

Lauseet järjestyksen kourissa

Lauseopin arkiston kaksi vuosikymmentä

JUSSI KALLIO

Alussa oli lause

»Kansallisten tieteiden kehittämisohjelma 1966–80» sisälsi laajan jakson kielentutkimuksen kehittämishankkeita, joista yhtenä luonnosteltiin suomen murteiden sanaston ja muoto-opillisten aineiden keruuta. Heti ohjelman julkistuttua syksyllä 1965 laativat professorit Osmo Ikola ja Matti Liimola tämän ohjelmakohdan täydennykseksi muistion, jonka johdanto-osassa he totesivat:

Olemme asiaa pohdittuamme päätyneet siihen käsitykseen, että kielentutkimusta käsittelevä ohjelman osa vaatisi erään täydennyksen, joka kokonaisuuden rinnalla on pieni, mutta joka kuitenkin on merkitykseltään olennainen.

Tässä lauseessa itää lauseopin arkiston siemen. Myöhemmin samassa muistiossa kirjoittajat hahmottelevat selkeät toimintalinjat kansankielen lauseopillisen aineiston hankinnalle ja käytölle:

Päämääräksi on asetettava deskriptiivisen *suomen kansankielen lauseopin* aikaansaaminen. Valmistuttuaan tämä teos tulisi olemaan ensiarvoisen tärkeä tutkimuksen ja myös nyky-suomen viljelyn apuneuvo kansankielen sanakirjan ja muiden siihen verrattavien suurteosten rinnalla. Aluksi olisi saatava aikaan kansankielen lauseopin arkisto, joka jo ennen kansankielen lauseopin valmistumista loisi tutkimukselle paljon uusia mahdollisuuksia ja joka lauseopin valmistuttuakin säilyttäisi merkityksensä sen täydentäjänä; onhan selvää, ettei läheskään kaikkea kertyvää aineistoa voitaisi sisällyttää painettavaan lauseoppiin.

Ehdotus kansankielen lauseopillisen aineksen keruusta sai nopeasti ilmaa alleen. Kevään mittaan 1967 käytyihin neuvotteluihin perustui valtion humanistisen toimikunnan kanssa solmittu nelivuotinen tutkimussopimus, joka allekirjoitettiin 25. toukokuuta, ja työ alkoi heti kesäkuun alusta. Tutkimussopimus koski yhteishanketta, jonka suhteellisen itsenäisiä osakkaita oli toi-

Lauseet järjestyksen kourissa

saalta muoto-opin, toisaalta lauseopin aineksen keruu. Lauseopin hankkeen asemapaikaksi tuli Turun yliopiston suomen kielen laitos, ja työtä on koko ajan johtanut professori Osmo Ikola.

Tietureitse vai käsipelillä?

Tutkimussopimuksen liitteeksi jätetyssä tutkimussuunnitelmassa kuuluu lauseopin osuutta luonnehtiva peruslause näin:

Lauseopillisen aineiston keruun ja arkistoinnin tavoitteena on luoda kaikkia suomen murrealueita edustava laaja kansankielen lauseopin arkisto, joka on siten suunniteltu ja järjestetty, että sitä voidaan helposti käyttää mahdollisimman monenlaisten lauseopin kysymysten tutkimiseen.

Helpon ja monipuolisen käytön vaatimus heitti heti vastasilmiä kysymyksen arkistojärjestelmän luonteesta. Oliko aihetta tyytyä käsivaraiseen kortistojärjestelmään vai tuliko rohkeasti heittäytyä tietokoneiden armoille? Tutkimussuunnitelmassa oli varovin sanakääntein otettu kantaa jälkimmäisen vaihtoehdon puolesta:

Asiaa tutkittaessa on käynyt yhä ilmeisemmäksi, ettei lauseopillista arkistoa luotaessa ole järkevää jättää käyttämättä automaattisen tietojenkäsittelyn tarjoamia etuja, jotka ilmeisesti moninkertaistavat kokoelman käyttömahdollisuudet.

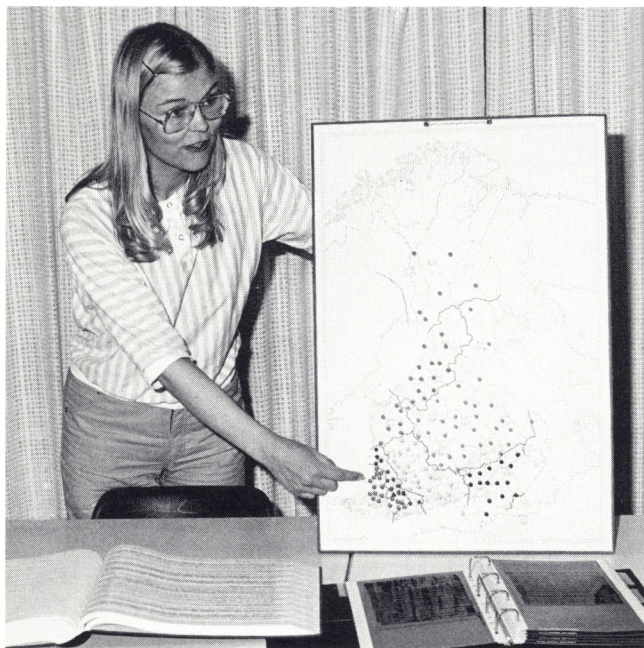
Silti asiasta käytiin perusteelliset neuvottelut tietokonealan tuntijain kanssa. On muistettava, että tuolloiset tieturit olivat nykyisiin verrattuna kömpelöjä ja heikkotehoisia. Eivätkä ne olleet jokamiehen omuutta, sillä mistään mikroista ei osattu edes uneksia! Ikolan ja Liimolan muistiokin tyytyy tässä kohdin viittaamaan työn mahdolliseen rationaalistamiseen »esim. reikäkorttijärjestelmän avulla» mainitsematta itse tietokonetta nimeltä.

