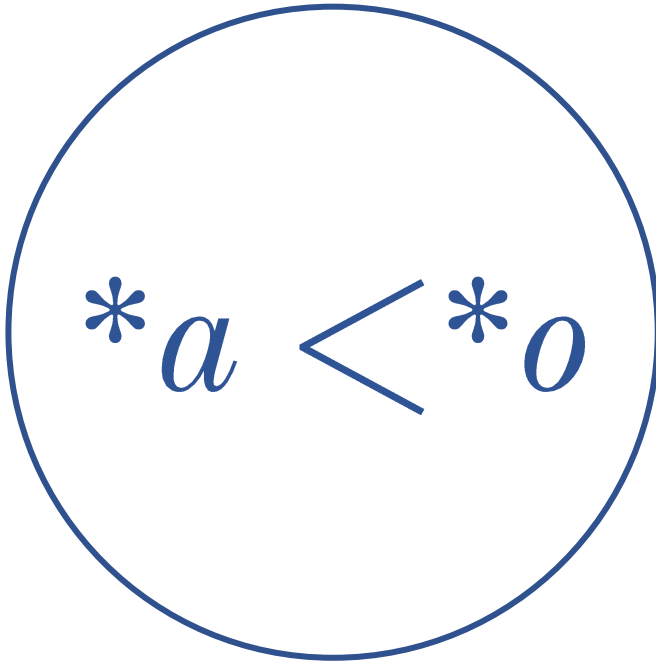


Miten historiallinen kielitiede tutkii menneisyyttä?

Historiallis-vertaileva metodi lyhyesti selitettynä

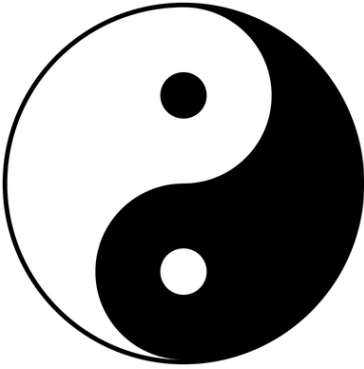
© Santeri Junttila 2018



Historiallinen eli diakroninen kielentutkimus on vanhin esihistoriaa tutkiva tiede. Sen perusta, historiallis-vertaileva metodi kehitettiin jo 19. vuosisadan alkupuolella.

Verrattomasti suurin osa tähänastisesta historiallisesta kielentutkimuksesta on ollut indoeuropeistiikkaa eli kohdistunut indoeurooppalaisiin kieliin. Fennougristiikka eli uralistiikka on kuitenkin ollut tutkimuksen kärjessä alusta alkaen. Alan menetit kehittyvät jatkuvasti, mutta perusta pysyy.

Etymologia ja äännehistoria



Historiallisessa kielentutkimuksessa on kaksi puolta, jotka etenevät rinnan tuottaen jatkuvasti uutta tietoa toisilleen: etymologia ja äännehistoria. Etymologia tutkii sanojen alkuperää, äännehistoria taas kielten menneisyydessä vaikuttaneita äänne muutoksia ja niiden ehtoja. Historiallis-vertaileva metodi on äännehistoriallinen, mutta aineistonsa se saa etymologialta.

Äännehistoria

Kielihistoriasta tuli tiedettä, kun ymmärrettiin, että jokainen äänne muutos vaikuttaa **vain tietynä ajankohtana tietyissä kielessä**, ja silloin se vaikuttaa periaatteessa **kielen koko sanastoon** eli kaikkiin sanamuotoihin, joissa tietty äänne esiintyy tietyissä äänneympäristössä.

Kunkin kielen historiassa on siis omat äänne muutoksensa, ja ne ovat tapahtuneet tietyissä järjestyksessä. Tämä järjestys eli **suhteellinen kronologia** selviää **ehdollisia äänne muutoksia** tutkimalla. Ehdollinen äänne muutos on tapahtunut vain tietynlaisessa äänneympäristössä.

Ehdollisten äänne muutosten tulosta kielessä ovat sellaiset äänneelliset erot (eli äänne vaihtelut), joilla ei ole omaa merkitysfunktiota. Esimerkiksi suomessa taivutetaan eri tavoin sanoja *lasi*, *kuusi* ja *käsi*. Taivutusmuodot *lasin*, *kuusen* ja *käden* ovat funktioltaan aivan samanlaisia genetiivejä, joten niiden äänneainesten *-si-*, *-se-* ja *-de-* ero eli vaihtelu perustuu menneisyydessä tapahtuneisiin ehdollisiin äänne muutoksiin.

lasi : *lasin* (*i*-vartalo)

kuusi : *kuusen* (*e*-vartalo)

käsi : *käden* (*e*-vartalo, *-te*-tyyppi)



Sanoissa *kuusi* ja *käsi* on *-i* vain niissä muodoissa, joissa se on sanan loppuäänteenä. Kaikkien päätteitten edellä sen tilalla on joko *-e-* tai *kato*. Monikkumuodoissa *kuusia*, *kuusissa* jne. on toki *-i-*, mutta se on etymologisesti eri *-i-*, monikon tunnus (ja alun perin *-j-*).

