

Jaakko Häkkinen

Kieltenvälinen vertailu historiallisessa kielitieteessä

0. Esipuhe

En pyri tekemään tästä kirjoituksesta perinpohjaista historiallisen kielitieteen teorian, metodin ja tutkimushistorian esittelyä, koska näistä aiheista on jo saatavana ajantasaista tietoa (Campbell 1998; Fox 1995). Lähestymiskulmani on käytännönläheinen: kerron esimerkein havainnollistaen, kuinka omalla alallani eli uralilaisten kielten diakronisessa tutkimuksessa sovelletaan vertailevan menetelmän erilaisia tyyppejä käytäntöön. Esimerkkien avulla esittelen samalla alan uusia tutkimustuloksia.

Kirjoitus koostuu neljästä luvusta: Ensimmäisenä on johdantoluku, jossa lukija tutustutetaan historiallisen kielitieteen olennaisimpiin merkintätapoihin, käsitteisiin ja käytäntöihin. Toisessa luvussa esitellään tyypiltään vertikaaliset (pystysuuntaiset), yhden kielen sisällä pysyttelevät vertailumenetelmät: sisäinen rekonstruktio, jossa yhden kielen sisäisen vaihtelun perusteella pyritään rekonstruoimaan menneisyyteen vaihteluton alkutila, ja kielenmuistomerkkien (kohteet joissa on kirjoitusta) tutkiminen, jolla saadaan ensikäden tietoa menneisyyden kielestä.

Kolmannessa luvussa esitellään tyypiltään horisontaaliset (vaakasuuntaiset) varsinaiset eli kielten väliset vertailumenetelmät: vertaileva rekonstruktio, jossa sukukieliä vertailemalla pyritään rekonstruoimaan menneisyyteen kantakieli, ja kontaktikielten välinen vertailu, jonka avulla voidaan hyödyntää toisesta kielilinjasta vertailumenetelmillä saavutettua tietoa. Neljäs luku on lyhyt yhteenveto. En käsittele erikseen tutkimuskohteen perusteella määrittyviä erikoisaloja kuten etymologiaa (sanojen alkuperän tutkimusta) tai onomastiikkaa (nimistöntutkimusta), koska näissä voidaan soveltaa tarpeen mukaan erilaisia vertailutyyppejä.

1. Johdanto: historiallinen kielitiede

1.1. Merkintätavat

Kaikki kielenaineokset kursivoidaan silloin kun ne eivät ole osa selittävää tekstiä, esimerkiksi ”suomen kielen sana *sana*” ja ”suomen kielen äänne *ä*”. Kielenaines ilman lisämerkkejä merkitsee

sellaisenaan nykykielestä tavattua sanaa, äännettä tai muotoa. Kulmasulkeissa oleva kielenaines merkitsee kirjakielen mukaista asua: <*kengät*>. Esimerkiksi gootinkieliset sanat ilmaistaan näin, koska kielellä ei ole enää puhujia, vaan kaikki sen sanat tunnetaan vain kirjoitetussa muodossa. Pystyviivojen välissä ilmaistaan arkkifoneemi tai arkkifoneeminen muoto: |-ssA|. Arkkifoneemi A merkitsee tässä, että suomen inessiivin pääte voi esiintyä joko asussa /-ssa/ tai /-ssä/, riippuen siitä onko sana vokaalisoinnun suhteen takainen vai etinen. Hakasulkeissa oleva kielenaines merkitsee ääntämystä eli foneettista asua: [ä], [k^hejät] (uralilaisia kieliä varten on kehitetty oma foneettinen tarkekirjoituksensa: ks. Posti & Itkonen 1973). Vinoviivojen välissä oleva kielenaines merkitsee foneemista eli fonologista asua: /kengät/.¹ Sanan tai muun kielenaineksen merkitys ilmaistaan apostrofien (heittomerkit, puolilainausmerkit) sisällä kursivoimattomana: *ketara* 'reen kaplas; sepä'.

Asteriskia * käytetään historiallisessa kielitieteessä merkitsemään sukukielten vertailun perusteella luotua rekonstruktioita. Kaksoisasteriskia ** puolestaan voidaan tarvittaessa käyttää erotukseksi asteriskista silloin kun ilmaistaan (1) hypoteettista, (2) kadonnutta tai (3) analogisen yleistyksen syrjäyttämää äännelaillista sanaa. Ensimmäisen tyyppin esimerkki: jos kantauralista periytyvällä pohjoissaamen sanalla *njoammil* 'jänis' olisi vastine suomessa, se olisi ***nomala*. Toisen tyyppin esimerkki: paikannimi Tornio on peräisin sanasta ***tornio* 'keihäs', jota Lapissa ei enää tunneta mutta joka on säilynyt myöhäisiin aikoihin kaakkoishämäläisissä murteissa. Kolmannen tyyppin esimerkki: suomen sana *lahti* on analoginen eli taivutusvartalon perusteella yleistetty (paikannimiä taivutetaan yleensä paikallissijoissa: Lahdessa, Lahdesta, Lahteen), ja se on syrjäyttänyt äännelaillisen nominatiivimuodon ***laksi*, joka tunnettiin vielä vanhassa kirjasuomessa ja joka on edelleen käytössä karjalan kielessä (vrt. *kaksi* : *kahden*).

Kielenpiirteiden muuttumista ilmaistaan nuolenkärjillä > ja <, esimerkiksi: *š > h merkitsee että aikaisempaan kielentasaan rekonstruoitu *š on kehittynyt nykykielen h:ksi. Myös h:n eteen voidaan laittaa asteriski, jos halutaan ilmaista että muutos on tapahtunut jo selvästi nykykieliä varhaisemmissa rekonstruoituissa kantakielitasoissa. Johtosuhdetta tai lainautumista voidaan merkitä nuolilla → ja ←, esimerkiksi: suomen *lasi* ← ruotsin *glas*. Merkillä ~ erotetaan eri kielten tai murteiden vastineet toisistaan, esimerkiksi: suomen *terva* ~ viron *tõrv*, tai vaihtoehtoiset rekonstruktioita: **kakta* ~ **kakti*.

1.2. Ääntenmuutos ja rekonstruktio

Kielen muuttumisen kannalta tärkeä käsite on ääntenmuutos, jollainen on esimerkiksi myöhäiskantasuomeen tultaessa tapahtunut $*š > h$. Ääntenmuutokset voivat olla ehdottomia, jolloin äänne muuttuu kaikissa asemissa (kuten edellisessä esimerkissä), tai ehdollisia, jolloin äänne muuttuu vain tietyssä asemassa (esimerkiksi $*m > n$ sanan lopussa) tai äänneympäristössä (esimerkiksi $*ti > si$ eli $*t$ muuttuu s :ksi vain i :n edellä). Ehdoton ääntenmuutos muuttaa kielen äännejärjestelmää, joko rikastuttaen (jos muutoksen tuloksena kieleen syntyy aivan uusi äänne) tai köyhdyttään sitä (jos äänneitä lankeaa yhteen). Ehdollinen ääntenmuutos puolestaan lisää kielessä paradigmaattista vaihtelua eli tekee sanojen eri muodoista erinäköisiä (taivutus *vesi* : *veden* eikä $**vesin$ on seurausta kolmesta ehdollisesta ääntenmuutoksesta, ks. luku 2.2).

Rekonstruktio on prosessi, jossa pyritään löytämään se kantakielen äänne, josta kaikkien tytärikielten äännevästineet sujuvimmin ja uskottavimmin (eli vähimmin ääntenmuutoksin) selittyvät. Ääntenmuutokseksi rekonstruoidaan $*ti > si$ eikä $*s > *t$ muiden vokaalien kuin i :n edellä, koska ensin mainittu on perustellumpi sekä *vesi*-tyypin sanojen taivutusmuotojen ja kielen muun sanaston perusteella (ks. luku 2.2) että sukukielten edustuksen perusteella (ks. luku 3.1).

1.3. Periytyminen, lainautuminen ja syntyminen

Historiallis-vertailevaan kielentutkimukseen kuuluu olennaisesti ajallinen syvyys; tämä ominaisuus tekee kielten vertailusta diakronista ja erottaa sen synkronisesta kielentutkimuksesta, esimerkiksi yksittäisen nykykielen tutkimisesta. Oikeastaan aina kun vertaillaan kahden tai useamman kielen välisiä yhtäläisyyksiä, mukana on automaattisesti aikasyvyys: yhteinen piirre ei esiintyisi vertailtavissa kielissä, ellei se olisi aikaisemmin periytynyt, lainautunut tai syntynyt niihin. Nämä kolme alkuperämahdollisuutta liittyvät kaikkiin kielen ilmiöihin, mutta havainnollistan niitä lyhyesti etymologian eli sanojen alkuperän ja sen tutkimuksen osalta.

Suomen sanan *kala* etymologia on, että se periytyy jo uralilaisesta kantakielestä, missä sana oli aivan samassa asussa ja merkityksessä $*kala$; tähän tulokseen päädytään sukukielten vertailun perusteella. Suomen sanan *rengas* etymologia on, että se on lainattu varhaisesta germaanisesta kielimuodosta; tähän tulokseen päädytään kielikontaktien ja lainasanojen tutkimuksen perusteella. Sanojen ”syntyminen” on sen sijaan mutkikkaampi asia. Sanat eivät tietystikään synny tyhjästä, vaan niillä on jonkinlainen vähintään assosiatiivinen suhde kielen muuhun sanastoon. Johdokset ja yhdyssanat voidaan lukea niihin liittyvien perussanojen yhteyteen, mutta jäljelle jää vielä suuri

joukko sanoja, jotka ovat ilmestyneet kieleen lähes tyhjästä: äänteellisesti motivoidut eli ekspressiiviset sanat.

Ekspressiivisana voi olla joko ääntä jäljittelevä eli onomatopoeettinen tai ääntä kuvaileva eli deskriptiivinen, ja ekspressiivisen sananmuodostuksen perusyksikköä voidaan nimittää fonesteemiksi. Esimerkiksi fonesteemi *#hVri-* (# on tässä fonesteemin merkitsijä ja *V* on mikä tahansa vokaali) liitetään tietynlaiseen äänentuottamiseen: *hyristä, hurista, horista, höristä*. Ekspressiivisanoihin liittyy usein myös affekti eli miellelyhtymiä tuottava tunnelataus: esimerkiksi suomen vokaali *ö* on usein voimakkaasti affektinen, ja tuottamalla jonkin fonesteemin pohjalta *ö*:llinen sana tavoitetaan usein juuri tietynlainen sävy (esim. *örveltää, mökeltää, höpöttää, sönköttää, röhöttää*). Ekspressiivisanat on ryhmä, johon voi syntyä suuri määrä uusia sanoja lyhyessäkin ajassa, mutta toisaalta tällaiset sanat saattavat olla lyhytikäisiä eivätkä aina vakiinnu koko kieliyhteisön käyttöön. (Jarva 2003; Kulonen 2010.)

1.4. Kantakieli

Selvimmän historiallisen kielitieteen aikaulottuvuus näkyy kantakielen käsitteessä: siihen menneisyyden pisteeseen, jossa erillisten nykykielten linjat vertailevassa menetelmässä kohtaavat toisensa, muodostuu uusi kielimuoto, rekonstruoitu kantakieli. Rekonstruoitu kantakieli ei ole täysin sama asia kuin se todellinen kantakieli, josta kielikunnan kielet ovat todellisuudessa polveutuneet, johtuen eräistä vertailevan menetelmän rajoituksista. On esimerkiksi havaittu, että vertaileva rekonstruktio tuottaa kantakielen, joka on aina säännöllisempi kuin tytärikielien: todellisen kantakielen morfofonologisen vaihtelun tavoittaminen on vertailevan rekonstruktion avulla hankalaa (Korhonen 1974). Joka tapauksessa kielen muuttumista tutkittaessa kantakieli on välttämätön käsite – ennen tämän käsitteen (ja historiallisen kielitieteen) keksimistä kielten samankaltaisuutta pyrittiin selittämään johtamalla kielet suoraan toisista kielistä, esimerkiksi suomi hepreasta tai kreikasta.

Kantakielen kohdalla on painotettava, että ilmiönä se on todellinen, mutta rekonstruktiona vain viitteellinen. Suomi ja unkari polveutuvat epäilyksettä yhteisestä kantakielestä, vaikkei sitä voidakaan yksityiskohtaisesti tavoittaa: vain sellaiset ainekset, jotka ovat säilyneet kummallakin taholla, voidaan ongelmitta palauttaa kantakieleen. Tästä seuraa, että rekonstruoitu kantakieli on aina aineksiltaan niukempi kuin sen nykyiset tytärikielien (ja todellinen, joskus olemassa ollut kantakieli). Jos ajatellaan kaavamaisesti, että on neljä yhtä todennäköistä vaihtoehtoa (piirre säilyy

molemmissa kielissä, piirre katoaa suomesta mutta säilyy unkarissa, piirre katoaa unkarista mutta säilyy suomessa, piirre katoaa molemmista kielistä), niin rekonstruoituun kantakieleen tavoitetaan vain neljäsosa todellisessa kantakielessä esiintyneistä piirteistä. Tämä ilmiö, jota voidaan kutsua vaikkapa nimellä *diakroninen kato*, kertautuu aina syvemmän tason kantakielessä, olettaen että sukupuu (kielikunnan taksonomisen rakenteen kuvaaja) on binaarisesti jakautuva eli kukin kantakielitaso hajoaa aina kahtia, ne edelleen kahtia jne. Silloin jos kantakieli hajoaa suoraan kolmia, diakroninen kato on pienempi, olettaen että piirteen rekonstruoimiseksi kantakieleen riittää sen esiintyminen kahdessa tytärikielissä kolmesta. Tällöin diakroninen kato tapahtuu suhteessa 4 : 8 eli 1 : 2, kun suhde kahtia jakautuvan kantakielen tilanteessa on 3 : 4. Tarkastelukulmaa kääntäen voidaan sanoa, että *diakroninen säilymä* on kolmia jakautuvassa kantakielessä 1 : 2 ja kahtia jakautuvassa kantakielessä 1 : 4. Neljiä jakautuvassa kantakielessä diakroninen säilymä on joko 5 : 16, jos kantakieleen palautettavalta piirteeltä vaaditaan esiintyminen vähintään kolmessa kielihaarassa, tai 6 : 16, jos hyväksytään myös sellaiset piirteet, jotka esiintyvät vain kahdessa kielessä jotka kuitenkin ovat kauimpana toisistaan, mutta peräti 11 : 16 jos hyväksytään missä tahansa kahdessa kielessä esiintyvät piirteet.

