhdeluku on 0,04 ja ylin 0,70. Sitä vastoin lounaismurteissa yksilöiden väliset erot ovat pienempiä: persoonapäätteen suhdeluvun vaihteluväli on 0,25–0,41. Subjektipronominin käyttö ei sen sijaan vaihtelee puhujittain niin paljon kuin päätteen käyttö. Puhujien pronomisuhdeluvut vaihtelevat yleensä murteesta riippumatta 0,71:n ja 0,98:n välillä (poikkeuksena tosin Rautu: 0,42). Nämä tulokset eroavat melko paljon Paunosen (2001, 159) havainnoista: hänen tutkimiansa helsinkiläisten pronomini-indeksit vaihtelevat 0,19:n ja 0,94:n välillä. Paunosen (mas. 161) mukaan eniten pronominia käyttävät helsinkiläisistä

nuoret puhujat. Tätä eivät omat havaintoni suomen muista murteista näytä tukevan: Lau-seopin arkiston informantit ovat olleet jo haastatteluhetkellä iäkkäitä.

Puhujittaisten suhdelukujen tarkastelu osoittaa myös, että minimaalisen referenssin keinoja ei käytä kukaan puhujista. Toisin sanoen ei ole yhtäkään sellaista puhujaa, jolla sekä persoonapäätteen että subjektipronominin suhdeluvut olisivat molemmat huomattavan pieniä. Diskurssifunktionaalisesti kovin vähäisten referentiaalisten keinojen käyttäminen saattaisi vaarantaa referentin tunnistamisen, joten tätä voi pitää hyvin odotuksenmukaisena tuloksena. Sen sijaan referenssin keinojen vastakkaista ilmiötä – redundanssia – esiintyy jonkin verran. Taulukon 4 mukaan on nimittäin joitakin puhujia, jotka käyttävät melko paljon sekä persoonapäätettä että subjektipronominia. Esimerkiksi luumäkeläisen puhujan persoonapääteindeksi on peräti 0,92, mutta hän käyttää kuitenkin subjektipronomiakin selvästi yli puolessa tapauksista (pronomini-indeksi 0,71). On toisaalta syytä huomata myös, että hänen subjektipronominilukunsa on koko vertailun pienimpiä.

Yksittäisten suhdelukujen kovin tarkkaan analysointiin ei tässä ole syytä ryhtyä, koska verbien määrä puhujittain jää taulukossa usein muutamaankin kymmeneen. Kovin pienissä osa-aineistoissa sattumallakin voi olla vaikutuksensa tuloksiin. Taulukon suhdelukuja voi kuitenkin vertailla kokonaisvaltaisesti ja pyrkiä etsimään puhujien välisestä vaihtelusta laajempia yhteyksiä. Taulukko 4 osoittaa nimittäin selvän tendenssin persoonapäätteen ja subjektipronominin välisessä suhteessa. Jos puhuja käyttää päätettä vähän, hän käyttää yleensä suhteellisen paljon pronominia tai toisinpäin. Erityisesti tämä on huomattavissa kahdella lounaisten siirtymämurteiden puhujalla: sekä Ahlaisten että Porin puhujilla on hyvin matala päätteen suhdeluku (0,04 ja 0,07), mutta molemmat käyttävät lähes aina subjektipronominia (0,98). Kaakkoismurteissa matalimman päätteindeksin (0,17) saaneella, Lappeesta kotoisin olevalla puhujalla, on murreryhmänsä korkein pronomini-suhdeluku (0,90). Samalla tavoin lounaismurteiden matalimman päätteindeksin (0,25) saaneella, kalantilaisella puhujalla, on vastaavasti ryhmänsä korkein pronomini-indeksi (0,90). Päätteen ja pronominin käänteinen suhde näyttää jossain määrin pätevän myös toiseen suuntaan. Kaakkoismurteista rautalaisella puhujalla on huomattavan matala pronominisuhdeluku (0,42), mutta hänellä esiintyy persoonapäätettä toiseksi eniten murreryhmässään (0,73). Myös lounaismurteissa ja lounaisissa siirtymämurteissa matalimman pronominiiluvun saaneella puhujalla on vastaavasti ryhmänsä korkein persoonapääteluku (Luvia: pronomini-indeksi on 0,74 ja päätteindeksi 0,41; Merikarvia B: 0,74 ja 0,70).

7. Lopuksi

Tässä artikkelissa olen pyrkinyt löytämään yksikön ensimmäisen persoonan verbinpäätteen (*n*) katoa tai säilymistä selittäviä tekijöitä lounaismurteissa, lounaisissa siirtymämurteissa ja kaakkoismurteissa. Tarkastelu osoitti, että sananloppuisen *n:n* katoa esiintyy persoonapäätteessä varsin yleisesti: kaikissa tutkituissa murteissa katotapaukset ovat enemmistönä, vaikka kaakkoismurteissa kato- ja säilymätapausten esiintymäerot ovatkin

pieniä. Tutkimus osoitti joka tapauksessa, että persoonataivutuksen päätteen koostuminen pelkäästä loppu-*n*:stä ei välttämättä vaikuta päätteen säilymiseen.