<i>kuusi</i>	<i>kuuse t</i>	<i>kuuse nsa</i>	<i>kuus ta</i>	<i>kuus i a</i>	<i>kuus i ssa</i>
<i>käsi</i>	<i>käde t</i>	<i>käte nsä</i>	<i>kät tä</i>	<i>käs i ä</i>	<i>käs i ssä</i>

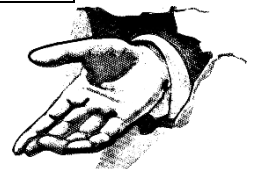
Tästä on päätelty, että aiemmin on sanassa ollut *-e*, joka on muuttunut *i*:ksi sananlopussa (ja kadonnut *-t-n* edellä: *kuus|ta*). Saadaan siis kantamuoto **kuuse* ja äänteenmuutos *-e > -i* (huomaa viiva ennen kirjainta merkitsemässä, että äänne on ollut sanan lopussa).

Toinen vaihtoehto on, että sanassa on ollut *-i*, joka on muuttunut *e*:ksi kaikkialla muualla kuin sananlopussa (ja kadonnut *-t-n* edellä), eli kantamuoto on ollut **kuusi* ja äänteenmuutos *-i > -e-* (huomaa viivat molemmin puolin kirjainta merkitsemässä, että äänne on ollut sanan keskellä).

Jälkimmäinen vaihtoehto *-i > -e-* herättää kuitenkin jatkokysymyksen, kun tiedämme, että jokainen äänteenmuutos vaikuttaa kielen koko sanastoon: jos sanassa **kuusi* on loppuvokaali muuttunut *e*:ksi aina ennen päätettä, miksi sitten ei sanota myös *lasi* : ***lasen* (yksi tähti merkitsee rekonstruoitua kantamuotoa, kaksi tähteä kontrafaktuaalista eli olematonta muotoa)?

Tästä syystä on yleensä ajateltu, että *-e*-vartaloisissa sanoissa alkuperäinen loppuvokaali on ollut *-e*: **kuuse*, siis

	<i>*kantamuoto</i>	>	nykymuoto						
	<i>*kuuse</i> :	<i>*kuusen</i>	>	<table border="1"><tr><td><i>kuusi</i></td><td>:</td><td><i>kuusen</i></td></tr></table>	<i>kuusi</i>	:	<i>kuusen</i>		
<i>kuusi</i>	:	<i>kuusen</i>							
	<table border="1"><tr><td><i>*lasi</i></td><td>:</td><td><i>*lasin</i></td></tr></table>	<i>*lasi</i>	:	<i>*lasin</i>	>	<table border="1"><tr><td><i>lasi</i></td><td>:</td><td><i>lasin</i></td></tr></table>	<i>lasi</i>	:	<i>lasin</i>
<i>*lasi</i>	:	<i>*lasin</i>							
<i>lasi</i>	:	<i>lasin</i>							



Kielihistoriallisissa päätelmissä tarvitaan kuitenkin myös **sukukielten** apua. Itämerensuomalaisesta kantakielestä eli kantasuomesta ensin haarautunut kieli on eteläviro eli ukala. Siinä on sanat

kuus : *kuusõ* 'kuusi'

klaas : *klaasi* 'lasi'

Ukalan *kuus* vastaa äänteellisesti täysin suomen sanaa *kuusi*, kun tiedetään, että pitkän ensitavun jälkeinen toisen tavun vokaali on ukalasta kadonnut (merkitään $V > 0$), ja jos kadonnut vokaali oli *i*, se palatalisoi edeltäneen konsonantin: $-si > ś$. Jälkitavun *õ* vastaa tässä suomen *e*:tä.

Mutta sana *klaas* ei vastaa etymologisesti suomen sanaa *lasi*, eli sitä ei voi palauttaa samaan kantamuotoon säännöllisten äänne muutosten kautta. Tämä johtuu siitä, että ukalan *klaas* on laina saksasta ja suomen *lasi* laina suomenruotsista. (Tästä emme vielä voi varmuudella päätellä, että kantasuomi jakautui ennen kuin lasi tunnettiin meillä päin, sillä lasilla on periaatteessa voinut olla jokin toinenkin nimi, joka nyt on kadonnut.)

Valtaosa suomen, ukalan ja muitten itämerensuomalaisten kielten *i*-vartaloisista sanoista onkin lainoja historialliselta ajalta, siis kantasuomen jälkeiseltä kaudelta. Lopuillakaan ei ole yhteistä, kantasuomeen palautettavaa muotoa.

Niinpä kaikki *i*-vartaloiset nominit ovat kantasuomea uudempia sanoja.

Nyt kuitenkin tiedämme, etteivät *i*-vartaloiset sanat estä rekonstruoimasta *e*-vartaloisten sanojen alkuperäiseksi taivutukseksi **kuusi* : **kuusin*, josta $-i > -e$ -muutoksen kautta tuli **kuusi* : **kuusen*. Nykyisiä *i*-vartaloisia sanoja vain **ei ollut vielä olemassa** tämän muutoksen tapahtumahetkellä.

