

Itämerensuomen seuraava etymologinen sanakirja

1. Viimeisten parin vuosikymmenen aikana on suomalaisia hemmoteltu kolmella suurella etymologisella hakuteoksella. Niitten valmistumiseen johtanut toimitustyö alkoi viimeistään Emil Nestor Setälän perustaessa Tutkimuslaitos Suomen Suvun vuonna 1930, aineiston keruu paljon aiemmin, jo Henrik Gabriel Porthanin kerätessä kiinnostavia sanoja suomesta ja sen sukukielistä. Suomalaisilla etymologeilla ei ole vielä pienintäkään syytä hakeutua muihin töihin, vaan tämä on vasta alkua. Vertailun vuoksi Klugen saksan etymologisesta sanakirjasta on ilmestynyt jo 24., korjattu ja laajennettu painos.

Kun laajennamme näkökulmaa suomen kielestä koko itämerensuomeen, näyttävät saavutetut tulokset kutistuvan jo varsin pieniksi tekemättömän työn rinnalla. Virostakaan ei ole vielä julkaistu viimeisteltyä etymologista sanakirjaa, saati sitten pienemmistä lähisukukielistä. Voisiko seuraava suurhanke olla etymologinen kuvaus itämerensuomalaisten kielten sanastosta? Entä millaista esitystä lähisukukielet tarvitsisivat – olisiko kustakin kielestä laadittava oma etymologinen sanakirjansa vai kaikista yhdessä? Miten itämerensuomalaisen kielikunnan sisäinen äänteellinen ja semanttinen variaatio olisi parhaiten kuvattavissa etymologisina hakuteoksina?

2. Vanhimmat itämerensuomalaisia sanastokerrostumia etymologisesti kuvaavat esitykset, esimerkiksi balttilaisia lainoja käsittelevät Thomsen 1890 ja Kalima 1936, eivät voineet pyrkiä kielikunnan sisäisen variaation kuvauksen suhteen minkäänlaiseen kattavuuteen, koska lähdeaineisto oli vielä hyvin hajanaista ja osa peruskenttätöystäkin tekemättä. Vielä *Suomen kielen etymologisen sanakirjan* (SKES 1955–81) toimittaminen kesti puoli vuosisataa ennen kaikkea lähdemateriaalin keruun työläyden takia. Toimitustyön loppuvaiheessa aineisto oli kasvanut paljon alkuaikojaa laajemmaksi, mikä johti sarjan jälkipuolen osien huomattavaan paisumiseen. Vasta *Suomen sanojen alkuperä* (SSA 1992–2000) sekä Groningenin yliopiston tutkimusryhmän kokoama *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen* (LÄGLOS 1991–) ovat voineet luoda systemaattiset toimitusperiaatteet lähdeaineistosta löytyvän variaation kuvaamiseksi.

Koska suomen murteista löytyy suurin osa siitä sanastosta, joka kantasuomesta on periytynyt nykyisiin itämerensuomalaisiin kieliin, toimivat SSA ja sen edeltäjä SKES paremman puutteessa kaikkien itämerensuomalaisten

kielten etymologisina hakuteoksina. LÄGLOS ottaa askelen pitemmälle kuvattessaan koko kielikunnan muinaisgermaanisen lainakerrostuman, sijoittuvathan vanhimmat germaanikontaktit aikatasolle, jonka yhteydessä ei ole mielekästä puhua suomen kielestä. Koska SSA:nkin aineisto painottuu voimakkaasti kielen vanhimpiin sanastokerrostumiin, sen rajaaminen jää vasta muutaman viime vuosisadan aikana muodostuneen kansalliskielen pohjalta väkinäiseksi.

Uusin etymologinen kokonaisuus suomen kielestä on Kaisa Häkkisen 2004 julkaisema *Nykysuomen etymologinen sanakirja* (NSES). Siinä missä SSA pyrkii kattavuuteen etymologisesti tutkitun suomen sanaston kuvaamisessa, Häkkinen keskittyy sana-artikkelien luettavuuteen ja teokseen valitun materiaalin oleellisuuteen nykysuomalaisen kielestä kiinnostuneen maallikkolukijan kannalta. SSA:ssä on runsaasti sekä materiaalia lähi- ja etäsukukielistä että yleiskielelle tuntematonta murteitten ja vanhan kirjakielen sanastoa. Ei olekaan yllättävää, että hakuteoksena suomalaisen kodin hyllyssä Häkkisen sanakirja toimii paljon paremmin ja täyttää näin myös Setälän SKES:lle vuonna 1924 asettamat tavoitteet (Tanner 2007, 14–15) paremmin kuin SKES tai SSA. NSES on kuitenkin monessa suhteessa myös kielitieteilijän kannalta esimerkiksi helppokäyttöinen. Tässä pienessä kirjaelmassa pohdin, kuinka sen periaatteikseen ottamat selkeys ja luettavuus voisivat toteutua myös kielihistoriallista kattavuutta tavoittelevissa suomen ja sen lähisukukielten etymologisissa hakuteoksissa.

3. SSA:ää on arvosteluissa yleisesti kiitetty, mutta suurteos on herättänyt myös huomattavasti kritiikkiä, ennen kaikkea liiallisesta varovaisuudesta etymologiain todennäköisyyden arvioinnissa, sanarinnastusmerkintöjen tulkinnanvaraisuudesta ja etymologisten ongelmien kiertämisestä ekspressiivisysselytyksen avulla (Häkkinen 1993, Nikkilä 1997 ja 2002, Aikio 2001 sekä NSES:n arvostelu Saarikivi 2006). Teoksen olemassaolon mielekkyyttä sen sijaan ei ole asetettu kyseenalaiseksi, koska SKES:n jatkajaa pidetään kansallisesti tärkeänä suurhankkeena. Etymologisen teoksen merkitystä olisi kuitenkin arvioitava myös kielitieteen kannalta.