Kantakieli hajoaa yleensä aikojen kuluessa tytärikieliksi, joista voi edelleen tulla uusien tytärikielten kantakieliä. Samasta kantakielestä polveutuvien sukukielten kokonaisuus on nimeltään *kielikunta*, ja muusta kielikunnasta erottuvien lähisukuisten kielten kokonaisuus on nimeltään *kielihaara*. Suomi kuuluu itämerensuomalaiseen kielihaaraan yhdessä karjalan, inkeröisen, lyydin, vepsän, vatjan, viron ja liivin kanssa, ja kaikki itämerensuomalaiset kielet polveutuvat välikantakielestä nimeltään (myöhäis)kantasuomi. Välikantakielet kantasuomi, kantsaame, kantaamordva, kantaamari, kantaapermi, kantaunkari, kantaamansi, kantaahanti ja kantasamojedi puolestaan ovat kantaauralin tytärikieliä, joista kantaunkaria lukuun ottamatta jokaisella on vähintään kaksi nykypäivään säilynyttä tytärikieltä.

Edellä listatut välikantakielet ovat kiistattomia yksikköjä, mutta uralilaisten kielten sukupuussa on nähty myös erilaisia näiden ja kantaauralin väliin sijoittuvia kielentasoja, jotka ovat kuitenkin nykyään kiistanalaisia. Siinä missä kantasuomi, kantsaame ja muut mainitut eroavat kantaauralista kymmenien äänne muutosten perusteella (kantsaamen osalta Korhonen 1981; Sammallahti 1998; kantasuomen osalta Lehtinen 2007), ei esimerkiksi itämerensuomalais-saamelaisen välikantakielen tai ugrilaisen välikantakielen (> kantaunkari, -mansi ja -hanti) tueksi ole voitu esittää parhaimmillaankaan kuin yksittäisiä äänne muutoksia. Itämerensuomalais-saamelaisen kantakielen osalta ks. Sammallahti (1999) vs. T. Ikonen (1997); ugrilaisen välikantakielen osalta ks. Honti

(1998) vs. Salminen (1998). Lähinnä näitä kiistanalaisia kantakieliä on perusteltu runsaalla yhteisellä sanastolla, mutta yhteistä sanastoa voi ilmaantua kielihaaroihin muutenkin kuin yhteisestä kantakielestä perittyinä – esimerkiksi lainautumalla joko kielihaarasta toiseen tai kolmannelta kielihaarasta kumpaankin. Toistaiseksi ei siis ole varmaa, onko esimerkiksi itämerensuomalais-saamelaisista kantakieltä ollut olemassa, ja jos onkin, olisi vaiheen pitänyt olla erittäin lyhyt verrattuna kantasuomen ja kantasaamen erilliskehitysten kausiin. Välikantakielistä ja erilaisista sukupuista on enemmän luvun 3.1 lopussa.

Kantakieleen liittyy olennaisesti myös käsite *divergenssi*, jolla tarkoitetaan yhtenäisen kielimuodon eriytymistä, esimerkiksi kantakielen hajoamista tytärikieliksi. Tätä hajoamista kuvataan yleisesti sukupuulla. *Konvergenssi* puolestaan voi merkitä divergenssin vastakohtaa yleensä eli kielten lähentymistä pitkään jatkuneiden kontaktien seurauksena, mutta se voi merkitä myös riippumatonta rinnakkaiskehitystä, eli tilannetta jossa kaksi kieltä kehittyi samaan suuntaan sisäisten tendenssiensä vuoksi, ilman kontaktivaikutusta. Jälkimmäistä tilannetta voi nimittää tarkentaen esimerkiksi riippumattomaksi konvergenssiksi.

1.5. Rekonstruktiot symboleina

Rekonstruoitu kantakieli on symbolinen järjestelmä, jota ei pidä lukea liian konkreettisesti.

Kantauraliin rekonstruoidaan liudentunut *ś selittämään vastaavuussarjaa suomen *s* ~ saamen *š* ~ mordvan *ś/ž* ~ udmurtin *ś/ž* jne. siksi, koska tällainen lähtöäänne sopii hyvin kantakielen äännejärjestelmän osaksi ja koska se on foneettisesti uskottava lähtökohta kyseiselle vastaavuussarjalle. Nimittäin kantakielessä oli kaksi muutakin sibilanttia, *s (> suomen *s* ~ saamen *s* ~ mordvan *s/z* ~ udmurtin *s/z*) ja *š (suomen *h* ~ saamen *s* ~ mordvan *š/ž* ~ udmurtin *š/ž*), joten kolmas sibilanttivastaavuussarja voidaan perustellusti palauttaa liudentuneeseen *ś:ään.

Koska meillä ei kuitenkaan voi olla varmaa tietoa kyseisen kantauralin äänneen laadusta vaan kaikki mitä tiedämme on nykykielten edustus, voitaisiin rekonstruoida äännettä periaatteessa yhtä hyvin merkitä vaikka numerolla 4 tai symbolilla #. Toki tutunnäköinen kirjain on abstrakteja symboleja havainnollisempi, etenkin jos rekonstruoidun äänneen foneettinen arvokin on erittäin suurella todennäköisyydellä oikein (kuten *s:n kohdalla), mutta silti on hyvä muistaa sen symboliarvo: rekonstruoitu *s on merkki joka edustaa nykykielten vastaavuussarjaa suomen *s* ~ saamen *s* ~ mordvan *s/z* ~ udmurtin *s/z*. Rekonstruktiot voivat muuttua, mutta nykykielten vastaavuussarjat eivät.

1.6. Historiallisen kielitieteen vertailun tyypit

Historiallisessa eli diakronisessa kielitieteessä tutkitaan kielen muuttumista, joten se perustuu jo lähtökohdiltaan vertailuun: kielen ajallinen muuttuminen nähdään vain kun verrataan kielen kahta eri-ikäistä vaihetta toisiinsa. Diakronisuus voidaan saavuttaa kahdella eri tavalla. *Vertikaalisen vertailun* piiriin kuuluvat historiallisessa kielitieteessä kielenmuistomerkkien (vanhat kirjoitukset, mukaan lukien riimukivet, brakteaatit jne.) tutkiminen ja sisäinen rekonstruktio. Tällainen vertailu pysyy tyypillisesti yhden kielimuodon piirissä. *Horisontaalisen vertailun* piiriin kuuluu varsinainen vertaileva menetelmä sekä sukukielten (vertaileva rekonstruktio) että kontaktikielten välillä (kontakti- ja lainasanatutkimus). Tällaisessa vertailussa ikään kuin ”kolmiomitataan”² menneisyyden kieltä kahden tai useamman nykykielen avulla. Kaikissa menetelmissä on epävarmuustekijöitä, joita käsitellen kunkin menetelmän esittelyn yhteydessä.

Kokonaiskuvan kannalta kaikki mahdolliset vertailun tasot on otettava huomioon, muuten saatetaan päätyä puutteelliseen lopputulokseen. Yhden kielen sisäisen rekonstruktion avulla saadaan jonkinlainen kuva kielen muutoksista, ja usein myös muutosten keskinäisestä järjestyksestä. Kahden eri kielen sisäiset muutokset voivat olla yhtenevät, jolloin saadaan erotettua kielikunnan sisäinen yksikkö, kielihaara. Kieltenvälisen vertailun avulla selvitetään kielten ja kielihaarojen tarkemmat suhteet eli kielikunnan taksonomia (jota havainnollistetaan sukupuulla) sekä rekonstruoidaan kantakielen aineksia kielen eri tasoilla. Lainasanatutkimuksen avulla voidaan tarkentaa kielen muutosten järjestystä eli suhteellista kronologiaa, ja lisäksi lainasanojen avulla voidaan saada selville mikä edustus on alkuperäinen, mikäli tästä on epäselvyyttä. Lainasanojen avulla saatetaan myös saada selville, mitä reittiä kielen muutos on tapahtunut, jos esimerkiksi äänneen muuttuminen toiseksi voisi teoriassa tapahtua useammankin vaihtoehdoisen välittävän äänneen kautta.

2. Vertikaalinen vertailu

Historiallisen kielentutkimuksen kaksi päävertailumenetelmää ovat siis vertikaalinen vertailu ja horisontaalinen vertailu. Vertikaalinen vertailu ei tyypillisesti ole kieltenvälistä vaan kielensisäistä, ja siinä tavoitetaan yhden kielen rajoissa kielen varhaisempi vaihe. Käsitellen seuraavassa aluksi kielenmuistomerkkejä ja sen jälkeen sisäistä rekonstruktioita.

Vertikaalisessa eli pystysuorassa vertailussa tavoitetaan yhden kielen rajoissa kielen varhaisempi vaihe, kun nykykieltä verrataan siihen mitä tiedetään kielen aikaisemmasta vaiheesta, ilman muihin kieliin vertailemista. Vertikaalista vertailua voidaan kuvata ideaalitapauksessa muodoltaan I-tyyppiseksi: yhden kielen kehitysjanalta valitaan kaksi eriaikaista pistettä, joita vertaillaan. Kuitenkin vain nykykielen vertaaminen vanhoihin kielenmuistomerkkeihin edustaa puhdasta I-tyyppiä, ja sekin vain sillä edellytyksellä, että kielenmuistomerkkiin tallennettu kielimuoto edustaa täsmälleen samaa kielilinjaa kuin vastaava nykykieli. Esimerkiksi nyky-suomi kirjakielenä ei ole Agricolan kirjoittaman vanhan kirjasuomen suora jatkaja, koska Agricola kirjoitti lähinnä Turun murteella (tosin kotimurteellaan höystettynä), kun taas nyky-suomi on itä- ja länsimurteiden kompromissi (Koivusalo 1988; K. Häkkinen 1994). Niinpä nyky-suomen ja vanhan kirjasuomen suhdetta voidaan kuvata pikemminkin Y-tyyppiseksi: ne edustavat rinnakkaisia joskin lähekkäisiä kielilinjoja, eivät täysin samaa kielilinjaa.

Sisäinen rekonstruktioakaan ei edusta puhdasta I-tyyppiä. Siinä yhden kielimuodon nykyisen vaihtelun taustalle menneisyyteen oletetaan vaihtelematon lähtötaso. Tätä menetelmää voidaan kuvata T-tyyppiseksi: vaakaviiva edustaa nykytasossa ilmenevää paradigmaattista vaihtelua, ja pystyviiva edustaa tämän vaihtelun heijastamista menneisyyden vaihteluttomaan tilaan. Joka tapauksessa Y- ja T-tyypit voidaan luokitella ideaalin I-tyypin reaalitumiksi, koska niissä vertailu pysyy yhden kielen sisäisenä.

2.1. Kielenmuistomerkkien tutkiminen

Kirjoitustaito keksittiin jo yli 4 000 vuotta sitten, ja varhaisimmat kirjoitukset tunnetaan Sumerista (nuolenpääkirjoitus) ja Egyptistä (hieroglyfikirjoitus). Eräitä indoeurooppalaisia kieliäkin on alettu kirjoittaa jo yli 3 000 vuotta sitten (heetti, kreikka), samoin kiinaa ja seemiläisiä kieliä. Germaania kieliä alettiin kirjoittaa riimuilla pian ajanlaskun alun jälkeen. Uralilaisista kielistä vanhimmat kielenmuistomerkit ovat muinaisunkaria (1000-luvulta jaa.; Kulonen 1993) ja muinaispermiä (1300-luvulta; Bartens 2000); suomea alettiin kirjoittaa 1500-luvulla, vaikka satunnaisia sanoja tai pätkiä oli jo aikaisemmin muunkielisen kirjoituksen seassa (K. Häkkinen 1994). Novgorodin tuohikirjeistä yksi on kirjoitettu ilmeisesti karjalansukuisella itämerensuomalaisella kielellä 1200-luvulla (Laakso 1999).

Kielenmuistomerkkien tärkeys on siinä, että ne näyttävät kuvan siitä todellisesta kielestä, jota tiettyyn aikaan puhuttiin (tai kirjoitettiin) – kyseessä ei siis ole nykytilanteen pohjalta päätelty rekonstruktio lukuisine epävarmuustekijöineen. Toisaalta vanhojen kielenmuistomerkkien ongelmana on vakiintumaton oikeinkirjoitus: esimerkiksi Mikael Agricola saattoi kirjoittaa samankin sanan tai äänteen monella eri tavalla. Vaikka suomea pidetään hyvin säilyneenä kielenä ja vaikka sitä on alettu kirjoittaa vasta noin 500 vuotta sitten, eroavat vanhimmat kielenmuistomerkit silti melko selvästi nyky-suomesta. Esimerkiksi klusiilien *k* ja *t* heikossa asteessa esiintyvät spirantit olivat vielä jossain määrin säilyneinä Agricolan kielessä, ja niitä merkittiin kirjoituksessa vastaavasti <gh>:lla tai <g>:llä ja <dh>:lla tai <d>:llä: <wighat> 'viat' (<*viyat : *vika), <wlgos> 'ulos' (<*ulyos : *ulko-), <tiedhe> 'tiedä' (<*teeðäk : *teetä-). (K. Häkkinen 1994.)

Kielenmuistomerkkien tutkiminen voi olla puhtaasti synkronista eli vertailematonta, ainakin aluksi, yritettäessä saada käsitys kirjoituksen takana olevasta sisällöstä ja kielestä usein horjuvasta oikeinkirjoituksesta ja vaihtelevista kirjakielen konventioista huolimatta. Sitten kun kielenmuistomerkkin kieli on avattu ja tunnistettu, voidaan sitä alkaa vertailemaan oletettuun saman kielilinjan nykykieleen. Toisaalta vertailu voi olla mukana jo kielen tunnistamisessa: esimerkiksi Novgorodin tuohikirjeen 292 (ainoa kokonaan itämerensuomalaisella kielellä kirjoitettu) sisällön tulkintaan saattaa vaikuttaa se, mitä kielimuotoa sen katsotaan edustavan.

Suomen kielenmuistomerkkejä tutkiva saa tutkia vanhoja kirjoja joiden kirjoitus on laadittu latinalaisella aakkostolla (tosin erilaisilla kirjasintyypeillä), mutta venäjän kielenmuistomerkkeihin kuuluu esimerkiksi tuohikirjeitä, ja aakkostojakin oli käytössä kaksi erilaista: glagoliittinen sekä sen ja kreikkalaisen aakkoston pohjalta luotu kyrillinen, jonka jatkaja on venäjän nykyinen aakkosto. Skandinaavisiin kielenmuistomerkkeihin kuuluu kivipaasia ja brakteaatteja (metallisia riipuksia tai kolikoita), joihin molempiin kaiverrettiin kantaskandinaavin sanoja riimukirjoituksella (futhark; kirjoitusmerkkien lukumäärä ja ulkomuoto vaihteli ajan ja paikan mukaan).

