

Aleksis Kiven sanakirjan hanke

Suomen ensimmäinen kirjailijasanakirja, Aleksis Kiven sanakirja, on jo pitkään kuulunut ajankohtaisimpiin suuriin sanakirjahankkeisiimme. Suomalaisen Kirjallisuuden Seuran pyynnöstä antoi prof. Terho Itkonen tästä hankkeesta v. 1978 seuraavan lausunnon. Vaikka olot ovat sen jälkeen muuttuneet (ks. jälkkirjoitusta), Virittäjä on katsonut yhä aiheelliseksi julkaista lausunnon sekä Aleksis Kiven 150-vuotisjuhlan muistamiseksi että samalla sanakirjatyön vauhdittamiseksi.

Suomalaisen Kirjallisuuden Seuralle

Seura on pyytänyt minulta lausuntoa professori E. A. Saarimaan jälkeensä jääneen käsikirjoitusluonnoksen ”Aleksis Kiven sanakirja” (1082 konekirj.sivua) julkaisukelpoisuudesta sen jälkeen kun tohtori Seppo Räsänen on sitä täydennellyt ja hiukan tasotellutkin. Lausuntonani esitän kunnioittaen seuraavan.

1. Tiedossani ei ole, onko Kiven-sanakirjan, ensimmäisen kirjailijasanakirjamme, julkaisuhanketta koskaan joudut-

tu varsinaisesti perustelemaan. Jos perusteluja kaivataan, ne voidaan esittää lyhyesti ja varauksettomasti. Kivi on kansalliskirjailija, suomalaisen kirjallisuuden varsinainen aloittaja ja eräiltä osin sen yhä ylittämätön mestari. Kielenkäyttäjänä hän on omaa luokkaansa. Hänen kieltensä on poikkeuksellisen monikerroksista: siinä soivat rinnan, lukijan kannalta lähes erittelemättömänä kokonaisuutena, vanhan kirjasuomen rekisteri, sitä seuranneen ripeän kehitysvaiheen — varhaisnykysuomen — rekisteri ja hahmotumassa olevan nykysuomen rekisterin ensi alkeet, ja kaiken tämän läpi ovat kuultavissa kansankielen ainekset, jotka ovat peräisin kirjailijan kotimurteesta Nurmijärven murteesta. Kiven sanaston spektri on jo niiden kerrosten lukuisuuden takia monipuolinen ja vivahteikas. Varsinaisen elävyytensä hänen sanastonsa kuitenkin saa sanoja laajemmissa ilmauskokonaisuuksissa. Kiven sanat on kuultava tekstiyhteydessään; siinä vasta ilmenee hänen kielenkäyttönsä luonteisuus ja toisaalta sen luovuus. Sellainen sanakirja, jossa Kiven sanastoa valaistaisiin mm. riittävin kontekstinäyttein, olisi kirjallisuuden harrastajalle ja tutkijalle selitysteos ja retkiopas Kiven sanataiteeseen, kielen harrastajalle ja tutkijalle taas moneltakin kannalta arvokas lähde- ja hakuteos. Jos jostain suomalaisesta kirjailijasta yleensä aiotaan sanakirjaa, Kivi sen ansaitsee ensimmäisenä. Se on yhtä luonnollista kuin että norjalaiset ovat ottaneet sanakirjatyön kohteeksi Ibsenin, venäläiset Puškinin, unkarilaiset Petőfin, saksalaiset Goethen ja Suomen ruotsalaiset Runebergin.

En tiedä, onko Kiven-sanakirjaa koskaan verrattu näihin läheisiin vertauskohtiin. Kirjailijasanakirjan käsitteen selvittämiseksi katson aiheelliseksi lyhyesti esitellä nämä viisi sanakirjaa, ennen kuin siirryn Kiven-sanakirjaan.

2.1. Vuonna 1957 ilmestyi Henrik Ibsenin koottujen teosten yhtenä osana kaksi nidosta käsittävä »Orbok», jonka oli kolmen vuosikymmenen kuluessa (1928—57) toimittanut professori Ragnvald Iver-

sen. Se ei ole varsinaisesti kuitenkaan yhden miehen työsuoritus. Sen pohjana on kollektiivinen poimintatyö: n. 360 000 Ibsenin teoksista tehtyä sanalippua, jotka kuuluvat kirjanorjan sanakirjatyön kokoelmiin. Tekijällä on sitä paitsi ollut apunaan tieteellistä työvoimaa, useita nuoria norjan kielen tutkijoita, jotka ovat tämän hankkeen palveluksessa tehneet alkulauseesta päättäen yhteensä ainakin viisitoista miestyövuotta.

Iversenin Ibsenin-sanakirja ei ole täydellinen siinä mielessä, että se kävisi Ibsenin teosten konkordanssista; sen pohjana oleva 360 000 lipun kokoelmakin on näet vajaa. Joka tapauksessa sanakirja sisältää 1 194 suurella kaksipaltaisella petiittisivullaan yht. 28 650 sana-artikkeliä (jos erisnimet jätetään luvusta, n. 27 000). Tilan säästämiseksi artikkelit on jouduttu tekemään niin tiiviitä kuin suinkin. Niiden ytimenä ovat yleensä lausesitaatit, jotka esitetään alkuperäisestä vahvasti lyhennetyssä muodossa. Merkityksiä ei sana-artikkeleissa ole suorastaan ryhmitelty; artikkeleilla näyttää kuitenkin olevan semmoinen sisäinen loogiikka, että samantapaisten käyttötapojen sitaatit ovat aina peräkkäin ja sitaatin jälkeen usein vielä viitataan (jfr. X 130 -tyyppisellä viitteellä) kohtiin, joissa samanlaista käyttöä lisäksi tavataan. Kaikkia käyttötapojakaan ei ole kirjattu; johdannossa mainitaan, että aineistossa sanasta *hånd* 'käsi' esiintyi 90 käyttötapaa (ordalag eller ordforbindelser) ja että näistä oli lopullisessa toimitustyössä karsittava noin kaksi kolmannesta. Kaikista sanoista ei ole lausesitaatteja ollekaan vaan tyydytään vain viittaamaan yhteen esiintymäpaikkaan; jos esiintymäpaikkoja on useita, ei kaikkia luetella vaan viitteen jäljessä on lyhenne ofl. (= ym.).

Iversen itsekin pahoittelee alkulauseeseen työn puolinaisuutta. Kuinka edustavaksi sanakirja on kaiken karsinnan jälkeen saatu, sitä on ulkopuolisen vaikea sanoa, mutta silmään pistävää teoksessa sen typografisesta asusta lähtien on sen eräänlainen silppuisuus ja rankomaisuus. Sanakirja on varmasti hyödyllinen haku-

teos kielentutkijalle, mutta sitaattien lyhentelyn ja puristeisen esitystavan takia se ei voine olla kovin kätevä, ja kirjallisuutta harrastavan suuren yleisön (tai valioyleisönkään) käteen se tuskin on erityisen omiaan. Olisi uskaliaasta väittää, että Ibsenin kieli elää tämän kirjan sivuilla. Toisaalta siinä ei näytä selvästi tavoitetun edes semmoista tieteellisen analyttisyyden astetta, jota tämänkaltaiselta haku-teokselta voisi odottaa.

2.2. Vuotta ennen kuin Iversenin Ibsenin-sanakirja ilmestyi, oli Neuvostoliitossa alkanut ilmestyä Neuvostoliiton tiedeakatemian kielitieteen laitoksen teettämä Puškinin kielen sanakirja (Slovar' jazyka Puškina). Tämmöistä sanakirjaa oli alusteltu jo 1800-luvulla, mutta yritys oli jäänyt kesken. Työ oli päässyt uudestaan käyntiin 1933; sitä johti nelimiehinen toimitus (päätoimittajana ensin prof. G. O. Vinokur, sitten akateemikko V. V. Vinogradov), ja leksikografeina mainitaan ensimmäisessä osassa seitsemän nimeä. Ensimmäinen osa ilmestyi 1956, neljäs ja viimeinen 1961; yhteensä teoksessa on 3 817 aihealateista kaksipalstaista sivua.