Hankkiakseen oppia ja kokemusta mittavista aineistonkeruuhankkeista sekä tietoa suuren maailman virtauksista aineistojen arkistointimenetelmissä lauseopin työn johtaja teki syksyllä 1967 opintomatkan kahdeksaan sellaiseen Länsi-Euroopan yliopistoon, jossa oli toteutteilla vastaavanlaisia hankkeita. Kiertomatkan tutustumiskohteisiin kuuluivat 1) Lontoon yliopiston »Survey of English Usage», 2) Leedsin yliopiston »Survey of English Dialects», 3) Liègen kansainvälisen tutkimuslaitoksen »L'Étude des Langues Anciennes par Ordinateur», 4) Saarbrückenin yliopiston »Automatische syntaktische Analyse des modernen Deutsch», 5) Freiburgin yliopiston »Automatic Syntactic Analysis and Synthesis of English Sentences», 6) Münchenin yliopiston germanistiikan laitos, 7) Göteborgin yliopiston nyky-

JUSSI KALLIO

ruotsin taajuussanastohanke ja 8) Tukholman yliopiston KVAL-ryhmä ja ruotsin puhekielen tietokonetutkimus. Matkansa aikana prof. Ikola vakuutui tietokoneiden merkityksestä ja tulevaisuudesta kielitieteellisen tutkimuksen apuna, joten lopullinen päätös lauseopin arkistointijärjestelmän rakentamisesta automaattisen tietojenkäsittelyn pohjalle syntyi retken jälkeen kuin itsestään.

Eivät vaikeudet silti siihen loppuneet. Esimerkiksi ATK-asiantuntijain palvelut saatiin ensi alkuun Turun yliopiston sovelletun matematiikan laitokselta tilapätyönä, mistä ja asian outoudesta luontojaan aiheutui molemmipuolisia viestintäongelmia. Suunnitteluvirheitä ja harharetkiäkään ei aina voitu välttää. ATK-palvelut vakinaistuivat syyskuun alusta 1969, kun yliopiston palvelukseen otettiin matemaatikko humanistisia tutkimushankkeita varten. Alusta saakka lauseopin arkisto on ollut matemaatikon ankarin työnteettäjä. Toimi on myöhemmin muutettu humanistisia tutkimushankkeita palvelevaksi yliopiston laskentakeskuksen amanuensuuriksi.



Lauseopin arkiston aineiston murrealuekatetta esittelemässä maist. Ulla Palomäki, joka on pitkään työskennellyt arkiston tutkijana. Kuv. Jarmo Koskinen.

Lauseet järjestyksen kourissa



Arkistoinneksen käsittely käy kätevästi tietokoneen päätteestä; takana lauseopin tutkimushanketta alusta loppuun johtanut prof. Osmo Ikola ja edessä lauseopin arkiston ATK-asiantuntijana viimeksi ollut Turun yliopiston laskentakeskuksen amanuenssi maist. Pekka Porri. Kuv. Jarmo Koskinen.

ATK-tehtävän korvaukseksi heikohko palkkaus on vaihduttanut toimenhaltijoita nopeasti, mikä puolestaan on heikentänyt ATK-palvelujen täystehoista hyödyntämistä tutkimushankkeiden hyväksi.

Teknisiä vaikeuksia jouduttiin ensi vuosina kiertämään jopa Kööpenhaminan kautta, kun yliopiston oman tietokoneen muistitilavuus ei riittänyt kontekstisanakirjoja tuottavan KWIC Index -ohjelman vaatimaan aineiston lajitteluun. Pulmaan toi avun Kööpenhaminassa toimiva Pohjois-Euroopan yliopistojen tietokonekeskus (NEUCC), jonne silloisen matemaatikon oli matkustettava pari viikoksi itse paikalle tilannetta selvittämään. Retki oli tuloksekas. Niinpä arkiston tutkijoille tuli parina seuraavana vuonna tutuksi Turun tullikamari, josta he tämän tästä saivat noutaa Kööpenhaminassa kahdeksana kappaleena tuotettuja murretekstien kontekstisanakirjojen normaali- ja käänteisversioita. Laitteistotilanteen myöhemmin parannuttua tämäkin työvaihe siirtyi takaisin omaan yliopistoon, ja lopulta kontekstisa-

JUSSI KALLIO

nakirjojen paperikopioista kokonaan luovuttiin siirtymällä mikrofilmikopioihin, jotka ratkaisevasti säästävät arkistotilaa.

Jälkikäteen on helppo todeta, että tietokonepohjaisen arkistointijärjestelmän valitseminen oli puhtaasti tekniseltä kannalta etulinjan ratkaisu, joka merkittävästi on lisännyt arkiston käytöllisyyttä. Laitteistojen nopea kehitys on johtamassa siihen, että tutkija voi pikapuoliin ottaa lauseopin arkiston aineistoon suoran yhteyden omasta kotimikrostaan. Valittu arkistojärjestelmä on toisaalta omiaan suuntaamaan tutkimusta toisenlaisille urille, kuin alun pitäen oli ajateltu. Tästä ei välttämättä koidu vain etua, vaan siitä voi aiheutua omat haittansa ja rajoituksensa. Keskittyminen jonkin solidin tutkimussuunnitelman toteuttamiseen saattaa kärsiä.

