

Suomen sanaston dynamiikkaa

ANTERO NIEMIKORPI *Suomen kielen sanaston dynamiikkaa*. Acta Wasaensia No 26; Kielitiede 2. Universitas Wasaensis. Vaasa 1991. 414 s. + liitteet 1, 2.

Dynamiikka juontuu kreikan sanasta *dýnamis*, joka merkitsee voimaa. Väitöskirjansa komealta kalskahtavan nimen valintaa Antero Niemikorpi perustelee sillä, että työn perimmäisenä tarkoituksena on kieltä ohjailevien voimien ja niiden sanastollisten heijastumien tutkiminen (s. 9). Terminologisena esikuvana tässä yhteydessä on ollut Vilho Setälä, joka on jo aiemmin puhunut kielen dynamiikasta suomen sanojen kvantitatiivista rakennetta selvittäessään.

Pelkän nimen perusteella erehtyy kuitenkin odottamaan liikoja. Niemikorpi on tutkinut lähinnä erityyppisissä teksteissä esiintyvien lekseemien, sananmuotojen ja saneiden pituuksia ja esiintymisfrekvenssejä sekä näiden korreloitumista mm. lauseiden ja virkkeiden pituuksien, sanaluokkien sekä lauseenjäsenfunktioiden kanssa. Sen sijaan sanaston tutkimuksen piiriin olennaisesti kuuluva sananmuodostus sekä lekseemien välisten morfologisten suhteiden samoin kuin morfosyntaktisen kompleksisuuden selvittäminen on jäänyt kokonaan käsitteilyn ulkopuolelle.

Väitöskirjassa on kahdeksan lukua. Ensin tulee lyhyt (5 s.), johdannonomainen katsaus ongelmanasetteluun sekä muutamiin teknisiin yksityiskohtiin. Toisessa luvussa esitellään leksikkoa ja

sen perusyksiköitä. Kolmannessa tekijä luo katsauksen kielen kuvauksessa käytettyihin tunnuslukuihin sekä sanantajuutta ja -pituutta koskeviin hypoteeseihin (erit. s. 55–57). Seuraavassa luvussa (IV) hän esittelee tutkimusaineistoaan ja yleensä korpuksen valintaan liittyviä ongelmia. Viides luku on omistettu saneiden, sananmuotojen ja sanojen pituuksille. Kuudennessa luvussa tarkastellaan, miten nämä korreloivat lauseiden ja virkkeiden pituuksien sekä eri sanaluokkien sanemäärien kanssa. Lisäksi tarkastellaan ilmausten pituuksien sekä sanaluokkien rikkautta ja tekstin luettavuutta kuvaavien indeksien välisiä suhteita. Seitsemännessä luvussa tutkitaan, miten eripituiset sanat jakautuvat eri lauseasemiin. Kahdeksas valaisee sanantajuuteen liittyviä ominaisuuksia, mm. eri sanaluokkien ja eri viestinlajien sanojen esiintymistiheyksiä.

Väitöskirjan aineistona on käytetty ns. Oulun korpusta, jonka luomisessa Niemikorpi on alusta pitäen ollut mukana ja jonka pohjalta on jo aiemmin julkaistu huomattavan runsaasti tutkimuksia sekä laadittu eritasoisia oppinäytetöitä. Lähdeluettelossa on kolmisenkymmentä kyseiseen korpukseen tai sen osa-aineistoihin perustuvaa tutkimusta. Niemikorven omia töitä on lähdeluettelossa kymmenen nimikettä eli enemmän kuin kenenkään muun tutkijan töitä. Näistä uusien on Suomen kielen sanaston frekvenssianalyysia -nimisenä ilmestynyt nide (155 s.), joka sisältää väitöskirja-aineistoon kuuluvia taulukoita (Vaasan korkeakoulun julkaisuja. Tutkimuksia No 150. Kielitiede 18. Vaasa 1990). Tä-

män työn läheinen yhteys väitöskirjaan käy ilmi ainoastaan lähdeviitteestä (Niemikorpi 1990, 9), jossa väitöskirja on optimistisesti ajoitettu vuoteen 1990. Ilmestyminen lykkäytyi kuitenkin vuoden 1991 puolelle, ja tästä viivästyksestä koitui se ilo, että Niemikorven väitöskirjasta tuli ensimmäinen Vaasan yliopistossa tarkastettu väitöskirja.