Johdannossa tähdennetään Puškinin asemaa sanataiteilijana, jonka tuotanto on perustana yhä nykyiselle »venäjän kansalliselle kirjakielelle», ja korostetaan erityisesti hänen taitoaan tuoda kirjakieleen kansanomaisia aineksia. Sanakirjan tavoite ilmoitetaan kolmenlaiseksi: se on tarkoitettu apuneuvoksi venäjän kirjakie len historian tutkijalle, edelleen Puškinin teosten lukijalle ja niiden tutkijalle. Nämä tavoitteet teos ilmeisesti täysin saavuttaa. Sen pohjana näyttää olleen täydellinen Puškinin teosten sanalipusto; jokaisen artikkelin (paitsi jonkin yltyleisen muotosanan, kuten konjunktion *a* ja preposition *v*) alussa on mainittu kyseisen sanan esiintymien määrä, ja kunkin artikkelin lopussa on nonparellilla täydellinen luettelo kaikista kyseisen sanan eri taivutusmuotojen esiintymäpaikoista (esim. verbistä *dat'* 'antaa' partisiippien sijamuodot mukaan lukien yli 40 muotoa, yht. 628 esiintymäviitettä, jotka vievät hiukan toista isokokoista sivua). Sanakir-

Katsauksia

ja ei kuitenkaan ole pelkkä kondordanssi, vaan artikkelien runkona ovat lause-esimerkit, jotka on huolellisesti järjestetty merkitysryhmittäin ja merkitysryhmien sisässä vielä käyttötavoittain ja vivahteittain. Myös kunkin merkityksen frekvenssi mainitaan. Lause-esimerkit ovat useimmiten kokonaisina virkkeinä; joskus on tehty kolmen viivan ryhmillä osoitettuja poistoja, mutta tarvittaessa kontekstia on usein otettu mukaan naapurivirkkeidenkin puolelta. Erityistä huomiota on, kuten neuvostolaisessa leksikografiassa yleensäkin, omistettu fraseologialle. Monet sana-artikkelit ovat huomattavan pitkiä, mutta typografisin keinoin ne yleensä jäsentyvät selvästi. Teos näyttää olevan laadukasta leksikografiaa, todellinen monikäyttöteos siihen tapaan kuin alkulauseessa toivotaan.

Puškinin-sanakirja sisältää yht. 21 290 sana-artikkelia. Käsiteltävien sanaesiintymien määrää en huomaa ilmoitetun, mutta ilmeisesti se nousee satoihintuhansiin.

2.3. Sándor Petőfin syntymän satavuotisjuhlaaksi valmistui v. 1973 Unkarin tiedakatemian kielitieteen laitoksessa laadittu Petőfin sanakirjan (Petőfi -szótár: Petőfi Sándor élelművének szókészlete) ensimmäinen osa, kirjaimet A—F (1160 kookasta kaksipalstaista petiittisivua). Sanakirja on suunniteltu kolmiosaiseksi. Nelimiehisen toimituksen johdossa oli kuolemaansa saakka prof. László Gáldi; työntekijöitä mainitaan I osassa lisäksi neljä ja teoksen toimittamiseen osallistuneita henkilöitä viisi.

Tämänkin teoksen pohjana on täydellinen kokoelma kirjailijan tuotannon sanaesiintymistä. Alkulauseen mukaan »vain täydellisen korpuksen tutkimisen pohjalta voidaan päästä käsiksi kielen täydelliseen rakenteeseen». Sanakirjan julkaisemisen yhtenä perusteena mainitaan se, että Petőfin kieli on varsin lähellä unkarin nykyistäkin kirja- ja yleiskieltä; siten siitä voidaan saada runsaasti oppia nykykielestä. Lisänä mainitaan Petőfin kirjallinen ja historiallinen merkitys, joka ulottuu kauas Unkarin rajojen ulkopuolelle. »Vakaumuksemme on, että tämän

teoksen avulla onnistumme paremmin valaisemaan paitsi Petőfin kieltä ja tyyliä myös hänen ajatus- ja tunnemaailmaansa ja inhimillistä yksilöyttään.»

Rakenteeltaan tämä sanakirja muistuttaa melkoisesti Puškinin-sanakirjaa. Jokaisen hakusanan frekvenssi ilmoitetaan sana-artikkelin alussa; sen jälkeen seuraa tiedot kaikkien taivutusmuotojen frekvensseistä. Täydelliseksi konkordanssiksi teosta ei ole kuitenkaan tarkoitettu. Sen artikkelit jäsentyvät merkitysryhmiin samaan tapaan kuin Puškinin-sanakirjassa; mm. kappalejaon ansiosta on typografinen vaikutelma saatu erittäin selkeäksi. Esimerkkivirkeitä on runsaasti, ja niitä lyhennettäessä ja kontekstia täydennettäessä on menetelty suunnilleen samaan tapaan kuin Puškinin-sanakirjassa. Konkordanssin puutetta lievittää melkoisesti se, että eri merkitysryhmien (ja merkitysryhmien sisässä käyttötappaa yms. osoittavien alaryhmien) lopussa mainitaan runsaasti muitakin esiintymien sivunumeroina kuin ne, jotka koskevat itse esimerkkivirkeitä. Usein tällainen luettelo on täydellinen; toisinaan tyydytään kuitenkin vain mainitsemaan läsäsiintymien määrää. Petőfin jokainen sanaesiintymä on jo näiden laskelmien takia jouduttu käymään läpi merkitysjaotustakin varten, kuten toisaalta sanamuotojen frekvenssiluetteloja varten. Kaikkiaan teos tuntuu täyttävän samat ankarat laatuvaatimukset kuin Puškinin-sanakirja.

Teoksen mittasuhteista antavat käsityksen I osan lopussa mainitut tilastotiedot. I osassa on itsenäisiä hakusanoja 6 918, viittaushakusanoja 531 ja alihakusanoja 1 588; osaa varten läpikäytyjä sanaesiintymä on 145 066 (johon lukuun sisältyy määräisen artikkelin *a*, *az* osuus 30 501). Tästä päätellen sanakirja valmiina tulee sisältämään yli 20 000 hakusanaa, jotka pohjautuvat ehkä vajaan 400 000 sanaesiintymään. II osaa ei liene vielä ilmestynyt.

2.4. Ibsenin-, Puškinin- ja Petőfin-sanakirjat, varsinkin kaksi jälkimmäistä, ovat keskenään varsin yhteismittaisia. Melko erilaisia ovat saksalainen Goethen-

sanakirjan ja suomalainen Runebergin-sanakirjan hanke.

Goethen-sanakirjaa on valmisteltu vuodesta 1946 lähtien kolmen saksalaisen tiedeakatemian yhteistyönä, tieteellisinä johtajina Wolfgang Schadewaldt, Werner Simon ja Wilhelm Wissmann. Pelkästään I. vihkoa (A—abrufen) valmistelemissa mainitaan olleen kolmattakymmentä leksikografia. Työ on valtavasti laajempi kuin kolme tähän asti esittelemääni sanakirjaa; sen pohjana ovat Goethen koko tuotannon kaikki sanaesiintymät, noin kolme miljoonaa sanalippua. Ensimmäinen vihko ilmestyi 1966, viimeisen arvioidaan tulevan julki vuosisadan loppuun mennessä.

Tavoitteiltaan tämä sanakirja mainitaan todelliseksi monikäytöteokseksi:

Als synchronische Darstellung der Sprache eines Dichters, der zugleich Gelehrter, Forscher und Denker war, ist das Goethe-Wörterbuch von vornherein als ein *Mehrzweck-Instrument* geplant und angelegt. Das Wörterbuch soll den verschiedenartigsten wissenschaftlichen wie allgemein geistigen Aufgaben dienen, die sich aus der Art Goethes ergeben, und damit zugleich den verschiedenartigen Interessen entgegenkommen, die Goethe über die engeren Bezirke der Wissenschaft hinaus in den weitesten Kreisen des Publikums erweckt. Das Goethe-Wörterbuch soll wissenschaftlich, nicht »populär« sein, jedoch es soll seine Wissenschaftlichkeit auch in der hohen Forderung des Praktischen und der Einfachheit bewahren. Als ein Instrument wissenschaftlicher Information soll es, wie Goethes Sprache selbst, zugleich fasslich und gleichsam verständlich für jedermann, ja lesbar sein.

Pyrkimyksenä on antaa käsitys siitä »kielikosmoksesta», johon »yhden ihmisen kielenkäytössä yhtyy sanontojen ja ilmaisutapojen syvään kerrostunut moninaisuus» ja jota tämänveroisena ei saksalaiselta kielialueelta tavata muilta kuin Goetheltä. Sanakirja on tosin synkroninen, mutta sen selitetään olevan historiallinen siinä mielessä, että siitä voidaan lukea Goethen kielenkehityksen eri vaihei-

ta; se on sitä paitsi kielihistorian ja kulttuuri- ja hengenhistorian tutkimuksen tärkeä apuneuvo. Se kuvastaa paitsi Goethen sanamaailmaa myös hänen asia-maailmaansa; artikkelien kaikessa lakonisuuudessaan sanotaan olevan »piileviä Goethe-monografioita» ja tekevän sanakirjasta Goethen-tutkimuksen tärkeän välineen. Mainitaan myös Goethen kielenkäytön suuntaa-antavasta merkityksestä yhä nykyisille saksan kielen käyttäjille ja huomautetaan, että jokaiselle Goethen harrastajalle tällainen sanakirja on todellinen »löytöinstrumentti».