Tämä ei kuitenkaan tarkoita, ettei muutoksen suunta voinut olla $-e > -i$. Kysymystä muutoksen suunnasta ei itse asiassa ole ollut helppo ratkaista, mutta kielen kronologian kannalta ei ole oikeastaan merkitystä, kumpi vokaali kantamuotoon merkitään. Edempänä tekstissä kutsun sitä lyhyden vuoksi perinteisesti $-e > -i$ -muutokseksi.

Olennaisempaa on tietää, missä kielen historian vaiheessa äänteenmuutos tapahtui ja *e*-vartaloisten nominien äännevaihtelu *-i* ~ *-e-* (*kuusi* : *kuuse-*) syntyi. Sama äännevaihtelu on kaikissa itämerensuomalaisissa kielissä, joten *-e* > *-i*-muutos on tapahtunut jo kantasuomessa. Mutta onko se voinut tapahtua jo aiemmin?

Kantasuomen lähimmät sukulaiskielet olivat kantasaame ja kantamordva. Niitten tytärikielissä kuten pohjoissaamessa tai ersässä ei ole tällaista vokaalivaihtelua:

suomi	pohjoissaame	ersämordva
<i>kuusi</i> : <i>kuusen</i>	<i>guossa</i> : <i>guosa</i>	<i>куз</i> : <i>кузонь</i>
<i>nimi</i> : <i>nimen</i>	<i>namma</i> : <i>nama</i>	<i>лем</i> : <i>лемень</i>
<i>veri</i> : <i>veren</i>	<i>varra</i> : <i>vara</i>	<i>верь</i> : <i>верень</i>

Niinpä *-e* > *-i*-muutos ei ole kantasuomea vanhempi. Taulukosta yllä näemme sen sijaan muitten äänteenmuutosten seurauksia: pohjoissaamessa sanan loppuvokaali on muuttunut *a*:ksi ja ersässä kadonnut sananlopussa, mutta säilynyt genetiivin päätteen edellä (sanan muista äänneistä riippuen joko *-o*-na tai *-e*-nä).

Entäpä *käsi* : *käden*? Miksi *s* vaihtelee *t*:n ja *d*:n kanssa? Tälläkään vaihtelulla ei ole merkitysfunktiota, joten sekin perustuu aiempaan vaihtelemattomuuteen.

<i>käsi</i>	<i>käde t</i>	<i>käte nsä</i>	<i>kät tä</i>	<i>käs i ä</i>	<i>käs i ssä</i>
-------------	---------------	-----------------	---------------	----------------	------------------

Jos olemme koulussa tai suomen kielen kurssilla oppineet, että suomen *t* ja *d* ovat keskenään astevaihtelussa, tiedämme jo, että ne kuuluvat yhteen, ja historiallinen kielentutkimus on myös selvittänyt, että vahvan asteen muoto on aina alkuperäinen ja heikon asteen muoto äänteenmuutoksen (tässä tapauksessa *-t* > *-d-*) tulos. Merkitsemme tässä *-t*-n ja *-d*-n isolla kirjaimella *T*. Isoilla kirjaimilla merkitään historiallisessa kielentutkimuksessa arkkifoneemeja eli useamman äänteen joukkoa, jota yhdistää jokin yhteinen piirre.

Kumpi tässä siis on vanhempi, *-s-* vai *-T-*? Vastaus on nyt helpompi, koska tiedämme jo, että sanassa *kuusi* on kaikissa vastaavissa asemissa *-s-*:

<i>kuusi</i>	<i>kuuse t</i>	<i>kuuse nsa</i>	<i>kuus ta</i>	<i>kuus i a</i>	<i>kuus i ssa</i>
<i>käsi</i>	<i>käde t</i>	<i>käte nsä</i>	<i>kät tä</i>	<i>käs i ä</i>	<i>käs i ssä</i>

Samoin tiedämme, että *kuusi* on kantasuomea vanhempi sana, koska sillä on vastineet esimerkiksi pohjoissaamessa (*guossa*) ja ersämordvassa (*кыэ*). Jos *-s-* \sim *-T-* on uudempi kuin sana *kuusi*, se ei voi perustua muutoksen *-s-* $>$ *-t-*, sillä tällöin muodoista *kuuset*, *kuusensa* ja *kuusta* olisi tullut ***kuudet*, ***kuutensa* ja ***kuutta*. (Juuri näinhän taipuu numero *kuusi* '6!'). Onko muutos siis ollut *-t-* $>$ *-s-*?

Oletusta näyttää vastustavan, että suomessa on myös *-ti-*päätteisiä vanhoja *e*-vartaloisia sanoja, joista vaihtelu puuttuu:

lahti *lahde|t* *lahte|nsa* *lahte|a* *laht|i|a* *lahd|i|ssa*
tähti *tähde|t* *tähte|nsä* *tähte|ä* *täht|i|ä* *tähd|i|ssä*

Näissä sanoissa on kuitenkin aina *-ht-*yhtymä, joten voidaan ajatella, että *-h-*n jäljessä *-t-* ei ole muuttunut *-s-*ksi, onhan *-hs-* vaikea ääntää. On siis todennäköistä, että alkuperäinen sisäkonsonantti *käsi*-tyypin sanoissa on *-t-*, joka on muuttunut *s:*ksi joissakin muodoissa. Lopullisen varmuuden saamme vasta sukukielistä: missään niistä ei näitten sanojen vastineissa esiinny *-s-*ää.