Useissa tapauksissa aikaisempien, sekä omien että muiden laatimien tutkimusten tuloksia on ollut mahdollista hyödyntää luontevasti väitöskirjaan sisältyvien jatkotutkimusten pohjana (esimerkiksi sanaluokkien korrelaatioita eri yksiköiden pituuksien kanssa käsitteleviä tutkimuksia s. 198 alk. tai lauseen rakennetta koskevia töitä s. 283 alk.). Kuitenkin eräät aiemmat tutkimukset ovat osittain päällekkäisiä Niemikorven väitöskirjan kanssa, erityisesti Pauli Saukosen tekstilajeja ja tyylin osatekijöitä koskevat tutkimukset sekä Pauli Saukosen johdolla laadittu Suomen kielen taajuussanasto. Tällaisissa tapauksissa on kokonaan uusia ja yllättäviä tuloksia ymmärrettävästi vaikea tarjota.

Ennalta valmiin ja moniin muihinkin tarkoituksiin laaditun aineiston käyttäminen on joissakin suhteissa vaikuttanut epäedullisesti käsillä olevaan tutkimukseen. Esimerkiksi lekseemien ja sanaluokkien rajaamista koskevat ratkaisut on täytynyt suorittaa alkuperäisen korpuksen laadinnassa noudatettujen periaatteiden mukaan, mikä käytännössä merkitsee näitä koskevien lingvististen ongelmien lähes täydellistä sivuuttamista tässä työssä. Tämä näkyy esimerkiksi siten, että keskeisiä peruskäsitteitä on jätetty määrittelemättä. Mm. sivulla 19 tekijä tyytyy toteamaan, että »— — sanakäsitettä lienee mahdotonta määritellä yleispätevästi mistään lingvistiksestä oppirakennelmasta käsin». Käytännössä jokaisella lingvistisellä oppirakennelmalla on omat aksioomansa, oma terminologiansa, omat tavoitteensa ja oma sisäinen logiikkansa. Yhden rakennelman osia ei tietenkään voi muutta mutkitta siirtää toiseen tai pitää yleispätevinä,

mutta tämä ei suinkaan merkitse sitä, ettei sanaa voisi määritellä tietyn teorian puitteissa. Niemikorpi ei kuitenkaan esitä yhtään määritelmää. Myös varsinainen leksikologinen lähdekirjallisuus on jätetty huomiotta.

Muitakaan esimerkkejä lingvistiksesti merkitsevien ongelmien sivuuttamisesta ei ole vaikea löytää. Esimerkiksi eri sanaluokkiin kuuluvien lekseemien pituuksia tutkittaessa todetaan alle 4-grafeemisiä verbejä olevan vain kaksi, mutta mitään huomiota ei ole kiinnitetty näiden lemmatun muodon poikkeuksellisuuteen (EI ja LIE; yleensä verbeistä on käytetty I infinitiivin lyhyempää muotoa). Merkityksen kuvaamisessa on tyydytty toisaalta muutamiiin peruskurssitason merkityspiirteisiin (komponenttianalyysin kuvausvoimaa ja pätevyyttä systeeminä ei pohdita) ja Nykysuomen sanakirjan merkitysryhmiin ja näiden määrään yksityistä lekseemiä kohti. Kysymys kuvaustapojen suhteesta jää kuitenkin avoimeksi, sillä vaikka merkitysryhmien määrään perustuva hierarkia ei edes muutamien esillä olevien esimerkkien valossa näytä korreloivan piirrehierarkian kanssa (vrt. esim. NAINEN, PALVELIJA, PROFEETTA s. 342—343), asiaa ei mitenkään kommentoida. Merkitysryhmiä koskeva analyysi perustuu Raimo Jusilan tutkimuksiin, mutta näihin viitataan vasta myöhemmin.