Johdannossa todetaan, että teos ei tule olemaan pelkkä sanaindeksi vaan että siinä yhdistyvät leksikko ja konkordanssi. Miten konkordanssin idea on siinä toteutettu, ei johdannosta eikä sana-artikkelleistakaan oikein riidattomasti selviä. Teoksen valtavat sana-artikkelit sisältävät merkitysryhmissään paitsi suuren määrän esimerkkivirkkeitä myös viitteitä saman merkityksen tai käyttötavan muihin esiintymäpaikkoihin; onko viitteet tarkoitettu täydellisiksi, jää kuitenkin tietymättömäksi. Puhtaasti leksikografisena työnä sana-artikkelit ovat joka tapauksessa mitavia suorituksia, oikeaa vaivoja kaihtamatonta saksalaista tarkkuustyötä. Mainitsen vain artikkelin *Abend* (yli kolme hyvin isokokoista sivua, jotka vetävyyydetään vastaavat kukin lähes kolmea Puškinin- ja Petöfin-sanakirjojen sivua). Siinä on pitkä johdantokappale, jossa selvitetään, mitä vuorokauden osaa Goethen *Abend* hänen tuotantonsa eri vaiheissa tarkoittaa (tässä on saksalaisella kielialueella eroja, jotka osaksi liittyvät vuorokaudenaikojen jaotukseen romaanisissa kielissä). Sitaatilla valaistaan mm. sitä, kuinka italialainen vuorokauden jako vastaa Goethen ajantajua. Itse artikkeli on jäsennetty seuraavasti (siteeraan merkitysryhmien edellä omana kappaleenaan olevaa sisällysluetteloa):

A1 als Zeitangabe *a* präp Vbdgn (= Verbindungen) *b* adv Gen u Akk *u* Vorabend *d* im Zshg (= Zusammenhang) des ganzen Tages gesehen *2* als Erscheinung

Katsauksia

in der Natur *a . . b* mit charakterisierenden Adj 3 als Situation persönlicher wie gesellschaftlicher Art *a . . b* in der Grussformel *c* mit charakterisierenden Adj *d* bes verbale Vbdgn — verselbstständig 4 übertr: Ende des Lebens *B* Westen.

Vertauksen vuoksi mainitsen, että Ibsenin-sanakirjassa sanat *aften* ja *kveld* on kuitattu yht. 13 rivillä ja että Puškinin-sanakirjassa *vetšer* (243) esiintymää) saa lähes kaksi palstaa, Petöfin-sanakirjassa *este* (186 esiintymää) kolme palstaa. Kiven-sanakirjan tämänhetkisessä versiossa on *ehtoo* saanut 10 riviä ja *ilta* 2 riviä.

2.5. Suomessa on professori Lars Huldénin johdolla valmisteltu jo kymmenkunta vuotta Runebergin-sanakirjaa. Toisin kuin tähän mennessä mainituissa hankkeissa siinä on otettu avuksi automaattinen tietojenkäsittely. Runebergin täydellinen sanaluettelo on nyt jo valmis, ei hankalatekoisena ja -käyttöisenä lipustona vaan tietokoneliuskoina, jotka sisältävät aakkosjärjestyksessä frekvenssiedot Runebergin teoksissa esiintyvistä sanoista (n. 340 000 esiintymää n. 17—18 000 sanasta). Tämän työn tavoitteeksi ei ole kuitenkaan asetettu täydellistä sanakirjaa siinä mielessä kuin esim. Puškinin- ja Petöfin-sanakirjat. On katsottu, että se automaattisen tietojenkäsittelyn antamasta pohjasta huolimatta on yhä liian raskas tehtävä; varsinaiset sana-artikkelithan olisi tehtävä kuitenkin asiantuntijan käsityönä. Sen sijaan on tarkoitus valmistaa konkordanssit Runebergin eri teoksista. Sellainen on jo valmiina runokokoelmista; tulossa on toinen Värikki Stoolin tarinoista ja kolmas virsirunoudesta. Konkordanssityötä on huojennettu siten, että muutaman kymmenen tavallisimman muotosanan esiintymät jäävät pois. Lisäksi on suunniteltu teoksittaisia frekvenssiluetteloja, jotka täydentäisivät frekvenssien yleisluetteloja. Yksi sellainen on jo valmis.

Se enemmän tai vähemmän määräkaavainen artikkelinkirjoitus, joka tästä hankkeesta jää pois, tavallaan korvautuu sillä, että aineistoa heti ryhdytään käsittelemään valituilta osin syvällisemmin.

Tarkoituksena on laatia tutkielmia ja tutkimuksia Runebergin sanaston eri puolista. Tätä osaa hankkeesta ei ole sidottu mihinkään kiinteään suunnitelmaan, vaan tutkielmien on tarkoitus ilmestyä avoimena sarjana.

3. Katsauksestani toivon käyneen selville, kuinka suuri työ tasokkaan kirjailijasanakirjan toimittaminen on ja millaisia seikkoja tämäntyyppisessä sanakirjassa on nykyään tapana ottaa huomioon. Kun Saarimaa aikoinaan optimistisesti lähti laatimaan Aleksis Kiven sanakirjaa, hänellä ei ehkä ollut täysin selvää käsitystä siitä, millaiseen urakkaan oli ryhtynyt. Vaikuttaa siltä, että hänen työnsä on lähtenyt liikkeelle nimenomaan Kiven idiomien selvittelystä, alalta jolla hän on suuresti ansioitunut mm. kirjasen »Selytyksiä Aleksis Kiven teoksiin» laatijana ja useiden merkittävien tutkielmien kirjoittajana. Muistettava on, että Saarimaa jo 1910-luvulla pyrki selvittämään Kiven sanaston nurmijärveläisen pohjakerroksen osuutta paikkakunnalla muistiinpanoja tekemällä. Vanhemmiten häntä taas kiinnosti erityisesti se, miten Kivi järjestelmällisesti rikasti sanavarastoaan sanakirjojen (ennen muuta Europaeuksen ruots.-suom. sanakirjan) avulla. Vasta vähitellen sanakirjatyö näyttää laajentuneen käsittämään Kiven koko sanavaraston. Tämä on vaatinut ilmeisesti moninkertaista Kiven teosten läpikäymistä aikana, jolloin sen kopiointiin täydelliseksi sanalipustoksi ei ollut varoja ja jolloin automaattista tietojenkäsittelyä sanakirjatyön apuna ei vielä tunnettu. Tulos on sikäli kunnioitettava, että Saarimaa onnistui saamaan haravaansa lähes kaikki Kiven teoksissa esiintyvät sanat, mainitsemaan kustakin vähintään yhden esiintymäkohdan ja useimmista monia. Seppo Räsäsen tarkistusten ansiosta sanakirja on tältä osin — hakusanastoltaan — nyt luultavasti täydellinen; ainakaan en ole pistokokein löytänyt yhtään Kivellä esiintyvää sanaa, joka sanakirjan nykyisestä käsikirjoitusversiosta puuttuisi. Yhteensä käsikirjoituksessa on nyt n. 13—14 000 sanaa ja Räsäsen lisäysten jälkeen n. 50 000 sivu-

numeroin mainittua näiden sanojen esiintymäkohtaa. (Laskelmat olen perustanut tekemääni 5 %:n otokseen koko käsikirjoituksesta; tarkat arvot keskivirheineen ovat $13\,582 \pm 715$ ja $49\,092 \pm 2\,263$.) Karkean laskelmani mukaan Kiven teoksissa on kaikkiaan kolmisensataatuhatta sanasiintymää, joten suunnilleen joka kuudes esiintymä on päässyt näihin viitteisiin.

Kuten edeltä on käynyt ilmi, nykyaikaisissa kirjailijasanakirjoissa on usein pyritty suurempaan täydellisyteen kuin pelkästään toteamaan se, että sana varmasti esiintyy kirjailijan tuotannossa vähintään kerran. Tavanomaista on ilmoittaa hakusanojen frekvenssit, taivutuskielissä (esim. venäjässä ja unkarissa) sitä paitsi yksityisten sananmuotojen frekvenssit. Jos Kiven sanakirjan teko aloitettaisiin nyt, olisi kai selvää, että pohjatyö olisi tehtävä automaattisen tietojenkäsittelyn avulla. Silloin sananmuotojen frekvenssit saataisiin tietokoneiluskoihin melkein valmiina (homonymiatapaukset olisi tosin tarkistettava) ja sanojen frekvenssit olisivat laskettavissa siltä pohjalta. ATK tuottaisi monia muitakin etuja. Sen avulla saataisiin helposti myös kontekstisanakirja, jossa kukin juokseva sana olisi tekstiyhteydessään; tulos vastaisi täydellistä Kiven kielen lipustoa, ja sen pohjalta voitaisiin lopulliseen sanakirjaan saada frekvenssitietojen lisäksi aukoton konkordanssi. Kun aineisto olisi täydellinen, ei liioin mikään merkitys- tai muototutkimus, mikään käyttötapa voisi välttyä tutkijan lajittelulta hänen kirjoittaessaan lopullisia sana-artikkeleita.

