

Indo-Iranian borrowings in Uralic

Critical overview of sound-substitutions and distribution criterion

SAMPSA HOLOPAINEN

Väitöksenalkajaisitelmä Helsingin yliopistossa 14. joulukuuta 2019

Uralilaisten kielten indoiranilaiset eli arjalaiset lainasanat on aihe, jolla on pitkä tutkimushistoria. Indoeurooppalaiseen kielikuntaan kuuluvia indoiranilaisia kieliä puhutaan nykyään Keski- ja Etelä-Aasiassa ja Kaukasiassa. Kielikunta jakaantuu ainakin kahteen päähaaraan, indoarjalaisiin kieliin, joita ovat monet Intiassa ja lähialueilla puhuttavat kielet, kuten hindi-urdu, bengali ja gujarati, sekä iranilaisiin kieliin, joihin kuuluu Iranissa ja lähialueilla puhuttavan persian eli farsin kielen lisäksi useita muita Keski-Aasian tai Lähi-idän kieliä, kuten kurdi, paštu, wahan, Georgiassa ja Venäjällä puhuttava osseetti ja monia muita. Mahdollisesti oman kielihaaran sa indoiranilaisten kielten joukossa muodostavat nuristanilaiskielet koillisessa Afganistanissa. Sekä indoarjalaisilla että iranilaisilla kielillä on pitkä, jopa tuhansia vuosia käsittävä kirjallinen historia, ja muinaisindoarjalaisella sanskritin kielellä sekä muinaisiranilaisilla avestan ja muinaispersian kielillä on keskeinen asema koko indoeurooppalaisen kielikunnan tutkimuksessa. Pitkä kirjallinen historia on myös myöhemmillä, myöhäisantiikin ja keskiajan aikana atestoiduilla keski-iranilaisilla kielimuodoilla, kuten keskipersialla, sogdilla ja baktrialla samoin kuin keski-indoarjalaisilla paalilla ja prakriteilla, ja ne ovat myös olleet tärkeitä filologisen ja lingvistisen tiedon lähteitä. Lisäksi kielitieteen tulosten sekä antiikin kirjallisten lähteiden perusteella tiedetään, että näiden kielten varhaiset puhuma-alueet ulottuivat paljon pohjoisemmaksi kuin nykyään.

Indoiranilaisen kielihaaran historiallisella tutkimuksella on vielä pidemmät perinteet kuin lainasanatutkimuksella, ja voidaankin paljoa liioittelematta todeta, että indoiranilaisten kielten historia on yksi parhaiten tutkittuja ja tunnettuja maailman kielten joukossa, kiitos juuri sen aseman, mikä erityisesti sanskritin kielellä on ollut

indoeurooppalaisten kielten historiallis-vertailevassa tutkimuksessa tutkimussuunnan alkuvaiheista alkaen. Pitkästä tutkimushistoriasta huolimatta indoiranilaisten kielten äännehistoriassa ja sanaston kehityksessä on kuitenkin monia epäselvyyksiä ja lisäselvitystä kaipaavia seikkoja, erityisesti kieliryhmän iranilaisen haarassa. Tällaisia ovat esimerkiksi monet etymologioimattomat sanat sekä eräät kantairanin äännejärjestelmään liittyvät kysymykset. Tämänkaltainen tilanne on tietenkin ominainen kaikille rekonstruoiduille kielille, ja kuvailemani kaltaiset ongelmat koskevat erityisesti juuri uralilaista kielikuntaa, josta tunnetusti puuttuvat samanlaiset vanhat kirjalliset lähteet kuin indoiranilaisilta kieliltä. Vaikka uralilaisten kielten äänne- ja sanastohistorian tutkimusta pidetään yleisesti ottaen korkealaatuisena, on selvää, että uralilaisen etymologian saralla on vielä paljon tehtävää.