Jotta (haku)teos olisi selkeä, sillä pitäisi olla lyhyesti määritelty päätavoite, jolle kaikki muu on alisteista. SSA:llä on yhden päätavoitteen sijasta kaksi: ensinnä se pyrkii suomen kirjakielen vakiintuneen¹ sanaston kattavaan etymologiseen kuvaukseen – kuten tavallaan NSES:kin (s. 6) – mutta toisaalta myös tavoittelee samanlaista kattavuutta murteittemme etymologisesti käsitellyn sanaston osalta. Nämä kaksi tavoitetta saattavat ensilukemalta vaikuttaa helpostikin yhteen sovitettavilta, mutta niitten välillä on vaikea ristiriita, jonka aiheuttaa suomen nuoruus itsenäisenä kielimuotona. Historiallisen kielentutkimuksen kannalta kielimme murteet ovat pelkkä sattumanvarainen katkelma yhtenäistä

1. Vierasperäiset sivistyssanat, samoin kuin muut suomeen myöhään kotiutuneet kansainväliset sanat, joitten äänneasukin osoittaa ne vieraiksi, on yleensä jätetty pois, ellei niillä ole levikkiä kansankielessä (SSA I, 7).

itämerensuomalaista murrejatkumoa. Niitten sanaston käsitteleminen muun itämerensuomen leksikosta eroavin periaattein luo hakuteoksen esitykseen sotkuisuutta. Murrejatkumon sisäinen variaatio tarvitsisi yhtenäiset kuvauseriaatteet.

Millainen sitten pitäisi olla itämerensuomen variaation kuvaus etymologisessa sanakirjassa tai vaikkapa jonkin tietyn lainasanakerrostuman kuvauksessa hakuteoksessa? Tässä artikkelissa esittämiäni ajatuksia aion soveltaa käytäntöön valmistellessani kokonaistutkimusta itämerensuomen balttilaisista lainoista. Tässä haen vastausta lähtökohtanani SSA:n ja LÄGLOS:n toimituseriaatteet sekä niitten toteutuminen käytännössä. Pohdin, mitkä periaatteet ja käytänteet kannattaa omaksua olemassa olevista teoksista sellaisinaan, mitä kannattaa soveltaa laajemminkin ja mistä olisi syytä luopua kokonaan.

4. SSA:n periaate esitellä ”murteissa esiintyviä sanan äänteellisiä variantteja – – vain siinä tapauksessa, kun vaihtelu ei selity normaalin murrevariaation perusteella” on erinomainen lähtökohta selkeälle ja tiiviille etymologiselle sana-artikkelille, varsinkin kun siihen yhdistetään vielä SKES:sta puuttunut mutta SSA:ään sisällytetty periaate ilmoittaa myös hakusanain murrelevikit. Itämerensuomalaisen kielikunnan etymologisessa esityksessä tätä ”murreperiaatetta” olisi kuitenkin järkevää soveltaa kaikkiin kielimuotoihin niitten statuksesta riippumatta, mitä SKES, SSA ja LÄGLOS eivät tee.

Suomen ulkopuoliseksi määritellyn itämerensuomalaisen kielialueen sanaston kuvaamiseen on SSA:lla (1992, 13) aivan toisenlainen toimituseriaate, jota nimitän tässä lähisukukieliperiaateeksi: ”Lähisukukielistä on otettu mukaan tavallisesti yksi muoto kustakin tai muutama kyseistä sanuetta hyvin edustava sana. Levikkiä tai saantipaikkaa ei esitetä, mutta erikoistapauksessa voidaan kielen päämurteisiin viitata karkeahkoilla ilmansuuntaa osoittavilla lyhenteillä.” LÄGLOS (1991, XXVII) taas ilmoittaa esittävänsä lähisukukielten muodot pääsääntöisesti tärkeimpien sanakirjain (InkS, KKS, LyMS, SVJ, VKSR, VKKMS, VKMMS, EEW, VMS ja LivWb) mukaan. Näistä teoksista puuttuvat sanat esitetään SKES:n mukaan. Muita murremuotoja LÄGLOS mainitsee, jos niitten merkitys tai äänneasu antaa rekonstruktion kannalta olennaista lisäinformaatiota.

Käytännössä SSA ja LÄGLOS esittelevät lähisukukielten aineiston varsin samalla tavalla. Eroja näkyy lähinnä toimituskuntain huolellisuudessa sekä niissä valintatilanteissa, joihin toimituseriaatteet väistämättä ajavat: toimittajan on päätettävä, mitkä sanat edustavat sanuetta hyvin tai antavat olennaista lisäinformaatiota. Esimerkkinä olkoon koko kielikunnassa tavattava *aura*:

(SSA)

aura¹ (Agr; yl.) ’kyntölaite / Pflug’, rinn. murt. *arra*, *aar(r)a*, *a(a)tra*, *aarta*, *aira* ~ ink *atra* ’sahra’ | ka *atra* ’(kaski)aura, sahra’ | ly *adre*, *adru* ’aura, sahrat’ | ve *adr* | va *adra* | vi *ader* (g. *adra*), *adr*, *ahr* (g. *ahra*) | li *a’ddarz*, *a’draz* id.