Kansainväliset mitat täyttävä, sano-kaamme Puškinin- tai Petöfin-sanakirjaa muistuttava teos olisi kaikissa suhteissa Kiven-sanakirjan varsinainen ihanne. Suomen oloissa sellaiseen ei lähiaikoina kuitenkaan ole mahdollisuuksia. Saari-*maa* ja Räsänen ovat joka tapauksessa tehneet suppeamuotoisempaa leksikkoo varten suuren pohjatyön, ja koetan nyt selvittää, miten sen varaan voitaisiin rakentaa kotimaisen julkaisukynnyksen selvästi ylittävä Kiven-sanakirja, kansainvälisistä mitoista mutta samalla kustannuk-

sista tinkien. Kun näinkin suuren yrityksen markkinointi meillä on perustettava paitsi ammattifilologien ja kirjallisuuden-tutkijain myös ja ennen kaikkea kirjallisesti valistuneen yleisönosan ostohalukkuuteen, koetan samalla tähdätä siihen, että sanakirjasta voitaisiin saada mahdollisimman laajojen käyttäjäpiirien teos.

4.1. *Frekvenssien* poisjäämistä pidän tutkimuksen kannalta valitettavana mutta en niin pahana puutteena, että sen takia kannattaisi enää alistaa Kiven koko tuotantoa automaattiseen tietojenkäsittelyyn.

4.2. *Konkordanssin* idean sanakirja täyttää vain osittain. Harvinaisimpien sanojen esiintymät nähtävästi kyllä on saatu mainituksi kaikki tai miltei kaikki. Kaikkein tavallisimmat taas ovat varsin vähien esiintymäviitteiden varassa; tämä ei olisi kovin suuri puutos, jos siihen ei liittyisi se, että myös merkitysryhmitus ja esimerkiksi puuttuvat (ks. 4.3.3). Valitettavaa on, että melko sattumanvaraisten viittausten varaan on jäänyt suuri määrä myös näiden ääritapausten väliin sijoittuvaa sanastoa, ts. keskimääräistä jonkin verran harvinaisempia ja varsinkin keskimääräistä jonkin verran tavallisempia sanoja. Tämän voi pistokokein todeta jokaiselta Koottujen teosten sivulta, ja samanhan osoittaa jo viitteiden määrä (n. 50 000) suhteessa Kiven juoksevien sanojen kokonaismäärään (n. 300 000). Puute olisi helpompi kestää, jos olisi takeita siitä, että mikään olennainen sanan käyttötapa ei ole jäänyt viitteistä pois. Nuo takeet puuttuvat. Siltä kannalta käsikirjoitus jää jälkeen Ibsenin-sanakirjastakin; sehän on voitu perustaa verraten täydelliseen aineistoon, jonka vasta spesiaalisti on harkiten karsinut. Ilmeisesti tämä on kuitenkin lähes välttämätön paha, jota vain hiukan voidaan lievittää (ks. 4.3.3).

4.3. *Sana-artikkelien rakenteessa* ja *esitystavassa* on sellaisia puutteita, jotka olisi voitava korjata: a) esimerkkivirkkeiden vähyys ja sitaattien suppeus, b) merkitysryhmitysten vähyys, c) sitaattia vailla olevien esiintymäviitteiden sijoitus artikkeleissaan.

Katsauksia

4.3.1. Esimerkkivirkkeitä on kaikkiaan viljelty säästeliäästi. Lukuisiin hakusanoihin liittyy pelkkä esiintymäviite; on myös sellaisia hakusanoja, joissa esiintymäviitteitä on useita, ilman yhtäkään havainnollistavaa esimerkkivirkettä. Tämä tekee kirjan yleisilmeen hiukan kuivakkaaksi; tutkijan kannalta kirja jää epäkäytännölliseksi, ja tavallisen lukijan se helposti vieroitaa. Monen monissa sanartikkeleissa Kiven kieli elää vielä laihemmin kuin Ibsenin kieli Ibsenin-sanakirjassa. Kiven suosikkiadjektiiveihin Seitsemässä veljeksessä kuuluu *huikea*; se kuvaa monenlaisten aistimusten intensiteettiä ja moninaisia kvaliteetteja varsin verevissä virkkeissä, mutta sanakirjassa luetellaan vain 19 esiintymäkohdan sivunumerot ilman ainoaakaan sitaattia. Tämä tuntuu tuttujen sanakirjaperiaatteiden vastaiselta erityisesti sellaisissa sanoissa, joihin liittyy tekstiyhteyttä myöten vaihtelevia konnotatiivisia aineksia; niitä voi tuskin muuten havainnollistaa kuin esimerkein. *Huikean* lisäksi voidaan mainita vaikka seuraavat esimerkkivirkettä vaille jääneet (sulkeissa esiintymien määrä): *ilo* (13), *ilta* (13), *kaikua* (20), *kaino* (8), *kajahdella* (9), *kajahtaa* (3), *kakara* (4), *kallista* (8), *kallis* (10), *kaltainen* (7), *tirkistää* (4), *viskaista* (5), *viskasta* (13). Tapauksia on kaikkiaan hyvin paljon. Epäjohdonmukaistakin esimerkkivirkkeiden puute tällaisissa tapauksissa on, kun aivan vastaavanlaisissa sanoissa toisaalta on voitu viljellä esimerkkivirkkeitä kohtuullisesti, jopa runsaasti: *ankara* (sanakirjan laajin artikkeli: n. 125 esimerkkiä, useat tosin vain virkkeestä irrotettuja kaksisanaisia konneksioita), *ehtoo* (12 esimerkkiä, lisäksi 6 esimerkitöntä sivunumeroa), *kallistua* (15 esimerkkiä, 3 esimerkitöntä sivunumeroa), *kohina* (vastaavasti 4 + 9), *kohista* (7 + 3), *kohtu* (28 + 7), *muoto* (19 + 14), *virittää* (7 + 8).

Esimerkkivirkkeiden lisääminen suurentaisi tietysti kirjan kokoa, mutta kun se samalla veisi siitä kuivanlaisen luette-
lon tuntua ja lisäksi sen luettavuutta, en epäröisi suosittaa aika vahvaakin lisäystä. Yleisperiaatteena voisi olla, että jokaisen

hakusanan yhteydessä on ainakin yksi esimerkkivirke; tästä voitaisiin tinkiä enintään sellaisten selvätarvoitteisten substantiivien kohdalla, joiden käyttöä ja merkitystä konteksti ei mitenkään täsmäntäisi. Sellaiset Kivelle luonteenomaiset ekspressiivisanat kuin *huikea* ja *kaikua* voisivat hyvin saada jopa kymmenkunta esimerkkivirkettä, jos niiden avulla voidaan selvittää sanojen konnotaatioita. Olisi myös huolehdyttävä siitä, että kontekstia on riittävästi, keskimäärin enemmän kuin tähänastisissa esimerkkivirkeissä. Virkkeistä voi toisinaan poistaa epäolennaisia aineksia, mutta lukuisissa tapauksissa tekisi mieli täydentää sanakirjan niukanlaista esitystapaa uusimpien ulkomaisten mallien mukaiseksi. Sekin lisäisi sanakirjan luettavuutta ja sen yleistä käyttökelpoisuutta ja ennen kaikkea päästäisi Kiven kielenkäytön tässä haku-teoksessakin oikeuksiinsa.

4.3.2. Samaan suuntaan vaikuttaisi myös merkitysten ja käyttötapojen selvempi ryhmittely. Toisinaan ryhmitys on tehty asianmukaisesti ja havainnollisesti; esim. *kohdella*-sanan merkitysryhminä ovat 1. tavata, kohdata, 2. käsitellä, suh-
tautua; behandla, bemöta, 3. esittää, näyttellä osaa, *tarkoittaa*-sanan merkitysryhminä 1. menä, 2. pyrkiä, tähdätä; syfta ja *tila*-sanan vastaavasti 1. maatila, talo; gård, hemman, lägenhet, 2. olotila, sija; tillstånd, rum, 3. kieli. Mutta usein sanartikkelit ovat jääneet tältä kannalta hahmottomiksi: on vain lueltu sitaatteja ja sitaattittomia sivunumeroita järjestyksessä Koottujen teosten I osasta IV:een edeten. Esim. *koettaa*-sanan merkityksinä mainitaan sana-artikkelin alussa 'yrittää; koetella, kokea, kokeilla', minkä jälkeen eri merkitysten esiintymät tulevat sekaisin sivujärjestyksessä. Olisi ollut suotavaa erottaa merkitysryhmät esim. seuraavasti: 1. yrittää, 2. koetella, kokeilla, 3. kokea. Samaan tapaan *kohtaus*-sanan merkityksiksi mainitaan 'tapaus', 'taudin kohtaus', 'näytelmän kohtaus' ja aineisto seuraa erittelemättä. *Muoto*-sanan pitkäkköön erittelemättömään esimerkkisarjaan up-
pooa näkymättömiin fraasi *kaikella muo-*

toa, jne.