suomi	pohjoissaame	ersämordva
<i>käsi</i> : <i>käden</i>	<i>giehta</i> : <i>gieda</i>	<i>кедь</i> : <i>кедень</i>
<i>sata</i> : <i>sadan</i>	<i>čuohti</i> : <i>čuođi</i>	<i>сядо</i> : <i>сядонь</i>

Äänteenmuutoksen *t* $>$ *s* ehdotkin on helppo selvittää. Ensimmäinen ehto on, ettei *t:n* edellä ollut *h-*äännettä. Toinen ehto näkyy tämän sivun ylimmästä taulukosta: *t* $>$ *s* on tapahtunut aina *i:n* edellä. Niinpä sitä kutsutaankin *ti* $>$ *si* -muutokseksi.

Entä kumpi muutos on vanhempi, *e* $>$ *i* vai *ti* $>$ *si*?

Vastaus on taas yksinkertainen. Muutos $e > i$ on luonut ehdot $ti > si$ -muutokselle. Sen täytyy siis olla vanhempi. Muutenhan sanoisimme *käsi* : ***käsen* (jos alkuperäinen vartalovokaali oli *i*) tai ***käti* : *käden* (jos alkuperäinen vartalovokaali oli *e*, niin ennen sen muuttumista *i*:ksi ei koko $ti > si$ -muutosta voinut tapahtua).

Näitten esimerkkien valossa ymmärrämme, että ääntenmuutoksen luonteen ymmärtäminen merkitsi historiallisen kielentutkimuksen syntyä tieteenä. **Kukin ääntenmuutos vaikuttaa kielen koko sanastoon** eli kaikkiin sanamuotoihin, joissa tietty äänne esiintyy tietyssä äänneympäristössä, ja siksi suomessa on paitsi *käsi* : *käden*, myös *vesi* : *veden*, *viisi* : *viiden*, *varsi* : *varren* (< *varden*), *kansi* : *kannen* (< *kanden*) jne. Tämän lainalaisuuden ansiosta ääntenmuutokset voidaan asettaa keskinäiseen järjestykseen.

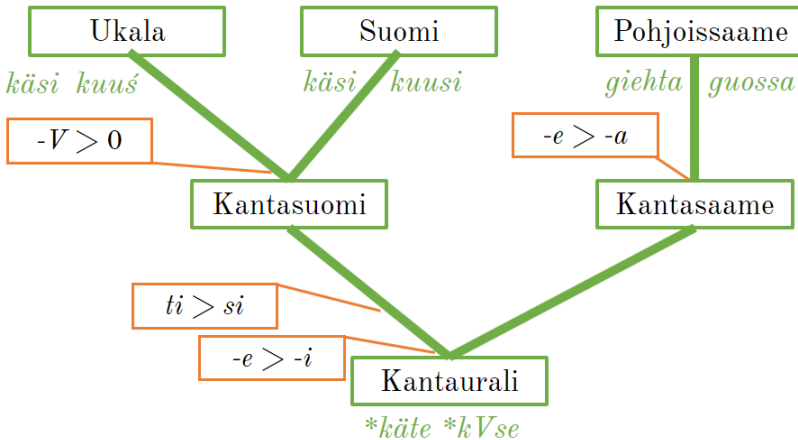
Vielä perustavanlaatuisempaa ääntenmuutosten kronologian kannalta on ymmärtää, että **kukin ääntenmuutos vaikuttaa vain tietyssä ajankohtana ja vain tietyssä kielessä**. Niin $-e > -i$ -muutos kuin $ti > si$ -muutoskin ovat kantasuomalaisia, eivätkä liity mitenkään vaikkapa pohjoissaamen kielen kehitykseen tai suomen kielen kantasuomen jälkeiseen kehitykseen. Tämän vuoksi kantasuomen jälkeen suomeen saadut sanat *tila* ja *peti* eivät ole muuttuneet muotoihin ***sila* ja ***pesi* ($ti > si$ -muutos), ja vaikkapa pohjoissaamassa *käsi* : *käden* on *giehta* : *gieda* (pohjoissaamen $-e > -a$ -muutos), eikä esimerkiksi ***giehti* : ***giede(n)* (kantasuomen $-e > -i$ -muutos) tai ***giesi* : ***giede* ($ti > si$ -muutos).

Tähän perustuu kielihistorian arvo historiallisia päätelmiä tehtäessä.

Suhteellinen kronologia ja sukupuu

Ääntenmuutosten suhteellisen kronologian avulla hahmotetaan kielten **sukupuu** sijoittamalla ääntenmuutokset haarautumiskohtiin. Näitä haarautumiskohtia kutsutaan **kantakieliksi**. Kukin kantakieli on jakautunut tytärkieliksi, kun jokin ääntenmuutos on tapahtunut osassa sen aluetta, mutta jäänyt muualla toteutumatta. Ääntenmuutosten suhteellinen kronologia siis määrää, millaisten välikantakielten kautta koko kielikunnan kantakieli on kehittynyt nykykieliksi.