Tutkittavina olivat äänneypäristön, verbityypin, aikamuodon, referentiaalisen etäisyyden ja puhujien välisen eron vaikutukset persoonapäätteen esiintymiseen. Menetelmänä käytin satunnaismetsäanalyysia, joka asettaa selittäviä tekijöitä tärkeysjärjestykseen sen mukaan, millainen vaikutus niillä on tutkittavan ilmiön esiintymistodennäköisyyteen. Satunnaismetsä osoitti murteiden välisiä eroja. Lounaismurteissa vain äänneypäristö vaikuttaa selvästi persoonapäätteen esiintymiseen, kun taas muilla tekijöillä ei näytä olevan merkitystä. Siirtymämurteissakin äänneypäristö on merkityksellinen tekijä, mutta tässä murreryhmässä puhujien väliset erot ovat suuria, joten puhujittainen ero nousi analyysissa äänneypäristöä merkitsevämmäksi. Puhujien välinen tarkastelu paljasti, että erityisen vähän persoonapäätettä siirtymämurteissa käyttävät Porista ja Ahlaisista kotoisin olevat puhujat. Lyytikäinen ym. (2013, 80) huomauttavat juuri Porin murteen esittelyn yhteydessä, että porilaisten puhetta kuunnellessa huomiota kiinnittää sananloppuisen *n*:n katoaminen lähes aina. Porin seudun erilaisuus näyttää ilmenevän selvästi persoonapäätteenkin esiintymisessä, mutta toisaalta kaksi muuta saman alueen paikkakuntaa, Noormarkku ja Merikarvia, eivät eroa tässä suhteessa muista tutkituista murteista.

Kaakkoismurteissa katotapauksia on kaikkiaan vähemmän kuin kahdessa länsimurteessa. Satunnaismetsä osoitti, että kaakkoismurteissa päätteen esiintymiseen vaikuttaa useampi tekijä kuin kahdessa muussa murreryhmässä. Pitäjä/puhuja-tekijä osoittautui selvästi tärkeimmäksi, mutta myös subjektin muoto, äänneypäristö, verbityyppi ja aikamuoto selittävät analyysin mukaan päätteen esiintymistä. Koska äänneypäristö ei kaakkoismurteissa näytä olevan niin kattava selittäjä kuin lounaismurteissa, voi ajatella muiden tekijöiden silloin pääsevän analyysissa enemmän esiin. Jos pääte ei katoa systemaattisesti tietyissä äänneypäristöissä, voi esimerkiksi frekventti verbilekseemi (kuten *sanoa*) osoittautua vastaavasti merkitykselliseksi.

Kaakkoismurteiden alue ei ole murteellisesti yhtä yhtenäinen kuin lounaismurteiden alue (esim. Wiik 2004, 27, 285). Sitä onkin murretutkimuksissa jaettu edelleen alaryhmiin (esim. Keski- ja Etelä-Karjala, Kettunen 1940). Myös Leskinen (1971) on tarkastellut loppu-*n*:n esiintymistä erikseen useissa eri kaakkoismurteiden alaryhmissä. Se, että pitäjä/puhuja-tekijä nousi analyysissa tärkeimmäksi, voisikin siten kaakkoismurteiden tapauksessa selittyä eri pitäjien välisillä murre-eroilla. Eri puhujien lähempi tarkastelu (taulukko 4 luvussa 6) ei sitä kuitenkaan vahvista. Luumäki ja Lappee kuuluvat Kettusen (1940) kartoissa molemmat samaan, Etelä-Karjalan ryhmään, mutta persoonapäätteen esiintymisessä juuri näiden kahden pitäjän välinen ero on koko vertailun suurimpia. Analyysien tuloksia voi tulkita myös niin, että persoonapäätteen kato tai säilyminen on kaakkoismurteissakin enimmäkseen idiolekteihin sidoksissa oleva ilmiö.