2.2. Sisäinen rekonstruktio

Sisäisessä rekonstruktiossa havainnoidaan ja pyritään selittämään kielessä nykytasolla ilmeneviä vaihteluja. Esimerkiksi suomen kieltä pidetään morfologiselta tyypiltään agglutinatiivisena, eli sijapäätteet on liitetty mekaanisesti sanan vartaloon: *auto* : *auto-ssa*. Kuitenkin kielessä on myös epäodotuksenmukaista sanavartalon sisäistä vaihtelua, sellaista kuin sanassa *vesi* : *vede-ssä*. Samaa tyyppiä, jossa vain nominatiivissa on *s* ja kaikissa muissa sijamuodoissa *t* tai *d*, kuuluu muitakin

sanoja (*susi*, *käsi*, *kaksi*, *laksi* 'lahti', *kuusi* '6'). Tällaisen kielensisäisen vaihtelun taustalle rekonstruoidaan vaihtelematon lähtötilanne: **veti* : **vete-n* : **vete-nä*. On kaksi syytä, miksi lähtömuodot rekonstruoidaan *t*:lliseksi eikä *s*:lliseksi: ensiksikin meillä on jo ikivanhoja pelkästään *s*:llisiä sanoja (*kuusi* : *kuusen*, *jousi* : *jousen*) muttei ikivanhoja pelkästään *t*:llisiä sanoja (toki nuoria lainasanoja kuten *vati* : *vadin* jonka jo *i*-vartaloisuus osoittaa nuoreksi), ja toiseksi nominatiivi on ainoa sija, jossa esiintyy *s*:llinen asu. *Vesi*-tyyppisten sanojen perusteella voidaan siis päätellä, että jossain vaiheessa suomen kielen menneisyydessä on tapahtunut muutos **ti* > *si* (alkuperäinen **t* on muuttunut *s*:ksi vain *i*:n edellä). Periaatteessa sisäinen rekonstruktio on abstraktio ja viittaa vain epämääräiseen menneisyyteen, mutta käytännössä sen avulla saavutettavat rekonstruktioasot ovat yhdistettävissä tiettyyn konkreettiseen menneisyyden kielentagoon, kuten jäljempänä selviää.

Kun nyt sisäisen rekonstruktion perusteella voidaan olettaa johonkin suomen kielen esimuotoon äänne­muutos **ti* > *si*, saadaan tälle vahvistus lainasanoista (ks. luku 3.2): niinpä balttilaisissa lainasanoissa liettuan sanaa *tiltas* vastaa suomen *silta* ja liettuan sanaa *martì* vastaa suomen sana *morsian*, mutta liettuan sanaa *dervà* vastaa suomen sana *terva* (muutos on siis tapahtunut lainasanoissakin vain *i*:n edellä muttei foneettisesti läheisen *e*:n edellä; Kalima 1936). Samaa todistaa myös vertaileva rekonstruktio (ks. luku 3.1), koska sukukielissä tätä muutosta ei ole tapahtunut: suomi *siivota* ~ pohjoissaame *divvut* 'korjata'; suomi *sika* ~ ersämordva *tuvo*.

Sisäisen rekonstruktion keinoin saadaan myös selville monien äänne­muutosten suhteellinen kronologia eli keskinäinen järjestys. Optimaalisessa tapauksessa jopa yhden ainoan sanan avulla selviää useiden äänne­muutosten keskinäinen järjestys. Esimerkiksi sana *kaksi* taipuu *kahden* : *kahta* : *kahtena*, ja sanan sijaparadigman nykyinen morfofonologinen vaihtelu selittyy neljän eri äänne­muutoksen pohjalta (Lehtinen 2007):

– *Yksinäisklusiilien astevaihtelu* – radikaalissa eli sanavartaloa koskevassa astevaihtelussa *t* säilyy avotavun (vokaaliloppuinen tavu) edellä, ja sen heikon asteen vastineena esiintyy *d* umpitavun (konsonanttiloppuinen tavu) edellä. Radikaalin astevaihtelun syntyä voidaan siis kuvata ehdollisena äänne­muutoksena: **t* > **δ* (> kirjasuomen *d*, murteiden *δ*, *r*, *l* tai kato) umpitavun edellä.

Suffiksaalissa eli taivutuspäätteitä koskevassa astevaihtelussa tavun laatu ei ollut ratkaiseva vaan tavuasema: painollisen tavun (ensimmäinen, kolmas paitsi jos viimeinen) jälkeen oli vahva aste ja painottoman tavun (toinen, neljäs; viimeinen tavu aina painoton) jälkeen heikko aste (**maa-ta*, **kala-δa* > *maa-ta*, *kala-a*).

– **kt* > *ht* – ehdollinen äänne­muutos: *t*:tä edeltävä *k* on muuttunut *h*:ksi.

- **ti > si* – ehdollinen äänne­muutos: *t* on muuttunut *s*:ksi *i*:n edellä.
- **i > e* – ehdollinen äänne­muutos: sanan­loppuinen *i* on muuttunut *e*:ksi jäätyään sanan sisälle.

Ennen näitä muutoksia sana taipui puhtaan agglutinatiivisesti, ilman morfofonologista eli sanavartalossa ilmenevää vaihtelua, niin että eri sijapäätteet lisättiin aina samanlaiseen sanavartaloon: **kakti : *kakti-n : *kakti-ta : *kakti-na*. Tätä sisäisen rekonstruktion keinoin saavutettua kielentasa­oa voidaan irrallisesti nimittää vaikkapa *esikantasuomeksi*; nimitys viittaa kielentasa­oon, jossa yksikään kantasuomalainen äänne­muutos ei vielä ollut tapahtunut, ottamatta lainkaan kantaa siihen, polveutuuko samasta kielentasa­osta muitakin kielihaaroja kuin itämerensuomi. Vertailevan menetelmän avulla (ks. luku 3.1) saadaan sitten selville, että äännetason perusteella samasta kielentasa­osta voidaan johtaa ainakin saame ja perustellusti myös mordva. Vertailevan menetelmän yhteydessä tästä samasta kielentasa­osta voidaan käyttää nimitystä *suomalais-saamelais-mordvalainen välikantakieli*. Metodisen läpinäkyvyyden kannalta olisi tärkeää käyttää eri nimityksiä eri menetelmällä rekonstruoidusta kielentasa­osta (sisäinen vs. ulkoinen eli vertaileva rekonstruktio).

Sanan *kaksi* taivutusmuodoista nähdään, että muutos **ti > si* on tapahtunut ennen muutosta **kt > ht* (muutosketju on **kakti > *kaksi*), koska jos järjestys olisi päinvastainen, muutosketju olisi **kakti > **kahti > **kaksi*. Muutos **ti > si* on kuitenkin tapahtunut myöhemmin kuin sanan­loppuinen muutos **i > e* (muutosketju on **kakti-n > *kakte-n > *kahte-n > kahde-n*), koska jos järjestys olisi päinvastainen, muutosketju olisi **kakti-n > **kaksi-n > **kakse-n*. Edelleen voidaan päätellä, että astevaihtelun synty on myöhäisempi kuin muutos **kt > ht* (muutosketju on **kakti-n > *kakte-n > *kahte-n > kahde-n*), koska jos järjestys olisi päinvastainen, edellytykset *ht*:ksi muuttumiselta olisivat kadonneet ja muutosketju olisi **kakti-n > *kakte-n > **kakde-n > **kakde-n*.

Näin sisäisen rekonstruktion avulla saadaan äänne­muutosten keskinäiseksi järjestykseksi seuraava:

1. **i > e*
2. **ti > (*ci >) > si*
3. **kt > ht*
4. yksinäisklusiilien astevaihtelun synty

Tässä vielä *kaksi*-san­an taivutusparadigman kehitys vaihe vaiheelta:

0. **kakti : *kakti-n : *kakti-ta : *kakti-na* (esikantasuomi)

1. **kakti : *kakte-n : *kakte-ta : *kakte-na*
2. **kakci : *kakte-n : *kakte-ta : *kakte-na*
3. **kakci : *kahte-n : *kahte-ta : *kahte-na*
4. **kaksi : *kahδe-n : *kahte-δa : *kahte-na*
5. *kaksi : kahde-n : kaht-a : kahte-na* (suomi)

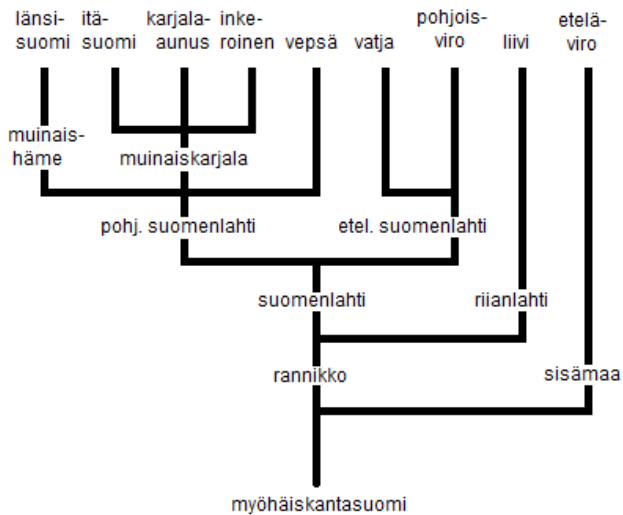
Partitiivimuoto on epäsäännöllinen: odotuksenmukainen olisi ***kahte-a* (vrt. *laksi : lahde-n : lahte-a : lahte-na*). Muutoksen **ti > *ci > *si* välivaihe **ci* on selvyuden vuoksi merkitty oikeisiin kielentasoihin, vaikka suomen kielen sisäisen rekonstruktion avulla tätä välivaihetta ei tavoiteta – se tavoitetaan vasta vertailtaessa suomea muihin itämerensuomalaisiin kieliin.

3. Horisontaalinen vertailu

Toisin kuin edellä havainnollistettu vertikaalinen vertailu, edustaa horisontaalinen vertailu varsinaista vertailevaa menetelmää, jossa eri kieliä vertaillaan toisiinsa. Tarkastelen aluksi luvussa 3.1 sukukielten välistä vertailua (vertaileva rekonstruktio), jota voidaan kuvata V-tyyppiseksi: nykykielet ovat sakaroiden kärjissä ja niiden pohjalta rekonstruoitava kantakieli V:n kannassa. Luvussa 3.2 käsittelem lainasanoihin perustuvaa vertailua.

3.1. Sukukielten välinen vertailu

Historiallisessa kielitieteessä sukukieliä vertailemalla pyritään rekonstruoimaan niiden yhteinen kantakieli. Vertailupisteitä syntyy teoriassa niin monta kuin on haarautumia kielikunnan sukupuussa. Viime vuosikymmeninä on perusteltu useita uusia vertailupisteitä eli välikantakieliä. Suomen kielen kaikkien länsimurteiden taustalle voidaan rekonstruoida kieli nimeltä muinaishäme, ja suomen kielen itämurteiden, karjalan ja inkeröisen taustalle voidaan rekonstruoida muinaiskarjala. Se piste, josta muinaishäme ja muinaiskarjala voidaan johtaa, voidaan nimetä myöhäiskantasuomen Suomenlahden pohjoismurteeksi. Tämä kielimuoto palautuu yhdessä mm. viron (pohjoisviron) kanssa myöhäiskantasuomen Suomenlahden murteeseen, joka yhdessä liivin ja eteläviron kanssa palautuu myöhäiskantasuomeen. Seuraava itämerensuomalaisen kielihaaran sukupuu perustuu Petri Kallion tutkimuksiin (Kallio 2007; tulossa):



Myöhäiskantasuomi on nykyarvioiden mukaan vajaan kahden vuosituhannen ikäinen, mutta haarakohdat eivät ole mittakaavassa: myöhäiskantasuomesta muinaiskarjalan tasoon kului noin vuosituhat, ja siitä nykyaikaan samoin noin vuosituhat. Kun mennään vielä kauemmas taaksepäin kohti koko uralilaisen kielikunnan sukupuun tyveä, voidaan kantasuomi, kantsaame ja kantamordva voidaan palauttaa länsiuralilaiseen murteeseen (kantaauralin länsimurre), joka puolestaan yhdessä keski- ja itäuralilaisen murteen kanssa voidaan palauttaa kantaauraliin (J. Häkkinen 2007). Jäljempänä tässä luvussa esitellään erilaisia uralilaisen kielikunnan sukupuita; niissä aikaulottuvuus kantaauralista myöhäiskantasuomeen on uusien käsitysten mukaan myös noin kaksi vuosituhatta. Kaikki edellä mainitut sukupuun haarakohdat eli kantakielitasot voidaan soveltuvilta osin rekonstruoida, eli niitä voidaan käyttää vertailussa apuna. Kontaktikielten vertailuun yhdistäen saadaan näin moniportaisen asteikon avulla muodostettua erittäin tarkka käsitys kielten kehityksestä.

Vertaileva rekonstruktio on vuorovaikutuksessa sekä eteenpäin että taaksepäin ajassa: esimerkiksi myöhäiskantasuomen rekonstruktio elää sen mukaan, mitä uutta selviää sen esimuodoista (edeltävistä kantakielistä) ja jälkeläismuodoista (tytärkielistä). Eräs havainnollinen esimerkki viime vuosilta on samojedikielistä. Havaittiin, että pohjoisessa Keski-Siperiassa Taimyrin niemimaalla puhuttavassa nganasanissa on säilynyt sellaisia vokaalieroja kantaauralista, joita ei ole säilynyt muissa samojedikielissä (Helimski 2005). Koska kyseiset vokaalit edustuvat toisistaan poikkeavasti sekä nganasanissa että kantaauralissa, ne on pakko rekonstruoida erillisinä myös väliin jäävään kielentasaan eli kantasamojediin, vaikka pelkästään samojedikieliä toisiinsa vertaamalla nganasanin edustus jäisi vähemmistöön eikä erillisiä vokaaleja siksi osattaisi rekonstruoida kantasamojediin.

Lähisukukielten vertailu

Edellä kielensisäistä vertailua käsittelevässä luvussa 2.2 oli esimerkki siitä, kuinka sisäisen rekonstruktion keinoin voidaan rekonstruoida esikantasuomalainen kielentaso ja kehitysvaiheet siitä nykysuomeen. Vertailevassa menetelmässä verrataan suomea yhtäältä itämerensuomalaisiin lähisukukieliin (karjala, inkeroinen, lyydi, vepsä, vatja, viro, liivi) ja toisaalta etäsukukieliin eli muihin uralilaisiin kielihäaroihin (saame, mordva, mari, permi, unkari, mansi, hanti ja samojedi). Luonnollisesti myös muiden kielihäarojen kohdalla pitää tehdä vastaavanlainen lähisukuisten kielten välinen vertailu: kantasamojedi rekonstruoidaan vertailemalla nenetsiä, enetsiä, nganasania, selkuppia, kamassia ja matoria (eräillä näistä on tai oli suurestikin toisistaan eroavia murteita tai jopa kieliä). Itämerensuomalaisten kielten perusteella rekonstruoitua kantasuomea verrataan sitten kantasamojediin ja muiden kielihäarojen kantakieliin, jotta tavoitettaisiin koko kielikunnan taustalla oleva kantakieli, kantaurali.