Yllä esitettyjen päätelmien pohjalta saamme jo yksinkertaisen sukupuun kolmelle nykykielille, suomalle, ukalalle ja pohjoissaamelle. Ersämordvan jätämme pois tästä puusta, koska sen ja pohjoissaamen keskinäisiä äännesuhteita emme ole käsitelleet – emme siis voi sanoa, onko sillä yhteinen välikantakieli kantasuomen, kantasaamen, molempien tai ei kummankaan kanssa.



Etymologia ja kielten sukulaisuus

Tässä tehdyt äännehistorialliset päätelmät sukupuun hahmottelemiseksi on esitetty hyvin pelkistetyksi. Ennen kaikkea pois jäi kaikki **etymologinen tutkimus**, joka on osoittanut meille, mitkä eri kielten sanat ovat sukua keskenään eli polveutuvat samasta näitten kielten yhteisen kantakielen sanasta (on olennaista ymmärtää, että tämä on sama asia), jolloin niitä kutsutaan toistensa (etymologisiksi) **vastineiksi**. Pelkistäminen oli mahdollista, koska sellaiset sanat kuin *kuusi*, *guossa* ja *ryz* ovat niin saman kuuloisia, että niiden yhteenkuuluvuus on jo intuitiivisesti selvää.

Etymologia etsii kielten sanojen välisiä **säännöllisiä äännevastaavuuksia** ja vastaavuuksien laadun pohjalta päättelee, ovatko sanat etymologisia vastineita vai liittyvätkö ne toisiinsa lainauksen kautta. Jos jonkin kieliparin sanojen välillä ei ole minkäänlaisia säännöllisiä äännevastaavuuksia, ei kieliparin välille voi osoittaa mitään yhteyksiä.

Tällöin sanojen samankuuloisuus on luultavimmin vain sattumaa – minkä tahansa kahden sattumanvaraisesti valitun kielen välillä löytyy useita saman kuuloisia sanoja, koska kaikissa kielissä on suppea määrä äänteitä, joita yhdistelemällä ne muodostavat tuhansia lyhyehköjä sanoja.

suomi	pohjoissaame	ersämordva
<i>kuusi</i> : <i>kuusen</i>	<i>guossa</i> : <i>guosa</i>	<i>куз</i> : <i>кузонь</i>
<i>nimi</i> : <i>nimen</i>	<i>namma</i> : <i>nama</i>	<i>лем</i> : <i>лемень</i>
<i>veri</i> : <i>veren</i>	<i>varra</i> : <i>vara</i>	<i>верь</i> : <i>верень</i>

Säännölliset vastaavuudet löytyvät vasta, kun kahdesta kielestä verrataan toisiinsa useita sanapareja (tai useammasta kielestä useita sanasarjoja), joissa on sama äänne. Yllä olevan taulukon *kuusi*, *nimi* ja *veri* sopivat tähän tarkoitukseen, koska niissä on sama loppuvokaali, mutta pitävät etymologiset päätelmät luonnollisesti tehdään paljon laajemman aineiston pohjalta.

Kielet, joissa on samansukuisia sanoja ja kieliopillisia aineksia, ovat **sukukieliä** eli nekin polveutuvat yhteisestä kantakielestä. Etymologia siis osoittaa myös kielten sukulaisuuden. Äännehistorian tehtävänä on sitten selvittää kielten sukulaisuuden aste eli niitten yhteisen kantakielen paikka suhteellisessa kronologiassa. Se tehdään, kuten edellä on kuvattu, selvittämällä keskinäinen järjestys äänteenmuutoksille, jotka ovat aikaansaaneet sukukielten väliset äänne-erot.

Suomen, ukalan, pohjoissaamen ja ersämordvan keskinäinen sukulaisuus on niin ilmiselvää, ettei sitä ole koskaan jouduttu todistamaan. Tämä johtuu siitä, että niissä on paljon yhteistä sanastoa ja niitten yhteisen kantakielen – jota nykyään nimitetään länsikantauraliksi – jälkeen ei niissä ole tapahtunut niin paljon äänteenmuutoksia, että isoa osaa kielten keskeisestä sanastosta (esimerkiksi numeroita) ei tunnista samaksi.

Paljon vähemmän intuitiivinen ja usein kyseenalaistettu on unkarin sukulaisuus muitten uralilaisten kielten kanssa. Etymologia ja äännehistoria osoittavat sen kuitenkin kiistämättömäksi.

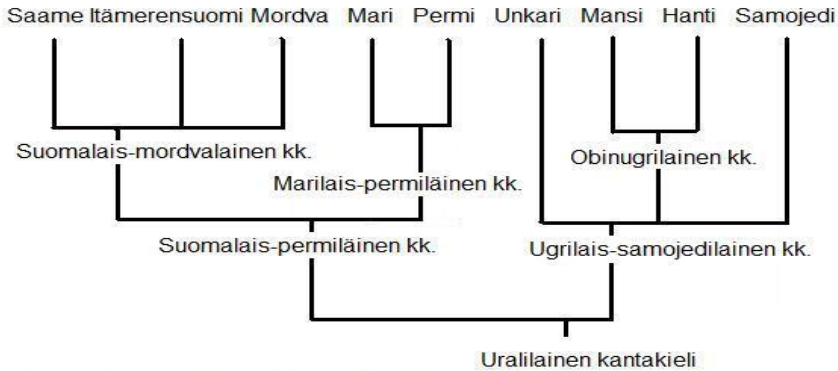
suomi	pohjoissaame	unkari
<i>käsi</i>	<i>giehta</i>	<i>kéz</i>
<i>kyynel</i>	<i>ganjal</i>	<i>könny</i>
<i>kepeä</i>	<i>gehpis</i>	<i>kevés</i>
<i>kolme</i>	<i>golbma</i>	<i>három</i>
<i>kota</i>	<i>goahti</i>	<i>ház</i>
<i>kuusi '6'</i>	<i>guhtta</i>	<i>hat</i>