Erityisesti lainasanatutkimuksen ja laajemmin kielikontaktien tutkimuksen alalla juuri uralistien voidaan katsoa tehneen pioneerityötä, sillä uralistiikan piirissä on jo erityisen varhain, Thomsenin (1869) ajoista asti, tunnistettu se todistusarvo, joka lainasanoilla voi olla kielihistoriaa tutkittaessa, ja lainasanatutkimuksen perinnettä on jatkanut erityisesti Jorma Koivulehto ja hänen oppilaansa. Erityisesti niissä kielissä, joissa ei ole kirjallisia lähteitä, lainasanat auttavat pääsemään käsiksi kielen kerrostumiin ja sitä kautta menneisyydestä kertoviin tietoihin, jotka muuten olisivat ulottumattomissamme.

On kuitenkin huomautettava, että juuri indoiranilaisten lainasanojen osalta tilanne ei ole yhtä hyvä kuin esimerkiksi balttilaisten tai germaanisten lainasanojen. Tähän on monia syitä, keskeisimpänä kenties se, että indoiranilaislainojen tutkimus on jäänyt vähemmälle huomiolle viime vuosikymmenten urauurtavassa lainasanatutkimuksessa. Toisaalta balttilais- ja germaanilainat ovat pääsääntöisesti nuorempia kuin indoiranilaiset lainasanat, mikä helpottaa niiden tutkimista. Osasyynä on varmaankin myös uralistiikkaa pitkään vaivannut itämerensuomikeskeinen lähtökohta: koska monet lainasanat esiintyvät vain itäisemmissä uralilaiskielissä, ovat ne olleet itämerensuomalaisten kielten tai suomen kielen lähtökohdista ponnistavien etymologien kannalta vähemmän kiinnostavia, ja koska monien itäisempien kielten (esim. permiläisten tai ugrilaisten kielten) äännehistoria on huonommin tunnettua, on selvää, että näissä kielissä esiintyvien lainasanojen tutkimus on haastavaa. Toisaalta myös vain itämerensuomeen rajoittuvat indoiranilaiset lainasanat ovat vasta hiljattain päätyneet tutkimuksen mielenkiinnon kohteiksi, kiitos erityisesti Jorma Koivulehdon (ks. esim. Koivulehto 1999, 2001) ja Asko Parpolan (ks. esim. Parpola 1999) tutkimusten.

Yllä kuvaamani tilanne voi selittää sen, että monet indoiranilaislainojen yksityiskohdat ovat pitkään olleet epäselviä tutkijayhteisölle ja väärinymmärrykset näistä lainakontakteista ovat varsin yleisiä, vaikka lainasanatutkimusta on aina pidetty keskeisenä uralilaisten kielten historiallis-vertailevassa tutkimuksessa. Harhaanjohtavaa tietoa löytyy käsikirjoista ja muista yleisesityksistä. Lainasanojen suhteelliseen kronologiaan, levikkiin ja ajoituksiin liittyy suurta sekaannusta, mitä on varmasti osaltaan aiheuttanut hankala terminologia. Termejä arjalainen ja indoiranilainen käytetään rinta rinnan, ja indoiranilaista voi terminä olla hankala erottaa suppeammasta iranilaisesta. Samoin indoarjalainen ja arjalainen menevät helposti sekaisin keskenään. Kaiken lisäksi sekaannusta on aiheuttanut vanhemmassa tutkimuksessa, erityisesti Rédein (1986) monografiassa, esitetty näkemys siitä, että uralilaisen kielikunnan vanhimmat

indoeurooppalaiset lainat ovat nimenomaan indoiranilaisesta kielihaarasta saatuja, ja termiä indoiranilainen tai arjalainen on tästä syystä käytetty lainasanoista, joilla ei voi olla mitään tekemistä koko kielihaaran kanssa.

Sen lisäksi, että lainasanatutkimus on tärkeä lähde uralilaisten kielten esihistorian tuntemukselle, on lainasanatutkimuksen tuloksilla luonnollisestikin arvoa myös indoiranilaisten kielten tutkijoille. Erityisesti äännehistorian kysymyksiin uralilainen lainasanaevidenssi tuo parhaimmillaan keskeistäkin evidenssiä, ja varhainen lainautuminen voi olla tärkeä seikka myös suppealevikkisen sanan tai johdoksen ikää punnittaessa. Myös tuntematonta alkuperää olevien iranilaisten ja indoiranilaisten sanojen varhainen lainautuminen kantauraliin tai sen tytärkieliin voi olla avuksi indoiranilaisten kielten sanaston eri-ikäisiä kerrostumia tutkittaessa.