Suomen vertaaminen itämerensuomalaisiin kieliin vahvistaa sen mitä sisäisen rekonstruktion avulla saatiin selville. Esimerkiksi viro on osallistunut aivan samoihin äänteenmuutoksiin kuin suomi:

Suomi: *kaksi* : *kahde-n* : *kaht-a* : *kahte-na*

Viro: *kaks* : *kahe* : *kaht* : *kahe-na*

Myöhäiskantasuomi: **kaksi* : **kahde-n* : **kahte-δa* : **kahte-na*

Erot kielten välillä johtuvat siitä, että virossa *t:n* heikossa asteessa on katoedustus myös sanan vartalossa, virossa on tapahtunut sananloppuisten äänteiden kulumista ja lisäksi essiivi on alettu muodostaa genetiivivartalon pohjalta, toisin kuin suomessa. Viime mainittua piirrettä lukuun ottamatta viro voitaisiin esimerkkisanan paradigman osalta jopa johtaa suomesta, eli konservatiivisempi suomi voisi periaatteessa olla innovatiivisemmän viron esimuoto, kantakieli. Huomioitaessa kielen muut tasot tämä oletus on kuitenkin hylättävä: viroa kokonaisuutena ei voida johtaa suomesta, vaan ne molemmat voidaan johtaa vain yhteisestä kantakielestä.

Kun vertailuun otetaan kirjakielten lisäksi myös murteet, käy ilmi että eteläviro eroaa kaikista muista itämerensuomalaisista kielimuodoista: se näyttää alkaneen kehittyä omaan suuntaansa jo hyvin varhain eikä ole osallistunut kaikkiin niihin muutoksiin, jotka yhdistävät suomea ja viron kirjakieltä (joka perustuu viron kielen pohjoismurteisiin).

0. **kakti* : **kakti-n* : **kakti-ta* : **kakti-na* (esikantasuomi)
1. **kakti* : **kakte-n* : **kakte-ta* : **kakte-na*
2. **kakci* : **kakte-n* : **kakte-ta* : **kakte-na*
3. **kakci* : **katte-n* : **katte-ta* : **katte-na*
4. **katsi* : **kate-n* : **katte-δa* : **katte-na*
5. **kats̥* : **kat̥e-n* : **kat̥e-δa* : **kat̥e-na*
6. *kats̥* : *kat̥ō* : *katt̥ō* (eteläviro) (essiivi ei produktiivinen)

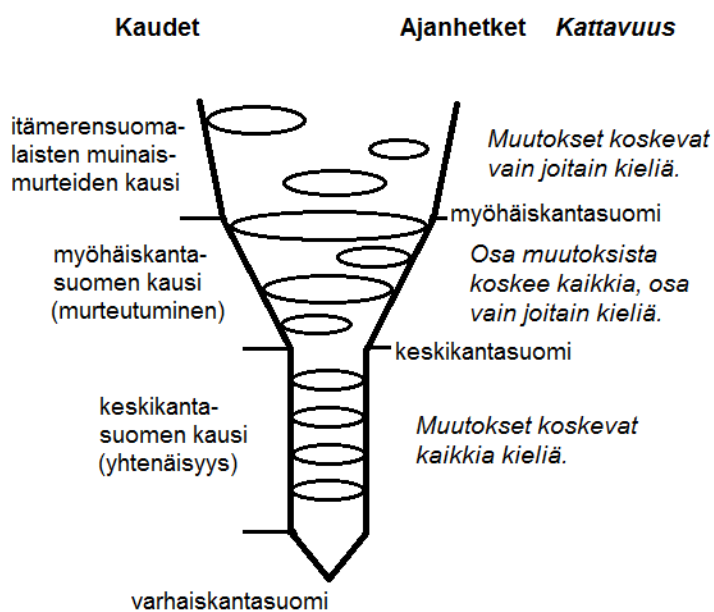
Suomen kehitys esitettiin sisäisen rekonstruktion kohdalla (luku 2.2), ja siihen vertaamalla nähdään erot eteläviron ja suomen välillä. Vaiheeseen 2 saakka eteläviro seuraa yleisitämerensuomalaista kehitystä, mutta sitten alkaa eriytyminen:

- Vaiheessa 3 etelävirossa tapahtuu muutos **kt* > *tt* ja muualla itämerensuomessa muutos **kt* > *ht*.
- Vaiheessa 4 astevaihtelu (joka foneettisena ilmiönä on alkanut jo aikaisemmin) foneemistuu³ etelävirossa (tässä avotavun edellä olevaa geminaattaa *tt* vastaa umpitavun edellä yksinäis-*t*) kuten jossain vaiheessa kaikkialla muuallakin itämerensuomessa paitsi ääri-laidoilla liivissä ja vepsässä, ja lisäksi etelävirossa tapahtuu muutos **kc* [*kts*] > *ts* ja muualla itämerensuomessa muutos **kc* > *ks*.
- Vaiheessa 5 etelävirossa loppu-*i* katoaa liudentaen edeltävän konsonantin, ja toisen tavun *e* velaaristuu⁴ *õ*:ksi [*e*] ensimmäisen tavun takavokaalin jäljessä (progressiivinen palataaliharmonia eli vokaalisointu).
- Vaiheessa 6 etelävirossa sanamuodot kuluvat eli sananloppuisia äänneitä katoaa.

On huomattava, että vaikka eteläviro on jo varhain alkanut kehittyä omaan suuntaansa, on se silti osallistunut myös yhteisvirolaisiin muutoksiin, kuten **δ*:n ja sananloppuisen **n*:n katoaminen. Erilaiset katoamiset tietysti voisivat olla tapahtuneet myös itsenäisesti eri kielissä, mutta eteläviro on osallistunut myös myöhäiskantasuomalaiseen muutokseen **š* > *h*, joka on tapahtunut hyvin myöhään (Kallio 2007). Nykytasolla siis puhutaan yhdestä viron kielestä, vaikka kielihistoriallisesti syvin itämerensuomen sisäinen kuilu eli vanhin murreraja sijaitsee pohjoisviron ja eteläviron välillä: eteläviron esimuoto (kantasuomen sisämaamurre) on eronnut murteutunut erilleen kantasuomesta jo siinä vaiheessa kun pohjoisviron, suomen ja kaikkien muidenkin itämerensuomalaisten kielten esimuodot olivat vielä yhtä ja samaa kieltä, myöhäiskantasuomen rannikkomurretta.

Tämä tilanne aiheuttaa määritelmäongelmia. Määritelmällisestihän myöhäiskantasuomi on itämerensuomalaisten kielten viimeinen yhteinen kantakieli. Kielimuodot ovat kuitenkin eriytyneet

eri aikoina, ja eteläviro näyttää alkaneen eriytyä jo hyvin varhain, ikään kuin jo keskikantasuomen kaudella. Kuitenkin eteläviro on ollut pitkään murteenomaisessa suhteessa muuhun kantasuomeen, joten se voidaan laskea saman kielen piiriin kuuluvaksi. Voidaan siis puhua myöhäiskantasuomen murteista, kunhan hahmotetaan että myöhäiskantasuomen määritelmä on tällöin muuttunut: se ei enää ole pelkästään se *hetki* aikajanalla johon kaikki itämerensuomalaiset kielet palautuvat, vaan siitä tulee se *kausi*, jonka aikana yhtenäinen kantasuomi alkoi murteutua eli jolloin tapahtui sekä sellaisia muutoksia, jotka eivät enää koskettaneet koko kieliä, että sellaisia muutoksia, jotka koskettivat koko kieliä. Myöhäiskantasuomen kausi alkaa siis ensimmäisestä muutoksesta joka ei enää kattanut koko itämerensuomalaista kieliä, ja se päättyy viimeiseen muutokseen joka vielä kattoi koko itämerensuomalaisen kieliä. Myöhäiskantasuomen kautta edeltää keskikantasuomen kausi, mutta myös keskikantasuomi-käsitteellä voidaan viitata kauden sijasta pisteeseen aikajanalla: joko siihen pisteeseen, johon tullessa tiettyjen konsonantifoneemien eliminoituminen oli tapahtunut (Kallio 2007), tai siihen pisteeseen, joka välittömästi edeltää ensimmäistä eteläviron muusta kantasuomesta erottavaa muutosta. Käsitteiden monitulkintaisuuden vuoksi on tärkeää määritellä se, käytetäänkö nimityksiä ajanhetkistä vai kausista.



Perinteisesti itämerensuomen sisäisessä jaotellussa on keskitytty vain kieliin ja ohitettu murteet, mutta taksonomisesti merkittävien murteiden mukaanotto tuo kuvaan olennaista lisätietoa. Eteläviron varhaisen eriytymisen lisäksi murteiden avulla tarkentuu myös muinaiskarjalan jakautuminen (Salminen 1998).

Etäsukukielten vertailu

Sisäisen rekonstruktion avulla tavoitettu esikantasuomalainen kielentaso oli puhtaasti agglutinatiivinen, eli kieliopilliset tunnuksot liitettiin vaihtelemattomaan vartaloon: **kakti* : **kakti-n* : **kakti-ta* : **kakti-na*. Suunnilleen tästä samasta kielentasosta voidaan johtaa myös saamen ja mordvan edustukset, tosin *kaksi*-sanon toisen tavun vokaalissa on eroja. Koska kyse on nyt vertailevasta rekonstruktiosta eikä sisäisestä rekonstruktiosta, käytän esikantasaame-nimityksen sijasta nimitystä varhaiskantasaame, vaikka konkreettinen kielentaso onkin sama.

0. Varhaiskantasaame: **kakta* : **kakta-n*

1. Myöhäiskantasaame: **kuoktē* : **kuoγtē-n*

2. Pohjoissaame: *guokti* : *guovti* 'kaksi : kahden'

0. Varhaiskantamordva: **kakta* : **kakta-n* ~ **kakti* : **kakti-n*

1. Myöhäiskantamordva: **kaytâ* : **kaytâ-ń*

2. Ersämordva: *kavto* : *kavto-ń* 'kaksi : kahden'

Saame viittaa yksiselitteisesti toisen tavun **a*:han, kun taas mordvan toisen tavun vokaali voi olla peräisin kummasta tahansa, joko **a*:sta tai **i*:stä: ersämordvan *o* esiintyy liudentumattoman konsonantin jäljessä, *e* liudentuneen (Bartens 1999). Saamelle ja itämerensuomelle yhteinen piirre on astevaihtelu, mutta sitä ei voi johtaa yhteisestä kantakielestä: saamessa astevaihtelu heikentää konsonanttiyhtymän ensimmäistä konsonanttia (**kt* > **γt*), itämerensuomessa puolestaan konsonanttiyhtymän jälkimmäistä konsonanttia (**ht* > **hδ*) (Korhonen 1981). Mordvassa ei ole lainkaan astevaihtelua, vaan **k* on heikentynyt tämän konsonanttiyhtymän osana (samoin kuin vokaalien välissä) kaikissa asemissa.

Kaikissa kielihaaroissa genetiivin päätte **-n* on säilynyt: eteläsaamessa se on säilynyt, vaikka pohjoissaamesta se on kadonnut, ja suomessa se on säilynyt vaikka virossa se on kadonnut (*kaks* : *kahe-Ø*). Astevaihtelu on mahdollistanut sijapäätteen katoamisen pohjoissaamessa ja virossa, koska genetiivi eroaa nominatiivista jo sanavartalon osalta. Näissä kielissä taivutus on siis genetiivin osalta siirtynyt agglutinoivasta tyyppistä fusionaaliseen (flekteeeraavaan) tyyppiin.⁵

Ottamalla vertailuun vielä kaukaisempia etäsukukieliä voidaan havaita, että itämerensuomen *kaksi*-sanon ensimmäisen tavun *a* on myöhäsyntyinen: itäisemmät uralilaiset kielihaarat viittaavat siihen,

että alkuperäinen asu olisi ollut **kākta* tai **keкта* (toisessa tavussa esiintyi vain kaksi vokaalia, väljä **a* ja suppea **i* tai **ə*, vaikka perinteisesti rekonstruktioissa onkin suomen vokaalisointua seuraten käytetty **ä*:tä myös jälkitavuuissa; Korhonen 1988), joten läntisten kielihaarojen edustus voidaan selittää sanan vokaaliston epäsäännöllisellä velaaristumisella eli takaistumisella (> **kakta*). Vain itämerensuomessa täytyy olettaa myös vartalovokaalin vaihtuneen (**kakta* > **kakti*), mikä voidaan selittää esimerkiksi lukusanan *yksi* vaikutuksella: viereisillä lukusanoilla on katsottu olevan taipumusta alkaa muistuttaa toisiaan, koska ne luetellaan peräkkäin (Kulonen 2002).

Toisen tavun vokaalin erilaisuus itämerensuomessa ja saamessa osoittaa, että sisäisen rekonstruktion avulla tavoitettu esikantasuomi ei ole täysin sama kielentaso kuin vertailevan rekonstruktion avulla tavoitettu länsiurali (suomalais-saamelais-mordvalainen välikantakieli). Tässä tapauksessa sisäisen rekonstruktion keinoin ei saavuteta aivan yhtä vanhaa kielentasoa, vaan sillä tavoitetaan ajanhetki sen jälkeen kun itämerensuomalainen kielihaara oli eriytynyt länsiuralilaisesta kantakielestään ja sana **kakta* oli jo epäsäännöllisesti muuttunut asuun **kakti*.

Välikantakielet

Etäsukukieliä vertaamalla saatetaan päätyä huomaamaan, että jotkin kielihaarat ovat käyneet läpi yhteisiä uudennoksia. Tällöin rekonstruoidaan välikantakieli perimmäisen kantakielen (kantaurali) ja kiistatonta kielihaaraa edustavan kantakielen (kantasuomi) väliin. Noin sadan vuoden ajan oletettiin olleen olemassa (itämeren)suomalais-saamelainen kantakieli, mutta 1980-luvulta lähtien siihen kuten muihinkin välikantakielisiin on kohdistettu kritiikkiä (K. Häkkinen 1984; T. Itkonen 1997; Salminen 2002).

On totta, että uudennosten määrässä välikantakielet eivät vastaa kiistattomien kielihaarojen kantakieliä: esimerkiksi kantsaamelaisia tai kantasuomalaisia (vain itämerensuomalaisille kielille yhteisiä) muutoksia on paljon enemmän kuin sellaisia muutoksia, jotka koskevat vain kahta tai kolmea kielihaaraa (ks. luku 1.3). Jos tällaisia muutoksia kuitenkin on kielen joka tasolla, voidaan välikantakielen rekonstruoimista pitää perusteltuna. Tietysti muutosten lukumäärää voidaan pitää vihjeenä muutoskauden pituudesta: lyhyen muutoskauden erottamaa kielimuotoa voidaan kutsua vaikkapa murteeksi ja pitkän muutoskauden erottamaa kielimuotoa kieleksi. Vaatimalla kaikilta välikantakieliltä yhtä suurta muutosten määrää kuin kiistattomien kielihaarojen kantakieliltä kadotetaan olennaista tietoa kielikunnan rakenteesta koska erottelukyky silloin heikkenee.