(LÄGLOS)

AURA¹, (*a*)*atra*, *arra* 'Pflug'ingr. *atra*; karel. *atra*, *oatra*; lüd. *adr*; weps. *adr*; wot. *adra*; est. *ader*(G. *adra*), li *a'ddörz*

Graafiselta kannalta olennaisin ero sana-artikkelien välillä on SSA:n tapa toistaa sukukielten sanojen välillä niitten lähes identtisetkin merkitykset. Lisäksi SSA esittää enemmän äänteellisiä variantteja kuin LÄGLOS ja jää näin tekstin luettavuudessa siitä jälkeen. Molempien artikkelien tulkintaa rasittavat kuitenkin samat ongelmat, jotka aiheuttaa itämerensuomalaisten kielten perinteisessä statushierarkiassa pitäytyminen. Tämä hierarkia ei pohjaudu tutkimustuloksiin vaan on muotoutunut tutkimushistorian aikana vallinneitten sisä- ja ulkopoliittisten paineitten alla, eikä sen seuraamiselle ole nykyään enää kielitieteellisiä tai yhteiskunnallisia perusteita.

Perinteisen hierarkian mukaisten kahdeksan ims. kielen, suomen, inkeröisen, karjalan, lyydin, vepsän, vatjan, viron ja liivin lisäksi viimeistään parin viime vuosikymmenen kuluessa neljää murteena pidettyä kielimuotoa, meänkieltä, kveeniä, võroa ja settoa on alettu kutsua kieliksi niitten puhujain omasta vaatimuksesta. Näille kullekin on lyhyessä ajassa myös kehitetty yhtenäinen standardikirjakieli, joka ”vanhoista” kielistä puuttuu niin vatjalta kuin karjalaltakin. Itämerensuomen kirjakieliä on tätä nykyä 12 elävää ja yksi kuollut: meä, kveeni, suomi, vienankarjala, tverinkarjala, aunuksenkarjala, lyydi, vepsä, viro, võro, setto ja liivi sekä kirjakielenä jo sammunut inkeröinen.

Itämerensuomalaisia kirjakieliä on viime aikoina useita sekä syntynyt että kuollut, eikä ole mitenkään taattua, että määrä pysyy lähitulevaisuudessakaan samana. Etymologisen esityksen kannalta niitten statuksen määrittelyä paljon olennaisempaa on valottaa kielikunnan esihistoriallisen murteisiin jakautumisen prosessia. Esihistoriakaan ei tue perinteistä kielijakoa. Kielihistorialliset perusteet olla luokittelematta võron ja seton kieliä vanhaan tapaan eteläviron murteeksi ovat erittäin painavat. Kuitenkin myös suomen kielen päämurreerot ovat selvästi vanhempia ja niin ollen etymologille olennaisempia kuin esim. inkeröisen ja karjalan tai lyydin ja vepsän välinen eroavuus. Tämä näkyy valitussa sana-artikkelissakin: diftongillinen *aura* esiintyy vain suomen länsimurteissa, sen sijaan itämurteissa ja muualla itämerensuomessa on vallalla *tr*-yhtymän säilyttänyt muoto. Kummankaan sanakirjan artikkeli ei tee nähtäväksi tätä kahtiajakoa, joka monen lainan iänmäärittämyksen kannalta olisi kuitenkin olennainen kriteeri.

5. Murteitten ja kirjakielen statuseron häivyttäminen murreperiaatetta kaikkiin itämerensuomalaisiin kielimuotoihin soveltamalla päästäisi sanakirjan toimittajan valintatilanteista, joihin lähisukukieliperiaatteen noudattaminen johtaa. Hänen ei tarvitsisi päättää, edustaako murremuoto sanuetta hyvin tai antaako se rekonstruktion kannalta olennaista lisäinformaatiota, kun mukaan

pääsisivät vain hakusanasta äännesuhteiltaan poikkeavat muodot. Äännelaillisten suhteitten vallitessa pelkkä kielten listaaminen riittäisi.

Olennessa hyöty lähisukukieliperiaatteen hylkäämisestä murreperiaatteen tieltä olisi luettavuuden huima paraneminen. Muodoltaan ja merkitykseltään vähän tai ei ollenkaan vaihtelevien, ennen kaikkea itäisten ims. kielten muotojen toisto jäisi kokonaan pois. Esimerkkinä olkoon SSA:n hakusana *kirja* (1992, 369, ilman lähdeviitekappaletta):

kirja (yl.) ’kirja (Agr), kirje, asiakirja; murt. myös piirto, merkki, kuvio, koriste, kirjonta (mm. puvussa) / Buch, Brief, Dokument, Urkunde; Stickerei’, *kirjava* (Sorol 1625), *kirjata* (Gan 1786 ’kirjailla; kuvata’), *kirjasto* (Gottlund 1828), *kirje* (nykymerk. Kilpinen 1844), *kirjallinen* (Sorol 1614), *kirjallisuus* (Lönnr 1831), *kirjailija* (Kilpinen 1845), *kirjapaino* (Juteini 1810), *kirjanpainaja* (1673), *kirjain* (EIngman 1834; varhemmin mm. *kirjansauwa* eli *bokstawi*), *kirjoittaa* (ks. tätä) ~ ink *kirja* ’kirja, kirje; puvun koristekirjonta’ | ka *kirja* ’kirja, kirje, paperi; asiakirja, sopimus; kirjoitus, kirjoittaminen; kirjonta, kirjailu, kuvio, koriste, raita, juova, täplä’, *kirjo(i)*, *kirjikki* ’kirjava (lehmä)’, *kirja(h)ine* ’koristeltu, kirjottu’, *kirjava* ’kirjava, monivaiheinen’ | ly *kird’(e)* ’ompelu-, kudontakoriste, koristepiirros; kirjain’, *kird’aine* ’kirje’, *kird’au* ’kirjava’ | ve *kird’*, *kirj* ’kirja, kirje’, *kird’aine*, *kirjāne* ’kirje’ | va *tširja* ’kirje’, *tširjova* ’kirjava’ | vi *kiri* (g. *kirja*) ’kirje, kirjoitus; todistus; kirja; (neulottu) kuvio, koruompelus’, *kirjak*, *kirjes* ’kirjava lehmä’ | li *kēra* ’kuvio, koruompelus; kirjoitus, kirjain; kirje’, *kērabi* ’kirjava’. — Sm > lp *gir’je* ’kirja, kirje; täplä’; ims > latv *ķiris* ’valkojuovainen musta härkä’, ven *kirjāk* ’mereen palaava punertava kutulohi’