Useissa kohdin laskelmat ovat jonkin verran epätarkkoja aivan teknisistä syistä. Esimerkiksi radion kieltä koskevissa näytteissä epäselviksi jääneet erisnimet on korvattu x-symboleilla, jotka toisaalla on laskettu yhden merkin pituisiksi sanoiksi. Myös esimerkiksi matemaattiset kaavat on korvattu kirjainsymboleilla. Lukusanat lienee eri kohdissa laskettu eri tavoin sen mukaan, onko ne tekstiin merkitty kirjaimin vai numeroin. Erisnimiä ja yleisnimiä ei ole voitu erottaa toisistaan, koska Oulun korpuksessa on kauttaaltaan käytetty kapiteelikirjaimia. Lyhenteet on laskettu mukaan sanojen joukkoon sellaisinaan, mutta pisteettöminä. Myös verbien käsittely on

epäyhtenäistä siinä mielessä, että useissa laskelmissa on otettu verbinä huomioon vain aktuaaliset finiittimuotoiset verbisaneet. Näin ollen partisiippien osuus finiittisten liittomuotojen osana jää hämäräksi, eikä sitä esim. lauseenjäsenten positioita selvitetessä ole mitenkään otettu huomioon. Kaikki työssä esitetyt lukumäärät ja laskelmat eivät myöskään perustu täsmälleen samoihin ja samalla tavoin analysoituihin osakorpuksiin. Esimerkiksi osa sivulla 60 esitetyistä saneiden, sananmuotojen ja sanojen suhteita kuvaavista luvuista ei vastaa edellisillä sivuilla esitettyjä absoluuttisia määriä, ja sama pätee kauttaaltaan sivun 73 asetelmaan.

Niemikorpi korostaa työssään useassa kohdin kielen kumulatiivisia, samaan suuntaan vaikuttavia ominaisuuksia. Toisaalta kvantitatiivisen kielentutkimuksen nimekkäimpien edustajien, esim. Gabriel Altmannin ja hänen koulukuntansa mukaan kielen dynamiikan kannalta olennaista on vastavoimien välinen vuorovaikutus (Altmann tutkijaryhmineen esitteli näkemyksiään mm. Oulussa kansainvälisessä kielen synergetiikkaa käsitelleessä symposiumissa lokakuussa 1990). Niemikorpi on valinnut tutkimuskohteensa niin, että tämä aspekti jää selvästi taustalle. Syynä lienee osittain se, että samansuuntaiset voimat ja niiden vaikutus on helpompi havaita ja laskennallisesti osoittaa yksinkertaisten korrelaatioiden avulla. Toisaalta pelkien grafemaattisten pituuksien ja esiintymistiheyksien selvittämiseen keskittyvä näkökulma on jo sinänsä liian kapea koko sanaston rakenteen ja/tai käytön dynamiikkaa koskevien taustatekijöiden selvittämiseksi.

Tutkimuksessa käytetty terminologia on epäyhtenäistä mm. siksi, että lähdekirjallisuudessa käytetyt termit on sellaisenaan siirretty käsillä olevaan tutkimukseen niiden keskinäistä suhdetta tarkemmin selvittämättä (ks. esim. *kanta*, *juuri*, *vartalo* ja *perussana* s. 42, 55, 90, 347, 348, 349). Finiittimuotoisen verbin asemesta puhutaan välillä modusmuo-

toisesta verbistä (s. 266, 280). Hankalia sekaannuksia saattaa työn lukijoille aiheuttaa se, että termiä *frekvenssi* (tai ilmeisesti tästä saatua lyhennettä F) on käytetty toisaalta sekä esiintymisfrekvenssin (-tiheyden) että tiettyyn luokkaan kuuluvien tapausten määrän merkityksessä (esim. s. 53, 66, 150). Tähän on tietysti syynä se, että käytäntö tilastotieteessä ja lingvistiikassa on erilainen, mutta juuri tästä syystä olisi pitänyt valita selväsanaisesti määritelty ja johdonmukainen menettelytapa: työn potentiaalisissa lukijoissakin on sekä kieli-tieteilijöitä että статистиikan edustajia.