Sanakirjan täydentäjä on lisäämiensä esimerkkien perusteella koettanut joskus selventää merkitysjakoa mutta ei ole aina täysin onnistunut. Saarimaan alkuperäisessä käsikirjoituksessa on mainittu *omistaa*-verbin merkitykseksi 'tunnustaa omakseen, syykseen'. Täydentäjä on antanut tälle merkitykselle numeron 1 ja lisännyt »2. perusmerkityksessä». On Kiven sanakirjan kannalta anakronistista sanoa vuonna 1847 käyttöön tullutta uutta merkitystä 'besitta, ha' »perusmerkitykseksi», vaikka se nykyisessä kielenkäytössä on hallitseva. *Kalleus*-sanana Saari-
maa on määritellyt 'aarre, skatt' (3 sitaattia + 4 sivunumeroa). Täydentäjä on lisännyt tämän merkityksen edelle merkityksen 'arvokkuus' ja pujottanut joukkoon yhden sivunumeron lisää mutta ei ole silti jakanut artikkeleita kahteen merkitysryhmään. Lukijan on työlästä etsiä merkitystä 'arvokkuus' viidestä KT:n sivunumerosta, kunnes se lopulta löytyy. *Kuullella*-verbin ainoana merkityksenä Saarimaalla on 'kuunnella'. Tarkistaja on antanut sille kaksi merkitystä — 1. 'parantua', 2. 'kuunnella' — ja sijoittanut kolme esiintymää 1. merkityksen kohdalle, loput 2. merkityksen. 1. merkityksen kohdalla mainituista esiintymistä kaksi kuitenkin on selvästi merkityksessä 'kuunnella' ja kolmas (Juhanin repliikissä KT I 106 »Mutta luulempa, tuo silmä-kulta kuultelee hieman») on merkitykseltään epäselvä, ehkä pikemmin 'tuntua (kipeältä)' kuin 'parantua'. Lukija saa oudon käsityksen tämän keskeisen verbin merkityksistä ja siitä, mikä niistä on ensisijainen.

Yksimerkityksiset sanat voidaan usein jättää ilman selitystäkin, milloin merkitys on sama kuin nykyisessä yleiskielessä. Liian taloudellisesti on kuitenkin menetelty, kun merkityksenselitystä vaille on jätetty esim. *salko* (sanan merkitys ei liene kaikille nykykielisille selvä) ja kun *sahrat*-sana ja eräistä muista viitataan vain »Sel.»-teokseen (näin ei pitäisi mielestäni tehdä, kun kyseessä kerran on itsenäinen sanakirja eikä »Selitysten» kylkiäinen).

Vailla merkityksenselitystä on myös esim. *sakea*. Se ei ehkä olisi kovin haitallista, ellei artikkeli sattuisi jatkumaan hakusanan jälkeen seuraavasti: »hajusta. I 153. s. ukkospilvi 209. piimää s:ta kuin terva 233 --.» Lukija käsittää asian niin, että *sakea* olisi Kiven kielessä vain olfaktorisia aistimuksia kuvaava adjektiivi. Kyseessä on kuitenkin vain yksi perusmerkityksen pohjautuva käyttötap; konteksti löytyy Juhanin Turun-matkan kuvauksesta: »-- istuu tyttöjä kuin posliinivauvoja, tuokuttaen kauas ympärillensä sakean hajun kalleista öljyistä ja rasvoista.» Selvintä olisi ollut selittää *sakea* esim. synonyymeillään *tiheä*, *sankka* ja luetella sen jälkeen esimerkit tavanomaisimmista konteksteista lähtien ja harvinaisempiin edeten, jolloin tämä esimerkki jäisi pikemmin loppupuolelle artikkeleita.

4.3.3. Sekä esimerkkien lisääminen että merkityksenselitysten lisääminen ja ryhmittely käynee harvinaisempien sanojen osalta yleensä tyydyttävästi käsikirjoitukseen jo sisältyvän aineiston ja sivunumeroviitteiden perusteella. Sitä on nähtävästi yritettävä samalta pohjalta myös sellaisten sanojen kohdalla, jotka sijoittuvat jonnekin tavallisten ja harvinaisten välimaille. Aineisto ja viitteistö on kuitenkin keskimäärin sitä niukempaa, mitä tavallimpi sana. Kaikkein suurtaajuusimpia sanoja koskevat artikkelit ovat yleensä kaikkein vähäantisimpia: jokunen esimerkki (joskus ei yhtään) eikä minkäänlaista jaotusta. *Antaa*-verbistä mainitaan 19 esiintymää mutta ei yhtään esimerkkivirkettä. Esiintymiä on Kiven tuotannossa luultavasti satoja (vrt. Puškinilla *dat'* 628, Petöfillä *ad* 720), ja jos koko aines olisi ollut käsillä, siitä ehkä olisi syntynyt 4—5 painosivun mittainen artikkeli kuten vastaavista verbeistä Puškinin- ja Petöfin-sanakirjoissa. Jopa Ibsenin-sanakirjan melko karua *gi'*-artikkelia on valaistu yli 30 sitaattilla. Muita vastaavanlaisia: *kaksi* (14 sivunumeroa, ei esimerkkejä), *tuoda* (5), *täällä* (8), *vielä* (13). Olennaisesti parempi ei ole niiden suurtaajuisia sanoja koskevien artikkelien laita, joissa on jokunen esimerkkivirke ja lisäksi jokunen

Katsauksia

sivunumero mutta ei liioin merkitysryhmitystä: *minä* (14 esimerkkiä + 1 muu esiintymä), *nuori* (7 + 3), *nähdä* (8 + 10), *olla* (25 + 0), *saada* (7 + 15), *saattaa* (20 + 4), *toinen* (8 + 10), *tämä* (2 + 21), *täytyä* (5 + 1).

Näissä tapauksissa ei artikkelia voi helposti laatia käsikirjoituksessa olevien esiintymäviitteiden perusteella. Tavallisten sanojen käyttöspektri on kuitenkin olennainen osa kirjailijan kielenkäyttöä; sanakirja jää varsin vajaaksi, jos ei tuota spektriä yritetä valaista. En näe muuta keinoa kuin että kokenut leksikografi poimii varta vasten laadittavan sanaluettelon mukaan huomattavan runsaan määrän esiintymiä paristasadasta suurtaajuisimmasta sanasta. Tämä lipusto (parhaiten kai kserokopioita rastituista KT:n sivuista) olisi sitten pohjana, kun sana-artikkelit kirjoitetaan. Sanakirja ei saisi näin keskeiseltä osalta jäädä aukkoiseksi.

4.3.4. Sana-artikkelien yleisessä rakenteessa on ymmärtääkseni vielä yksi puute: esimerkivirkkeiden sivunumerot ja ilman sitaattia mainitut esiintymien sivunumerot ovat kussakin sana-artikkelissa toistensa lomassa aivan Koottujen teosten sivujen mukaisessa juoksevassa järjestyksessä. Artikkelien typografia ja luettavuus paranisivat suuresti, jos sitaattittomat esiintymäviitteet siirrettäisiin kunkin artikkelin (merkitysryhmiin jaetuissa artikkeleissa kunkin merkitysryhmän) loppuun ja niiden väliset pisteet jätettäisiin pois. Myös sivunumeroiden kursivointi lienee tarpeetonta. Näin tavallinen lukija voisi ensi näkemältä saada artikkelista irti keskeisen ja häntä kiinnostavan aineksen, ts. hyvin valitut ja lajitellut esimerkivirkkeet, ja pelkästään spesialisteja varten lisätyt esiintymäviitteet jäisivät pois tuon taipaleen näkökuvaan pirstomasta. Tämä merkitsee melkoista korjausta artikkelien tähänastiseen rakenteeseen. Korjaus kävisi kai kuitenkin kivuttomasti puhtaaksi-kirjoituksen yhteydessä; uusi puhtaaksi-kirjoitus jouduttaisiin joka tapauksessa teettämään.

4.4. Saarimaan viittaukset Kiven käyttämien sanojen nurmijärveläiseen murrepohjaan ovat semmoinen arvokas lisä, jota vastaavaa en ole ulkomaisissa kirjailijasanakirjoissa tavannut. Näitä viittauksia on käsikirjoituksessa kai toistatuhatta, ja ne antavat vakuuttavan kuvan siitä, mikä osuus Kiven varsinaisella äidinkielellä on ollut hänen taideproosansa muotoutumiseen. Viittaukset on sanakirjaa uusittaessa tietenkin säilytettävä. Niiden johdosta kuitenkin pari huomautusta.