Nauhasta nauhaan

Käsiteltävä murreaineisto oli alun perin ajateltu valittavaksi ääninauhoista, ja tämä menettely osoittautui käytännön työssä kätevimmäksi. Turun yliopiston suomen kielen laitoksella on itsellään melkoinen äänitarkisto, ja puutteellisesti nauhoitetuilta alueilta saatiin täydennystä suomen kielen nauhoitarkistosta Helsingistä. Erityisongelmiin oli kuviteltu tarvittavan jonkin verran vasituista kenttäkeruuta, mutta valitun toteutuksen luonteen ja varojen jatkuvan niukkuuden vuoksi siihen ei lopultakaan liennyt aikaa.

Käsitteilyn alunalkaiseksi tavoitteeksi oli asetettu keskimäärin tunti puhuttua murretta pitäjää kohti eli yhteensä noin 500 tuntia valmiiksi eriteltyä murretekstiä. Kun työ pääsi käyntiin, osoittautui pian, että tosi tavoite voisi olla puolet tuosta eli arkiston kooksi tulisi yhteensä noin 250 murreainesyksikköä. Tästäkään määrästä ei loppuun saakka pystytty pitämään kiinni, vaan arkiston lopulliseksi laajuudeksi jäi 132 tuntia muoto- ja lauseopillisesti eriteltyä murretekstiä. Käytännössä tämä tarkoittaa, että yhden pitäjän murretekstin pituus karkean keskimääräisesti arvioiden on 6 700 murresanaa. Vertailuaineistoksi työstettiin lisäksi 30 vastaavan laajuista kirjoitetun ja puhutun yleiskielen aineistoerää.

On mahdoton arvioida, mikä määrä kieliopillisesti käsiteltyä aineistoa tämäntapaisen arkiston perustaksi sopisi tai riittäisi. Selvää lienee, että taaajaan esiintyvien ilmiöiden selvittelyyn juuri ja juuri riittää tuollainen parinsadantuhannen lauseen kokoelma, jollainen lauseopin arkisto nykyisellään on. Mutta murteiden lauseoppiin kokonaisuutena sisältyy piirteitä, joita tunnin näytteeseen osuu harvalti eikä joka tuntiin ollenkaan. Näiden kylläläiseen tavoittamiseen tarvittaisiin ehkä miljoonan lauseen määrä. Tästä nä-

kökulmasta tuo aikanaan esitetty joka toisen pitäjän tiheys eli 250 aineistoyksikön laajuus tosiaan vaikuttaisi tutkimuksen kannalta realistiselta ja muutenkin kohtuulliselta arviolta.

Ääninauhoista valitut tekstit ovat enimmäkseen tavanomaista murrehaastattelua, joka tuon tuosta väljenee soljuvaksi kerronnaksi. Mahdollisuutta myöten näytteisiin on pyritty saamaan keskustelunomaisuutta eli murreoppaiden keskinäistä seurustelua. On oletettu, että monipuolinen aiheisto loisi teksteihin samalla muoto- ja lauseopillista vaihtelua eli että monipuolinen aiheisto tuottaisi monipuolista aineistoa.

Ensimmäinen tehtävä ryhdyttäessä tekstiä nauhoista purkamaan oli sopivan tarkekirjoituksen valinta. Päädyttiin periaatteessa foneettiseen vaikka suhteellisen karkeaan merkintään, jonka katsottiin lauseopin tutkimuksiin riittävän. Sanojen suhteellinen lausepaino osoitettiin kolmiportaisella asteikolla, koska painolla saattaa lauseopillisessa tulkinnessa toisinaan olla merkitystä. Erilaisia puheen prosodiikkaan kuuluvia kangerteluilmiöitä kuten tankkausta, takertelua, epäröintiä, viivytystä ja muuta vastaavaa huomioitiin myös ja ne otettiin tarkekirjoituksessa huomioon jälkeinpäin arvostellen turhankin tarkkaan. Usein ne tekstiä eritellessä pikemmin hämmentävät kuin auttavat lauseen jäsenystä. Toisaalta ne kytkeytyvät puheen koko prosodiaan niin tiukasti, ettei niiden erillinen osoittelu karkeassa tarkekirjoituksessa hyödytä oikeastaan mitään eikä ketään.

Kaikki tekstit on varsinaisen litteroijan lisäksi toinen tutkija tarkistanut. Kun työssä oli jatkuvasti siirryttävä murrealueesta toiseen eikä litteroijia tai tarkistajia voitu sitä lyytä vaihtaa, ei syntyään murretta taitavien litteroijien käyttö tullut kyseeseen kuin harvoin, niin suositettavaa kuin se muuten olisi. Tästä johtuu, etteivät tekstit kaikesta tarkistuksesta huolimatta voi olla täysin virheettömiä – vaikka yleensä sentään aika luotettavia.

Huvittaviakin tapauksia sattui, kun tutkimusapulaisia aluksi oli vain kaksi ja nämä tarkistivat ristiin toistensa litteroinnit. Oikeista tulkinnoista käytiin välillä kiivaita taisteluja. Pahin pattitilanne syntyi Muolaan näytteestä, joka oli kummallekin vierasta murretta. Niinpä olin joutunut tekemään kollegani siitä laatimaan tunnin litterointiin peräti noin 700 korjausmerkintää. Koska valtaosa korjauksista koski toistuvia tyyppivirheitä, toverini saattoi suoralta kädeltä hyväksyä viitisensataa korjausta, mutta parisataa merkintää jäi yhä riitteeseen. Tämän jälkeen työn johtaja otti tavakseen ratkaista kiistat arvovallallaan itse, ettei niiden selvittelyyn turhaan käytettäisi liikaa aikaa. Periaateratkaisuista keskusteltiin aina viikkokokouksissa, joten ne koituivat tuoreeltaan litterointityön hyväksi. Tätä menetelyä sovellettiin kaikkiin työssä vastaan tullessiin ongelmiin hankkeen koko kestoajan. Nuo viik-

