

Lingvistisen tiedon perusteista¹

ARTO MUSTAJOKI (Helsinki)

Tieteen yleisten pelisääntöjen mukaan väitteet, joita esitetään, on perusteltava. Tieteellinen tieto eroaa näin esimerkiksi auktoriteettitiedosta, jonka perusteeksi riittää arvostetun henkilön mielipide. Tieteellisenä ei myöskään pidetä intuitiivista tietoa, koska se perustuu yksinomaan puhujan sisäiseen näkemykseen. Miten sitten kielitieteen piirissä noudatetaan näitä tieteellisen tutkimuksen yleisiä ehtoja? Mihin viime kädessä perustuvat eri kielten kuvaukset?

Tutkiessaan kieliä lingvistit joutuvat ottamaan kantaa eri muotojen ja rakenteiden oikeisuuteen, ratkaisemaan, ovatko ne hyväksyttäviä kyseisen kielen yksiköitä vai eivät. Väitteidensä tueksi venäjän kieltä tutkiva lingvisti voi esittää mm. seuraavanlaisia todisteita. Ensiksikin hän saattaa viitata kuuluisan kielenhuoltajan akateemikko Petrovin mielipiteeseen. Toiseksi hän voi osoittaa, että tunnettu kirjailija Ivanov on käyttänyt kyseistä kielen yksikköä. Edelleen tutkijan on mahdollista vedota yleiseen kielenkäyttöön, esimerkiksi lehdistön käyttämään kieleen. Vielä hän voi tutkimuksiinsa nojaten väittää, että useimmat venäjää äidinkielenään puhuvat ihmiset pitävät kyseistä muotoa tai raken-

netta kieliopillisesti oikeana. Ja jos tutkijan äidinkieli on venäjä, hän voi vedota myös omaan kielelliseen intuitioonsa.

Edellä esitetyt lingvistisen tiedon perusteet voidaan luokitella kahdeksi ryhmäksi. Kirjailija Ivanovin kieli ja yleisesti esimerkiksi lehdistössä käytettävä kieli edustavat tuotosta, kielimateriaalia, josta muodostuu tutkimuksen *korpus*. Jos tutkija taas vetoaa akateemikko Petrovin, omaan tai kaikkien kielenkäyttäjien kielitajuun, käsitykseen siitä mikä on oikein, mikä väärin kielessä, hän menettelee periaatteessa täysin toisin kuin korpustutkija. Hän ei perusta kuvaustaan konkreettiseen kielimateriaaliin vaan sellaiseen, mikä on suoran havainnon tavoittamattomissa, eli kielelliseen *kompetenssiin*.

Ovatko korpus- ja kompetenssitutkimus sitten tieteellisiä tarkastelutapoja siinä mielessä kuin alussa määrittelin asian? Eikö ole arveluttavaa, epätieteellistä vedota kielelliseen intuitioon ja auktoriteetteihin? Pohtiessani vastausta tämänkaltaisiin kysymyksiin tarkastelen ensin näiden kahden tutkimusotteen, korpus- ja kompetenssitutkimuksen, asemaa lingvistiksessä traditiiossa yleensä ja neuvostoliittolaisessa kielitieteessä eritoten.

Kielitieteen varsinainen läpimurto tapahtui viime vuosisadalla. Tutkimuksen pääpaino oli silloin kielten menneisyyden ja sukulaisuussuhteiden selvittämisessä.

¹ Venäjän kielen ja kirjallisuuden professorin virkaanastujaisesityelmä Helsingin yliopistossa 2. helmikuuta 1983.

Havaintoja

Tällaisessa tutkimuksessa kaiken perustana on luonnollisista syistä tietty korpus, ne kirjalliset muistomerkit, jotka ovat tutkijan käytettävissä. Hän voi tosin mennä päätelmissään tämän primaariaineiston ulkopuolelle, ja näin tapahtuukin. Kuitenkin hänen deduktiivisen ja induktiivisen päättelynsä lähtökohtana on loppujen lopuksi tuo aineisto. Näin ollen diakronista, kielen historiaan kohdistuvaa tutkimusta tulee pitää tyypillisenä korpustutkimuksena.