Näyttää siltä, että mordva voidaan johtaa aivan samasta välikantakielestä kuin itämerensuomi ja saame. Äännetason perusteella voidaan sanoa, että vain näitä kolmea kielihaaraa yhdistää esimerkiksi se muutos, että kantaauralin takainen **ɛ* (aiemmin rekonstruoitu **j*:ksi) lankesi yhteen **a*:n kanssa. Esimerkiksi suomen sanat *ala* ja *kala* sekä mordvan vastineet *alo* ja *kal* sisältävät saman ensitavun vokaalin, kun taas vaikkapa marissa (*iil*, *kol*) ja hantissa (*il*, *kuul*) sanoissa on erilaiset ensitavun vokaalit. Ensimmäinen sana palautetaan kantaauralin asuun **ɛ*la (**j*la) ja jälkimmäinen asuun **k*ala (Sammallahti 1988). Toinen vain läntisiä kielihaaroja yhdistävä äänteellinen uudennos on sanansisäisen liudentuneen dentaalispirantin **ʃ*:n lankeaminen yhteen **ʃ*:n kanssa, ja kolmas on sananalkuisen **w*:n katoaminen lyhyen pyöreän vokaalin edeltä (koskee ainakin jossain määrin myös maria). Voidaankin sanoa, että itämerensuomi, saame ja mordva voidaan johtaa yhteisestä länsiuralilaisesta äännejärjestelmästä, eikä mitään muita uralilaisia kielihaaroja voida johtaa tästä samasta äännejärjestelmästä (J. Häkkinen 2007).

Muoto-opin tasolla voidaan niin ikään rekonstruoida yhteisiä länsiuralilaisia (itämerensuomalais-saamelais-mordvalaisia) uudennoksia. Tällainen on esimerkiksi *s*-paikallissijasjarja: suomi *kodissa* : *kodista* : *koti(h)in* ~ eteläsaame (’tyttö’) *niejtesne* : *niejteste* : *niejtese* ~ ersämordva (’maa’) *modaso* : *modasto* : *modas*. Myös suomen illatiivin *h* palautuu aikaisempaan **s*:ään. Lisäksi yhteistä kaikille näille kielille on vanhan erosijan (suomen kieliopissa partitiivi) kehittyminen objektin merkitsijäksi (vrt. *poistua jonkun luo-ta* vs. *kaivella maa-ta*). Translatiivin päätte **-ksi* rajoittuu näihin kielihaaroihin, ellei sitten kantasamojedin tunnus **-tʃ* ole sen vastine; äänteellisesti rinnastus on moitteeton (U **-ksi* > samoj. **-tʃ*), mutta toisaalta samojedin **t* voi olla peräisin niin monesta muustakin äänneestä ja äänneyhtymästä, että tunnus voi olla sattumaltakin sopivannäköinen. Tämä aines ei myöskään esiinny samojedissa sijapäätteenä vaan esimerkiksi tundranenetsissä nominien predestinatiivisen taivutuksen (’jotakuta varten’) tunnuksena johon voi liittyä sijapäätte (Salminen 1997); rakennetta on hankala suomentaa lyhyesti.

xər^o-də-da

veitsi-varten-px.3sg

’veitsi, joka on häntä varten’

xər^o-də-m-da

veitsi-varten-acc.sg-px.3sg

’veistä, joka on häntä varten’

Sanaston tasolla saame ja itämerensuomi muodostavat selvemmin yksikön, johon mordva ei kuulu: ne jakavat paljon yhteisiä sanoja, jotka muista kielihaaroista puuttuvat. Tämä tilanne on oikeastaan odotuksenmukainenkin, koska saamea ja itämerensuomea on puhuttu ”aina” lähekkäin. Näiden kieliharojen esimuodot ovat edenneet yhdessä kohti länttä, kun taas mordva on jäänyt kauas taakse, Okan itäpuolelle lähelle Volgan suurta mutkaa. On kuitenkin kaksi syytä, miksei yhteiselle sanastolle voida antaa paljonkaan painoa pohdittaessa suomalais-saamelaisen välikantakielen olemassaolaisuutta.

Ensiksi on huomioitava äännetaso: sanojen suomalais-saamelainen rekonstruktio- taso on useimmiten aivan sama kuin länsiuralilaisen murteen (suomalais-saamelais-mordvalainen välikantakieli) rekonstruktio- taso. Onkin mahdollista, että osa tällaisista sanoista on sellaisia, joilta on vain kadonnut vastine mordvasta, koska on olemassa myös sellaisia sanoja, joiden levikki kattaa vain itämerensuomen ja mordvan tai vain saamen ja mordvan. Toiseksi lainasanojen avulla on saatu selville, että äännetasolla itämerensuomi ja saame alkoivat merkittävästi eriytyä vasta luoteisgermaanin vaiheen jälkeen eli vähän ennen ajanlaskun alkua. Silti kielihaarat ovat olleet alueellisesti eli kieliyhteisöinä erillään jo paljon aikaisemmin, minkä todistaa seuraavan paleogermanisen lainasanan erilainen johdos itämerensuomessa ja saamessa (Kylstra 1991, 1996):

suomi *kavio* < myöhäiskantasuomi **kapja* < varhaiskantasuomi **kapa-ja*

~ eteläsaame *gvehpere* < myöhäiskantasaame **kuopērē* < varhaiskantasaame **kapa-ra*

← paleogermani⁶ **kāpa-s* ’kavio, sorkka’ > kantagermaani **χōfa-z* > englanti *hoof*

(Saamen sana on myöhemmin lainattu suomeen asussa *kopara* ’poron sorkka’.)

Esikantasuomea (varhaiskantasuomea) ja esikantasaamea (varhaiskantasaamea) on näin ollen puhuttu naapureina useita vuosisatoja, parhaassa tapauksessa jopa vuosituhat, ennen kuin erottavia äänne- muutoksia oli tapahtunut niin paljon että keskinäiset lainasanat pystytään useimmiten erottamaan. Vielä nykyäänkään niitä ei pystytä kaikissa tapauksissa erottamaan: suomen *kieli* ja pohjoissaamen *giella* ovat yleisen käsityksen mukaan kanta-uralista perittyjä sanoja, mutta ne voisivat äänne- asunsa perusteella olla yhtä hyvin vasta sata vuotta sitten kielestä toiseen lainattuja, johtuen lainasanojen sopeuttamisesta oman kielen äänne- järjestelmään (*k* ↔ *g*) ja jälkitavun vokaalin kohdalla yleisestä etymologisesta nativisaatiosta (*i* ↔ *a*) (Aikio 2007).⁷ Vaikka äänne- erojen puuttuessa näyttääkin siltä, että tällaiset sanat palautuisivat yhteiseen suomalais-saamelaiseen välikantakieleen, ne voivat yhtä hyvin olla mainitun pitkän naapuruuden aikana suuntaan tai toiseen omaksuttuja lainasanoja, joita vain ei ole tunnistettu lainasanoiksi koska kielet eivät vielä olleet

eronneet toisistaan riittävästi. Sanaston tasolla vaikuttaa erilaisia vääristäviä prosesseja, jotka tekevät siitä epäluotettavan tason kielihaarojen välisiä suhteita arvioitaessa: näkymätön konvergenssi saa esimerkiksi itämerensuomen ja saamen näyttämään yhteisen sanaston perusteella läheisemmiltä kuin ne ovatkaan, ja näennäinen divergenssi saa samojedinin ja unkarin näyttämään sanaston perusteella varhimminkin eriytyneiltä uralilaisilta kieliltä (J. Häkkinen, tulossa).

Yksin sanaston perusteella ei siis voi luotettavasti päätellä mitään suomalais-saamelaisen välikantakielen todellisuudesta, kun taas äänne- ja muototason perusteella mordva voidaan johtaa aivan samasta kantakielestä. Tällä perusteella ei ole tarvetta rekonstruoida erillistä suomalais-saamelaista välikantakieltä, vaan itämerensuomi, saame ja mordva voidaan johtaa länsiuralilaisesta murteesta. Lisäksi on huomioitava diakroninen kato, josta jo alunpana on ollut puhetta. Kun länsiurali näyttää jakautuneen kolmiksi, voi länsiuralissa olleella piirteellä olla teoriassa kahdeksanlainen levikki: (1) piirre esiintyy itämerensuomessa, saamessa ja mordvassa; (2) piirre esiintyy itämerensuomessa ja saamessa; (3) piirre esiintyy itämerensuomessa ja mordvassa; (4) piirre esiintyy saamessa ja mordvassa; (5) piirre esiintyy itämerensuomessa; (6) piirre esiintyy saamessa; (7) piirre esiintyy mordvassa; (8) piirre on kadonnut kaikista kieliharoista.

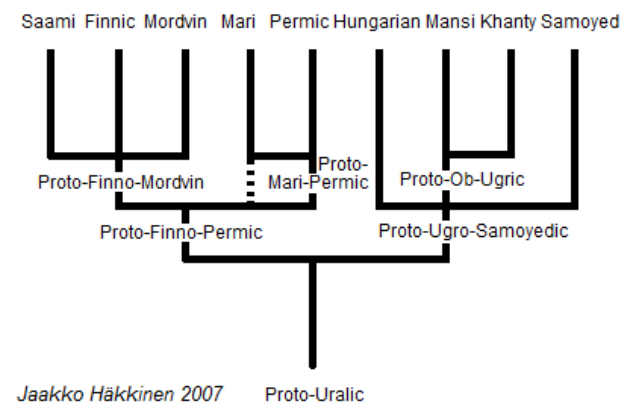
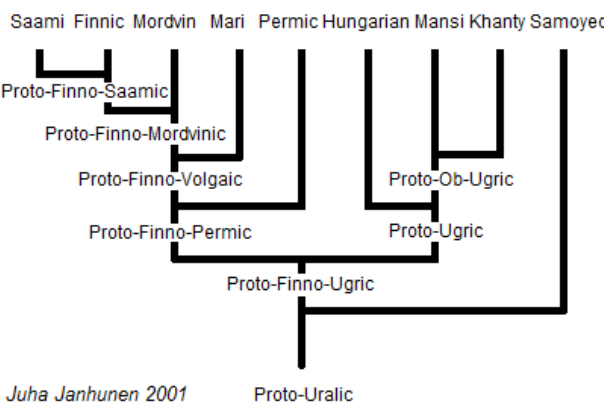
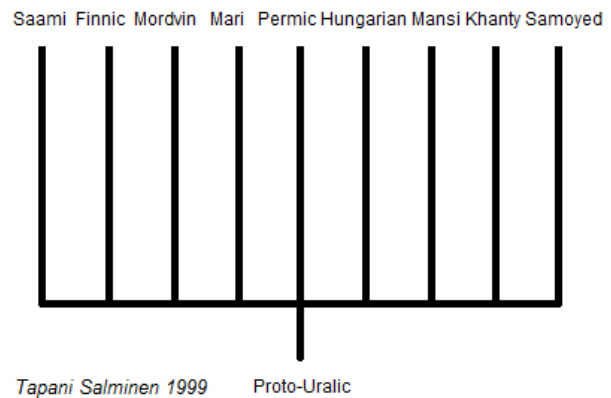
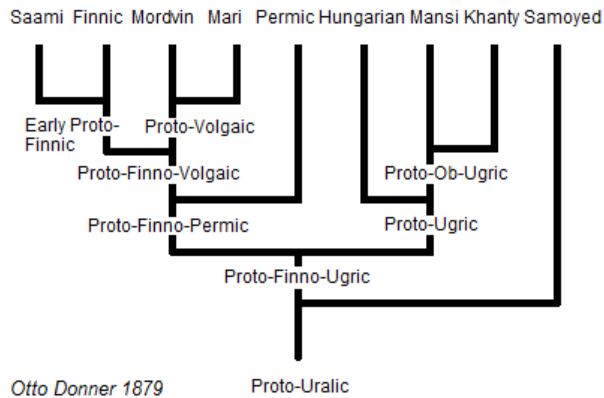
Vain yksi piirre kahdeksasta löytyy kaikista länsiuralin tytärielistä (olettaen että kaikkien vaihtoehtojen todennäköisyys on yhtä suuri), kun taas kolmella piirteellä kahdeksasta on kahteen kieliharaan rajoittuva levikki. Loput neljä piirrettä kahdeksasta jäävät ikuisesti vertailevan menetelmän ulottumattomiin, eikä niitä osata palauttaa länsiuralilaiseen murteeseen ellei jokin toinen menetelmä anna siihen mahdollisuutta. Näin ollen tilanne, jossa piirteitä löytyy eri yhdistelmillä aina kahdesta kieliharasta kerrallaan, sopii oikein hyvin kuvaan länsiuralilaisesta murteesta. Tällaisia piirteitä ei siis ole välttämätöntä olettaa jo eriytyneiden kieliharojen väliseksi kontaktivaikutukseksi vaan ne voidaan edellä esitetyn laskennallisen periaatteen perusteella palauttaa yhteiseen kantakieleen eli länsiuraliin. Jos löytyy riittävästi sellaisia piirteitä, jotka löytyvät kaikista kolmesta kieliharasta, on taloudellistakin olettaa myös kahteen kieliharaan rajoittuvat piirteet kantakielestä periytyneiksi. Tietysti silloin jos kahden kieliharaan (esimerkiksi saamen ja itämerensuomen) jakamia piirteitä on selvästi enemmän kuin on kolmen kieliharaan jakamia piirteitä, on syytäkin epäillä kontaktivaikutusta, koska kantakielestä kummankinlaisten piirteiden pitäisi periytyä suunnilleen samassa suhteessa.

Sukupuu ja alkukoti

Sukukielten vertailun avulla saatuja tuloksia kielikunnan rakenteesta havainnollistetaan sukupuun avulla. Tärkeää on kuitenkin ymmärtää mitä sukupuu kuvaa ja mitä se ei kuvaa: sukupuu kuvaa pelkästään kantakielen divergenssiä eli hajoamista tytärkieliksi. Sukupuu ei sen sijaan ota kantaa konvergenssiin eli jo eriytyneiden kielihaarojen välisiin kontakteihin. Sukupuu ei ole ristiriidassa kontaktiteoreettisen aaltomallin kanssa – se vain esittää eri asioita. Sukupuussa näkyvä divergenssi voi hyvinkin johtua siitä, että yksi eriytyvistä kielihaaroista on päätenyt kontaktiin eri kielikuntaan kuuluvan kielen kanssa (E. Itkonen 1998). Kantakielen hajoamisen ja kielikunnan taksonomisen rakenteen havainnollistajana sukupuu on edelleen lyömätön.