JUSSI KALLIO

73701123 ((PV. NO MITENKAS SE VIITAKET TEROTETTII,)) PV
 73701124 *NO SE/ **LIIPPA *OLI *SEMMON,EN *KIVESTÄ *TEHTY
 *SANOTTIIN ***KUUSAMOL.../ ***KUUSAMOL **LIIPOKS KR
 NIITA.
 73701125 SE OL, ***KUUSAMOSSA SE **KIVI.../ **KIVIPAEKKA
 *SEMMON,EJ JOSTA *TEHTIIL **LIIPPA SEMMON,EM **PATKA KR
 VUAN./
 73701126 SANO*TAAN *NYT **KAKSKYMMENTÄ/ SENTTIA
 KAKSKYMMENTÄVIIS SENTTIA **PITKÄ JA *PARIN SENTIN
 VAHVUINEN/ *KIVIPATKA SI... *SE TEHTIIN SIITE.// KR
 73701127 JA *SIITÄ OL, VIELÄ *SEMMON,IA **JUTTUJAES SIITÄ
 ***KUUSAMOL **LIIPASTA KUN SIELLÄ ***SOTKAMOLLA//
 **KYSSYY/ **MUU... *MUUVAM *MIES JOTTA SIT
 ***KUUSAMOSTA OL, SE ***SOTKAMOSSA SE
 **SAVOTTALAENE- TTA **MINKÄLAENEN/ **LIIPPAVUOS, SE
 *TUJL, NYT ***KUUSAMMOON./ KR
 73701127 ((SUORAA ESITYSTÄ. MINKÄLAENEN --- KUUSAMMOON.))
 73701128 NI/ HANO- TTA *HUONOLLAENEM MUUTEM MUT KYL
 **HAAN.../ **HAAMPA *ANTAA **MAESTOOJ JA VETÄS,
 **KORVALLEN *NYRKILLÄÄN SIITÄ SANASTA. KR
 73701129 SE **LOUKKAANTU NIIM *PAL,JO// Z Z. KR

Kappale juoksevaa tekstiä lauseopin arkiston kokoelmista; Nilsin murrettä.