Tämän vuosisadan alussa kielitieteessä tapahtui voimakas murros, joka liitetään yleensä sveitsiläisen lingvistin Ferdinand de Saussuren nimeen. Hän korosti diakronisen ja synkronisen tutkimuksen eroa. Synkronisessa, nykykieleen kohdistuvassa tarkastelussa tutkijalla on käytettävänä ei ainoastaan kielenpuhujien tuottama kielimateriaali vaan myös heidän kielellinen tietonsa, kompetenssinsa.

Kun pohditaan korpuksen ja kompetenssin asemaa kieltä koskevassa tutkimuksessa, ei voida sivuuttaa toista de Saussuren dikotomiaa, *langue*—*parole*-erottelua (käytän jäljempänä »suomennettuja» muotoja *langi* ja *parolli*). Vaikka tutkijat edelleenkin kiistelevät siitä, mitä de Saussure näillä käsitteillä oikein tarkoitti, ei liene väärin luonnehtia niitä seuraavin epiteetein: *langi* on sosiaalinen ilmiö, *parolli* individuaalinen; *langi* muodostaa systeemin, *parolli* ei; *langi* on kielen ilmene-mismuotona pysyvä, primaari, *parolli* puolestaan enemmän tai vähemmän satumanvarainen, toissijainen. Vaikka *langin* ja *parollin* määritelmät ovat siis edelleen jonkin verran diffuusit, näillä käsitteillä on varsin keskeinen sija lingvistisissä teorioissa.

Neuvostoliittolaisessa kielitieteessä mainittu dikotomia on ollut vilkkaan, osin kii-vaankin keskustelun kohteena. Jo 1920-luvulla maineikas lingvisti Štšerba otti kantaa esittämällä kolmijakoisen kielen aspektien malliin. Hänen mukaansa erotetaan *puhetoiminta* (on sama kuin puhumis- ja ymmärtämisaktit), *kielisysteemi* (on periaatteessa kaikkien puheaktien perusteella muodostunut päätelmä kielen rakenteesta)

ja *kielimateriaali* (on kaikki jossakin kieli-yhteisössä tiettyinä aikana puhuttu ja ymmärretty).

Kiintoisa on myös Aleksejevin äskettäin esittämä konseptio. Hän liittää *langi*—*parolli*-dikotomiaan eräitä lisäkäsitteitä: uusuksen, normin ja systeemin. *Uusus* on se miten yleensä puhutaan, *normi* se miten tulee puhua ja *systeemi* se miten periaatteessa voidaan sanoa. Uusuksen, joka siis on statistinen yleistys kielimateriaalista, Aleksejev sijoittaa *parolliin*; *normi* ja *systeemi* puolestaan ovat hänen mukaansa *langin* alakäsitteitä.

Näin olemme sitoneet *langin* ja *parollin* käsitteet kompetenssi—korpus-problematiikkaan. *Langissa* tutkija selvittää kielen normia, sitä mikä on kielessä oikein, mikä väärin; silloin hän nojautuu ensisijaisesti kompetenssiin. *Parollissa* taas on kyse uusukselta, siitä miten todellisuudessa puhutaan tai kirjoitetaan, ja se käy tietenkin parhaiten ilmi kielimateriaalista, korpuksesta.

Langin ja *parollin* vaiheilla käytävän keskustelun keskeisyyttä neuvostoliittolaisessa lingvistiikassa kuvastaa se, että vuonna 1962 järjestettiin pelkästään tätä ongelmaa käsittelevä konferenssi. Periaatteellista pohdintaa on ollut erityisesti siitä, tuleeko *langin* ja *parollin* tutkimus katsoa kahdeksi kokonaan erilliseksi tieteenalaksi vai voidaanko asia tulkita siten, että kyseessä on vain yhden asian kaksi eri puolta. Useimmat tutkijat ovat nykyään sitä mieltä, että ei ole mielekäästä puhua kahdesta erillisestä tutkimuskohteesta vaan yhdestä objektista, jolla on kaksi erilaista olomuotoa.