Persoonamuotoisen verbin ominaisuuksiin kuuluu kyky viitata subjektin tarkoitteeseen yksinäänkin. Tästä syystä tutkittavana oli myös referentiaalisen etäisyyden vaikutus persoonapäätteen esiintymiseen. Analyysit osoittivat yhdenmukaisesti, että missään murreryhmässä etäisyydellä edelliseen samaviitteiseen mainintaan ei ole yhteyttä päätteen esiintymisen. Koska päätteen esiintymiseen vaikuttavat niin paljon muut tekijät,

pääteellisen tai päätteettömän verbin diskurssiasemalla ei ole merkitystä. Tulos on kuitenkin kiinnostava suhteessa siihen, mitä tiedetään subjektipronominin diskurssikäyttämisenestä. Pronominin esiintymisen on nimittäin havaittu olevan ensimmäisessäkin persoonassa sidoksissa referentiaaliseen etäisyyteen: pronomini jää todennäköisemmin pois, jos etäisyys on pieni. Jos ilman ilmipantua subjektipronominia oleva verbi esiintyy tavallisesti vain diskurssin kohdissa, joissa edeltävässä kontekstissa on samaviitteinen muoto, on viittauskohde kontekstin perusteella joka tapauksessa selvä ilman persoonapäätettäkin.

Tutkimuksen toisena päätavoitteena oli selvittää persoonapäätteen ja subjektipronominin esiintymisen yhteyttä. Luvussa 6 tarkastelin pronominin ja päätteen suhteellisia osuuksia puhujittain. Vertailu antoi selviä viitteitä siitä, että persoonapäätteen kato on yhteydessä lisääntyneeseen pronominin käyttöön: huomattavan vähän päätettä suosivat puhujat käyttävät yleensä vastaavasti enemmän pronominia verrattuna murreryhmänsä muihin puhujiin. Subjektipronominien tarkastelu paljastaa toisaalta sen, että pronominin käytössä ei kaiken kaikkiaan ole yhtä suuria puhujien välisiä vaihteluita kuin päätteen käytössä ja useimmilla pronominin ilmipaneminen on paljon yleisempää kuin sen poijättäminen. Puhutussa kielessä on siten varsin tavallista ilmaista persoonaa kaksinkertaisesti sekä verbinpäätteellä että subjektipronominilla.

Lauseopin arkiston litteraateissa esiintyvät erikoismerkit

*, **	paino	...	kesken jäänyt sana, takeltelu
'	liudennus	/	lyhyt tauko
-	sidonta	//	pitkä tauko

Lähteet

Aineslähteet

LaX = Lauseopin X-arkisto. Turun yliopiston kieli- ja käännöstieteiden laitos, Lauseopin arkisto. Turku.

Kirjallisuus

- BAAYEN, HARALD 2008: *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press, Cambridge.
- BYBEE, JOAN 2007: *Frequency of use and the organization of language*. Oxford University Press, Oxford.
- BYBEE, JOAN – SCHEIBMAN, JOANNE 1999: The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37–4, 575–596.
- GIVÓN, TALMY 1983: *Topic continuity in discourse: A quantitative cross-language study*. John Benjamins, Amsterdam.
- DUVALLON, OUTI 2006: Milloin pronominisubjekti jää pois puhutussa suomessa? ANNELI PAJUNEN ja HANNU TOMMOLA (toim.): *Kielitieteen päivät Tampereella 19.–20.5.2005*, 203–217. Tampere University Press, Tampere.
- HELASVUO, MARJA-LIISA 2014a: Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. *Journal of Pragmatics* 63, 63–78.
- 2014b: Searching for motivations for grammatical patterns. *Pragmatics* 24, 453–476.