Artikkelissa esiintyvä samojen tai samankaltaisten ainesten toisto on vältettävissä vain luopumalla johdosten niputtelusta kielten – eli statukseltaan kieliksi määriteltyjen murteitten – mukaan. Tiiviiden ja luettavuuden saavuttaminen edellyttäisi siis kantasanan ja kunkin sen johdoksen esittelyn eriyttämistä.

Ensimmäiseen hakusanan jälkeiseen kappaleeseen voisi jättää vain kantasanan merkitykset² levikkeineen, kaikki johdokset taas esitellä perässä luettelona ennen varsinaista etymologiaa. Esitystapa olisi graafisesti miellyttävä eikä vaatisi lukijalta suurta asiantuntemusta:

***kirja** sm ink ka *kirja* ly ve *kird’* va *tširja* vi *kiri* li *kēra*

1 ’kirja’ sm (Agr) ink ka ve (va vi); 2 ’kirje’ va vi (sm ink ka ve li); ’asiakirja, sopimus’ (sm ka), ’paperi’ (ka), ’todistus’ (vi); 3 ’ompelu-,

2. Olen poistanut myös saksankieliset merkityksenselitykset, koska ne on SSA:ssa annettu vain suomen sanalle.

kudontakoriste, koristepiirros' ly, 'koriste, neulottu kuvio, koruom-
pelus' (sm ka vi li); kirjonta, kirjailu (mm. puvussa, sm ink ka);
4 'kuvio' li, 'piirto, merkki, (sm ka vi), raita, juova (ka), täplä (ka);
5 'kirjoitus, kirjoittaminen' (ka vi li); 'kirjain' (ly li)

kirja(h)ine ka 'koristeltu, kirjottu', ly ve 'kirje'

kirjailija sm (Kilpinen 1845)

kirjain sm (EIngman 1834; varhemmin mm. *kirjansauwa* eli *bokstawi*)

kirjak vi 'kirjava lehmä'

kirjallinen sm (Sorol 1614)

kirjallisuus sm (Lönnr 1831)

kirjasto sm (Gottlund 1828)

kirjata sm (Gan 1786 'kirjailla; kuvata')

kirjava sm (Sorol 1625) ka ly 'kirjava, (ka myös) monivaiheinen';
va *tširjova*; li *kēra*

kirje sm (nykymerk. Kilpinen 1844)

kirjes vi 'kirjava lehmä'

kirjikki ka 'kirjava (lehmä)'

kirjo(i) ka 'kirjava (lehmä)'

kirjoittaa (ks. tätä)

kirjanpainaja sm (1673)

kirjapaino sm (Juteini 1810)

Sm > saa *girji* 'kirja, kirje; täplä'; ims > lv *ķiris* 'valkojuovainen
musta härkä', ven *kirják* 'mereen palaava punertava kutulohi'

Hakusanaksi olen valinnut kantasuomen rekonstruktion, joka useimmissa tapauksissa on varsin yksinkertaisin säännöin johdettavissa etenkin suomesta mutta myös useimmista muista tytärkielistä. Toinen vaihtoehto olisi asettaa kielimuodot keskinäiseen hierarkiaan Setälän ja LÄGLOS:n tavoin: haku-
muotona olisi suomenkielinen asu, tämän puuttuessa vironkielinen ja viron-
kielisenkin puuttuessa jonkin pienemmän ims. kielen asu. Eri kielten sanoja on
kuitenkin hankalaa aakkostaa keskenään, onhan jo suomen ja viron aakkos-
järjestyksissäkin huomattavia eroja.

Kirjakielten muodot olen merkinnyt sana-artikkeliin omaksi kappaleek-
seen heti hakusanan alle ennen merkityskappaletta. Täydennettynä ”uusilla”
kirjakielillä³ luettelo näyttäisi tältä:

***kirja** me kv sm ink ka *kirja* ly ve *kird'* va *tširja* vi vs *kiri* li *kēra*

3. Merkitsen: me = meä, kv = kveeni, vs = eteläviro (võro-setto). Nämä kirjakielet olisi luonnollisesti otettava huomioon samassa määrin kuin muutkin kirjakielet kautta koko etymologisen hakuteoksen.

Kirjakielten muotoja ei kannata kokonaan jättää pois sana-artikkelista, koska suuria eroja aikaansaava säännönmukainenkin foneettinen vaihtelu olisi ilman niitä usein hankala hahmottaa. Johdinaineiksissa säännöllinen vaihtelu ei yleensä aiheuta yhtä suuria eroja, joten johdosten kohdalla eri kirjakielten muodot voisi kirjoittaa auki vain johdinaineuksen vaihdellessa epäsäännöllisesti kuten yllä *kirjava*-sanan kohdalla. Koko kielialueella esiintyvät ikivanhat johdokset olisi syytä ottaa omiksi hakusanoikseen. Kirjakielten muodot on myös helppo sijoittaa teoksen loppuun kielikohtaisiin hakemistoihin, joissa viitataan kyseiseen artikkeliin. Esimerkiksi liivin kielen hakemistossa olisi merkintä:

kēra *kirja
kērabi *kirja, johdos 9

Merkityksen selitykset ovat SSA:ssä ”luonteeltaan identifioivia”, eikä etymologisen sanakirjan voi olettaakaan kuvaavan sanan koko merkityskirjoa eri murteissa edes pääpiirteittäin. Olen yllä rypästännyt SSA:n antamat kantasanan erikieliset merkitykset viiteen ryhmään, mutta oikean sana-artikkelin laatiminen vaatisi luonnollisesti merkitysten täydentämistä eri lähteistä, lähinnä suurista murre-sanakirjoista. SSA:n mainitsemista merkityksistä voisi monet esittää lyhyemmin tai jättää kokonaan pois. Havainnollisinta olisi esittää merkitykset yleisyysjärjestyksessä, ensinnä merkitys, joka on useimmissa murteissa ensisijainen tai yksi päämerkityksistä. Nämä kielet olen maininnut merkityksen jälkeen ilman sulkua, perässä tulevat suluisia kielet, joissa sama merkitys esiintyy toissijaisena. Hakusanan **kirja* viimeinen, viides merkitys ei ole missään kielessä SSA:n perusteella ensisijainen.

Johdosten tiivis esittely on hankalaa. Jo Yrjö Toivonen SKES:n esipuheessa (1955, XI) piti ”johdannaisia” pääsyylisinä uumoilemaansa teoksen paisumiseen. SSA:n toimitus on karsinut ison osan SKES:ssa luetelluista johdoksista, muttei ole noudattanut johdonmukaisesti sinällänsä erinomaista toimitusperiaatettaan (1992, 13): ”Johdoksia – – ei ole otettu mukaan tyhjentävästi, vaan kussakin tapauksessa on rajoitettu edustavaan valikoimaan, nimenomaan sellaisiin käyttötapoihin, jotka saattavat olla sanan alkuperän kannalta valaisevia.” Missä mielessä *kirjasto*, *kirjapaino* ja *kirjanpainaja* ensiesiintymävuosineen ovat *kirjan* alkuperän kannalta valaisevia, semminkin kun *kirja* saman artikkelin mukaan esiintyy nykyisessä päämerkityksesssänsä jo Agricolalla? Tietenkin on hyvä, että tällaiset tiedot ovat jossain hakuteoksessa saatavilla, ja olkoon se SSA niin kauan kuin suomalaisilla ei ole omaa Klugen tai EWUng:in kaltaista kirjakielen kulttuurihistoriallista etymologista sanakirjaa. Tässä on toimitus kuitenkin ottanut itsellensä ylimääräisen tehtävän, jota se ei ole edes yrittänyt täyttää mitenkään johdonmukaisesti.

Johdosten esittely erillisinä luettelokappaleinaan ajaisi tavoittelemaan jonkinlaista kattavuutta johdettujen sanojen kuvauksessa – yhdyssanoja voi luonnollisesti ottaa mukaan vain poikkeustapauksissa. SSA:n käytäntö jättää epäselväksi, onko sama johdos voitu yhden kielen kohdalla ottaa tärkeänä mukaan

mutta muualla jättää huomiotta. Näin on käynyt ainakin suomen *kirjon* ja sen viron äännevastineen *kirju* 'kirjava' kohdalla. Niitten sijasta artikkelissa mainitaan vain karjalan vastine *kirjo(i)*. Useimpien hakusanojen kohdalla kattava johdosluettelo varmasti venyisi kattamaan yli puolet koko artikkelista. Kappaleitten selkeän erottelun ansiosta pituus tuskin häiritsee lukijaa, mutta kasvattaisi etymologiseen kokonaisuuteen tarvittavan työmäärän valtavaksi. Niinpä johdokset voisi itämerensuomen etymologisessa esityksessä rajata kansankieliin ja jättää ulkopuolelle varmat yleiskieleen sepitetyt uudissanat kuten *kirjalija*, *kirjain*, *kirjallinen*, *kirjallisuus*, *kirjasto* ja *kirjata*.

Lekseemissä *kirja* ei esiinny lainkaan epäsäännöllistä äännevaihtelua. Etymologin harmiksi useimmat sanat eivät käyttyä yhtä mukavasti. Huomatavasti vaihtelua sisältävä sana-artikkeli näyttää aika toisenlaiselta. Alla SSA:n *aarto* ilman varsinaista etymologiakappaletta:

aarto (Jusl 1745 *aartopuu*; Peräp Länsip Ink), *arto* (LönnrLis 1886 »Au»; paik. KarjKann), *arta* (Lönnr 1874), *aaro*, *aarre* (haj. murt.) 'lyhteiden, verkkojen yms. kuivausteline t. -riuku / Gestell zum Trocknen von Garben, Netzen u. ä.' ~ ka *aarto* 'verkon vapeet', *arto-*: ruisarto 'ruislyhteiden kuivausorsi' | ly *ard(o)* 'vaatteiden, nuotan tms. kuivausriuku' | ve *ard* (g. -on) id. | va *arto* 'viljan-kuivaushaasia' | vi murt. *ardu* 'varastosuoja asuinrakennuksen päässä, aitan t. navetan sivussa oleva koppi olkia ja heiniä varten', *artu* 'rakennuksen päässä oleva navetta'

Etymologille olennaista on erottaa epäsäännöllinen äännevaihtelu säännöllisestä. Niinpä etymologinen hakuteos voisi koota keskenään säännölliset variantit selvästi yhteen, rekonstruoida niille mahdolliset kantamuodot, esittää jokaisen omana levikkikappaleenaan ja niputtaa merkitykset kaikkien kappaleitten alle. Säännöllisistä äännesuhteista poikkeavia variantteja ei enää tarvitsisi poimia pakollisten sukukielitoistojen, johdosten ja semanttisesti merkityksellisten poikkeamien seasta. Äänneasultaan yksinkertaisin tai vanhakantaisin rekonstruktio muoto voisi aakkostaa artikkelin.