Onko tällaisessa pohdinnalla sitten käytännön merkitystä? Onko tutkimuksen tulosten kannalta relevanttia se, mihin tutkija väitteensä perustaa, kompetenssiin vai korpukseen? Monesti on niin, että suurtaakaan ristiriitaa ei eri tavoin saatujen tulosten välillä ole. Kuitenkin on huomattava, että juuri niissä tapauksissa, jotka ovat kuvauksen kannalta kaikkein kiintoisimpia ja kiistaisimpia, kielimateriaali ei aina vastaa kielenkäyttäjien käsitystä kielen yksiköiden oikeisuudesta. Merkittävä on myös se

seikka, että kaikki, mikä kielessä mieles-tämme on oikein, ei välttämättä esiinny puheessamme. Normin ja uusuksen selvää periaatteellista eroa himmentää tosin se, että kirjallinen viestintä yhtenä uusuksen lajina noudattaa varsin pitkälle kielitajun norminkäsitystä, mutta olisi väärin kuvitella, että kirjoitettu kielimateriaali olisi erehtymätön kompetenssin kuvastin.

Teoreettisesti on siis niin Neuvostoliitos-sa kuin muuallakin kiinnitetty vakavaa huomiota käsitepareihin langi—parolli, kompetenssi—korpus, mutta kuinka on yksityisten tutkimusten osalta? Usein, varsinkaan kielten kokonaisesityksissä, kirjoit-tajat eivät mainitse sitä, mihin kuvaus viime kädessä perustuu. Tällainen epä-tarkkuus haittaa myös tieteellistä keskuste-lua. Kun kaksi russistia kiistelee siitä, tulisi-ko venäjän kieliopissa mainita muoto *mnoga apel'sin* vai *mnogo apel'sinov*, ei väitte-lyssä ole paljonkaan mieltä, jos toinen pu-huu uusuksesta, toinen normista.

Eräissä kieliopissa sen sijaan mainitaan selvästi, mihin kuvaus perustuu. Kaikkein painokkaimmin kompetenssintutkimuk-seksi on tunnustautunut chomskylainen generatiivi kielioppi. Sen tarkoituksena on antaa säännöstö, jonka avulla voidaan tuottaa periaatteessa kaikki kieliopillisesti oikeat lauseet ja vain ne. Tämäntapaisen, joskus autonomiseksi luonnehditun kielitieteen periaatteita pohdittaessa on myös herännyt kysymys kielitieteen empiirisyydestä yleensä.

Generatiivia tutkimusta harrastetaan myös Neuvostoliitossa; siellä on eräitä omiakin oppi-isiä. Kuitenkin perinnäiselle venäläiselle kielitieteelle on vierasta kysei-selle teorialle ominainen kompetenssikes-keinen ajattelupata. Tätä kuvaavat hyvin maassamme äskettäin vierailleen N. Ju. Švedovan, kahden tuoreimman akateemi-sen kieliopin päätoimittajan, mielipiteet. Kun häneltä kysyttiin, mihin kieliopissa esitettävät väitteet perustuvat, hän vastasi yksiselitteisesti: kortistoon, siis kielimateri-aalista kerättyyn korpukseen. Švedova ko-rosti, että hänellä ei ole oikeutta panna omiaan, luottaa omaan kielitajuunsa; kaikki esimerkit on pystyttävä näyttämään

tosiksi kortiston avulla. Sen verran hän kuitenkin antoi periksi korpuksen yksin-valtiudesta, että hyväksyi kortistossa ole-vien ei-normatiivien esimerkkien karsin-nan intuition perusteella.

Tutkimuksen lähtökohdan valinta — korpus vai kompetenssi — ratkaisee vain yhden kielitieteen periaateongelmista. Toinen, yhtä visainen on kysymys, miten tutkijan tulee valita korpus tai miten hän pääsee käsiksi kielenpuhujien kompetenssiin.