Kuinka tätä nykyistä tutkimustilannetta tulisi lähestyä? Jo Thomsen (1869) painotti äännesubstituutioiden tärkeyttä esihistoriallisten lainasanakerrostumien tutkimuksessa. Lainanantajakielen foneemien substituutio antaa keinon erottaa toisistaan eri-ikäisiä lainasanakerrostumia, ja tämä metodi onkin hiottu huippuunsa itämerensuomalaisten kielten balttilaisten ja germaanisten lainasanojen tutkimuksessa, Thomsenista alkaen. Indoiranilaisiin lainoihin tällaista metodologiaa ei ole sovellettu täysmittaisesti. Hartmut Katzin (2003) postuumina julkaistu habilitaatiotyö on kyllä metodisesti oikeilla jäljillä, sillä se lähestyy tutkimusaineistoaan nimenomaan äännesubstituutioiden kautta, mutta itsekritiikin puute ja liian monimutkaisten substituutioiden olettaminen on tehnyt Katzin tuloksista osittain käyttökelvottomia. Katzin ohella erityisesti Jorma Koivulehto sovelsi tiukkoja lainasanatutkimuksen kriteerejä myös indoiranilaisiin lainasanoihin sen ohella, että hän teki urauurtavaa tutkimusta erityisesti germaanilainoista, ja toi ne lähemmäs thomsenilaisia standardeja, mutta Koivulehdon tulokset koskivat vain osaa lainasana-aineistosta. Hänen huomattavimpana tuloksenaan voidaan pitää kantairanilaisen lainasanakerrostuman esittämistä: Mayrhoferin (1989: 5–7) kantairanin äännehistorian rekonstruktiota seuraten Koivulehto (1999: 219–229) osoitti, että kantairanilaiset lainasanat voidaan erottaa vanhemmista, kantaindoiranilaisista lainoista kantaindoeuroopan palataalisten klusiilien *k:n, *g:n ja *g^h:n jatkajien eli kanta-arjan *c:n, *ǰ:n sekä *ǰ^h:n ja kantairanin *ts:n sekä *dz:n substituutioiden perusteella, sillä nämä affrikaatat substituutiitiin uralilaisissa kielissä eri tavoin (esimerkiksi *śata ← kanta-arjan *cata- mutta *mača 'mato' ← kantairanin *matsa-ka-). Kantairanilaislainat voidaan myös erottaa myöhemmistä iranilaislainoista, joissa affrikaattojen sijaan tavataan sibilantteja.

Heidermanns (2001) on verrannut äännesubstituutiota äännelakiin, joka tunnetusti on historiallis-vertailevan kielentutkimuksen tärkein työkalu. Vaikka Heidermanns (2001: 50–51) päätyy lainasubstituutioita ja äännelakeja vertaillaan lopulta toteamaan, ettei äännesubstituutiota voida pitää yhtä tarkkana kriteerinä kuin äännelakia, on se silti kaikista tärkein kriteeri nyt kuvaamassani lainasanojen tutkimustilanteessa. Äänne muutoksen säännöllisyydestä ja äännelakien tarkkuudesta huolimatta on todettava, että tutkimuksessa päästään varsin harvoin tilanteeseen, jossa vastaesimerkkejä tai poikkeuksia ei jäisi ollenkaan jäljelle – tällaisessa tilanteessa täytyy tietysti pyrkiä selittämään vastaesimerkit analogian tai ennestään tuntemattomien äännekehitysten kautta, mutta jos se ei ole mahdollista, täytyy päätelmät perustaa nimen-

omaan varmojen esimerkkien varaan. Sama toimii tietysti lainasanatutkimuksessakin, jossa aivan samankaltaiseen tarkkuuteen ei ole mahdollista pyrkiä.