Saarimaa viittaa Kiven sanan nurmijärveläisyyteen pelkällä hakusanan jälkeisellä Nj-lyhenteellä, milloin sana semmoiseen on artikkelissa ilmenevässä merkityksessä Nurmijärven murteen mukainen. Jos hänellä on esitettävänä lisätietoja (esim. sanan muoto murteessa vähän toisenlainen tai siitä on käytettävissä murreasuinen esimerkkilause tai selitys), kommentti sen sijaan on artikkelin lopussa. Sanakirjan tarpeisiin pidän tätä järjestelmää varsin käyttökelpoisena. Tekstin jälkeisissä kommentteissa on kuitenkin pieniä muodollisia puutteita, jotka pitäisi korjata. Esim. sub *kaistapäinen, tillastaa, töyräs* ja *virskutella* nuo kommentit kuuluvat seuraavasti:

- Nj kaistapäinen hiukka mielisairaana vikaa
- Nj lapset jotain tahtoo
- Nj pellon töyräs
- Nj Orava tirskuu t. rirskuttaa [kirjoitusvirhe?], virskuu virskahteele

Tällaisissa tapauksissa näyttää alkuperäiset pikamuistiinpanot siirretyn sanakirjaan muokkaamattomina. Lukijalle ne jäävät helposti epäselviksi. Niitä sopisi täsmentää ja väljentää esim. seuraavaan tapaan:

- Nj sana kansanomainen (selitetty: »hiukka mielisairaana vikaa»)
- Nj sana kansanomainen (selitetty: »lapset jotain tahtoo»)
- Nj tuttu esim. sanallitossa *pellon töyräs*
- Nj samakantaisia verbejä oravan ääntelystä: *virskuu, virskahteele* (myös *tirskuu, tirskuttaa*)

Vielä harhauttavampia ovat kommentit sub *hetaleinen*, *ometta* ja *sompa*:

(Nj *hepaleinen*, *risainen*)
Nj »jossain muualla *ometto*»
Nj troka

Ensimmäisessä kommentissa on ajatuk-
sena luultavasti se, että sanakirjan tekijän
muistiinpanoissa Nurmijärveltä on vain
hetaleista muistuttava sana *hepaleinen*.
Kommentti voisi kuulua yksinkertaisesti
näin: Vrt. Nj *hepaleinen* 'risainen'; se jo
osoittaisi, että murteesta ei (toistaiseksi)
ole löytynyt läheisempää vertauskohtaa.
Ometta-sanan kommentissa taas on jää-
nyt selvään sanomatta se, mikä alkuperäi-
sen muistiinpanon nähtävästi on aiheut-
tanut: *ometta* kyllä on Nurmijärven mur-
retta, sen sijaan *ometto* (jota Kivi myös
käyttää ja joka sanakirjassa on omana
hakusananaan) ei ole nurmijärveläistä.
Sub *ometta*-sanasta riittäisi nyt lyhenne Nj heti
hakusanan jälkeen, sub *ometto* pitäisi lisä-
tää artikkelin loppuun: Nj outo, kertojan
mukaan käytössä »jossain muualla».
'Viulun jousta' merkitsevän *sompa*-sanan
kommentti on eksyttävän lakoninen.
Sompa tässä merkityksessä ei näytä kuu-
luvan Nurmijärven murteeseen, eikä sitä
tietääkseni tunneta mistään muustakaan
murteesta; mistä Kivi on sen saanut, jää
hämäräksi. *Troka* sen sijaan on nurmijär-
veläistä. Kommentti voisi kuulua: Nj ou-
to, sijalla *troka* 'viulun jousi'.

Tähän tapaan kohennettuina kommen-
tit palvelisivat sanakirjan käyttäjää entis-
tää paremmin. Niiden typografiaa olisi eh-
kä vielä harkittava mm. kursivointien ja
lainausmerkkien käytön osalta. Toivotta-
vaa tietysti olisi, että ne antaisivat samalla
lopullisen kodifikaation Kiven sanaston
nurmijärveläisyyksistä; tämän sanakirjan
ilmestyttyä sellaiseen tuskin ryhtyy enää
kukaan, varsinkin kun kenttätyö käy tällä
alalla aina vaikeammaksi. Jo oman vä-
häisen murteentuntemukseni perusteella
olen huomannut siellä täällä tapauksia,
joissa hyvinkin luonteenomaisen nurmi-
järveläisyyden murrelähteisyyden on jäänyt
toteamatta. Muutama esimerkki: *kimpus-*

taa 'painia' (Saarimaalla vain viittaus Eu-
rénin sanakirjaan, jossa *kimpustella* 'brot-
tas'), *liukua* (matkanteosta millä tahansa
kulkuneuvolla), *meiju* 'juhannuskoivu',
suvenihanainen 'sinivuokko', *toispuol* 'toi-
sella puolella', *toistain* 'toistamiseen, uu-
destaan'.

Olisiko tällaisia löydettävissä lisää?
Olen kokeeksi käynyt tältä kannalta läpi
sanakirjan käsikirjoituksen sivut 496—
528 monivuotisen tuusulalaisen kieleno-
paani Ester Kytökosken kanssa, jonka
murre ainakin sanastollisesti on samaa
kuin Nurmijärven Palojoen murre (Esterin
kotikylä Nahkela on Palojoen naapur-
rikylä ja hänen sukujuurensa osaksi nur-
mijärveläiset). Näillä sivuilla Saarimaa
viittaa yht. 30 sanan nurmijärveläisyy-
teen: *massahdella* (»Nj massutti suu-
taan»), *massauttaa*, *mehto*, *mehu*, *meukas*,
meinata, *meininki*, *mekastaa*, *mekastella*,
mekkari, *mellakka*, *melkoinen* (Saarimaan
mukaan vain adverbina: melkosen suuri),
melkotella, *menehtyä*, *menota* (mainitaan
jonkun Lohjalta muuttaneen käyttämäk-
si), *meriteerata*, *meteli*, *metelöitä*, *metsän-
haltija*, *metsäsahrat*, *mettäs*, *meuhata*,
miehevastus, *miehusta*, *mieliisti*, *mieliä*,
moikaa, *mojoo*, *mojova*, *molkahdella*,
molskahtaa. Murteen mukaisiksi ne yle-
ensä osoittautuivat tässä tarkistuksessakin;
mehto (pro *mettäs*) ei tosin missään ta-
pauksessa ole nurmijärveläistä tai tuusu-
lalaista, vaan Saarimaalle on tässä sattun-
nut erehdys. Mutta lisäksi ilmeni joltisen-
kin varmasti ainakin seuraavien sanojen
tai niitä muistuttavien murreperäisyys:
massahdella (»maassa massahtelee»: kie-
lenoppaalle tuttu tämäkin verbi eikä vain
Saarimaan saama *massuttaa*), *mehevä*
(mm. lihavasta ihmisestä), *meiju* 'juhan-
nuskoivu', *mekko* 'rohdinpaita', *mekosti*
'runsaasti, paljon' (on Saarimaallakin
nurmijärveläiseksi merkittynä mutta vain
sub *aika*: aika mekosti), *melkoinen* (myös
adj.: »melkosen summan sai kätees»),
melkska (sel.: »semmost oikee kauheet jy-
tyy ja semmost jytinä»), *menota* (vakuut-
taa kotoperäiseksi eikä vain muualta tul-
leiden käyttämäksi), *mesiäinen* 'kimalai-
nen', *metsäneito* (murt. »mettäneitty»; jo

Katsauksia

ennemmin olen tallentanut EK:lta tähän sanaan liittyvää uskomuksellistakin ainesta), *metsäsika* (murt. »mettäsiika» = 'mäyrä'), *mieltäkääntävä* (»mielenkääntö»), »kääntelee miältä» esim. kuvottamiseksi), *mieluisa* (»mialusa», sen sijaan Kiven käyttämää rinnakkaisjohdosta *mieläinen* kielenopas vierastaa), *miesvoima*, *miilu* (esim. sysimiilu), *moiskahdella* (Aaltoska sanonut vanhasta Nahkelan ukosta Perttiinistä: »sylkee semmottii suurii tupakkasylkii nii et moiskahteleo pitkin nurkkii»), *molskia*, *molskina*, *monaisti* (vähän epävarma tieto). Lisättävä olisi vielä *mato* 'käärme', joka muistaakseni on aiemmissa muistiinpanoissani. Siis noin kahden kolmanneksen lisäys Saarimaan mainitsemiin, kaikki periaatteessa maininnan arvoista sanakirjassa. (Sen sijaan sanakirjassa ei kannattane noteerata sellaisten sanojen murteenmukaisuutta kuin *marssia*, *matala*, *matka*, *melkein*, *menestyä*, *menettää*, *mennä*, *meno* jne., jotka ovat kirjakielen keskeisiä sanoja ja joko vanhastaan tai viimeistään kirjakielen leviämisenä tuttuja missä tahansa murteessa.)