Yhtäältä unkarin sanat näyttävät hyvin erilaisilta kuin suomen ja pohjoissaamen, mutta toisaalta taulukossa näkyy kaikkien kolmen kielen sanojen välillä sananalkuisen konsonantin säännöllinen suhde. Se vain on taas ehdollinen: suomen *k*-ta ja pohjoissaamen *g*-tä vastaa etuvokaalin edellä unkarin *k*- ja takavokaalin edellä unkarin *h*-. Kummastakin äänne-edustuksesta todistaa kaikkiaan vain toistakymmentä suomi–unkari-sanaparia, mutta se on jo riittävä määrä, sillä näitten sanojen välillä on muitakin säännöllisiä äännesuhteita. Kun esimerkiksi tiedetään, että suomen *käsi*-sanana aiempi muoto on **käte*, löytyy siitä sama äännesuhde kuin sanasta *kota*: suomen *-t* = pohjoissaamen *-ht* = unkarin *-z*.

Kielten välinen sukulaisuus voidaan havaita aina, kun niissä on säilynyt tarpeeksi paljon yhteisiä sanoja ja kieliopillisia aineksia säännöllisten äännevastaavuuksien löytämiseksi. Jos kielten välinen sukulaisuus on hyvin etäistä, eli niitten yhteinen kantakieli on jakautunut jo useita vuosituhansia ennen niitten vanhimpia kirjallisia muistomerkkejä, ei yhteistä ainesta yleensä ole riittävästi sukulaisuuden todistamiseksi. Tästä syystä esimerkiksi indoeurooppalaisten ja uralilaisten kielten keskinäinen sukulaisuus on jäänyt arvoitukseksi, vaikka niillä onkin muutamia vanhoja saman kuuloisia sanoja, joita ei voi varmasti lainoiksikaan osoittaa.



Uralilaisten kielten sukupuu



Uralilainen sukupuu Jaakko Häkkisen mukaan (2007).

Nykykäsitys uralilaisten kielten sukupuusta rakentuu 19. vuosisadalta lähtien kerättyjen tietojen ja niistä tehdyn kielihistoriallisen tutkimuksen pohjalle. Kielimateriaalin kerääminen ja järjestäminen on kuitenkin ollut sukupolvien mittainen työ, ja vasta nykytutkijoilla alkaa olla käytössään riittävä aineisto luotettavien kielihistoriallisten päätelmien tekemiseksi.

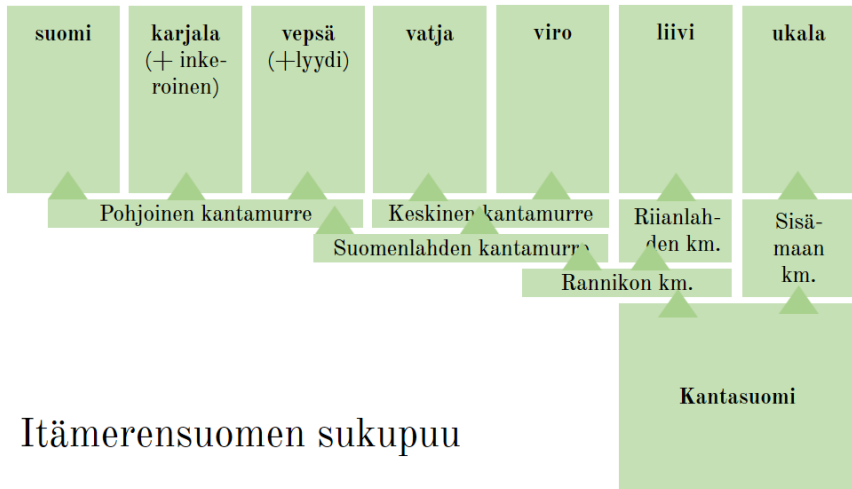
Jaakko Häkkisen kielipuussa jokaisen kantakielen haarautuminen liittyy todistettuihin äänne- ja sanamuutoksiin. Tässä se eroaa olennaisesti aiemmista puista, joissa oli useita lähinnä maantieteen ja sanojen levikin pohjalta hahmoteltuja, kielihistoriallisesti todistamattomia välikantakieliä.

Häkkisen puusta on jätetty pois uusimman tason yhdeksän välikantakielen haarautuminen edelleen: esimerkiksi itämerensuomi hänen puussansa kuvaa vain kantasuomea.