Vaikka binaarista eli aina kahtia haarovaa sukupuuta, jollainen perinteinen, lähes vuosisadan vallinnut uralilaisen kielikunnan sukupuu suuressa määrin oli, onkin kritisoitu, niin lähtökohtaisesti binaarisuus on kuitenkin looginen välttämättömyys: kun uusi kieli syntyy erottavien piirteiden kautta, se irtautuu vain yhdestä toisesta kielestä. Jos erottelukyky olisi riittävän tarkka, kaikki murteutumiset voitaisiin erottaa ajallisesti toisistaan, ja lopputulokseksi saataisiin puhtaan binaarinen sukupuu. Uralistiikassa kritiikki onkin kohdistunut siihen, ettei riittävää tarkkuutta ja kiistattomuutta kielten eroamishetkistä ole saatu, mutta silti on suosittu binaarisia sukupuuta, usein melko mielivaltaisesti eri vaihtoehdoista valiten. Kriitikot ovatkin suosittelleet heikosti perustelluista välikantakielistä luopumista (K. Häkkinen 1984; Salminen 1999; 2002). Voidaan kuitenkin perustella myös, että olennaista informaatiota jää puuttumaan, mikäli sukupuuhun hyväksytään vain sellaisia yksikköjä, joiden tueksi voidaan esittää kymmenittäin uudennoksia. Sellaiset välikantakielet, joiden tueksi on vain kourallinen uudennoksia (lähinnä äänne­muutoksia; yhteisiä sanoja löytyy toki enemmän), voivat olla aivan yhtä todellisia – ehkä nämä vaiheet vain kestivät lyhyemmän aikaa kuin vaikkapa arviolta tuhatvuotinen kantasuomen kausi (varhaiskantasuomesta myöhäiskantasuomeen), jolloin tapahtuneet äänne­muutokset lasketaan kymmenissä.

Tuloksena olevan sukupuun malli riippuu tietysti myös siitä, mitä aineistoa hyödynnetään. Alla on esimerkinomainen valikoima erilaisista uralilaista kielikuntaa kuvaavista sukupuista. Perinteisessä Otto Donnerin sukupuussa painotus oli muoto-opin piirteissä, kun taas Juha Janhusen sukupuu perustuu lukusanojen todistukseen ja Jaakko Häkkisen sukupuu äännetason todistukseen. Tapani Salmisen sukupuu perustuu siihen, että välikantakielille ei ole hänen mielestään riittävästi perusteita millään tasolla.



Uralilainen sukupuu on helppo visualisoida kartalle, koska uralilaisten kielihaarojen varhaisimmat alueet ovat muodostaneet lännestä itään ulottuvan nauhan. Alkukodin sijoittaminen aikaan ja paikkaan liittyykin usein sukupuun käsittelyyn: missä ja milloin sukupuun tyveä eli kantauralia on puhuttu? Visuaalisista syistä kantauralia on edellä esitetyissä sukupuissa piirretty kohdakkain keskimmäisen kielihaaran eli permin kanssa, mutta tällä ei ole tarkoitus ottaa kantaa alkukodin sijaintiin; esimerkiksi Janhunen sijoitti kantauralin Siperiaan. Silti kantauralin on toistuvasti perusteltu sijoittuvan juuri Volgan suuren mutkan (Volgan ja Kaman yhtymäkohta) ja Uralvuoriston väliin eli permiläisten kielten alueelle, ja sinne todisteet näyttävät tälläkin hetkellä viittaavan (Kallio 2006; J. Häkkinen 2009).

Kantauralin ajoituskin on vaihdellut suuresti. Virossa vakiintui jo 1950-luvulla ja Suomessa 1980-luvulla käsitys, että uralilaisen kielen leviäminen täällä lännessä liittyisi tyypillisen kampakeramiikan leviämiseen (Viron osalta ks. Moora 1956; Suomen osalta ks. Gallén 1984). Tällä perusteella uralilaista kieltä ajateltiin puhutun Suomessa jo noin 6 000 vuotta sitten. Tämä ns. jatkuvuusteoria perustui lähtökohtaisesti siihen, että arkeologisesti havaittavan jatkuvuuden oletettiin todistavan myös kielellisestä jatkuvuudesta. Samaa menetelmää seuraten syntyi vielä

radikaalimpi jatkuvuusteoria, jota oli väläytelty jo ainakin 1970-luvulla mutta joka saavutti suosionsa huipun 1990-luvulla: sen mukaan uralilaista kieltä olisi puhuttu Suomessa jo jääkauden jälkeisestä alkuasuksesta lähtien. Jo pitkään oli ollut tutkijoita jotka ajattelivat, ettei arkeologinen jatkuvuus itse asiassa voi todistaa kielellisestä jatkuvuudesta, ja indoeuropeistiikassa jatkuvuusteorioita vastustettiin jo 1980-luvulla, mutta Suomessa varsinainen jatkuvuusperustelujen kritiikki nousi pintaan vasta 2000-luvulla (Mallory 1989; Aikio & Aikio 2001; J. Häkkinen 2010a). Viime aikoina onkin perusteltu kantauralin leviämisen aiempaa myöhempää ajankohtaa: uralilainen ekspansio näyttäisi alkaneen Volgan mutkan tienoilta vasta suunnilleen 4 000 vuotta sitten (Kallio 2006; Häkkinen 2009). Näin kivikautisen Suomen kampakeraamisella kulttuurilla ei enää olisikaan yhteyttä uralilaisen kielen leviämiseen, eikä itämerensuomen ja saamen eroakaan olisi tapahtunut Suomessa jo kivikaudella vaan Suomen kaakkoispuolella vasta pronssikaudella (Aikio 2001; 2006; Saarikivi 2006; yhteenveto 2000-luvun tuloksista: J. Häkkinen 2010a; 2010b).

3.2. Kontaktikielten välinen vertailu

Myös kielikontaktien tutkimus kuuluu vertailevan menetelmän piiriin, jossa tarkastellaan erityisesti lainasanoja. Tällöin hyödynnetään kummankin kielilinjan vertikaalisen ja horisontaalisen vertailun avulla saatua tietoa vertaamalla kahden eri kielihaaraan tai kielikuntaan kuuluvan kielen kehityslinjoja tarkoituksena selvittää, missä vaiheessa kummankin kielen kehitystä kontaktit ovat tapahtuneet. Tämän menetelmän tyyppiä kuvaa H, jossa pystyviivat edustavat eri kielilinjoja ja poikkiviiva niiden välistä kontaktivaikutusta. Kontaktikielten välinen vertailu on hedelmällisintä silloin, kun kummankin kielen kehityksestä on jo jonkinlainen käsitys; tällöin kontaktivaikutus voidaan sijoittaa oikeaan kohtaan kielen kehitysjuoksuissa, jolloin kielen muutosten kronologiat voidaan ankkuroida toisiinsa.

Lainasanakerrostumien kronologia

Kronologia eli ajallinen järjestys voi olla suhteellista tai absoluuttista. Suhteellinen kronologia merkitsee keskinäistä aikajärjestystä, absoluuttisessa kronologiassa taas tämä aikajärjestys sidotaan kalenterivuosiin. Kronologioita käsitellään myöhemmin tässä luvussa. Lainasanakerrostumien suhteellinen kronologia saadaan selville vertaamalla sanojen äänneitä; tämännäköisessä vertailussa sanataso ja äännetaso siis yhdistyvät. Edellytyksenä sille, että eri kerrostumia yleensä voidaan erotella, on juuri eriävä äänne-edustus. Ero voi johtua kahdesta syystä: (1) äännesubstituutiosta eli

vieraan äänten sopeuttamisesta lainattaessa oman kielen äännejärjestelmään, tai (2) ääntenmuutoksesta.

Germaanisten lainasanojen kohdalta löytyy esimerkki kummastakin tyypistä. Lainanantajakielen sananalkuinen **st*-yhtymä on vanhemmissa lainasanoissa korvattu (itämeren)suomessa **s*:llä (kantagermaani **stūra-* → suomi *suuri*), uudemmissa **t*:llä (ruotsi *stol* → suomi *tuoli*).

Ääntenmuutos taas erottaa lainasanat sen mukaan, onko muutos tapahtunut kielessä ennen vai jälkeen sanan lainautumisen. Ennen itämerensuomalaista **ti* > *si* -muutosta on lainattu esimerkiksi sana *karsia* (< **karci-δak* < **karti-* ← kantagermaani **skardja-*), joka on osallistunut tähän muutokseen aivan kuten vanhimmat perintösanatkin (kantaauralista perityt sanat). Vasta muutoksen jälkeen puolestaan on lainattu sana *hartia* (← varhaiskantaskandinaavi **hardiō*), joka ei ole kyseiseen muutokseen enää ehtinyt osallistua.

Myös lainanantajapuolen ääntenmuutokset saattavat erottaa lainasanoja. Erittäin havainnollisia ovat sellaiset lainasanat, jotka on lainattu kieleen moneen kertaan. Tällaisten sanojen avulla saadaan suhteutettua lainanantaja- ja lainansaajakielen ääntenmuutosten kronologia toisiinsa. Esimerkiksi ennen luoteisgermaanin **z* > **R* -muutosta on lainattu sana *keihäs* (< **kaihas* < **kaizas* ← kantagermaani **gaiزاز*), ja muutoksen jälkeen samasta sanasta on lainattu sana *kaira* (← varhaiskantaskandinaavi **gaiRaR*). Jo tämän yhden ainoan kaksinkertaisen lainasanan perusteella saadaan kuva eri tahoilla tapahtuneiden ääntenmuutosten keskinäisestä kronologiasta: lainautumisten välissä ovat tapahtuneet myöhäiskantasuomen muutos **z* > **h*, myöhäiskantasuomen Suomenlahden murteen sporadinen muutos **ai* > **ei* (ei tapahtunut etelävirossa eikä liivissä) ja luoteisgermaanin muutos **z* > **R* (> **r*). Suomessa ei ole myöhemmän lainautumisen jälkeen tapahtunut enää muutoksia, mutta germaanisella puolella (skandinaavissa) ovat tapahtuneet vielä muutokset **ai* > **ei* ja vokaalin sisäheitto (> myöhäiskantaskandinaavi **geirr* > islanti *geir* 'keihäs').

Parhaassa tapauksessa sama sana on lainattu kieleen eri aikoina peräti neljä kertaa, kuten erään germaanista saameen lainatun sanan kohdalla on käynyt (Aikio 2006):

1. PG **χeuja-* → EPS **kepji* > LPS **keppje* > pohjoissaame *gavja* 'pöly'
2. NwG **heuja-* → MPS **evji* > LPS **jievje* > pohjoissaame *jievja* 'valkoinen (eläimen karva)'
3. EPN **hiuja-* → LPS **ipje* > eteläsaame *ibje* 'pöly'
4. MPN ?**hyuja-* → SaaW **hujve* ~ SaaE **uvje* > pohjoissaame *uvja* 'untuva'

Tällaisten moninkertaisten lainasanojen ansiosta käsityksemme kontaktissa olevien kielten äännehistoriasta ja äänne­muutosten kronologiasta on nykyään jo hyvin tarkka. Tästä sarjasta (muun lainasana-aineiston tarjotessa varmistuksen) voidaan päätellä seuraavat seikat:

- Kantagermaanin (PG) *χ on korvattu varhaiskantasaamen (EPS) *k:lla ja *uj [*wj] *pj:llä.
- Kantagermaanista luoteisgermaaniin (NwG) tultaessa *χ kehittyi vähemmän hälyiseksi *h:ksi, joka keskikantasaamessa (MPS) jäi ilman substituuttia, koska kielessä ei ollut vielä *h:ta. (sananalkuinen *j on proteettinen). Germaaninen *uj korvattiin tässä vaiheessa *vj:llä.
- Luoteisgermaanista varhaiskantaskandinaaviin (EPN) tultaessa *eu kehittyi *iu:ksi, ja myöhäiskantasaameen (LPS) tultaessa vokaalisto oli jo käynyt läpi suuren mullistuksen. Saamessa ei vielä­kään ollut *h:ta, ja *uj korvattiin jälleen *pj:llä.
- Keskikantaskandinaaviin (MPN) kehittynyt *y korvattiin saamen muinaismurteissa *u:lla kuten tapahtui myös saamen itämerensuomalaisissa lainasanoissa; läntiseen murteeseen oli jo ilmestynyt *h, mutta siellä keskikonsonantisto ilmenee eri järjestyksessä: *uj korvattiin *jv:llä.

Kielitieteellinen ”kolmiomittaus”

Koska (itämeren)suomella on ollut kontakteja sekä saamen että germaanin kanssa, voidaan edellä esitetyn kaltaisia tietoja hyödyntää myös itämerensuomalaisten äänne­muutosten ja lainasanojen äännekorvausten (substituutioiden) kronologian selvittämisessä. Saamessa germaanisella *χ:lla tai *h:lla on kolme eri vastinetta eri-ikäisissä lainasanakerrostumissa: *k, kato ja *h. Itämerensuomessa vastineita on vain kaksi: *k ja *h. Yhteisten samanikäisten lainasanojen perusteella nähdään esimerkiksi, että kantasuomeen *h ilmestyi aikaisemmin kuin saameen:

EPN⁸ *hiuja- → LPF *hipjä > suomi *hipiä* ’iho; aivina’

EPN *hiuja- → LPS *ipje > eteläsaame *ibje* ’pöly’

Myöhäiskantasuomi (LPF) ja myöhäiskantasaame (LPS) näyttävät myös olevan suunnilleen samanikäisiä kielimuotoja, mutta keskinäisten lainasanojen perusteella myöhäiskantasaamen taso on saavutettu hieman aikaisemmin. Myöhäiskantasaameen on omaksuttu lainasanoja vielä keskikantasuomesta (MPF):

MPF *lajša (> suomi *laiha*) → LPS *lāššē > pohjoissaame *láš’ši* ’ohut’

Näin kolmen eri kielen keskinäisellä vertailulla saadaan muodostettua tarkempi kronologia kielten ja ääntenmuutosten menneisyydestä kuin vain kahta kieltä vertailemalla. Lainasanat mahdollistavat myös aivan eri kielikuntiin kuuluvien kielten vertailemisen.

Absoluuttinen ajoitus

Lainasanoja vertailemalla saadaan selville sekä lainasanakerrostumien että ääntenmuutosten suhteellinen ja joissain tapauksissa jopa absoluuttinen kronologia. Varhaisten kirjoitettujen germaanisten sanojen (riimutekstit sekä roomalaisissa lähteissä esiintyvät germaaniset nimet) perusteella on esimerkiksi päätelty, että myöhäiskantagermaanista luoteisgermaaniin (länsi- ja pohjoisgermaanin yhteinen esimuoto) johtanut muutos $*\bar{e} > *\bar{a}$ on tapahtunut vasta hyvin lähellä ajanlaskun alkua (Koivulehto 1999a). Näin kielenmuutosten suhteellinen kronologia voidaan ankkuroida summittaisesti kalenterivuosiin, jolloin suhteessa aikaisemmat muutokset ovat kyseistä ankkurikohtaa vanhempia ja suhteessa myöhäisemmät muutokset ovat kyseistä ankkurikohtaa nuorempia.