Työssä on muutenkin tavallista runsaammin terminologista epätarkkuutta ja epäjohdonmukaisuutta. Esimerkiksi s. 19 on käytetty termiä *primitiivitermi* ikään kuin se olisi kirjoittajan oma luomus; tarkoitettua käsitettä paremmin vastaava termi *primitiivikäsite* on kuitenkin käytössä useissa kielitieteen perusteoksissa (esim. Lyonsin *Semantics-teoksessa* ja Karlssonin *Johdatuksessa* yleiseen kielitieteeseen). Zipfin *economy* ja *equilibrium* eivät selvästikään merkitse samaa, vaikka Niemikorpi näin sanookin (ks. sitaattia s. 32). Termiä *äänne-muutos* on käytetty äännevaihtelun asemesta (s. 143), paikallissijoista esimerkkeinä on mainittu muotoja, jotka eivät ainakaan nykykielessä eivätkä välttämättä edes historiallisesti ole tällaisia (s. 172). Termiä »plussana» on käytetty sivulta 353 alkaen, vaikka määrittely tulee vasta s. 376. Tietokirjallisuuden suosikkiverbejä luonnehditaan ylimalkaisesti passiivisiksi ja refleksiivisiksi, ja esimerkkeinä on mainittu mm. VOIDA, MERKITÄ ja VAIKUTAA (s. 366).

Suomen, viron ja unkarin sanantaa-juuksien korrelaatiota selvitellessään (330–334) Niemikorpi arvelee suomen ja viron samankaltaisuuden johtuvan näiden läheisemmästä sukulaisuudesta kajoamatta mitenkään niihin aktuaalisiin rakenteellisiin ominaisuuksiin, jotka eroja itämerensuomen ja unkarin välillä olisivat saattaneet synnyttää (unkari on mm. artikkelikieli!). Sivulla 348 jää epäi-

lyttämään, mitä kirjoittaja tarkoittaa »vanhimpien sanojen (kantojen)» kulu-neisuudella. Muiden kuin pronominivartaloiden osalta sillä yleensä tarkoitetaan lyhenemistä ja fonotaktista yksinkertaistumista, mutta pronominivartaloissa, jotka Niemikorpi omassa esityksessään nostaa etualalle, ei edes voi olla puhetta varsinaisesta lyhenemisestä. Niin pitkälti kuin pronominivartaloita voidaan tutkia, ne näyttävät olevan yksitavuisia. On kysymys erilaisten koaffiksien liittymisestä ja sulautumisesta vartaloon sekä ylipäänsä epäsäännöllisestä äännehistoriallisesta kehityksestä.

Työssä sovellettu yksinkertainen, mekaaninen käsittelytapa sopii hyvin numeerisen aineiston käsittelyyn. Laskelmien yksityiskohtainen esitleminen ja läpikäyminen, mikä tekee työstä varsin raskaslukuisen, onkin sinänsä korrektaa ja tarpeellista, koska tutkimusta jatkossa todennäköisimmin luetaan vain fragmentaarisesti. Tähän kirjoittajakin lienee varautunut mm. siten, että kuvioiden ja taulukoiden tekstit on kautta koko teoksen pantu näkyviin sekä suomeksi että englanniksi, vaikka muu teksti lyhyttä yhteenvetoa lukuun ottamatta onkin suomenkielistä.

Työn jäsenitys on selkeä, joten käsitellyt asiakokonaisuudet hahmottuvat lukijalle jo sisällysluettelon ja lukujen numeroinnin perusteella. Tekstiä lukiessaan ja sisäisissä viitteissä mainittuja kohtia haikiessaan tosin toivoo, että alalukujen numeroinnissa olisi koko ajan kuljetettu mukana myös pääluvun numeroa (tämä täytyy katsoa erikseen sivun alareunasta). Käsikirjan luonnetta silmällä pitäen on puolusteltavissa myös se, ettei tekijä ole pyrkinyt valikoimaan tai painottamaan käsiteltyjä osakokonaisuuksia. Hän on julkaissut lähes kaiken tutkimuksen yhteydessä syntyneen materiaalin joko itse väitöskirjassa tai sen sivutuotteena syntyneessä Suomen kielen sanaston frekvenssianalyysia -niteessä (ks. edellä).