Palaan alkupuolen esimerkkeihin. Kor-pustutkimuksesta mainitsin kaksi menette-lytapaa: otetaan tarkasteltavaksi joko tie-tyin perustein valittu materiaali — esi-merkiksi »tunnustetulta kirjailijalta» — tai sitten teoriassa mikä tahansa tietyt yleiset kriteerit täyttävä kielellinen tuotos riip-pumatta sen tuottajasta. Edellistä voisi kutsua *valikoidun*, jälkimmäistä *valikoimat-toman uusuksen* tutkimukseksi. Neuvostoliit-ossa käytetään monesti valikoitua uususta; suosittua on vedota tunnetuilta kirjaili-joilta poimittuihin sitaatteihin.

Hyvä esimerkki kahdesta erilaisesta suhtautumistavasta on Neuvostoliitossa ja Suomessa tehdyt sanojen yleisyyttä selvit-tävät tutkimukset. Venäjän kielen frek-venssisanakirjan esipuheessa mainitaan, että otokset on poimittu sellaisilta kirjoit-tajilta, jotka käyttävät mahdollisimman hyvää venäjän kieltä. Vastaavaa suomen kielen taajuussanastoa laadittaessa ei täl-laisia normatiivisuuskriteerejä ole käytet-ty vaan materiaali on valittu satunnais-otannalla kaikesta mahdollisesta kielima-teriaalista. Valikoidun kielimateriaalin käytölle on perusteluna se, että lingvisti tutkiessaan parollia, uususta saa tällä ta-voin välillisesti tietoa myös langista, nor-mista.

Kompetenssin tutkimisesta oli esillä kolme vaihtoehtoa: akateemikko Petrov, lingvisti itse sekä kielenpuhujien kollektiivi. Pelatkaamme ensin akateemikko pois tarkastelusta. Kielenhuoltajana hän edus-taa selvästi erilaista suhtautumistapaa kuin muut: hän antaa ohjeita kielenkäyttäjille, selvittää, mikä on oikein ja suositelta-va. Tällainen tarkastelutapa on meillä

joskus leimattu epätieteelliseksi tai ainakin lingvistiikan piiriin kuulumattomaksi. Näin ei toki ole. Pikemmin kyseessä on yksi soveltavan kielitieteen hyvin keskeinen ja merkittävä lohko. Neuvostoliitossa kielenhuoltoon tai kuten siellä sanotaan puhekulttuurin vaalintaan kiinnitetään aivan erityistä huomiota. Jätän kuitenkin kielenhuoltajien antamaa *virallista normia* koskettelevat ongelmat tämän tarkastelun ulkopuolelle.

Jäljelle jää kaksi kompetenssintutkimuksen periaatetta. Kummassakin tapauksessa tavoitteena on nähtävästi saada tietoa *kollektiivisesta normista*. Asia voidaan nimittäin tulkita siten, että myös omaan kielelliseen intuitioonsa tukeutuva tutkija haluaa kuvata muutakin kuin vain henkilökohtaista murretaan, idiolektiaan: hän luottaa siihen, että hänen oma kielitajunsa kuvastaa riittävän hyvin koko kieliyhteisön käsitystä kielen yksiköiden oikeisuudesta. Myös sellainen apukäsite kuin »ideaalinen kielenpuhuja homogeenissa ympäristössä» kätkee todellisuudessa taakseen tutkijan oman persoonan eli Chomskyn itsensä.

Kun tutkija luottaa vain omaan kielelliseen intuitioonsa, on vaarana tietysti se, että kieltä koskevat käsitykset ovat individuaalisia, muiden käsityksistä poikkeavia. Ainakin periaatteessa kompetenssintutkimuksen tulisi kaiketi perustua koko kieliyhteisön näkemykseen.