SUODENNIEMI
22201

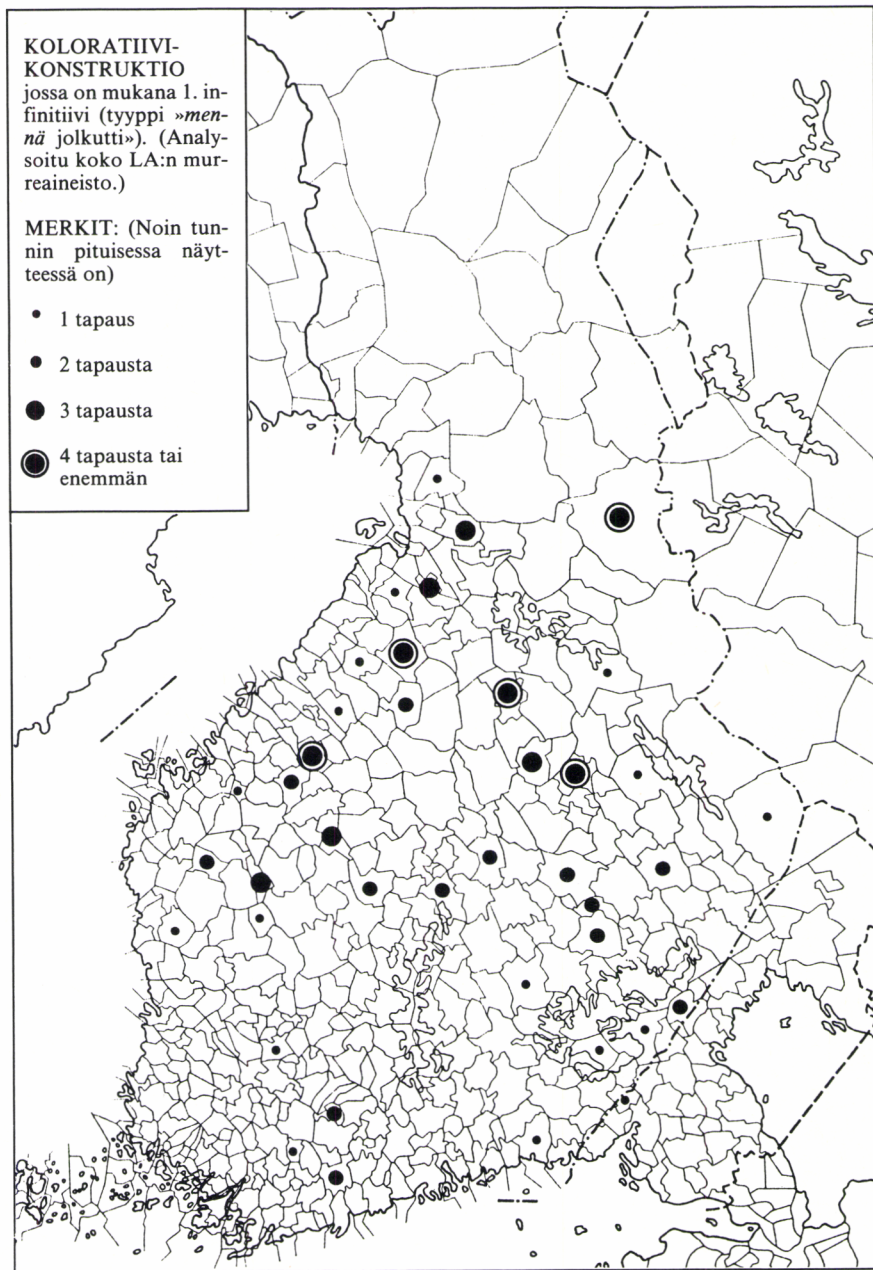
18. 9.1987

- 1 OS EM MINA *TIÄRÄJ JOS SE/ *JOS SE *TAPAHTU SITTEP **PÄIVISIN
 TAIKKA *KOSKA HX(M... *TAPAHTU <<MUTTA>> *EM MINA
 *KOSKAAN NÄHNY **EI *SILLON *KUMMINGAN *SILLON/ **SOTTIMEN
 AIKAAN NI- *EI/ **KOPPELOT- OLI **AINA *PUISA *SIALÄ
 **NE EI *OLLEIKAL *LASKENUM *MAAHA.
- 14 OS SITÄ *KUKERRUSTA EI *NIINKÄV VOIM **MATKIAS SE *TULLEE
 NIIM *MERKILLISELA *ANELL- ETI- EI SITÄ *OIKEEN *SUULLAAV
 VOIM **MATKIAM <<MUTTA>> SITÄ **KIHAUTTAMISTA SITÄ
 *MATKITA(A).
- 58 SI KO VAAN- *YKS *KUULA OLI *HAULIKOSSA MENEÄ **PALJO
 <<MUTTA>> NOIN....
- 65 SI SIINA TAYTY OLLA... <<MUTA>> SE- OLI **ITTES SEM
 *PYSSYN **TEHY/ HYVÄ *SILMÄ TAYTY *OLLA.
- 67 SI *SEMMONEN KO *NYKYNEM **PENNI TAIKKA/ **VIIS PENNIX
 *EI **NIIN SUURI... KO *NYKYNEV **VIIS PENNIA O/ <<MUTA>>
 **PENNIRRAHA SE./
- .
- .
- 502 OS **EI/ EI *NITÄ/ **PIÄNEMMÄT RATTAAT OLI **YKSPUISIA/
 <<MUTTA>> **ISOMMAK *KOOTTIIN/ **OSISTA.
- 511 OS *EI *TASÄ KYLÄSÄ O *OLLUT *TUULIMYLLYÄ *KOSKAAM <<MUTA>>
 *KYLÄ ***SUARENNIAMELLÄ *OLLU ON *USSEITA *TUALA OLI
 **PERÄKUNNASA/ JA *SE S... S... **YKS TUULIMYLLY ON
 *NYT TUALLA **KIRKOMMAELLÄ/ **MUSEOTARKOTUKSISA.

LÖYDÖKSIÄ 55
 KIRJAIMIA 31324
 SANOJA 6263
 VIRKKEITÄ 408

Kappale lauseopin arkiston hakutulostusta. Haun kohde *mutta*-sana tulostettu virke-
 konteksteineen.

Lauseet järjestyksen kourissa



Lauseopin arkiston aineistosta syntyy kätevästi murteiden lauseopillisia karttoja. Kuvassa koloratiivirakenteen tyyppiä *mennä jolkutti* levintä arkistoaineiksen mukaan. Osmo Ikolan johtaman Suomen murteiden lauseoppia -hankkeen tuotantoa.

JUSSI KALLIO

kokoukset ovat jääneet mieleen monin tavoin opettavina tilaisuuksina. Silti olen yhä vakaasti sitä mieltä, että erään kohdan tuossa kriittisessä Muolaan näytteessä pitäisi kuulua: »senko ol' "vee-et 'valuneet sen "silmit sel 'lehmä» eikä *veet*, kuten kahden äänellä yhtä vastaan kirjoihin merkittiin. Lauseoppiin ei tällaisella äänteellisellä tarkkuudella tietenkään ole mitään vaikutusta.

Toinen aste työssä oli litteroitujen tekstien muoto-opillinen erittely ja lävistys reikäkortteihin. Tämänkin vaiheen kehittämisessä oli omat mutkansa. Teksti lävistettiin reikäkortteihin alkuun juoksevana, mutta pian siirryttiin tätä selvästi kätevempään menettelyyn, jossa kuhunkin korttiin lävistettiin vain yksi sana ja samaan korttiin tätä sanaa koskeva muoto- ja lauseopillinen koodimerkintä. Käytännön syistä oli ensin rakennettu muoto-opillinen koodi ja vasta myöhemmin lauseopillinen, joten muoto- ja lauseoppi kooditettiin osin erikseen. Tästä aiheutui joskus tulkintahankaluuksia ja jopa pakollisia jälkikorjauksia. Työ järkeistyi kumminkin tältä osin, kun ruvettiin käyttämään erillisiä koodituslistoja, joihin itse teksti ja kustakin kortista näkyvät tekstin ja juoksevan sanan tunnistetiedot oli kirjoitettu valmiiksi ja joihin muoto- ja lauseopillinen kooditus merkittiin niitä varten tehtyyn taulukkoon. Näistä listoista kooditus lävistettiin kortteihin, ja koko lävistystyön – ensin tekstin lävistyksen tunnistekoodeineen ja sitten tekstiä koskevan kieliopillisen koodituksen – saattoi nyt tehdä alusta loppuun ammattiläivistäjä.

Vakiintuneimmillaan työ kulki neljänä suurena aaltona: 1) tekstin litterointi ääninauhasta ja litteroinnin tarkistus, 2) tekstin ja sen kehystiedoston lävistys reikäkortteihin, 3) tekstin kieliopillinen kooditus ja 4) koodituksen lävistys kortteihin. Käytetty koodi tai oikeammin sen perustana oleva kieliopillinen analyysi pohjautuu pääosin tutun kieliopin käsitteihin ja tulkintoihin sekä muoto- että lauseopissa. En huoli kronikoivassa esityksessäni sen tarkemmin selostaa arkiston koodia, koska siitä on tarkka selvitys lauseopin arkiston julkaisuun kuuluvassa Osmo Ikolan toimittamassa Lauseopin arkiston oppaassa (Turku 1985). Kirjanen opastaa samalla kieliopillisin haku- ja muin esimerkein lauseopin arkiston aineiston käyttöön.

