



KYSELYTESTIT MERKITYKSEN TUTKIMISESSA

Ulla Vanhatalo *Kyselytestit synonymien selvittämisessä. Sanastotietoutta kielenpuhujilta sähköiseen sanakirjaan.* Helsinki: Yliopistopaino 2005. ISBN 952-91-8645-2 (nid.), ISBN 952-10-2438-0 (pdf).

Tutkimusmetodin valinta on tieteessä aina tärkeää, mutta ehkä erityisesti merkityksen tutkijan täytyy pysähtyä pohtimaan metodien käyttöä ja kehittämistä: eiväthän merkitykset ole sellaisinaan näkyvissä missään luonnollisen kielen käyttöä sisältävissä aineistoissa, vaan ne pitää aina tulkita. Nykyisin jokaisen tutkijan käytössä on tietysti erilaisia korpuksia, mutta lopullinen päätös siitä, mitä kielellinen ilmaus merkitsee, perustuu kuitenkin tutkijan intuitioon. Voidaanko siihen luottaa — ja jos voidaan, niin kuinka pitkälle? Tätä kysymystä eivät merkityksen tutkijat ratkaisse aivan lähitulevaisuudessa.

Kaikilla kielenpuhujilla on käsityksiä kielestä, esimerkiksi leksikaalisista merkityksistä. Tämä tarkoittaa, että kielentutkijan on periaatteessa mahdollista kysyä myös maallikolta, mitä jokin kielellinen ilmaus merkitsee. Vastausten tulkinta ei kuitenkaan ole aina helppoa, koska kielenpuhujan intuitiot voivat näyttää kaoottisilta; niitä ei niin vain pysty käyttämään systemaattisessa tutkimuksessa. Asiat, joita semantiikan tutkija oikeasti haluaa tietää, ovat liian vaikeita ei-lingvistin suoralle intuitiolle. Kielenkäyttäjän monimutkaiset intuitiot on siis jollakin tavalla saatava yksinkertaiseen muotoon. (Ks. myös Cruse

1986: 10–15.) Tämä ongelmakenttä oli vahvasti esillä esimerkiksi tammikuussa 2004 Helsingissä Pohjoismaiden kielitieteen konferenssissa »On the necessity of experimental methods in semantics», jonka järjestivät Jarno Raukko ja Ulla Vanhatalo.

Ulla Vanhatalolta on nyt valmistunut väitöskirja *Kyselytestit synonymien selvittämisessä*, joka käsittelee juuri sitä, miten syntyperäisen puhujan luonteeltaan ei-systemaattisia tietoja erityisesti synonyymien merkityksistä olisi mahdollista selvittää ja hyödyntää. Tutkimus keskittyy ensisijaisesti intersubjektiiviseen kyselytestimenetelmään, jolla tarkoitetaan syntyperäisiltä puhujilta kyselyllä kerätyn tiedon systemaattista käyttämistä semantiikan tutkimisessa (ks. myös esim. Raukko 2002).

Tiedon sovellettavuus, erityisesti leksikografiassa ja toisen kielen opettamisessa, on kirjoittajalle hyvin tärkeää; sitä korostetaan monessakin väitöskirjan kohdassa.

Väitöskirja koostuu johdannosta ja kuudesta artikkelista. Artikkeleista neljä esittelee kukin yhden testisarjan ja sen tuloksia. Artikkelit keskittyvät erityisesti metodologiaan, mutta esittelevät tietysti myös saatua leksikaalista tietoa. Kaksi viimeistä artikkelia pohtii kyselytestien tulosten käyttöä elektronisissa sanakirjoissa.

Artikkelien yhteinen johdanto on perusteellinen, ja se onkin itse asiassa enemmän kuin pelkkä kooste. Aluksi Vanhatalo pohdii synonymiaa semanttisena ilmiönä ja käsittelee siihen liittyviä erilaisia näkökulmia. Synonymia on yksi leksikaalisen semantiikan vaikeimmista kategorioista: vaikka jo aika nuorilta koululaisilta odotetaan, että he pystyvät tuottamaan äidinkielen tunneilla synonyymejä, lingvistiikan sisällä vallitsee varsin kaoottinen yleiskuva siitä, miten synonymia pitäisi määritellä. Vaikka kaikkia tutkijoita tyydyttävää yksiselitteistä synonymian määrittelyä ei ole pystytty luomaan, Vanhatalon mielestä tärkeimpiä ovat (muun muuassa ei-syntyperäisen) kielenkäyttäjän tarpeet: sanakirjan on pystyttävä selittämään, mitä eroa synonyymeillä on syntyperäisen puhujan mielestä. Miten lopulta on mahdollista vetää raja synonyymien ja ei-synonyymien väliin? Tämä ei tietenkään ole ongelma kielenkäyttäjän kannalta vaan pikemminkin sanakirjan ja erilaisten testien laatijan näkökulmasta. Vanhatalo myöntää, että synonymia on asetteista ja jokaisessa tutkimuksessa on mietittävä erikseen, millaisia sanoja siinä käyttää.