- ***arta** sm (Lönnr 1874)⁴ *arta*
- ***arto(j)** sm (paik. KarjKann) ka *arto* ly ve *ard*
- ***artojV/artokV** vi (murt.) *ardu*
- ***arttojV/arttokV** vi (murt.) *artu*
- ***aarto(j)** sm (Peräp Länsip Ink) va *aarto* ka *aarto*
- ***aaro(j)** sm (haj. murt.) *aaro*
- ***aartek** sm (haj. murt.) *aarre*

4. Näin ryhmiteltynä näyttäisi kantamuoto **arta* pohjautuvan pelkästään Lönnrotin sanakirjamainintaan. Koska SSA ei anna siitä murrelevikkiä, jää kuva, että kyseessä voi olla vain yksi lukuisista sanakirjan kirjoittajan konstruoimista sanoista.

1) 'lyhteiden, verkkojen yms. kuivausteline t. -riuku' sm ka ly,
'viljankuivaushaasia' ve; 2) 'varastosuoja t. navetta rakennuksen
päässä, aitan t. navetan sivussa oleva koppi olkia ja heiniä varten' vi

aartopuu sm (Jusl 1745)

ruisarto ka 'ruislyhteiden kuivausorsi'

6. Kuinka suuri osa SSA:n kuvaamasta sanastosta olisi suoraan otettavissa itämerensuomen etymologiseen hakuteokseen? Pitäisikö mukaan hyväksyä vain todistetusti tai luultavasti kantasuomalaiset sanat? Perinteisen kriteerin mukaan kantasuomeen on palautettava jokainen sana, jonka levikkiin kuuluu vähintään yksi eteläinen ja yksi pohjoinen itämerensuomalainen kieli. Jos kuitenkin tahdomme omaksua uudemman käsityksen etelävirosta ensimmäisenä kantasuomesta lohjenneena kielimuotona, emme tiukasti ottaen voisi varmasti pitää mitään etelävirosta puuttuvaa sanaa kantasuomalaisena ilman vanhaa lainaetymologiaa tai etäsukukielten vastineita. Eteläviron lohkeaminen perustuu kuitenkin yhteen ainoaan, hyvin pienessä osassa sanastoa vaikuttaneeseen ääntenmuutokseen **kt > ht*, joka tuskin millään lailla vaikeutti ymmärtämistä kielenpuhujain kesken tai vähensi leksikaalisten innovaatioitten leviämistä kielialueen sisällä. Teknisesti ottaen tämä muutos toki lopetti kantasuomen äännteellisen yhtenäisyyden kauden, mutta sanastohistoriallisesti vedenjakajaksi siitä ei ole. Sanoja liikkuu viron ja eteläviron välillä yhä, ja on erittäin vaikeata määrittellä missä vaiheessa historiaa tämä ilmeisesti katkeamaton vaikutusvirta on muuttunut kielensisäisestä levikin laajentumisesta lainautumiseksi.

Toinen vaihtoehto olisi hyväksyä hakuteokseen kaikki lekseemit, jotka eivät kuulu minkään itämerensuomalaisen kielen ”omaan sanastoon”. Tätä kaavaili Setälä suuressa Etymologicum magnum fenno-ugricum -hankkeessansa (Tanner 2007, 17). Suomen osalta olen laskenut tällaisten hakusanain määrän tiedostomuotoisesta SSA:stä varsin yksinkertaisella tavalla. Yhdistämällä SSA:n kolme osaa valmistin taulukon kaikista teoksen hakusanoista. Sana-artikkelien kokonaismääräksi sain 10709. Näistä 1391 artikkelia muodostuu pelkästä hakusanasta ja viittauksesta toiseen artikkeliin. Jäljelle jäävistä 9318 varsinaisesta sana-artikkelista 5823:ssa eli 62 %:ssa SSA mainitsee vastineita vähintään yhdestä lähisukukielestä. SSA:n lähisukukielten tiedoissa on havaittu puutteellisuuksia ainakin lyydin ja vepsän osalta (Nikkilä 1997, 299 ja 2002, 133–134), mutta tietojen täydentäminen muuttanee kokonaiskuvaa korkeintaan muutaman sanan kohdalla.

3495 hakusanaa eli 38 % koko SSA:n materiaalista on siis vailla ainutakaan lähisukukielten vastinetta. Tähän lukuun sisältyy kuitenkin yhdeksän varmaa tai epävarmaa lainaa virosta sekä yksi karjalasta. Osalla sanoista taas on ikivanha lainaetymologia tai ainakin epävarmoja vastineita etäsukukielissä. Näitä en ole vielä ennättänyt joukosta erotella, mutta aavistelen määrän noussevan vain muutamaan kymmeneen. Niinpä runsaalla kolmanneksella kaikista SSA:n hakusanoista levikki rajoittuu nyky-suomen rajojen sisäpuolelle eikä muu-

takaan näyttöä kantasuomeen ulottuvasta iästä ole. Pitäisikö nämä sanat jättää suomen kielen erilliskehityksen vaiheeseen kuuluvina käsittelemättä itämerensuomen etymologisessa esityksessä? Vai otammeko huomioon, että iso osa tästä sanastosta esiintyy ainakin joko meän tai kveenin kielessä eli uusien kriteerien mukaan edustuu lähisukukielissä – ehkä raja tulisi vetää näitten ja ”suomen kielen oman sanastokerrostuman” väliin?