Arkiston käyttö perustuu tietokoneen tekemiin aineksenhakuihin ja kulloisenkin haun mukaisesti tulostuksiin. Hakuohjelmisto mahdollistaa useita tulostusmuotoja, jotka kaikki samoin kuin hakuohjelmisto yleensä ja sen käytön sisäiset ja ulkoiset ehdot selvitetään opaskirjassa perusteellisesti.

Arkiston koko aineisto on siis lävistetty reikäkortteihin, joita on yhteensä pitkästi toista miljoonaa. Tällainen arkistomuoto olisi laajuutensa tähden yksinään kovin hankalakäyttöinen. Niinpä se on nykyisellään vain aineksen

Lauseet järjestyksen kourissa

varmuusvarasto, sillä koko materiaali ohjelmineen kaikkineen on siirretty tietokoneen magneettinauhoihin, joissa se vaatii tilaa huomattavasti vähemmän ja on milloin tahansa kätevästi saatavilla. Näin arkistoinneksen muokkaus alusta loppuun tarkoittaa vain suurta tiedonsiirron ja tiedonkäsitteilyn sykliä ääninauhoista muistinauhoihin. Kaikessa mekaanisuudessaan se on melkein kuin inhimillisen ja ihmisaivoisen tiedonhankinnan kuva.

Apulaisia, tutkijoita, tutkimusta

Aluksi ei arkistolla muutama viikkoon ollut käytettävissään mitään omia tiloja, vaan työtä käynnisteltiin suomen kielen kokoelmahuoneen suuren pöydän ääressä Turun yliopiston päärakennuksessa. Pian saatiin kumminkin samasta päärakennuksen toisesta kerroksesta ihka omaan käyttöön pieni kattoikkunallinen huone, jossa tutkimusapulaisista toinen paukutti lävistintä, toinen kuulokkeet korvilla litteroi ääninauhoja. Tilaongelma ratkesi kerralla, kun suomen ja sen sukukielten sekä perinnetieteiden laitokset saattoivat vuoden 1969 alusta siirtyä yliopistonmäen kupeeseen omaan taloon, Fennicumiin. Lauseopin arkisto sai tilavan ja tarkoituksenmukaisen arkistihuoneen sekä vielä erikseen pikku huoneet litterointi- ja lävistystyötä varten. Neliöitä kyllä tarvittiin, kun Kotimaisten kielten tutkimuskeskuksen otettua arkiston haltuunsa siinä yhdessä vaiheessa työskenteli kolme tutkijaa ja pari litteroijaa; lisäksi oli löydyttävä työtilaa arkiston käyttäjillekin. Tupakuumeen noustua yhteistilassa ahertavien tutkijain kesken ajoin rauhoittamattomiin arkistihuone jaettiin myöhemmin väliseinällä kahtia. Silti arkistolla yhä on hallussaan viehättävä ja viihtyisä käyttötila.

Professori Osmo Ikola toimi tutkimussopimustyön lauseopillisen osan johtajana 1967–71, »suomen kansankielen lauseopin tietokonetutkimus- ja arkistointityön vastuullisena johtajana» 1972–76 ja »lauseopin arkiston tieteellisenä johtajana» 1976–84 eli johti hankkeen alusta loppuun, vaikka välillä toimi vaikeassa siirtymävaiheessa kaksi kautta (1974–80) yliopiston rehtorina.

Tutkimusapulaisina olivat aluksi Jussi Kallio (kokopäiväisenä) 1967–72 (sen jälkeen asiantuntijana ja »projektsihteerinä» arkistoinnin loppuun asti) ja Rauno Heinonen (osapäiväisenä) 1967–69. Pisimpään arkiston tutkimusapulaisena ja tutkijana työskenteli Olli Järvikoski 1969–84. Lyhempään tutkimusapulaisina (enimmäkseen osapäiväisinä) ovat olleet Aulikki Jalava 1970–71, Jorma Rakkolainen 1971–73, Marja-Berit Malmberg 1972–73, Markus Lahti 1973–74, Kirsti Tuominen 1973–75, Tuulikki Haanpää

+ E2 E1 +	81	82	83	84	85	91	92	93	
11	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	1 0.5	0 0.0	0 0.0	0.1 0.0
21	0 0.0	0 1.2	0 0.0	4 2.4	0 0.0	0 0.0	161 96.4	0 0.0	167 5.2
23	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	11 23.9	33 71.7	2 4.3	46 1.4
24	104 37.1	1 0.4	0 0.0	0 0.0	0 0.0	108 38.6	66 23.6	1 0.4	280 8.7
31	3 0.4	538 66.4	187 23.1	0 0.0	0 0.0	82 10.1	0 0.0	0 0.0	810 25.1
32	0 0.0	1 2.9	33 97.1	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	34 1.1
33	0 0.0	1 0.1	999 99.8	0 0.0	0 0.0	1 0.1	0 0.0	0 0.0	1001 31.0
41	0 0.0	0 0.0	161 96.4	0 0.0	2 1.2	2 1.2	0 0.0	2 1.2	167 5.2
43	0 0.0	0 0.0	0 0.0	0 0.0	0 0.0	2 100.0	0 0.0	0 0.0	2 0.1
51	1 0.4	1 0.4	259 98.9	0 0.0	0 0.0	0 0.0	0 0.0	1 0.4	262 8.1
53	0 0.0	450 97.8	8 1.7	0 0.0	0 0.0	0 0.0	1 0.2	1 0.2	460 14.2
	108 3.3	994 30.8	1647 51.0	4 0.1	2 0.1	207 6.4	261 8.1	7 0.2	3230

Lauseopin arkiston taulukkotulostusta: nominaalimuodot adverbiaaleina. E1 eli vasemmanpuolisin pystyriivi = sijamuotojen koodinumerot ja E2 eli ylin vaakariivi = infinitiivien ja partiippiin koodinumerot. Aineistona kirjoitettu ja puhuttu yleiskie-
li.