Näkisin mielelläni, että tämä puoli sanakirjassa hiottaisiin todella kuntoon. Ehkä puolentoista kuukauden pestillä saisi tottunut sanastaja Nurmijärven nykyisestäkin murteesta yhä melkoisen lisäsaaliin, joka arvokkaalla tavalla valaisisi Kiven sanaston juuria. Myös Saarimaan pikaisia muistiinpanoja, jotka yleensä kylä näyttävät olevan luotettavia, voitaisiin samalla tarkistaa ja varsinkin tynkämäisiä esimerkkilauseita korvata täyteläämillä. Jossain tapauksessa tulisi kyseeseen myös Sanakirjasäätiön kokoelmien hyväksikäyttö, mutta päämenettelynä tulisi olla kenttäkeruu edellä hahmottelemaani tapaan.

4.5. Kuten jo mainitsin, Saarimaa pystyi osoittamaan, että Kivi on aktiivisesti laajentanut sanavarastoaan silloisten ruotsalais-suomalaisten sanakirjojen avulla. Päälähteenä Kivellä on selvästi ollut D. E. D. Europaeuksen 1852 ilmestynyt ruots-suom. sanakirja, mutta merkkejä on myös siitä, että hän on ehtinyt hyödyntää

lisäksi Ferd. Ahlmanin 1865 ilmestynyttä sanakirjaa. Itse luulen todenneeni ainakin yhden tapauksen (*keidas* 'mätäs'), jossa hänen lähteenään on täytynyt olla Carl Heleniuksen sanakirja vuodelta 1838, ja asiasta v. 1962 Saarimaan kanssa keskustellessani hän mainitsi huomanneensa muitakin samanlaisia tapauksia. Heleniukselta ehkä tekemiään poimintoja Saarimaa ei ole käyttänyt Kiven-sanakirjansa hyväksi, mutta Europaeusta hän on seurannut varsin järjestelmällisesti: merkityksenselityksissä on suomalaisten synonyymien lisäksi tai sijastakin usein mainittu se ruotsin sana tai ne ruotsin sanat, joiden kohdalta kyseinen suomen sana löytyy Europaeukselta. Toisinaan viittaus Europaeuksen, joskus Ahlmaninkin sanakirjaan on artikkelin lopussa. Varsinkaan artikkelinalkuisia viittauksia ei liene syytä käsittää niin, että Kivi läheskään aina olisi suorastaan saanut kyseisen sanan Europaeukselta; joukossa on tavallisia sanoja, jotka hän on muutenkin voinut omaksua aikansa kirjakielestä. Viittaus Europaeukseen on tällöin kai käsiteltävä pikemmin viittaukseksi tuonaikaisen kirjakielen yleisempäänkin kantaan.

Sekä artikkelinalkuiset että -loppuiset viittaukset olisi yleensä nähdäkseni säilytettävä; ehkä niistä voi jonkin poistaa enintään silloin kun tavallisen sanan merkitys muutenkin on selvä eikä ole varsinaista syytä epäillä, että Kivi on tässä kohden turvautunut sanakirjan apuun. Myös Europaeuksen ruotsin oikeinkirjoitus olisi säilytettävä. Jostain syystä täydentäjä on ryhtynyt turhaan työhön: korjannut Europaeuksen sanakirjan oikeinkirjoituksen kauttaaltaan ruotsin nykyisen oikeinkirjoituksen mukaiseksi. Tällaisissa kohdin on tieteellisen sanakirjan oltava diplomaattisen tarkka; vanha oikeinkirjoitus olisi palautettava ja selvyden vuoksi lisättävä vielä artikkelinalkuisenkin ruotsinnosten eteen kirjain E (= Europaeus), kuten artikkelinloppuisissa on kauttaaltaan tehty. Artikkelinloppuisen viittausten rakenteesta ja typografiasta voidaan esittää osaksi samantapaisia huomautuksia kuin artikkelinloppui-

sista murre tiedoista. Tämä puoli olisi selvennettävä ja yhdenmukaistettava. — Sama koskee joskus ilmeneviä viittauksia Kalevalaan ja Kantelettareen.

4.6. Pitkin matkaa olen joutunut viittaamaan erilaisiin sana-artikkeleissa ilmeneviin *teknisiin epätasaisuuksiin*, jotka käsikirjoitusluonnoksessa ovat ymmärrettäviä mutta joita ei saisi olla enää painoon menevässä versiossa. Tältä kannalta äskeisvuosien viimeistely on ollut varsin osittaista, näyttääpä sinne tänne syntyneen uuttakin epätasaisuutta. Luettelen vielä muutamia korjausta vaativia seikkoja.

4.6.1. *Hakusanoissa* on pikku puutteita, jotkut niistä pikemmin tarkistustyön kuin Saarimaan käsikirjoitusluonnoksen jäljiltä. Hakusanana on esim. resiprookkipronomini *toinentoisemme* (mon. 1. p. omistusliitteinen muoto), vaikka sana-artikkelin aineistossa olisi ollut hakusanaksi luonnostaan sopiva *toinentoisensa* (yks. 3. p. omistusliite). Saarimaan työn jälkeen on hakusanaksi ilmaantunut *tilanne*, esiintymäviitteenä II 464. Viittauskohdasta löytyy Cononin repliikki »Margaretassa»: »Mutta kuinka luonnistui tilanne?» Se on kumminkin *tila*-sanan omistusliitteinen muoto; *tilanne* on suomessa vasta 1910-luvun uudissanaja, niitä joita paljon myöhempi kirjailijamestari F. E. Sillanpää varovasti kokeili ensimmäisissä teoksissaan.

Aiheellisesti on muutamia liitepartikkeleja otettu hakusanoiksi. Niiden käsittelyä tosin olisi syvennettävä. Esim. *-ko*-partikkelin esimerkit keskittyvät toistaiseksi vain siihen muotoerikoisuuteen, että tämä takavokaalinen variantti esiintyy etuvaalisissa sanoissa (Nurmijärven murteen mukaisesti; tätä taas ei mainita). Liitepartikkeleista puuttuvat kuitenkin *-s* ja *-sta(pa)*; sen sijaan tarkistaja on merkinnyt itsenäiseksi hakusanaksi sanan *mitätäpä*, joka on pronominin ja enklitin yhtymä. Hän on lisännyt myös hakusanan *mitämä*. Tässä on kysymys vain siitä, että Kivi on käsitellyt oikeinkirjoituksen kannalta enklitin tavoin kotimurteensa pronominin *miä* painotonta toisintoa *ma ~ mä*.

Jos tapaus halutaan säilyttää hakusanana, se on ainakin selitettävä: = 'mitä minä'.

Kiven sanaliitot on hakusanoituksessa käsitelty sanaliittoina (hakusanaksi vain liiton toinen jäsen) ja yhdyssanat yhdyssanoina (omiksi hakusanoikseen). Tämä nähdäkseni oikea periaate on tarkistustyössä venytetty vähän turhan pitkälle, kun omiksi hakusanoikseen on otettu paitsi erisnimi *Timo* (hämmästyttävää kyllä vain yksi esiintymäviite) myös yhdyssanat *Timo-parka*, *Timo-jukuri*, *Timo-poika*, *Timo-poloinen*, *Timo-veli* ja *Timo-veikkonen*. Tällaiset tekisivät painetussa sanakirjassa hiukan koomisen vaikutuksen. Kiven »sanojen» kokonaisuudrasta voidaan hakusanoituksessa tältä osin mieluummin tinkiä ja mainita tämän tyyppiset propri + appell. -tyyppiset yhdyssanat vain yhdysosien kohdalla.

Saman sanan ortografiset toisinnot ovat joskus turhaan eri hakusanoina. Esim. *viskasta* olisi yhdistettävä *viskaista*-artikkeliin; heti sen alussa voisi sulkeissa mainita esimerkkejä *i*:llisinä ja ilman *i*:tä kirjoitetusta tämän sanan muodoista (vrt. 4.6.2).

Homonyymit olisi pyrittävä erottamaan siten kuin Nykysuomen sanakirjassa, ts. hakusanaa edeltävin numeroin (tai sitten hakusanan jälkeisin numeroindeksin). Nyt käsikirjoituksesta voi tavata semmoisenkin artikkelin kuin »*voi* 1. subst. — 2. interj. —», ikään kuin kyseessä olisi polysemia. Myös NS:n mukaista esitystapaa on käytetty. (Sitä on viljelty joskus myös polysemiatapauksissa; Räsänen näyttää pyrkineen tasoittamaan tällaisia mutta ei huomatakseni ole päässyt täyteen johdonmukaisuuteen.)