Kontaktikielten horisontaalisen vertailun (H-tyyppi) avulla tämä vertikaalisen vertailun (I-tyyppi) avulla saavutettu germanistinen tulos voidaan hyödyntää myös uralistisella puolella. Niinpä kuluvan vuosituhannen tulokset ovat osoittaneet, että saame ja itämerensuomi olivat vielä niinkin myöhään kuin luoteisgermaanin aikaan melko lähellä toisiaan, melkoisen osan näitä kielihaaroja erottavista ääntenmuutoksista sijoittuessa vasta ajanlaskun alun jälkeisiksi (Koivulehto 2002; Aikio 2006; Kallio 2009). Alueellisesti saame ja itämerensuomi ovat kuitenkin eronneet jo aikaisemmin ja varsin idässä, koska mordva voidaan palauttaa samaan länsiuralilaiseen rekonstruktioitasoon josta saame ja itämerensuomi ovat eriytyneet ja koska yhteisten lainasanojen erilaiset johdinainekset osoittavat alueellisen eriytymisen tapahtuneen hyvin varhain (ks. edeltä alaluku Välikantakielet).

Lainautumissuunta

Aina ei ole helppo päätellä sanan lainautumissuuntaa. On esimerkiksi joukko sanoja, joilla on laaja levikki niin uralilaisessa kuin indoeurooppalaisessakin kielikunnassa. Lainasuunnan selvittämiseksi joudutaan tällöin turvautumaan tietoihin kielenulkoisesta maailmasta: tiedetään että nuoremmissakin kielikontakteissa lainasuunta on yleensä indoeurooppalaisista kielistä uralilaisiin kieliin, joten voidaan alustavasti olettaa tilanteen olleen samanlainen muinaisuudessaakin. Syy ei

tietystikään ole kielissä itsessään vaan siinä, että indoeurooppalaisia kieliä puhutaan uralilaisten kielten eteläpuolella. Euraasiassa innovatiiviset kulttuurikeskukset ovat sijainneet etelässä, joten usein myös kulttuurisanoja lainataan etelästä. Toinen syy on se, että etelässä on viljelty maata paljon kauemmin, mikä on mahdollistanut väestönkasvun ja sitä kautta aiheuttanut helpommin väestöliikkeitä etelästä pohjoiseen, ja ihmisten mukana kulkeutuu helposti sanoja. Kolmas peruste on se, että arkeologisesti on havaittu useita vaikutusaaltoja arolta metsävyöhykkeelle mutta vähemmän toiseen suuntaan; indoeurooppalaisten aluetta oli arovyöhyke ja uralilaisten aluetta metsävyöhyke (Koivulehto 1999b).

Toisaalta kielen levitessä toisen kielen alueelle ja syrjäyttäessä sen kielenvaihdon kautta voi uuteen kieleen jäädä substraatti aiemmasta kielestä. Yleensä tämä substraatti näkyy paikannimistössä sekä sellaisessa sanastossa johon sisältyvät esimerkiksi maastonmuotoihin, elinkeinoihin sekä paikalliseen kasvi- ja eläinkuntaan liittyvät sanat. Esimerkiksi venäjän kielen pohjoismurteissa on mittava suomensukuisista kielistä peritty substraatti, ja latvian kielessä on itämerensuomalainen (tarkemmin liiviläinen) substraatti (Saarikivi 2006; Balode 2008).

Usein lainautumissuunta voidaan päätellä myös sanan rakenteesta. Verrattaessa esimerkiksi ruotsin sanaa *glas* suomen sanaan *lasi* voidaan päätellä suomen sana lainatuksi ruotsista: sananalkuinen konsonanttiyhtymä on ollut suomen kielen fonotaksille (äänteiden yhdistämistä koskeva säännöstö) vieras, joten sananalkuinen *g* on pudotettu pois. Sen sijaan ei tunnu mahdolliselta, että ruotsalaiset olisivat lainanneet sanan suomesta ja lisänneet sen alkuun ylimääräisen *g*:n. Suomen sanan perään on toki lisätty ylimääräinen *i*, mutta tämä johtuu siitä, että yksitavuinen konsonanttiloppuinen sana on niin ikään ollut suomen fonotaksin vastainen (poikkeuksena *mies*; sellaiset sanat kuin *taos* ja *taus* ovat historiallisesti kaksitavuisia). Lisäksi sanan paljastaa nuoreksi jo sen taipuminen *lasi* : *lasin* eikä *e*:llisenä kuten vanhat sanat (*kuusi* : *kuusen*).

Vieläkin helpommin lainasuunta on selvitetävissä silloin, kun sanassa näkyy tietyn kielen kieliopillisia aineksia kuten sijapäätteitä tai johtimia. Esimerkiksi suomen sana *kuningas* on lainattu kantagermaanin sanasta **kuningaz* (> englannin *king*, ruotsin *konung*). Ensiksikin sanassa näkyy germaaninen nominatiivin päätte **-z*, toiseksi sana näyttää olevan johdos germaanisesta vartalosta **kuni-* 'suku; päällikkö', ja kolmanneksi itse *-ing*-johdinkin on germaaninen (Köbler 2007).

Muutoksen suunta

Aina ei sisäisen rekonstruktion eikä vertailevan menetelmän avulla saada selville, mikä on jonkin ääntenmuutoksen suunta. Tällöin ratkaisu saattaa löytyä lainasanatutkimuksen avulla. Uralilaisten kielten sibilanttien edustus on eräs tällainen arvoitus: yhtäältä eräissä sanoissa on kielikunnan läntisissä haaroissa *s ja itäisissä *t, *L tai kato, toisaalta toisissa sanoissa on läntisissä haaroissa *ś ja itäisissä *s. Kantaauraliin voitaisiin rekonstruoida yhtä hyvin läntisten kuin itäistenkin kielten edustuksen mukaiset äänteet. On ehdotettu, että muutos *t > *s olisi typologisesti yleisempi kuin päinvastainen muutos, mutta lainasanojen todistuksen mukaan muutos on silti epäämättä tapahtunut päinvastaiseen suuntaan: vanhoissa indoeurooppalaisissa lainasanoissa esiintyvä *s on säilynyt lännessä ja muuttunut idässä (> unkari Ø [kato] ~ mansi *t ~ hanti *L ~ samojedi *t), samoin indoeurooppalainen *ś on säilynyt lännessä ja muuttunut idässä (> unkari, mansi, hanti, samojedi *s). Esimerkkeinä kaksi kanta-arjalaista lainasanaa (J. Häkkinen 2009):

unkari *száz* ~ mansi KM *seēt* ~ hanti V *sàt* '100'
 < itäurali **se̊ta*
 < kantaaurali **śeta* (> ersämordva *śado* ~ suomi *sata*)
 ← kanta-arja **śata*- / kantaarja **śata*- '100'

mansi KM *uutər* 'herra, ruhtinas; sankari'
 < itäurali **aLəra*
 < kantaaurali **asira* (> ersämordva *azoro*)
 ← kanta-arja **asura* > avesta *ahura* 'herra'

4. Yhteenveto

Tässä kirjoituksessa olen avannut historiallisen kielitieteen käsitteitä ja lähtökohtia sekä esitellyt sen vertailumenetelmiä uralistiikan (≈ fennougristiikan) viime vuosien tuloksia esittelevin käytännön esimerkein. Historiallisessa kielitieteessä kieltenvälinen vertailu merkitsee aina menneisyyteen kurottamista: vertailu on keino varhaisempien kielentason tavoittamiseksi ja rekonstruomiseksi. Vertailu voi kohdistua kielessä minkä tahansa tason aineksiin: sanoihin tai nimiin, äänteisiin tai muotokategorioihin. Olen käsitellyt seuraavanlaisia historiallisen kielitieteen vertailumenetelmiä:

I-tyyppi = vertikaalisen vertailun ideaalityyppi, jossa nykykieli on verrattavan muinaiskielen suora ja puhdas jatkaja. Mahdollinen kielenmuistomerkkien tutkimuksessa.

Y-tyyppi = vertikaalisen vertailun tyyppi, jossa nykykieli ei ole verrattavan muinaiskielen suora tai puhdas jatkaja vaan joko kuuluu läheiseen sisarhaaraan tai on murteiden sekoittuma (kuten suomen kirjakieli). Kielenmuistomerkkien tutkimus edustaa yleensä tätä tyyppiä.

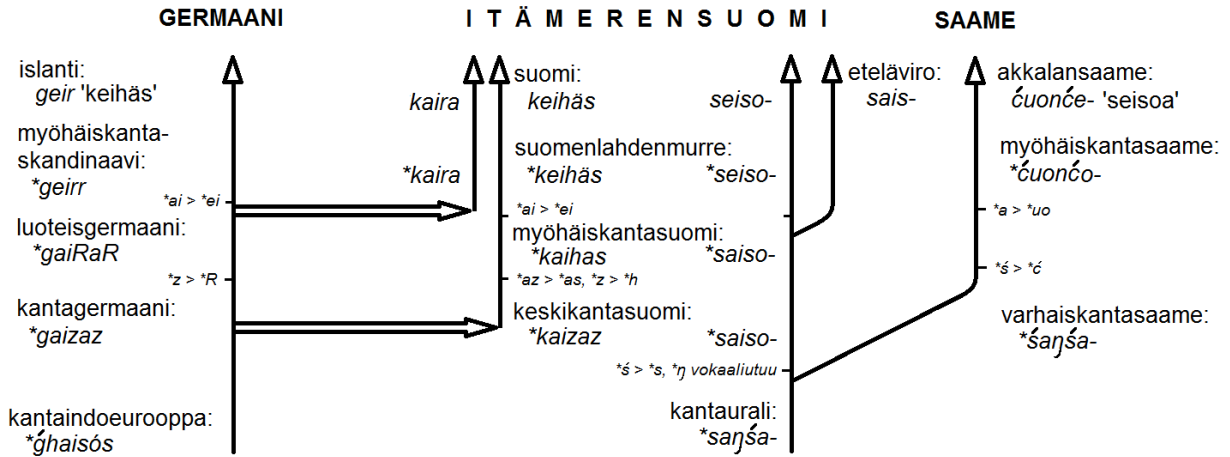
T-tyyppi = vertikaalisen vertailun tyyppi, jossa nykykielessä ilmenevän vaihtelun perusteella menneisyyteen rekonstruoidaan vaihtelematon alkutilanne. Sisäinen rekonstruktio edustaa tätä tyyppiä.

V-tyyppi = horisontaalisen vertailun tyyppi, jossa sukukieliä vertailemalla menneisyyteen rekonstruoidaan kantakieli. Vertaileva rekonstruktio edustaa tätä tyyppiä.

H-tyyppi = horisontaalisen vertailun tyyppi, jossa kahta eri kielilinjaa (ja usein kielikuntaa) edustavien kielten välisten kontaktien perusteella saadaan suhteutettua kielilinjojen itsenäiset kronologiat toisiinsa. Lainasanatutkimus edustaa tätä tyyppiä.

On tärkeää ymmärtää, että jokaisen vertailutyyppin kohdalla kyseessä on yksi ja sama kielellinen todellisuus – vain lähestymiskulma vaihtelee. Yhdessä nämä vertailutyyppit muodostavat kattavan verkoston, jossa I:n, T:n ja Y:n pystyviiva on yksi ja sama kielilinja; myös V:n toinen sakara ja H:n toinen pystyviiva on tämä aivan sama kielilinja. Mitä useampi kirjaintyyppi osuu kohdakkain, sitä vahvempi on viiva, eli sitä lähemmäs todellista kielen historiaa päästään. Aina eri vertailutyyppit eivät kuitenkaan anna samansuuntaisia tuloksia, ja silloin on selvitettävä, onko kyse sittenkin eri kielilinjoista tai aikasyvyyksistä. Yleisempi ongelma lienee kuitenkin, ettei kaikkia esiteltyjä vertailutyyppijä ole mahdollista hyödyntää, joten lopullisen rekonstruktion osuvuus jää epävarmaksi.

Seuraavassa kaaviossa havainnollistetaan historiallisen kielitieteen vertailumenetelmien verkostoa itämerensuomen näkökulmasta. Vasemmalla on lainasanatutkimuksen tuloksia: germaaninen sana 'keihäs' on lainattu kahdesti itämerensuomeen eri aikoina, ja lopputuloksena suomen kielessä on sanat *keihäs* ja *kaira*. Oikealla on vertailevan rekonstruktion tuloksia: kantaauralasta periytyvän sanan 'seisoa' perusteella hahmottuu ikään kuin suppea sukupuu, jossa ensin eroaa saame ja sitten eteläviro. Esimerkkisanoja yhdistää myöhäiskantasuomen suomenlahdenmurteessa muutamissa sanoissa tapahtunut äännevuoro **ai > *ei*. Aivan sama muutos on tapahtunut myös kantaskandinaavissa, mutta myöhemmin kuin suomenlahdenmurteessa – vasta sen jälkeen kun *kaira*-sana oli jo lainattu. Myöhäiskantasuomen suomenlahdenmurteessa muutos sen sijaan on tapahtunut jo ennen lainautumista, koska sana *kaira* ei ole ehtinyt osallistua kyseiseen muutokseen (tosin sporadisena muutoksena tämä ei koskenut koko sanastoa).



Lyhenteet

mansi KM = mansin Keski-Kondan murre

hanti V = hantin Vahin murre

SaaW = läntiset saamelaiskielet (etelä-, uumajan-, piitimen-, luulajan- ja pohjoissaame)

SaaE = itäiset saamelaiskielet (inarin-, koltan-, kildinin-, akkalan- ja turjansaame)

Kirjainyhdistelmät:

E- = varhais- (Early)

M- = keski- (Middle)

L- = myöhäis- (Late)

-P- = -kanta- (Proto)

-F = -suomi (Finnic)

-S = -saame (Saami)

-N = -skandinaavi (Nordic)

-U = -urali

-G = -germaani

Lähdeluettelo

Aikio, Ante 2006: On Germanic-Saami contacts and Saami prehistory. – *Suomalais-Ugrilaisen Seuran Aikakauskirja* 91, s. 9–55. Helsinki: Suomalais-Ugrilainen Seura.
<http://www.sgr.fi/susa/91/aikio.pdf> (2.7.2011)

Aikio, Ante 2007: Etymological nativization of loanwords: a case study of Saami and Finnish. — Ida Toivonen & Diane Nelson (toim.): *Saami Linguistics*, s. 17–52. Amsterdam: John Benjamins.
http://cc.oulu.fi/~anaikio/etym_nativ_preprint.pdf (2.7.2011)

Aikio, Ante & Aikio, Aslak 2001: Heimovaelluksista jatkuvuuteen. Suomalaisen väestöhistorian tutkimuksen pirstoutuminen. – *Muinaistutkija* 4 / 2001, s. 2–21. Helsinki: Suomen arkeologinen seura. http://cc.oulu.fi/~anaikio/Heimovaelluksista_jatkuvuuteen.pdf (2.7.2011)

Balode, Laimute 2007: Once again on Some Potential Finno-Ugrisms in Latvia. Juhani Nuorluoto (toim.): *Topics on the Ethnic, Linguistic and Cultural Making of the Russian North*. Slavica Helsingiensia 32.

Bartens, Raija 1999: *Mordvalaiskielten rakenne ja kehitys*. Suomalais-Ugrilaisen Seuran Toimituksia 232. Helsinki: Suomalais-Ugrilainen Seura.