1974–76 ja Kaija Korpela 1975. Tutkimuskeskuksen valtakautena ovat arkiston työssä urakoineet ensin osapäiväisinä apulaistutkijoina ja sitten kokopäiväisinä tutkijoina Anneli Pajunen 1976–82 ja Ulla Palomäki 1976–84 sekä apulaistutkijana Anna-Kaisa Koitto 1983–84.

Tutkimusapulaisten ja tutkijain lisäksi on arkistossa vuoteen 1980 asti aika ajoin työskennellyt tuntipalkkaisia litteroijia. Lähinnä lounaismurteita ovat käsitelleet Maija Grönholm ja Päivikki Suojanen, lounaisia välimurtei-

Lauseet järjestyksen kourissa

ta ja hämäläismurteita Kirsi Räsänen ja Leila Sulkko, eteläpohjalaismurteita Taina Karhu, keskipohjalaismurteita Silja Jutila sekä savolais- ja kaakkoismurteita Ulla-Maarit Heljasvuo, Liisa Turunen, Sirpa Neuvonen ja Kyllikki Torkkeli. Yhteensä siis 23 ihmistä on arkiston työssä apulaisina tai tutkijoina tahkonnut kannuksiaan lauseopin liipoin.

Arkiston ATK-asiantuntijoina ovat matemaatikon tai amanuenssin tointa hoitaessaan toimineet Markku Nurminen 1969–71, Yrjö Karjalainen 1971–72, Hannu Koivula 1973–75, Jussi Salmela 1975–77, Tuula Laaksonen 1977–78, Matti Rantanen 1978–79, Markku Nevaranta 1979, Hans Helenius 1979–80, Seppo Rantala 1980–81 ja Pekka Porri 1981–84. ATK-kirjoittajana oli työn loppuvaiheessa Taina Laine 1982–84.

Arkiston käyntipäiväkirjassa on vuosittain merkintöjä kymmenistä tutkijainvierailuista, joten on jokseenkin mahdotonta tarkkaan jäljittää, mihin kaikkiin tutkimuksiin arkiston aineistoa on hyödynnetty. Suuri osa vierailijoista on lisäksi ollut ulkomaisia tutkijoita. Koko joukko aineistoja on sitä paitsi pyynnöstä toimitettu tilaajille postitse valmiina paperitulosteina tai arkiston valmistuttua tietokoneen muistinauhakopioina.

Kokoelmia on arkiston koko toimiajan käytetty Turun yliopiston suomen kielen opetuksen tukena ja eriasteisten opinnäytetöiden aineistoina. Vasiten lauseopin arkiston ainekseen perustuvia opintoesitelmiä ja pro gradu -tutkielmia on vuotten mittaan kertynyt kymmeniä. Eri puolilla valmistuneiden murreaiheisten liseniaatintöiden ja väitöskirjojen tekijät ovat niin ikään aineistoa hyödyntäneet, jotkut laajaotteisemmin, jotkut vain kohdittain. Esteenä aineksen täystehoiselle käytölle ennen arkiston valmistumista olivat arkistojärjestelmän keskeneräisyys ja kehitteilläöolo sekä alituiset laitteiston muutokset. Teknisten rajoitusten poistuttua voisi odottaa lauseopin tutkimuksen uutta kukoistusta Turussa ja muuallakin.

Yksin jäät sinä lause

Vuoden 1972 alusta vuoden 1976 helmikuun loppuun lauseopin arkisto toimi prof. Ikolan valtion humanistiselta toimikunnalta saaman vuotuisen tutkimusrahan turvin. Kotimaisten kielten tutkimuskeskuksen perustaminen maaliskuun alusta 1976 merkitsi arkiston toimintaan ratkaisevaa muutosta: siitä tuli yksi keskuksen monista tutkimushankkeista tai oikeastaan tarkemmin Turun yliopiston ja keskuksen yhteishanke. Kansankielen lauseopin aineiston arkistoinnin loppuunsaattaminen tietyissä ohjelmallisissa rajoissa turvaantui. Keskuksen johtokunta antoi aineksen arkistoinnille uudet mää-

JUSSI KALLIO

rälliset tavoitteet, jotka merkitsivät tyytymistä murrealueittain noin joka neljännen pitäjän lopputiheyteen. Lauseopin arkistosta tahdottiin tehdä keskuksen mallihanke, joka valmistuisi asetettuun kohtuulliseen määräaikaan mennessä. Vuonna 1966 kangastelleille kansankielen lauseopin haaveille sai toistaiseksi jättää jäähyväiset.

Kansankielen lauseopillisen aineksen keruuhjelma valmistui vuoden 1984 loppuun ensimmäisenä ja tähän asti ainoana tutkimuskeskuksen pitkäaikaishankeista. Keskus tarvitsi tuolloin tutkijansa toisiin tehtäviin, joten lauseopin arkistoon ei jäänyt ketään vakinaista henkilökuntaa. Itse arkisto siirtyi Turun yliopiston haltuun. Yliopiston valtuusto oli marraskuun lopussa 1984 hyväksynyt arkistolle johtosäännön, jonka ensimmäinen pykälä totea:

Turun yliopistossa toimivan lauseopin arkiston tehtävänä on palvella suomen kielen lauseopin tutkimusta ja tältä osin tukea yliopisto-opetusta. — — palvelukset ovat resurssien sallimissa rajoissa kaikkien tutkijoiden käytettävissä.