- HOLLMANN, WILLEM B. – SIEWIERSKA, ANNA 2007: A construction grammar account of possessive constructions in Lancashire dialect: Some advantages and challenges. *English Language and Linguistics* 11(2), 407–424.
- ISK = HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN, RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- IVASKA, ILMARI 2014: Edistyneen oppijansuomen avainrakenteita. Korpusnäkökulma kahden kielimuodon tyypillisiin rakenteellisiin eroihin. *Virittäjä* 118, 161–193.
- KARLSSON, FRED 1977: Syntaktisten kongruenssijärjestelmien luonteesta ja funktioista. *Virittäjä* 81, 359–389.
- 1983: *Suomen kielen äänne- ja muotorakenne*. WSOY, Juva.
- KETTUNEN, LAURI 1940: Suomen murteet III A. Murrekartasto. SKST 188. Suomalaisen Kirjallisuuden Seura, Helsinki.
- LAPPALAINEN, HANNA 2004: *Variaatio ja sen funktiot. Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- 2006: Pronominisubjektin käytöstä ja poisjätöstä Kelan asiointikeskusteluissa. TARU NORDLUND, TIINA ONIKKI-RANTAJÄÄSKÖ ja TONI SUUTARI (toim.): *Kohtauspaikkana kieli. Näkökulmia persoonaan, muutoksiin ja valintoihin*, 37–64. Helsinki, Suomalaisen Kirjallisuuden Seura.
- LEHTONEN, HEINI 2006: Morfologinen variaatio maahanmuuttajataustaisten helsinkiläisnuorten puheessa. Juusela, Kaisu ja Nisula Katariina (toim.): *Helsinki kieliyhteisönä*, 255–274. Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos.
- LESKINEN, HEIKKI 1971: Tilastollisia havaintoja kaakkoismurteiden loppu-*n*:stä. *Virittäjä* 75, 343–369.
- LINDSTRÖM, LIINA – KALMUS, MERVİ – KLAUS, ANNELIIS – BAKHOFF, LIISI – PAJUSALU, KARL 2008: Ainsuse 1. isikule viitamine eesti murretes. *Emakeele Seltsi aastaraamat* 54, 159–185.
- LYYTIKÄINEN, ERKKI – REKUNEN, JORMA – YLI-PAAVOLA, JAAKKO 2013: *Suomen murrekirja*. Gaudeamus, Helsinki.
- PAJUNEN, ANNELI 2001: *Argumenttirakenne. Asiaintilojen luokitus ja verbien käyttäytyminen suomen kielessä*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- PAUNONEN, HEIKKI 2001: Yksikön 1. persoonan pronominin käyttö Helsingin puhemielessä. MARIANNE BLOMQUIST, MIRJA SAARI ja PETER SLOTTE (toim.): *Våra språk i tid och rum*, 154–168. Meddelanden från Institutionen för nordiska språk och nordisk litteratur vid Helsingfors universitet genom Mirja Saari B:21. Helsingfors universitet.
- RAPOLA, MARTTI 1966: *Suomen kielen äännehistorian luennot*. SKST 283. Suomalaisen Kirjallisuuden Seura, Helsinki.
- RUOPPILA, VEIKKO 1955: *Äyrämöismurteiden äännehistoria*. SKST 245. Suomalaisen Kirjallisuuden Seura, Helsinki.
- SKST = Suomalaisen Kirjallisuuden Seuran toimituksia.
- STRELLMAN, URPU 2005: *Persoonapronominin liikkäyttö: normin synty ja muotoutuminen*. Suomen kielen pro gradu -tutkielma. Helsingin yliopisto.
- TAGLIAMONTE, SALI – BAAYEN, HARALD 2012: Models, forests and trees of York English: *Was/were* variation as a case study of statistical practice. *Language variation and Change* 24, 135–178.
- WIIK, KALEVI 2004: *Suomen murteet. Kvantitatiivinen tutkimus*. Suomalaisen Kirjallisuuden Seura, Helsinki.
- 2006a: *Sano se murteella. Minkälaisia suomen murteet ovat ja missä niitä puhutaan?* Pilot-kustannus, Tampere.
- 2006b: *Suomen murteiden kartasto*. Turku 2006.
- VÄÄNÄNEN, MILJA (tekeillä): Yksikön 1. persoonan subjektin ilmaiseminen suomen vanhoissa murteissa.

MILJA VÄÄNÄNEN: Coding the first-person singular subject: the occurrence of a verbal agreement marker and subject pronoun in Finnish dialects

The article examines the means of coding first-person singular subjects in three regional dialects of Finnish, namely the southeastern, southwestern and southwestern transitional dialects. In Finnish, the first-person singular pronoun can be dropped because the verb agrees with the subject in person. When the subject is expressed by a pronoun, person is thus marked twice. If the pronominal subject is omitted, only the agreement marker (-*n*) refers to the subject referent:

(mä)	meni-n
(I)	go-SG1

Some Finnish dialects show variation in the form of the verbal person marker -*n*: in southwestern and southeastern dialects, this suffix is often omitted. However, the suffix can also be retained even in these dialects. The aim of the study is to find factors which explain this variation concerning the occurrence of the verbal suffix. Another question concerns the effect of suffix omission upon occurrence of the pronoun: is there a tendency to use the pronoun more if the suffix is omitted?

The dataset comes from the Syntax Archives of the University of Turku, from which a selection of 1,549 sentences with first-person singular verbs was extracted. The extracted sentences have been coded for a variety of factors, such as phonetic environment (the quality of the next phone), occurrence of the subject pronoun, verb type, referential distance to the previous coreferential form (see Givón 1983) and speaker.

The dataset is analysed using statistical methods (random forest). The analyses show that while phonetic environment is the most important factor in southwestern dialects, individual differences explain most of the variation in the other dialects. In southeastern dialects, however, the occurrence of the subject pronoun and the verb type also proves to be significant factors. Even if the occurrence of the verbal ending (-*n*) depends generally on individual differences, a detailed comparison between speakers reveals a reverse connection between the subject pronoun and the ending. Speakers who have a clear tendency to omit the verbal ending also have a tendency to use the pronoun more in comparison to speakers using the ending regularly.

Milja Väänänen
 mmvaan@utu.fi
 Suomen kieli ja suomalais-ugrilainen kielentutkimus
 20014 Turun yliopisto