Itämerensuomalaisten kielten sukupuu

Myös itämerensuomen sukupuu perustuu yli sadan vuoden ajan kerättyyn ja järjestettyyn aineistoon sekä siitä jatkuvasti tehtävään tutkimukseen. Alla oleva puumalli on hahmottunut viimeisten kolmen vuosikymmenen kuluessa ensin Pekka Sammallahten ja Tiit-Rein Viitson, sittemmin ennen

kaikkea Petri Kallion (2014) töissä. Siinäkin haarautumisten järjestys pohjautuu äänne muutosten keskinäiseen kronologiaan.



Itämerensuomen sukupuu

Soveltava etymologia ja absoluuttinen kronologia

Historiatieteenä diakroninen kielentutkimus pyrkii myös ikäämään tutkimansa ilmiöt. Ajoitusten antaminen kielen muutoksille on periaatteessa mahdollista niin kauas menneisyyteen kuin kielihistorian metodi ylipäänsä ulottuu, eli pari–kolmetuhatta vuotta vanhimpia kirjallisia lähteitä kauemmas.

Joskus äänne muutoksia voi iätä suoraan kirjallisten muistomerkkien perusteella. Näin on ajoitettu esimerkiksi latinan vähittäinen muuntuminen romaanisiksi kieliksi sekä kantaskandinaavin ja kantaslaavin murteutuminen.

Uralilaisten kielten kirjalliset lähteet ulottuvat vain muutaman vuosisadan taa, mutta niitten kosketukset paljon aiemmin kirjoitettuihin indoeurooppalaisiin naapurikieliin ovat jatkuneet vuosituhansia. Tästä syystä fennougristit (uralilaisten kielten tutkijat) ovat erityisen kiinnostuneita **lainasanoista**, ja esihistoriallisten lainasanojen tutkimus on fennougristiikassa kehittynyt metodisesti korkeammalle kuin minkään muun maailman kielikunnan tutkimustraditiossa.

Jos lainasana voidaan ajoittaa sen lähtökielessä kirjallisesti dokumentoidun äänne­muutoksen mukaan, niin tämän äänne­muutoksen ajoitus auttaa ikäämään myös lainan kohdekielen muutoksia ja jakautumisia.

Esimerkiksi kantagermaanista on peräisin suomen länsimurteitten *lieko* 'kaatunut puu', jonka lähtömuodon vastine ruotsissa on samaa merkitsevä *låga*. Ruotsin *å* palautuu muinaisskandinaavin pitkään *a*-vokaaliin, jonka taas kaikkein vanhimpien riimukirjoitusten ja germaanisten kielten latinalaisten lainojen (epäsuorasti siis antiikin latinankielisten tekstien) perusteella tiedetään kehittyneen kantagermaanin pitkästä *e*:stä (*ē*) jo joskus 100-luvulla aaj. Koska sanassa *lieko* on *-ie-*, sen on täytynyt tulla jo ennen tätä puhutusta germaanisesta kielestä, jossa oli vielä *-ē-* eikä siitä kehittynyt *-ā-*. Vuosisatoja myöhemmin saman sanan muinaisskandinaavisesta, *-ā-*llisesta muodosta lainattiin *laho*, joka edelleen esimerkiksi vepsässä merkitsee 'kaatunutta puuta'.

Tämän kautta tiedämme myös, että jo parituhatta vuotta sitten oli kantasuomessa sama pitkä *ōō*, joka virossa on säilynyt edelleen ja suomessa muuttunut *ie*-diftongiksi esimerkiksi sanassa *vieras*, viron *võõras*. Sanalla *lieko* ei ole vastinetta virossa eli sen on täytynyt jostain viron esimuodosta kadota, mutta tiedämme, että kantasuomessa se kuului **lōōko* – koska jokainen äänne­muutos vaikuttaa kielen koko sanastoon! *Lieon* ansiosta tiedämme edelleen, että 100-luvulla aaj. puhuttiin jo kantasuomea eikä enää länsikantauralia, jossa ei vielä ollut *ōō*-vokaalia.

Niinkin varhainen tapahtuma kuin kantaauralin jakautuminen on voitu ajoittaa lainasanojen äänne­piirteitten avulla. Uralilaisissa kielissä on nimittäin sanoja, joitten levikki ulottuu kielikunnan laidasta laitaan, joten niitten piti kuulua jo kantaauraliin. Muutamien äänne­piirteittensä vuoksi ne ovat kuitenkin voineet lainautua vasta arjalaisesta kantakielestä, joka indoeuropeistien mukaan erosi kantaindоеuroopasta viimeistään 1800 eaa., mutta ei paljon aiemmin. Tällainen sana on *sata*, jonka kanta-arjalainen lähtömuoto **sata* oli kehittynyt kantaindоеuroopan muodosta *k̑ntom*. Sen ja muutaman muun sanan ansiosta tiedämme, että kantaauralia puhuttiin vielä parituhatta vuotta eaa. jossain arjalaisten läheisyydessä,

eli luultavasti Volgan keskijuoksulla – ei siis esimerkiksi Suomessa ja Baltiassa jo kahta vuosituhatta aiemmin, kuten pitkään uskottiin.