Nykyään arkisto täyttää tehtävänsä mykkien lauseidensa kautta, sillä kohta ne ovat ainoat, jotka koko arkistossa ovat läsnä: yksin toistensa seurassa, hyvässä järjestyksessä ilman yhtään tutkijaa. Johtosäännön toinen ja kolmas pykälä määrittelevät arkiston johtokunnan tehtävät. Mutta tämä johtokunta se oikea fantomi on, koska ei ole mitään »arkistoa varten perustettujen virkojen ja toimien haltijoita tai hoitajia sekä muuta arkistossa työskentelemään määrättyä henkilökuntaa», joiden toimia johtokunta ohjaisi ja valvoisi. Huvin vuoksi sopii muistuttaa, että Ikolan ja Liimolan muistio lokakuulta 1965 kaavaili perustettavaksi erityistä lauseopin tutkimuslaitosta, jossa työskentelisi tohtorin arvoinen johtaja, kaksi vanhempaa ja yksi nuorempi assistentti sekä kaksi konekirjoittajaa.

Täysin lamassa ei lauseopin arkiston toiminta sen kokoelmien valmistuttua ole ollut, sillä aineiston kimpussa on arkiston tiloissa työskennellyt suomen murteiden lauseopin tutkimusryhmä, jonka työhön prof. Osmo Ikola sai humanistiselta toimikunnalta kolmivuotisen apurahan. Tutkittavana ovat ensi sijassa olleet lauseiden liittyminen toisiinsa sekä erilaiset upotteet ja muut lauserakenteet ja niihin kuuluva nominaalimuotojen käyttö. Tältä alalta tutkimusryhmällä on käsikirjoituksena valmiina noin 600 liuskan laajuinen lauseopin katkelma. Kun tämä lyhytaikainen hanke kuluvan vuoden lopussa päättyy, valtaa lauseopin arkiston kokoelmat hiljaisuus.

Voi ehkä päätellä, että aikoinaan edistysellinen arkistointijärjestelmä koituu lauseopin arkiston kohtaloksi. Sitä mukaa kuin tieturit yhä kehittyvät ja valtaavat alaa tutkimuslaitoksissa, jopa kodeissa ja tutkijankammiois-

sa, on helppo sanoa, ettei mitään määräpaikkaisia arkistoja enää tarvita, koska kokoelmat ja niitä hyödyntävät tietokoneohjelmat on kovin vaivatonta sellaisinaan kopioida ja siirtää mihin maailman kolkkaan hyvänsä. Tällainen puhe tulee vastasyntyneiden suusta. Turun yliopiston osalta tuollainen lyhytnäköinen ajattelu merkitsee sitä, että Paavo Ravilan professorikautena alkanut ja Osmo Ikolan koko pitkän professorikauden vahvana jatkunut lauseopin tutkimusperinne olennaisesti näivettyä taloudellisten edellytysten niukkuuteen. Tähänkin vaiheeseen kätkeytyy dramatiikkansa. Suomen kansankielen muoto- ja lauseopillisen keruun ja arkistoinnin tutkimussopimuksen allekirjoitti valtion humanistisen toimikunnan puolesta juuri Paavo Ravila. Ympyrä sulkeutuu.

Ei puuttunut luovuutta tieteenalamme pioneereilta, kun he sata vuotta sitten ideoivat murteiden sanakirjan, vanhan kirjasuomen sanakirjan ja nyky-suomen sanakirjan jättihankkeet. Ei puuttunut rohkeutta aloitteentekijöiltä eikä luottamusta päättäjiltä, kun kansankielen muoto- ja lauseopin tutkimussopimus kaksikymmentä vuotta sitten solmittiin. Mihin ovat kadonneet rohkeus ja luovuus ja luottamus nyt, kun uuden sukupolven olisi vastattava haasteisiin? Paraikaa puhutaan kovasti perustutkimuksen tukemisesta. Päättäjät on otettava sanoistaan kiinni. On ryhdyttävä aloitteihin. On tartuttava tilaisuuteen. Eihän sitä koskaan tiedä, missä puussa se piru istuu!

The Syntax Archives – 20 years old

JUSSI KALLIO

Research on the Finnish national languages, tradition and literature received a powerful boost in Finland 20 years ago, with the setting up of the National Sciences Development Programme for the period 1966–80. In order to supplement the programme's dialect projects, professors Osmo Ikola and Matti Liimola drew up a proposal for the collection and storage of syntactic material in the Finnish vernacular. The proposal was put into practice in June 1967, when the Syntax Archives were established, supported by

a four-year research agreement with the National Commission for the Humanities. The work continued to be financially supported by the Commission from 1972 to 1976, and from 1976 to 1984 under the auspices of the Domestic Languages Research Centre.

Right from the start the archive system was based on automatic data processing. The material is given a broad phonetic transcription from sound tapes, coded grammatically and written on punch cards, and finally transferred onto computer tape together with the

JUSSI KALLIO

relevant programs in order to serve as a data bank for future research. When the archive was finally ready for use, at the end of 1984, it had expanded to comprise 132 one-hour texts in local dialect, complete with morphological and syntactic analysis. In addition, as comparative material, a total of 30 equivalent sets of data have also been processed for the standard spoken and written language.

The Syntax Archives are located in the Finnish Language Department of

the University of Turku, where they are freely available to scholars. Detailed information on the archive material, the archival system and the use of the archives is given in *Lauseopin arkiston opas* ('Guide to the Syntax Archives'), edited by Osmo Ikola (Turku, 1985). Computer print-outs from the material have also been sent on request by post, but nowadays the availability of this service is restricted by the lack of permanent staff at the archives.