Toinen keskeinen työkalu lainasanatutkimuksessa on tietenkin lainasanojen levikin tutkimus. On kuitenkin käynyt varsin selväksi, että varhaisten lainasanojen tutkimuksessa levikillä on vain toissijainen rooli sanojen ikäämisessä, kuten Häkkinen (1983: 224–235) ja Koivulehto (1999: 208–209) ovat painottaneet. Juuri indoiranilaislajien tutkimuksessa on levikkikriteeriä seurattu liian orjallisesti, mikä on johtanut vinoutuneeseen kuvaan joidenkin uralilaisen kielikunnan haarojen, kuten permiläiskielten tai unkarin, lainasanojen tutkimuksessa. Esimerkiksi näillä kielihaaroilla on selvästi eri-ikäisiä lainakerrostumia, joiden ajoittaminen ei mitenkään onnistu pelkkää levikkiä tuijottamalla.

Lainasanojen levikkiongelmia kytkeytyy uralilaisen kielikunnan taksonomian ratkaisemattomiin kysymyksiin. Koska uralilaisen kielikunnan sisäistä rakennetta ja kieliryhmien keskinäistä suhdetta ei tunneta riittävän hyvin (Salminen 2002), ei voida antaa liikaa painoarvoa eri kielihaarojen jakamille lainasanoille. Varhaisemmassa tutkimuksessa on saanut liian paljon huomiota ajatus, että monet lainat ovat lainautuneet heikosti perusteltuihin välikantakieliin kuten suomalais-volgalaiseen kantakieleen tai kantaugriin. Samaan aikaan esimerkiksi ugrilaisen kielihaaran olemassaoloa perustellaan kehäpäätelmän omaisesti juuri yhteissanastolla, kuten jaetun lainasanakerrostuman olemassaololla. Tosiasiassa esimerkiksi monet levikiltään ugrilaiset lainasanat (esim. kantaindoiranin **saras*- ’järvi’ → kantauralin tai kantaugrin **sara* > unkarin *ár* ’virta’) on lainattu kielentasolle, joka on äänteellisesti sama kuin vanhemmilla kantauraliin saaduilla lainasanoilla. Monesti ei ole mitään selvää keinoa määrittellä, onko jokin sana oikeasti rajoittunut alkuperäiseltä levikiltään juuri niihin kieliin, joissa se nykyään esiintyy, vai ovatko sen vastineet vain kadonneet muualta. On tosiaan yleisesti tunnettua, että sanoja katoaa (ks. esim. Campbell 1999 [1998]: 180–183, 263–265), ja erityisesti tämä koskee niin sanottuja kulttuurisanoja, joita monien lainasanojen voidaan katsoa olevan.

Vielä olisi syytä mainita yksi levikkiä koskeva ongelma: joissakin tapauksissa on vaikeaa erottaa aitoja, jo kantaakieleen saatuja vanhoja lainoja myöhemmin rinnakkain lainatuista sanoista. Rinnakkaislajien ongelma on keskeinen kysymys uralilaisessa lainasanatutkimuksessa, mutta juuri indoiranilaisten lainasanojen osalta ongelmaa ei ole tutkittu riittävän laajasti, vaikka joitakin esimerkkejä on tunnettu jo vanhastaan (Korenchy 1972: 77). Esimerkiksi ’kultaa’ tarkoittavat sanat ugrilaisissa (esim. unkarin *arany*) ja niin sanotuissa volgalaisissa kielissä (esim. ersämordvan *sirñe*) ovat lainoja iranilaisesta sanasta **zaranya*- ’kulta’.

Jos kokonaisia lainasanakerrostumia koskevia tutkimustuloksia haluaa esittää, on selvää että uskottavat tulokset voivat perustua vain kriittisesti tarkasteltuun ja rajattuun tutkimusaineistoon. Yksi oman työni tavoitteista onkin ollut tutkimusaineiston tarkempi rajaaminen ja epävarmojen ja epäuskottavien etymologioiden karsiminen.