Yhtä mekaaninen käsittelytapa ei useinkaan sovi lingvististen ongelmien ratkaisuun. Sivulla 30 tekijä väittää, että

sanamuodot voidaan lajitella ja ryhmitellä korpuksesta täysin mekaanisesti. Saneiden osalta niin voidaan tehdä, mutta sanamuotojen lajittelemisessa on mahdotonta päästä täydelliseen tulokseen sanamuotojen homonymian takia. Parhaat tällä hetkellä käytössä olevat menetelmät antavat homonyymisille muodoille kaikki mahdolliset tulkinnat mutta eivät pysty sataprosenttisen luotettavasti valikoimaan niistä aktuaalisesti oikeata. Samalla sivulla Niemikorpi luettelee parin viime vuosikymmenen aikana kehitettyjä automaattisia analyysimenetelmiä, muttei mainitse erikseen Kimmo Koskenniemen maailmankuulua kaksitasomallia. Liian yksioikoinen on sivun 325 väite, jonka mukaan lyhyt lause ei muodostimiensa vähäisen määrän vuoksi teoreettisestikaan ottaen voi sisältää kovin runsaasti uutta tietoa. Ajateltakoon valaisevina esimerkkeinä vaikkapa ns. lauseenvastikkeita (*Lähdettyämmekin satoi* tms.). Myös konstituenttien sisäisellä rakenteella on merkitystä!

Tärkeä osa tutkimustulosten esittelystä ja havainnollistamisesta on erilaisilla kuvioilla ja taulukoilla. Näitä käytettäessä on kuitenkin pidettävä mielessä, millä ehdoin ne on laadittu. Esimerkiksi luvussa VI, jossa pohditaan sananpituutta tekstien sanastollis-kieliopillisten piirteiden heijastajana, on lähdetty siitä perusajatukselta, että tietyt asiat, esim. lauseiden keskipituus saneina, korreloivat tietyllä tavalla esim. tiettyyn sanaluokkaan kuuluvien saneiden prosentuaalisen osuuden kanssa. Havaitut poikkeavuudet ovat silloin poikkeuksia nimenomaan tästä oletetusta riippuvuussuhteesta.

Luvussa VI on kuvattu eri tekstiluokkien sanastollis-kieliopillisia piirteitä ja niiden välisiä korrelaatioita mm. regressiosuorien avulla. Kuvioiden havainnollisuutta olisi lisännyt se, että tekstiluokat olisi jotenkin erotettu toisistaan. Nyt kaikki luokat on merkitty täsmälleen samanlaisella ympyrällä, eikä tekstissäkään ole annettu mitään osviittaa siitä, miten luokat olisivat identifioi-

tavissa. Tämä olisi kuitenkin tarpeen, sillä poikkeavien luokkien keskinäiset erot saattavat olla varsin suuret.

Sivulla 276 on mielenkiintoinen ryhmittelyanalyysi, joka suurin piirtein odotuksenmukaisella tavalla (vrt. s. 282; myös Saukkonen on tutkinut vastaavia suhteita) kuvaa eri tekstilajien keskinäisiä suhteita 15 eri muuttujan avulla mitattuna (s. 275). Luokkia koskevat selitystekstit eivät kuitenkaan ole oikeilla kohdin, vaikkakin sinänsä oikeassa järjestyksessä. Eräissä muissakin kohdissa tekniset yksityiskohdat jättävät toivomisen varaa. Esimerkiksi sivujen 59, 81 ja 82 kuvioissa eivät mustat ja rasteroiduiksi tarkoitetut pallot eroa juurikaan toisistaan, joten lukijan omaksi tehtäväksi jää päätellä, mitä kuviot lopulta kertovat.