Kiintoisinaan menettelytapaan ovat päätyneet laajan kyselytutkimuksen neuvostoliittolaiset tekijät. Tutkimuksessa selvitetiin ihmisten kanta joihinkin venäjän kielen ääntämisen, muoto-opin ja lauseopin variantteihin. Valitessaan noin 4 000 informanttia tutkijoilla oli ensinnäkin seuraavat kolme kriteeriä: äidinkielen tuli olla venäjä, asuinpaikan kaupunki ja vastaajalla tuli olla vähintään keskiasteen koulutus. Näin kaikki maalla asuvat ja vain peruskoulun käyneet kaupunkilaiset putosivat pois mahdollisten koehenkilöiden joukosta. Tämän jälkeen vahvistettiin vielä kullekin yhteiskuntaryhmälle omat kiintiöt: lukeneisto miehitti koehenkilöpaikoista 40 %, opiskelijat 33 %, toimihenkilöt 18 % ja työläiset 9 %. Ryhmien suhteelliset

osuudet eivät siis lainkaan vastanneet niiden todellista osuutta väestöstä.

Tällainen menettely saattaa länsimaisesta tutkijasta tuntua epädemokraattiselta. Miksi kaikkien mielipide ei paina yhtä paljon? Kun asiaa pohtii syvemmalta, huomaa kuitenkin, että koehenkilöiden valintakriteerit ovat hyvin samantapaiset kuin lingvistien meilläkin monesti suosimat. Ero on pikemmin koehenkilöiden määrässä eikä laadussa: sen sijasta että tutkija turvautuisi oman kielitajunsa lisäksi kollegoidensa ja tuttaviansa muodostamaan rajalliseen otokseen kuten usein tapahtuu, mainitussa tutkimuksessa selvitetiin suuremman ihmisjoukon näkemys asiasta. Mielipiteen takana ei siis tässä tapauksessa ole joku yksityinen auktoriteetti mutta kyllä tietyllä tavalla valikoitu informanttijoukko. Voitaisiin kai puhua *painotetusta kollektiivisesta normista*.

Palaan alussa esittämiini kysymyksiin lingvistisen tiedon tieteellisyydestä. On käynyt selväksi, että kielitiede tutkii ensinnäkin kieliyhteisön tuottamaa kielimateriaalia. Se voidaan tietysti mielessä rinnastaa esimerkiksi luonnontieteiden tutkimuskohteisiin. Tällainen tutkimus ei kaiketi herätä vastaväitteitä tieteellisyyskriteerin suhteen. Perustuvathan väitteet selvästi havaittavissa olevaan, tutkijasta riippumattomaan materiaaliin, josta voidaan tarvittaessa antaa kiistattomia esimerkkejä ja myös numeerista tietoa.

Toisin on laita kompetenssintutkimuksessa. Siinä väitteet pohjautuvat viime kädessä kielitajuun, intuitioon, joka katsottiin riittämättömäksi tieteellisen tiedon perustaksi.

Ongelmallista on myös se, että tutkijan on mahdotonta saada välitöntä tietoa kieliyhteisön kollektiivisesta kielitajusta. Merkitseekö tämä sitten sitä, että tällaista tutkimusta tulee pitää epätieteellisenä ja siksi epäsuotavana? Tämäntapainen päätelmä olisi kohtalokas. Kieli-nimiseen ilmiöön liittyy nimittäin olennaisena osana normin käsite. Tämän seikan huomiotta jättäminen sulkisi tutkimuksen ulkopuolelleen sen, mikä kielessä on primaaria.

Tieteen tehtävänä on tutkia kaikkia il-

miöitä, myös niitä, jotka ovat välittömän havainnon tavoittamattomissa. Oikeaksi kielellistä kompetenssia koskevat väitteet tulevat, jos niille todella löytyy vastine ihmisten tajunnassa. Kielitieteen yleisen teorian tulee selvittää, millä keinoin voimme riittävän tarkasti ja riittävän objektiivisesti luoda valoa tähän lingvistien mustaan laatikkoon.

✓

✓