-sti-adverbit on sijoitettu toisinaan omiksi hakusanoiksi, toisinaan kantaadjektiivin yhteyteen (esim. *sukkelasti* sub *sukkela*, *tiuhasti* sub *tiuha*). Parasta olisi esittää ne aina omina hakusanoinaan.

Hakusanoina olevissa yhdyssanoissa on pientä epäselvyyttä. Esim. hakusanaa *virkaero* seuraavat samassa kappaleessa hakusanat *-mies* ja *-retki*. Olisi kirjoitettava mieluummin omille riveilleen täyd-

Katsauksia

liset yhdyssanat *virkamies* ja *virkaetki*, mutta jos lyhentäminen katsotaan välttämättömäksi (paljon sillä ei voiteta), olisi *virkaero*-sanana yhdyssosat erotettava pystyviivalla.

4.6.2. Silloin kun sanan *taivutuksessa* on merkkeille pantavaa, näytteet on upotettu esimerkkivirkkeiden joukkoon, ikään kuin niiden veroisina (esim. *kohtaissani*, *kohtanut* sub *kohdata*, *viskain* sub *viskata*). Tällaisia tapauksia on etenkin supistuma-verbien alalla, ja ne ansaitisivat selvemmän käsittelyn jo siksi, että kyseessä yleensä on Kiven murrepohjan suora tai välillinen vaikutus. Luontevinta kai olisi tämäntyyppisten tapausten esittäminen sulkeissa heti hakusanan jälkeen ennen merkityksenselityksiä ja esimerkkivirkeitä.

4.6.3. Turhan säästeliäästi ja harhauttavastikin on kirjoitettu esimerkkivirkeissä olevat hakusanan muodot sellaisissa melko lukuisissa tapauksissa kuin sub *kohtaus* (sitaateissa mm.: olvi-retkemme k:t, k:ilta), *tarpoa* (t:ttu). Olisi kolme tapaa saada tällaiset kohdat selviksi: 1) pystyviiva hakusanaan sitaateissa esiintyvien taivutusmuotojen yhteisen osan jälkeen ja kaikki sitä seuraava toistettava taivutusmuodoissa (siis *kohtaus/s: k:kset, k:ksilta, tar/poa: t:vottu*), 2) ei pystyviivoja hakusanaan ja taivutusmuoto lyhennettynä vain jos koko hakusanaan sisältyvä sekvenssi semmoisenaan toistuu taivutuksessa (siis *kohtaus*: kohtaukset, kohtauksilta mutta k. 'kohtaus' ja k:ta 'kohtaus-ta'), 3) ei pystyviivoja hakusanaan ja taivutusmuoto aina kokonaisena. Kolmannen tavan panisin etusijaan; kahdella muulla säästetään kirjaintilaa mutta aiheutetaan se haitta, että kussakin esimerkkivirkkeessä saa visuaalisesti vaatimattomimman ilmauksen se sana, johon huomion pitäisi kiintyä.

4.6.4. Asiakommentit ovat toisinaan paikallaan. Ne olisi kuitenkin erotettava jollain rakenneratkaisulla muusta artikkelista. Hahmoton tältä kannalta on mm. artikkeli *Olli*:

Olli 1. Raalassa on Olli-niminen torppa

(Klaukkalan kylässä samanniminen). Juuri niinkuin O:n muori viinahimossa pistää tuiskujuttuja Mäkelän emännälle I 114. Ollin muorista oli laulu, jossa toistui --. 2. miehennimi IV 198.

Tämä voitaisiin selventää vaikka seuraavanlaiseksi:

Olli 1. miehennimi IV 198. 2. asumuksen-nimi. Juuri niinkuin O:n muori viinahimossa pistää tuiskujuttuja Mäkelän emännälle I 114. (Raalassa — ja samoin Klaukkalan kylässä — oli Olli-niminen torppa. Ollin muorista oli laulu, jossa toistui --.)

5. Aleksis Kiven sanakirjan tämänhetkisestä käsikirjoitusversiosta olen joutunut yllä esittämään runsaastikin kritiikkiä. Nykyisellään se ei vielä ole kypsä painoon, mutta ehdotuksiani noudattamalla siitä luullakseni saataisiin julkaisukelpoinen teos. Vaikka se ei ylittäisi alan parhaiden ulkomaisten saavutusten rinnalle, siitä tulisi varsin hyödyllinen tutkijan ja kirjallisuudenharrastajan apuneuvo — ja samalla kunnianosoitus Aleksis Kivelle ja hänen hartaalle tutkijalleen ja tulkitsijalleen E. A. Saarimalle, jolta sanakirja jäi vuosikymmenien kypsyttelynkin jälkeen luonnosmaiseksi.

Tulen ydinkysymykseen, uuden käsikirjoituksen laadintaan. Kokenut leksikografi voisi sen valmistaa suurimmasta osasta artikkeleita nykyisen käsikirjoituksen pohjalta ehkä noin vuodessa. Tärkeimpänä työnä hänellä silloin olisi esimerkkivirkkeiden lisääminen jo kertyneiden sivunumeroviitteiden pohjalta, merkityksenselitysten parantaminen ja aineiston ryhmittely kunkin sana-artikkelin sisällä. Lisäksi tulisi kuitenkin suurtaajuisimpien sanojen käsittely, joka veisi varmasti useita kuukausia. Toinen olisi nurmijärveläisen murresanaston tarkistus, tottuneelta sanastajalta arviolta puolentoista kuukauden työ. Ulkomaisten kirjailijasanakirjojen vaatiman työ määrän rinnalla tämä on vähäistä, Suomen nykyoloissa taas aika raskas sijoitus, ellei siihen saada ulkopuolista rahoitusta. Ymmärtääkseni apua näin tärkeään hank-

keeseen olisi saatava. Yhteiskunnalle kuluja aiheuttamatta hanke on edistynyt jo pitkälle; nyt on kyseessä loppuvaihe, johon kuuluva työ tehtäisiin mahdollisimman taloudellisesti, laadusta silti ylen määrin tinkimättä.

TERHO ITKONEN

★

Vuodesta 1978 lähtien ei sanakirjahanke ole edistynyt. Toisaalta on tapahtunut tärkeä muutos, joka ei voi olla vaikuttamatta sanakirjan valmisteluun. Suomalaisen Kirjallisuuden Seura näet on julkaissut Kiven juhluvuotena 1984 uuden, valoladontaisen painoksen Kiven Kootuista teoksista, ja automaattisen tietojenkäsittelyn avulla voidaan ladelmasta saada aikaan täydellinen Kiven sananmuotojen konkordanssi, johon voidaan ottaa mukaan myös kunkin sananmuodon kontekstia. On selvää, että kun tällöinen apuneuvo on käytettävissä, sitä on myös käytettävä. Sanakirjatyön kokonaissuoritusta se kuitenkin pikemmin hidastaisi kuin nopeuttaisi; teoksen toimittajan olisi nyt käytävä läpi kaikki konkordanssiin kuuluvat sanaesiintymät (jotka eivät suomen runsastaivutteisuuden takia ole läheskään kaikki edes yhtenäisenä jonona

kulloisenkin hakusanan kohdalla), ja niin päädyttäisiin lopulta Puškinin- ja Petöfin-sanakirjoja muistuttavaan suoritukseen. Sanakirjan laadun ja täyteläisyyden hyväksi se kyllä koituisi, mutta sanakirjatyön vaatima aika, jonka varhemmassakin lausunnossani ilmeisesti arvioin liian vähäiseksi, lisääntyisi tästä huomattavasti.

Toisena pulmana on Nurmijärven-murteisen sanaston osoittaminen. Aika harventaa hyviä kielenoppaita armottomasti; silti lisäsanastus on nähdäkseni vielä mahdollisuuksien rajoissa. (Mm. Nurmijärven kulttuurilautakunnan vastikään järjestämä murreosanaston keruukilpailu osoitti, että hyviäkin murteentaitajia yhä löytyy.) Useissa tapauksissa voisi kenttätutkijan työn lähtökohtana olla lähimurteista kerätty sanasto, jonka voi olettaa kuuluneen Nurmijärvenkin murteeseen. Erityisesti tulee kyseeseen maist. Matti Vilppulan keräämä laaja ja arvokas Vihdin sanasto. Valitettavasti muiden lähipitäjien, kuten Nurmijärven omatkin, sanavarat ovat murteidemme sana-aineiston epätarkoituksenmukaisen järjestelyn takia lähes löytymättömissä arkistokoelmista — epäkohta, joka pahoin haittaa suomalaista dialektologiaa yleensäkin.

T. I.