Jo meän ja kveenin esimerkin valossa on selvää, että suomen sanaston jakaminen ”omaan” ja ”yhteiseen itämerensuomalaiseen” pelkän sukukielten vastineitten olemassaolon perusteella on mieletöntä. On harhaanjohtavaa puhua suomen kielen erilliskehityksestä esimodernina aikana, jolloin Suomea hallinnollisena tai valtiollisena kokonaisuutena ei ollut. Modernissa kansallisvaltiossa varttuneen ja koulunsa käyneen ihmisen – jopa etymologin – on vaikea hahmottaa, kuinka lyhytaikainen kehitysvaihe kielen historiassa on nykyinen kirja- ja yleiskielen dominanssi ja kuinka täysin toisenlaisten kielellisten vaikutusten alla sadat sukupolvet esivanhempiamme elivät. Suomen kielen erilliskehitykseksi voi kutsua ensinnäkin sitä tietoista kehittämistyötä, joka alkoi Agricolan *Abckiriasta* ja huipentui yleiskielen vakiintumiseen 1800-luvun kuluessa, toiseksi sitä kehitystä, joka suomen yleiskielessä sittemmin on tapahtunut. Tämän kehityksen tutkiminen kuuluu filologian alaan. Murteet ovat kehittyneet kantasuomesta aivan toisenlaisten lainalaisuuksien mukaan, juuri sellaisten, joita etymologi perinteisen fennougristiikan keinoin tutkii.

7. Kantasuomen pirstoutuminen vaikuttaa tapahtuneen vähittäin usean erisuuntaisen murteutumisprosessin tuloksena: vuoroin sisäiset innovaatiot ja ulkoiset vaikutteet ovat jakaneet kielialuetta, vuoroin murteitten väliset kontaktit ovat yhdistäneet uusia kokonaisuuksia. Kantasuomen jakautumiselle on turhaa yrittääkään määrittää mitään vuosilukua. En esitä tätä yleisenä totuutena kielikuntien jakautumisesta vaan päinvastoin itämerensuomen erityispiirtenä: esimerkiksi uralilaisen kantakielen jakautuminen vaikuttaa olleen varsin nopea ilmiö, jonka tuloksena on syntynyt yhdeksän selkeästi itsenäistä kielikuntaa.

Suurin murros itämerensuomen kehityksessä ei olekaan ollut kantasuomen pirstoutuminen murteiksi vaan modernisaatio, valtioitten muotoutuminen enemmistö- ja vähemmistökielineen sekä puhuttujen murteitten jääminen kirja- ja yleiskielten varjoon. Sukukielten keskinäistä ymmärtämystä hankaloittavat eniten sanastoerot, ja ehdottomasti suurin osa itämerensuomen sanastoeroista on syntynyt historiallisena aikana. Ristiretkistä lähtien vieraskielisten mahtien asettamat hallinnolliset rajat ovat synnyttäneet itämerensuomalaisten välille kielirajoja, ja kielten erot ovat sitä suuremmat mitä kauemmin raja on jakanut keskenään erikielisten mahtien hallintoalueita.

Merkittävin suomenkin sanastokerrostumien jakolinja kulkee esimodernin ja modernin leksikon välillä. Yhtäällä on puhuttujen murteitten kautta vähittäin levinnyt ja runsaasti varioiva sanasto, toisaalla kirja- ja yleiskielisen viestinnän sekä valtiollisen hallinnon kautta nopeasti levinnyt, laajalla kielialueella monoliittinen leksikko. SSA on yrittänyt silloittaa tätä leveätä sanahistoriallista

kuilua, mutta helpompaa olisi antaa sen jakaa kenttä kahtia esimodernin itämerensuomalaisen sanaston etymologisen hakuteoksen sekä modernin suomen kirjakielen kulttuurihistoriallis-etymologisen sanakirjan alueiksi.

En väitä, että jaolle itämerensuomen esimodernin ja suomen modernin sanaston välillä olisi mitään äänteellisiä kriteereitä. Erot ovat semantiikassa ja ennen kaikkea levikissä, joka esimodernissa leksikossa mitataan murteina, modernissa kielinä. Jakolinja ei myöskään sijaitse missään tarkassa tai itsessään selvässä kohdassa. Olisi ratkaistava, vedetäänkö se ehkä ristiretkiäikaan muinaisskandinaavisten ja novgorodinslaavilaisten kontaktien jälkeen, ennen muinaisruotsalaisia, alasaksalaisia ja muinaisvenäläisiä lainoja – vai johonkin myöhempään ajankohtaan. Olisi myös päätettävä, kuinka merkitä ne itämerensuomen hakusanat, joita ei voi palauttaa kantasuomeen asti.

Kielialue, jolle laaditaan vakiintuneen sanaston historialliseen kattavuuteen pyrkivä etymologinen kokonaisuus, olisi myös rajattava diakronisen kielentutkimuksen kriteerein – itämerensuomen tapauksessa siis ajallisesti, ei valtiollisten rajain mukaan. Seuraavan itämerensuomalaisen etymologian hakuteoksen ei pitäisi enää olla suomen vaan itämerensuomen etymologinen sanakirja, joka kokoaisi yhteen niin suomen, viron kuin pienempienkin lähisukukielten tutkimuksen tulokset. Toisaalta suomen kirjakieli, olkoonkin vielä alle puolen vuosituhannen ikäinen, ansaitsisi myös oman kulttuurihistoriallis-etymologisen sanakirjansa, jonka koostaminen kannattaisi aloittaa Häkkisen mainion hakuteoksen täydentämisestä. Tällaiseen teokseen sopiva sana-artikkeliformaatti olisi juuri NSES:n yleistajuinen pienoisessee. Itämerensuomalaisten vähemmistöjen nuorille kirjakielille ei ole tarkoituksenmukaista luoda vastaavanlaisia omia teoksia. Virossa taas tuskin kannattaa omin tutkijavoimin ponnistella kohti SSA:sta mukautettua versiota, virolainen lukijakunta kun odottaa omaa Häkkistään.