Joillakin lainasanoilla myös merkitys välittää tietoa esihistoriallisten kieli-kontaktien ajankohdasta. Tällaisia lainoja kutsutaan **kulttuurisanoiksi**. Esimerkiksi monet itämerensuomen vanhimmista slaavilaislainoista, kuten *pappi*, *raamattu* ja *risti* liittyvät kristinuskoon. Yhdessä historian-tutkimuksen ja arkeologian kanssa kielitiede on voinut ajoittaa nämä lainat ensimmäisen vuosituhannen lopulle. Niillä on vastineet kaikissa itämerensuomalaisissa kielissä, joitten koko alueella on kristinuskko siis opittu tuntemaan juuri slaaveilta. *Pappi* ja *risti* ovat myös itämerensuomen harvoja yhteisiä *i*-vartalaisia sanoja. Ehkä juuri slaavien vaikutus levitti itämerensuomalaisille paitsi kristinuskon, myös nykyään niin yleisen *i*-vartalotyypin. Joka tapauksessa niitten samanaikaisuus antaa *e*- ja *i*-vartalosisille sanoille hyvän ajoitusperusteen: edelliset sanat on omaksuttu ennen kristinuskon tuloa, jälkimmäiset sen jälkeen.

Varhaisempien kulttuurisanojen ikäysarvo on vielä suurempi, sillä niitten saapumisajasta ei ole säilynyt sellaisia historiallisia lähteitä, joitten perusteella voitaisiin tehdä päätelmiä kielen puhuma-alueista ja kontakteista. Esimerkiksi germaanisten lainojen antaman välillisen todistuksen kautta on kantasuomen $\check{s} > h$ -muutos voitu ajoittaa ajanlaskun pariin ensimmäiseen vuosisataan. Koska vanhimmissa balttilaisissa lainoissa liettuan konsonanteja \check{s} ja \check{z} vastaa aina itämerensuomen *h* eikä koskaan esimerkiksi *s*, on kantasuomalais-balttilaisten kosketusten täytyntä päätyä ennen tuota aikaa. Kun vielä tiedämme, että balttilaisista kielistä on peräisin paljon kaskiviljelyn ja sisävesikalastuksen sanastoa, voimme päätellä, että itämerensuomalaiset ovat harjoittaneet näitä elinkeinoja yhtäjaksoisesti – sanojen välillä unohtumatta – ainakin kahdentuhannen vuoden ajan.

Kielikunnan menneisyydestä kertovat sellaisetkin kulttuurisanat, joitten alkuperää ei tiedetä, jos niitten iän paljastaa levikki. Nykyisellään fennougristiikka pystyy nimittäin äännesuhteitten perusteella varsin luotettavasti erottamaan uralilaisesta kielestä toiseen lainatut sanat vanhoista, kantaauraliin tai länsikantaauraliin palautuvista perintösanoista.

Esimerkiksi monenlaisia säilytystiloja ja -kuoppia merkitsevää *purnu*-sanaa on pidetty kantauralilaisena, koska samanlaisia ovat pohjoissaamen *buorna* 'säiliösalvos', marin *purña* 'koppa', komin *бурня* 'laari' ja udmurtin *бурнĕ* 'sammio'. Nämä sanat kuitenkin muistuttavat toisiaan aivan liikaa voidakseen periytyä kantauralista. Marin, komin ja udmurtin äännekehityksen aikana kantakielen sanoista katosivat toisen tavun vokaalit. Suomen *-u*-ta taas pitäisi pohjoissaamassa vastata *-o*- eikä *-uo*-.

Purnu on siis varmasti lainautunut uralilaisesta kielestä toiseen, vaikka sen lähtökieli ja kulkusuunta jäävätkin epävarmoiksi. Kulttuurihistoriallisesti paljon kiinnostavampia ovat kuitenkin sanat *vaski* ja *uuhi* sekä suomeen asti säilymätön viljan, ohran tai vehnän nimitys **šępti*, jotka Jaakko Häkkinen rekonstruoi kantauralilaisiksi. Ne nimittäin poikkeavat huomattavasti useimpien kantauralin sanojen kulttuurikuvasta, jossa näkyy metsävyöhykkeen metsästäjä–keräilijäyhteisö.

Etenkin luontoon liittyvän sanaston pohjalta voi tehdä päätelmiä myös menneisyyden kielten puhuma-alueista. Itämerensuomalais-mordvalainen vastinepari, suomen *tammi* = ersän *түмо* osoittaa länsikantauralia puhutun suhteellisen etelässä, balttilaisperäiset *ankerias* ja *lohi* taas viittaavat kontakteihin Itämeren valuma-alueella. Kielten aiemmista puhuma-alueista kertovat myös niistä nykykieliin välittyneet **paikannimet**, joitten todistus ei kuitenkaan yllä kovin kauas menneisyyteen, sillä nimet vaihtuvat muuta sanastoa nopeammin.

Kirjallisuutta

Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. – *Suomalais-Ugrilaisen Seuran Aikakauskirja* 92, s. 9–56

Kallio, Petri 2014: The Diversification of Proto-Finnic. – *Fibula, Fabula, Fact. The Viking Age in Finland*, s. 155–68

Kallio, Petri 2015: The Stratigraphy of the Germanic Loanwords in Finnic. – *Early Germanic Languages in Contact*, s. 23–38.

Lang, Valter 2018: *Läänemeresoome tulemised*.