Esihistoriallisten lainasanakerrostumien tutkija joutuu monien haasteiden eteen: pitää pystyä operoimaan sekä lainanantaja- että -saajakielen tai -kielikunnan äänne- laeilla ja äännehistorialla, mikä voi erityisesti kokonaisia kielikuntia koskevassa tutkimuksessa olla varsin haastavaa. Lisäksi täytyy olla selvillä äännesubstituutioista, ja

tietysti myös semanttisten kehitysten mahdollisuuksia on punnittava paraleeleja etsien. Kaikki tämä tekee lainasanojen tutkimisesta kiinnostavan palapelin, joka hajoaa helposti yksittäisenkin palan puuttuessa. Haasteista huolimatta tutkimusta voi viedä eteenpäin vain tarjoamalla uusia synteesejä tasaisin väliajoin.

Lähteet

- CAMPBELL, LYLE 1999 [1998]: *Historical linguistics. An introduction*. Ensimmäinen MIT Pressin julkaisema laitos. Cambridge, Massachusetts: The MIT Press.
- HEIDERMANNS, FRANK 2001: Lautgesetz versus Lautersatz. Zur Mehrfachsubstitution bei lexikalischer Interferenz. – Heiner Eichner, Peter Arnold Mumm, Oswald Panagl & Eberhard Winkler (toim.), *Fremd und eigen. Untersuchungen zu Grammatik und Wortschatz des Uralischen und des Indogermanischen; in memoriam Hartmut Katz s. 39–53*. Wien: Praesens.
- HÄKKINEN, KAISA 1983: *Suomen kielen vanhimmasta sanastosta ja sen tutkimisesta*. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 17. Turku: Turun yliopisto.
- KATZ, HARTMUT 2003: *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*. Toim. Paul Widmer, Anna Widmer & Gerson Klumpp. Indogermanische Bibliothek. 3. Reihe: Untersuchungen. Heidelberg: Universitätsverlag Carl Winter.
- KOIVULEHTO, JORMA 1999: Varhaiset indoeurooppalaiskontaktit. Aika ja paikka lainasanojen valossa. – Paul Fogelberg (toim.), *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan s. 207–236*. Bidrag till kännedom av Finlands natur och folk 153. Helsinki: Suomen Tiedeseura.
- 2001: The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (toim.), *Early contacts between Uralic and Indo-European*. Suomalais-Ugrilaisen Seuran Toimituksia 242. Helsinki: Suomalais-Ugrilainen Seura.
- KORENCHY, ÉVA 1972: *Iranische Lehnwörter in den oburgischen Sprachen*. Budapest: Akadémiai Kiadó.
- MAYRHOFER, MANFRED 1989: Vorgeschichte der iranischen Sprachen; Uriranisch. – Rüdiger Schmitt (toim.), *Compendium linguarum iranicarum s. 4–24*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- PARPOLA, ASKO 1999: Varhaisten indoeurooppalaiskontaktien ajoitus ja paikannus kielellisen ja arkeologisen aineiston perusteella. – Paul Fogelberg (toim.), *Pohjan poluilla. Suomalais-ten juuret nykytutkimuksen mukaan s. 179–206*. Bidrag till kännedom av Finlands natur och folk 153. Helsinki: Suomen Tiedeseura.
- RÉDEI, KÁROLY 1986: *Zu den indogermanisch-uralischen Sprachkontakten*. Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, Heft 16. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- 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*. Päätoimittaja A.E. Кибрик. Москва: Издательство Московского университета.
- THOMSEN, VILHELM 1869: *Den gotiske Sprogklasses Indflydelse paa de finske*. København: Gyldendalske boghandel.

Sampsa Holopainen: *Indo-Iranian borrowings in Uralic: critical overview of sound-substitutions and distribution criterion*. Helsinki: Helsingin yliopisto 2019. Kirja on luetta-
vissa osoitteessa <http://urn.fi/URN:ISBN:978-951-51-5729-4>.

Kirjoittajan yhteystiedot:
etunimi.sukunimi@helsinki.fi