Lähteet

- Aikio, Ante 2001: Ajanmukainen sanakirja. – *Hiidenkivi* 3/2001, 34–35.
- EEW = Mägiste, Julius 1982–83: *Estnisches etymologisches Wörterbuch* I–XII. Moniste. Suomalais-Ugrilainen Seura, Helsinki.
- EWUng = *Etymologisches Wörterbuch des Ungarischen* 1. A–Kop, 2. Kor–Zs. Loránd Benkő (Hrsg.), Károly Gerstner, Antónia S. Hámori, Gábor Zaicz et. al. Budapest 1993–1995.
- Häkkinen, Kaisa 1993: Uutta ja vanhaa etymologian alalta. – *Sananjalka* 35, 165–170.
- InkS = Nirvi, R. E. (toim.) 1971: *Inkeröismurteiden sanakirja*. Lexica Societatis Fenno-Ugricae XVIII. Suomalais-Ugrilainen Seura, Helsinki.
- Kalima, Jalo 1936: *Itämerensuomalaisten kielten balttilaiset lainasanat*. Suomalaisen Kirjallisuuden Seuran Toimituksia 202. Suomalaisen Kirjallisuuden Seura, Helsinki.

- KKS = *Karjalan kielen sanakirja* I–VI. Lexica Societatis Fenno-Ugricae XVI:1–6. Kotimaisten kielten tutkimuskeskuksen julkaisuja 25. Kotimaisten kielten tutkimuskeskus – Suomalais-Ugrilainen Seura, Helsinki 1968–2005.
- Kluge, Friedrich 2002: *Etymologisches Wörterbuch der deutschen Sprache*. 23. durchgesehene und erweiterte Auflage. Bearbeitet von Elmar Seebold. Berlin, New York.
- LivWb = Kettunen, Lauri 1938: *Livisches Wörterbuch mit grammatischer Einleitung*. Lexica Societatis Fenno-Ugricae V. Suomalais-Ugrilainen Seura, Helsinki.
- LyMS = Kujola, Juho (toim.) 1944: *Lyydiläismurteiden sanakirja*. Lexica Societatis Fenno-Ugricae IX. Suomalais-Ugrilainen Seura, Helsinki.
- LÄGLOS = Kylstra, A. D. – Hahmo, Sirkka-Liisa – Hofstra, Tette – Nikkilä, Osmo 1991–1996: *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen*. I–II. Rodopi, Amsterdam – Atlanta, GA.
- Nikkilä, Osmo 1997: Suomen sanojen alkuperät *laahaamisesta pöveeliin*. – *Virittäjä* 101, 291–308.
- 2002: Suomen sanojen alkuperät päätökseen. – *Virittäjä* 106, 124–136.
- NSES = Häkkinen, Kaisa 2004: *Nyky-suomen etymologinen sanakirja*. WSOY, Juva.
- Saarikivi, Janne 2006: Vanhoja etymologioita uusissa kansissa. – *Virittäjä* 110, 111–123.
- SKES = *Suomen kielen etymologinen sanakirja* I–VII. Lexica Societatis Fenno-Ugricae XII:1–7. (Tutkimuslaitos ”Suomen suvun” julkaisuja III.) Suomalais-Ugrilainen Seura, Helsinki 1955–81.
- SSA = *Suomen sanojen alkuperä* I–III. Suomalaisen Kirjallisuuden Seuran Toimituksia 556. Kotimaisten kielten tutkimuskeskuksen julkaisuja 62. Suomalaisen Kirjallisuuden Seura, Helsinki 1992–2000.
- SVJ = Зайцева, М. И. – Муллонен, М. И. 1972: *Словарь венского языка*. Академия наук СССР. Карельский филиал. Институт языка, литературы и истории. Наука, Ленинград.
- Tanner, Satu 2007: *Salaperäinen Suomen suku. Tutkimuslaitos Suomen suvun historiikki*. Kotimaisten kielten tutkimuskeskuksen verkkojulkaisuja 3. Kotimaisten kielten tutkimuskeskus, Helsinki.
- Thomsen, Vilhelm 1890: *Beröringer mellem de finske og de baltiske (litauisk-lettiske) sprog: en sproghistorisk undersøgelse*. Blanco Lunos, København.
- VKKMS = Posti, Lauri 1980: *Vatjan kielen Kukkosen murteen sanakirja*. Toim. Seppo Suhonen. Lexica Societatis Fenno-Ugricae XIX. Kotimaisten kielten tutkimuskeskuksen julkaisuja 8. Suomalais-Ugrilainen Seura, Helsinki.
- VKMMS = Kettunen, Lauri 1986: *Vatjan kielen Mahun murteen sanasto*. Toim. J. Elomaa et al. Castrenianumin toimitteita 27. Helsingin yliopiston Castrenianumin laitokset – Suomalais-Ugrilainen Seura, Helsinki.

VKSR = Adler, Elna – Leppik, Merle et al. (toim.) 1990–2006: *Vadja keele sõnaraamat* 1–5. Eesti Teaduste Akadeemia – Keele ja Kirjanduse Instituut – Eesti Keele Instituut, Tallinn.

VMS = Pall, Valdek (toim.): *Väike murdesõnastik* I–II. Valgus, Tallinn 1982–89.

Santeri Junttila

santeri.junttila[a]helsinki.fi