Kuvioista ja taulukoista löytyy joitakin selviä virheitä (esim. s. 197 riveillä AIK UUT ja MAA; s. 245 AIK ASI, AIK PAK, AIK HAR ja NNÄY; on myös epätarkoituksenmukaista laskea eri suuntiin menevät poikkeukset suoraan yhteen). Sivujen 168 ja 169 taulukoissa vasen sarake kuvaa esiintymien määrää, vaikka selitysteksti viittaa esiintymien pituuteen grafeemeina (keskimääräiset pituudet on ilmoitettu varsinaisessa taulukossa kunkin sanaluokan osalta erikseen). Lisäksi samojen taulukoiden englanninkielisissä selitysteksteissä frekvenssit on ilmoitettu väärin. Sivujen 227, 228 ja 230 kuvioissa puhutaan pystyakselia koskevassa tekstissä pelkästään saneista, kun kysymys on ilmeisesti konjunktioista; kuvioiden mukaan lyhytsanaisimmissa teksteissä yli puolet saneista olisi kopulatiivikonjunktioita.

Yleisen väitöskirjakäytännön vastaisesti Niemikorpi paljastaa erään tärkeänä pitämänsä tutkimustuloksen työnsä ensimmäisellä tekstisivulla (s. 8; työ alkaa vasemmalta sivulta!): Suomenkielisen lauseen viimeinen sana on keskimäärin lauseen pisin sana. (Tosin hän myöhemmin (s. 174) toteaa, ettei sananpituutta tai muidenkaan ilmausten pituutta ole yleensä tarpeen sisällyttää tavan-

omaiseen kielenkuvaukseen.) Viskurilakia käsittelevässä luvussa VII tämä osoitetaan laskelmin. Tosin viskurilain käsittelyn tekee hieman hämäräksi se, että puhuttaessa lauseen »painavasta», mainitun lain mukaisesti loppuun sijoittuvasta osasta voidaan Niemikorven mukaan tarkoittaa »yksinkertaisesti lähinnä sananpituutta tai keskitaajuutta» (s. 284). Olisi ollut selkeämpää puhua erikseen pituudesta ja taajuudesta, koska pelkästään painavuudesta puhuttaessa on kussakin eri käyttötilanteessa muistettava tarkistaa, tarkoitetaanko näistä molempia vaiko ainoastaan jompaa-kumpaa. Nämä ominaisuudet tietysti korreloivat pitkälti, mutta eivät täydellisesti, ja hajonta erityisesti keskimääräistä pituutta lähellä olevissa luokissa lie-nee esiintymistiheyden osalta melkoinen.

Työn ulkoasu on selkeä ja miellyttävä. Kieliasu on paikoitellen tarpeettoman mutkikas, ja painovirheitä on aika tavalla (olen merkinnyt yli 150 kpl). Sivulla 207 viitataan samassa luvussa olevaan dialogiesimerkkiin, mutta tällaista ei työstä löydy. Lähdeviitteiden esittämisjärjestys ja muotoilu on paikoitellen epä johdonmukaista. Lähdeluettelo on siinä mielessä epäyhtenäinen, että läheskään kaikista sarjoihin kuuluvista teoksista (esim. SKS:n Tietolipas-sarjan niteistä) ei ole ilmoitettu sarjan nimeä ja/tai niteen numeroa. Vuosiluvuissa ja alaindekseissä on myös jonkin verran virheitä (esim. s. 18 ja 37: Laalo 1989: a vai b?; s. 16: Vesikansa 1978: a vai b?; s. 62 Johnson 1978, lähdeluettelossa 1979; s. 64 Valkonen 1973 lähdeluettelossa 1971 jne.).

Antero Niemikorven työn ansioksi on katsottava ennen kaikkea se, että se sisältää runsaasti kvantitatiivista perustietoa, joka perustuu tällä haavaa laajimpaan suomen kirjakieltä kuvaavaan ATK-korpukseen. Samoin siinä on voitu tämentää lukuisia suomen kieltä koskevia väitteitä, jotka aikaisemmin ovat perustuneet pelkkään intuition tai olen- naisesti suppeammista ja yksipuolisem- mista aineistoista suoritettuihin laskel-

miin. Ei liene esimerkiksi suurikaan yllätys, että keskimäärin pisimpiä saneita, sananmuotoja, sanoja, lauseita ja virkeitä käytetään tietokirjallisuudessa, mutta intuitionsa perusteella kukaan ei kuitenkaan voi täsmällisesti sanoa, miten pitkiä nämä keskimäärin ovat. Jos se halutaan tietää, se on varta vasten laskettava siihen tapaan kuin Niemikorpi väitöskirjassaan. Mikäli lasketaan niin laajasta aineistosta kuin Niemikorpi, urakkaa tuskin kannattaa ja tarvitsee enää kenenkään muun suorittaa uudelleen.