Bartens, Raija 2000: *Permiläisten kielten rakenne ja kehitys*. Suomalais-Ugrilaisen Seuran Toimituksia 238. Helsinki: Suomalais-Ugrilainen Seura.

Campbell, Lyle 1998: *Historical Linguistics: an Introduction*. Edinburgh: Edinburgh University Press.

Donner, Otto 1879: *Die gegenseitige Verwandtschaft der finnisch-ugrischen Sprachen*. Helsingfors: Acta Societatis Scientiarum Fennicae 11.

Fox, Anthony 1995: *Linguistic Reconstruction. An Introduction to Theory and Method*. Oxford / New York: Oxford University Press.

Gallén, Jarl (toim.) 1984: *Suomen väestön esihistorialliset juuret*. Tvärminnen symposiumi 17.–19.1.1980. Bidrag till kännedom av Finlands natur och folk, 131. Helsinki: Societas Scientiarum Fennica.

Helinski, Eugen 2005: The 13th Proto-Samoyedic Vowel. *Mikola-konferencia 2004*, Szeged: SzTE Department of Finno-Ugristics.

Honti, László 1998: Ugrilainen kantakieli – erheellinen vai reaalinen hypoteesi? *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii*. Suomalais-Ugrilaisen Seuran toimituksia 228, s. 176–187. Helsinki: Suomalais-Ugrilainen Seura.

Häkkinen, Jaakko 2007: *Kantauralin murteutumisen vokaalivastaavuuksien valossa*. Pro gradu -työ, Helsingin yliopiston Suomalais-ugrilainen laitos. <https://oa.doria.fi/handle/10024/7044> (2.7.2011)

Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. – *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, s. 9–56. <http://www.sgr.fi/susa/92/hakkinen.pdf> (2.7.2011)

Häkkinen, Jaakko 2010a: Jatkuvuusperustelut ja saamelaisen kielen leviäminen (OSA 1). – *Muinaistutkija* 1 / 2010, s. 19–36. <http://www.mv.helsinki.fi/home/jphakkin/Jatkuvuus1.pdf> (2.7.2011)

Häkkinen, Jaakko 2010b: Jatkuvuusperustelut ja saamelaisen kielen leviäminen (OSA 2). – *Muinaistutkija* 2 / 2010. <http://www.mv.helsinki.fi/home/jphakkin/Jatkuvuus2.pdf> (2.7.2011)

Häkkinen, Jaakko (tulossa): After the protolanguage: Invisible convergence, fake divergence, and boundary shift.

Häkkinen, Kaisa 1984: Wäre es schon an der Zeit, den Stammbaum zu fällen? – *Ural-Altäische Jahrbücher, Neue Folge* 4, s. 1–24.

Häkkinen, Kaisa 1994: *Agricolasta nykykieleen. Suomen kirjakielen historia*. Porvoo/Helsinki/Juva: WSOY.

Häkkinen, Kaisa 1996: *Suomalaisten esihistoria kielitieteen valossa*. Tietolipas 147. Helsinki: Suomalaisen kirjallisuuden seura.

Itkonen, Esa 1998: Sukupuu ja kontakti. – *Virittäjä* 1 / 1998.

Itkonen, Terho 1997: Reflections on Pre-Uralic and the 'Saami-Finnic protolanguage'. – *Finnisch-Ugrische Forschungen* 54, s. 229–266. Helsinki: Suomalais-Ugrilainen Seura.

Janhunen, Juha 2001: Indo-Uralic and Ural-Altaiic: On the diachronic implications of areal typology. – Carpelan, Parpola & Koskikallio (toim.): *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, s. 207–220. Suomalais-Ugrilaisen Seuran Toimituksia 242.

Jarva, Vesa 2003: *Venäläisperäisyys ja expressiivisyys suomen murteiden sanastossa*. Jyväskylä Studies in Humanities 5. Jyväskylä: Jyväskylän yliopisto.

Joki, Aulis J. 1973: *Uralier und Indogermanen. Die älteren Berührungen zwischen den Uralischen und Indogermanischen Sprachen*. Suomalais-Ugrilaisen Seuran Toimituksia 151. Helsinki: Suomalais-Ugrilainen Seura.

Kalima, Jalo 1936: *Itämerensuomalaisten kielten balttilaiset lainasanat*. Suomalaisen Kirjallisuuden Seuran Toimituksia 202. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kallio, Petri 2006: Suomen kantakielten absoluuttista kronologiaa. – *Virittäjä* 1 / 2006, s. 2–25. http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/2006_2.pdf (2.7.2011)

Kallio, Petri 2007: Kantasuomen konsonanttihistoriaa. – Jussi Ylikoski ja Ante Aikio (toim.): *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*, s. 229–249. Suomalais-Ugrilaisen Seuran Toimituksia 253. Helsinki: Suomalais-Ugrilainen Seura. http://www.sgr.fi/sust/sust253/sust253_kallio.pdf (2.7.2011)

Kallio, Petri 2009: Stratigraphy of Indo-European loanwords in Saami. – Tiina Äikäs (toim.): *Máttut – máddagat: The Roots of Saami Ethnicities, Societies and Spaces / Places*, s. 30–45. Giellagas-instituutin julkaisut 12. Oulu: Giellagas-instituutti. <http://www.mv.helsinki.fi/home/petkalli/mattut.pdf> (2.7.2011)

Kallio, Petri (tulossa): Suomen viikinkiaika kielihistorian valossa. (Luonnos)

Koivulehto, Jorma 1999a: *Verba mutuata. Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint*. Suomalais-Ugrilaisen Seuran Toimituksia 237.

Koivulehto, Jorma 1999b: Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa. – Paul Fogelberg (toim.): *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, s. 249–280. Bidrag till kännedom av Finlands natur och folk, 153. Helsinki: Societas Scientiarum Fennica.

Koivulehto, Jorma 2002: Contact with non-Germanic languages II: Relations to the East. – Oskar Bandle et al. (toim.): *The Nordic Languages. An International Handbook of the History of the North Germanic Languages*, s. 583–594. Berlin – New York: Walter de Gruyter.

Koivusalo, Esko (toim.) 1988: *Mikael Agricolan kieli*. Tietolipas 112. Helsinki: Suomalaisen Kirjallisuuden Seura.

Korhonen, Mikko 1974: Oliko suomalais-ugrilainen kantakieli agglutinoiva? Eli mitä kielihistoriallisista rekonstruktioista voidaan lukea ja mitä ei. – *Virittäjä* 1974.

Korhonen, Mikko 1981: *Johdatus lapin kielen historiaan*. Suomalaisen Kirjallisuuden Seuran Toimituksia 370. Helsinki: Suomalaisen Kirjallisuuden Seura.

Korhonen, Mikko 1988: Uralilaisten kielten jälkitavujen vokaaliston historiaa. – *Virittäjä* 1988, s. 8–24.

Kulonen, Ulla-Maija 1993: *Johdatus unkarin kielen historiaan*. Suomi 170. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kulonen, Ulla-Maija 2002: Suomalaiset ja suomalais-ugrilaiset luvut ja lukusanat. *Lukujen historia*, toimittanut Graham Flegg. Helsinki: Art House.

Kulonen, Ulla-Maija 2010: *Fonesteemit ja sananmuodostus. Suomen kontinuatiivisten u-verbijohdosten historiaa*. Suomi 197. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kylstra, A. D. & al. 1991, 1996: *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen*. Amsterdam: Rodopi.

Köbler, Gerhard 1989: *Gotisches Wörterbuch*. 2. painos.
<http://www.koeblergerhard.de/gotwbhin.html> (2.7.2011)

Köbler, Gerhard 2007: *Germanisches Wörterbuch*. 3. painos.
<http://www.koeblergerhard.de/germwbinw.html> (2.7.2011)

Laakso, Johanna 1999: Vielä kerran itämerensuomen vanhimmista muistomerkeistä. – *Virittäjä* 1999. <http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/vir99laakso.pdf> (2.7.2011)

Lehtinen, Tapani 2007: *Kielen vuosituhanet. Suomen kielen kehitys kantauralista varhaisuomeen*. Tietolipas 215. Helsinki: Suomalaisen Kirjallisuuden Seura.

Mallory, J. P. 1989: *In Search of the Indo-Europeans. Language, Archaeology and Myth*. London: Thames and Hudson.

Moora, Harri (toim.) 1956: *Eesti rahva etnilisest ajaloost*. Tallinn.

Posti, Lauri & Terho Itkonen (toim.) 1973: FU-transkription yksinkertaistaminen. *Castrenianumin toimitteita* 7. Helsinki: Suomalais-Ugrilainen Seura ja Castrenianumin laitokset.

Rapola, Martti 1990: *Johdatus Suomen murteisiin*. Tietolipas 4, neljäs painos. Helsinki: Suomalaisen Kirjallisuuden Seura.

Saarikivi, Janne 2006: *Substrata Uralica. Studies on Finno-Ugrian substrate in northern Russian dialects*. Väitöskirja. Tartu. <http://ethesis.helsinki.fi/julkaisut/hum/suoma/vk/saarikivi/substrat.pdf> (2.7.2011)

Salminen, Tapani 1997: *Tundra Nenets inflection*. Suomalais-Ugrilaisen Seuran toimituksia 227. Helsinki: Suomalais-Ugrilainen Seura.

Salminen, Tapani 1998: Pohjoisten itämerensuomalaisten kielten luokittelun ongelmia. – Riho Grünthal & Johanna Laakso (toim.): *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii*, s. 390–406. Suomalais-Ugrilaisen Seuran Toimituksia 228. Helsinki: Suomalais-Ugrilainen Seura.

Salminen, Tapani 1999: Euroopan kielet muinoin ja nykyisin. – Paul Fogelberg (toim.): *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, s. 249–280. Bidrag till kännedom av Finlands natur och folk, 153. Helsinki: Societas Scientiarum Fennica.

Salminen, Tapani 2002: Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. – *Лингвистический беспредел: сборник статей к 70-летию А. И. Кузнецовой*, s. 44–55. Москва: Издательство Московского университета.
<http://www.helsinki.fi/~tasalmin/kuzn.html> (2.7.2011)

Sammallahti, Pekka 1988: Historical phonology of the Uralic Languages. – Denis Sinor (toim.): *The Uralic languages*, s. 478–554. Leiden: Brill.

Sammallahti, Pekka 1998: *The Saami Languages. An Introduction*. Kárášjohka: Davvi Girji.

Sammallahti, Pekka 1999: Saamen kielen ja saamelaisten alkuperästä. – Paul Fogelberg (toim.): *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*, s. 249–280. Bidrag till kännedom av Finlands natur och folk, 153. Helsinki: Societas Scientiarum Fennica.

¹ Foneettisen ja fonologisen tason välinen ero on siinä, että fonologisella tasolla huomioidaan vain itsenäiset foneemit – e:tä edeltävä *k* on foneettisesti etisempi [*k̟*] kuin vaikka *u*:ta edeltävä *k*, mutta kyseessä ovat silti saman /*k*/-foneemin reaalistumat. Astevaihtelun vuoksi foneettisesti pitkä [*ŋ̃*] on syytä hahmottaa fonologisella tasolla kahden /*ŋ*/-äänteen yhtymäksi: /*ŋŋ*/ reaalistuu vain alkuperäisessä umpitavussa (konsonanttiloppuisessa tavussa), ollen täydennysjakaumassa avotavussa (vokaaliloppuisessa tavussa) esiintyvän /*ŋk*/-yhtymän kanssa: [*keŋk̟ä*] : [*keŋŋät̚*]. Yhtymän ensimmäinen [*ŋ*] on kuitenkin vain yksi /*n*/-foneemin allofoneista: seuraava velaariklusiili velaaristaa myös nasaalin, eli /*n*/:n ääntymäpaikka siirtyy *k*:ta vastaavaan kohtaan, jolloin se ääntyy [*ŋ*]:nä. Yhtymän toinen [*ŋ*] on oikeastaan *k*:n heikon asteen vastine, joka vain on suomessa assimiloitunut edeltävään nasaaliin ja muuttunut itsekin nasaaliksi [*ŋ*]. Jo myöhäiskantasuomessa /*k*/:n heikon asteen vastineen /*ŋ*/:n (soinnillinen velaarispirantti) allofoni nasaalin jäljessä on ollut [*g*] (**kenkä* : **kengät* = [*keŋkä*] : [*keŋgät̚*]), joten nykysuomen fonologisen syvätason asu /*kengät̚*/ olisi sattumoisin samannäköinen kuin kirjakielen asu <*kengät̚*>, jossa digrafeemi <*ng*> kuvaa [*ŋ*]-äännettä, ja myöhäiskantasuomen fonologinen rekonstruktio **kengät̚*.

² Kolmionmittaus on maanmittauksen menetelmä, jossa kolmion yhden sivun ja tunnettujen kulmien perusteella voidaan laskea puuttuvien sivujen pituudet. Historiallisessa kielitieteessä lähestytään arvoituksellista kantakieltä nykykielten ”kulmien” avulla.

³ Astevaihtelu syntyi foneettisena ilmiönä, klusiilin heikkenemisenä tietyin ehdoin (umpitavun alussa). Niin kauan kuin heikentynyt edustaja oli täydennysjakaumassa klusiilin kanssa eli sen allofoni, pysyi astevaihtelu foneettisena ilmiönä. Astevaihtelu foneemistui siinä vaiheessa, kun sen synnyttänyt ehto hämärtyi: virossa sananloppuisten konsonanttien katoaminen johti siihen, että niin klusiili kuin sen heikon asteen edustajakin (usein kato) saattoivat nyt edustua avotavun alussa, esimerkiksi **tupa* : **tuβan* > *tuba* : *toa*.

⁴ Liudentuminen tarkoittaa ilmiötä, jossa konsonantti saa *j*:mäisen lisävärin. Velaaristuminen tarkoittaa etuvokaalin takaistumista niin että äänteen muut ominaisuudet (korkeus ja labiaalisuus eli pyöreys) pysyvät ennallaan.

⁵ Morfologisen eli muoto-opillisen tyypologian luokittelussa agglutinoiva kieli on sellainen, jossa kielipilliset ainekset liitetään sanavartaloon ikään kuin liimaamalla (*talo-ssa-ni-kin*), kun taas flekteeraava eli fusionaalinen kieli on sellainen, jossa sanansisäisellä vaihtelulla ilmaistaan kielipillisiä funktioita.

⁶ Petri Kallio (2006) viittaa nimityksellä kantagermaania varhaisempaan kielentasaan.

⁷ Etymologinen nativisaatio on ilmiö, jossa tiiviissä kontaktissa ja mahdollisesti kaksikielisyydessä elävät sukukielten puhujat hahmottavat vanhan yhteisen sanastonsa äännesuhteita ja sopeuttavat nuoretkin keskinäiset lainasanat niiden mukaisiksi.

⁸ Lyhenteet selitetään kirjoituksen lopussa.