Teoreettisesti Niemikorven työtä ei voi pitää innovatiivisena: tutkimusmenetelmät eivät ole uusia eikä niitä ole kehitetty edelleen. Muoto-opillisen analyysin puuttuminen ja sananmuotojen morfologisen kompleksisuuden käsittelemättä jättäminen kaventavat näkökulmaa siinä määrin, ettei kaikkien havaittuja kvantitatiivisia suhteita selittävien lingvististen tekijöiden voi katsoa vielä tämän tutkimuksen myötä tulleen perin pohjin selvityksi, kielenulkoisista puhumattakaan. Saneiden ja sananmuotojen grafemaattisten pituuksien laskemisessa lienee sen sijaan otettu irti kaikki mahdollinen, ehkä hiukan enemmänkin.

KAISA HÄKKINEN

Saksan vaikutusta viron verbijärjestelmässä

CORNELIUS HASSELBLATT *Das estnische Partikelverb als Lehnübersetzung aus dem Deutschen*. Wiesbaden 1990. Veröffentlichungen der Societas Uralo-Altaica 31. 245 s.

Viron kielen harrastaja tulee herkästi kiinnittäneeksi huomiota sen tapaisiin verbirakenteisiin kuin *alla andma* 'antaa periksi' (sananmukaisesti »antaa alas»),

ette heitma 'moittia' (»heittää eteen»), *maha jätma* 'jättää, hylätä' (»jättää maahan»), *välja kannatama* 'sietää' (»sietää ulos»), *ära hellitama* 'hemmotella pilalle' (»hellitellä pois»), *üंबर mõtlema* 'ajatella uudelleen, muuttaa mieltään' (»ajatella ympäri»). Tyyppi vaikuttaa yleiseltä, kiinteältä ja selvärajaiselta, ja sen sanajärjestys ja oikeinkirjoituskonventiot ovat vakiintuneet (*alla andma*: *ma annan alla*; teonnimi *allaandmine*). Tämän tyyppisiä verbejä tai verbirakenteita ei suomessa tai monissa muissaakaan kielissä ainakaan yhtä laajalti esiinny, mutta saksan kielestä löytyvät tarkat vastineet useimmille. Saksa on Virossa vuosisatojen ajan ollut kulttuuri-, virka- ja prestiisikielenä. Cornelius Hasselblatt haluaa, Wolfgang Veenkerin uudistermiä käyttäen, puhua viron saksalaisesta *perstraatista*; tutummat termit sub- ja superstraatti viittaavat hänen mukaansa jo päättyneiden sulautumisprosessien jälkiin, adstraatti taas tasavertaisten naapurikielten suhteisiin, joista kummastakaan ei tässä ole kysymys. Niinpä on yksinkertaisinta olettaa viron tällaisten ns. partikkeliverbien olevan käännöslainoja saksasta. Tämä on ilmeisesti ollut oletuksena myös virolaisella kielenhuollolla.

Kysymystä viron partikkeliverbeistä on tähän mennessä laajemmin käsitelty ainakin Alo Raunin artikkelissa »Zu den Präverben im Estnischen» (Neuphilologische Mitteilungen 53) sekä muutamissa Tarton yliopistossa laadituissa opinnäytetoissa. Näiden esitysten lähtökohtana on kuitenkin saksan kieli, ei viron partikkeliverbityyppi sinänsä, ja ne keskittyvät luotteloimaan saksan prepositioiden ja prefiksien käännösvastineita virossa. Niinpä ilmeisen tuntuinen hypoteesi viron partikkeliverbien saksalaisperäisyydestä on oikeastaan jäänyt testamatta, kunnes Cornelius Hasselblattin Hampurin yliopistossa valmistunut väitöskirja syntyi tätä puutetta korjaamaan.

Hasselblattin väitöskirjan johdanto esittelee aluksi käännöslainan käsitettä. Tekijä on tutustunut kiitettävän laajalti