Toinen luku käsittelee testimenetelmiä. Kyselytestimenetelmä on selvästi yksi ratkaisu merkityksen subjektiivisuuden ongelmaan. Koska moni ihminen ei voi olla samalla tavalla väärässä, kyselytestien voi olettaa tuovan esiin kieliyhteisön yhteistä tietoa merkityksistä. Väitöskirjassa esittää sekä »koviensy»psykologististen että »pehmeiden» semanttisten kokeiden periaatteet ja lyhyt historia. Kolmas luku esittelee johdonmukaisesti alkuperäisartikkeleita. Taulukossa 1 sivulla 36 on esitetty tärkeimmät tulokset ja lueteltu kirjoittajan suorittamat testit.

Artikkeli (1) (»Naiset motkottaa aiheesta ja nalkuttaa syyttä: kyselytestit verbien semanttisten sisältöjen arvioinnissa») käsit-

telee kielteisen tunteen ilmaisevien puheaktiverbien merkityksiä. Informanteilta on eri kysymyksillä kerätty tietoa siitä, mitä oikeasti tarkoitetaan, kun jotakin puhetoimintaa sanotaan esimerkiksi *jurnuttamiseksi*, *valittamiseksi* tai *jäkättämiseksi*. Suomalaisilta lukiolaisilta on kysytty esimerkiksi, onko puhuja mies vai nainen, nuori vai vanha, kuinka vihainen hän on ja kuinka kovalla äänellä puhuu. Tieto, joka näin on saatu, puuttuu vielä sanakirjoista, mutta on täysin selvää, että se on tarpeellista sekä suomen kielen semantiikan tutkimisen että suomen kielen käyttäjän näkökulmasta.

Artikkeli (2) (»Kyselytestit vs. korpuslingvistiikka sanaston semanttisten sisältöjen arvioinnissa: mitä vielä *keskeisestä* ja *tärkeästä*?») vertailee korpustutkimuksen ja kyselytestimenetelmän tuloksia synonyymien *keskeinen* ja *tärkeä* (jossain määrin myös *merkittävä* ja *olennainen*) osalta. Juuri nämä sanat tuottavat ongelmia myös käännettäessä esimerkiksi suomesta viroon ja päinvastoin. Kyseessä siis on kielenkäytön kannalta vaikea sanapari, jota sanakirjat eivät riittävästi käsittele. Informanteilta on kysytty esimerkiksi sitä, kuinka luonteelta kuulostavat erilaiset lauseet, joihin tutkittavat sanat sisältyvät. Tuloksia on verrattu korpustutkimuksella saatuun tietoon (ks. Jantunen 2001).

Eniten kysymyksiä herättää kollokaation käsittely tässä artikkelissa. Kollokaatio on selvästi tilastollinen ilmiö, ja sitä voi selvittää korpustutkimuksessa. Kyselytestimenetelmän tulosten tulokinnassa kollokaateista puhuminen on kuitenkin jokseenkin ongelmallista: eihän tutkija voi tietää, millaisten sanojen perusteella informantti on arvioinut lausetta. Kirjoittaja on myös itse monessa kohdassa korostanut epäilyjään (esim. s. 361) siitä, mikä kollokaatti juuri on voinut vaikuttaa, mutta aika suuri osa artikkelia perustuu kuitenkin juuri näihin tulkintoihin. Tämä on ehkä ylipäänsä lau-

setta käyttävien semanttisten kokeiden suurin ongelma: lauseen arvioinnissa voi olla monta tekijää, jotka vaikuttavat informantin päätökseen. Toisaalta juuri lause näyttää olevan se kielen yksikkö, jota ei-lingvisti pystyy parhaiten arvioimaan.

Artikkeli (3) (»Kas ja kuidas aitab eksperimentaalsemantika võrrelda lähisünonüümide tähendust ning kasutust kahes keeles?») lisää väitöskirjaan kontrastiivisen näkökulman. Kyselytesteillä on tässä artikkelissa selvitetty virolaisten lukiolaisten intuitioita kielteisen tunteen ilmaisevien puheaktiverbien merkityksistä virossa. Tuloksia on verrattu artikkelissa (1) esitettyihin suomalaisten lukiolaisten arvioihin samasta puheaktiverbien ryhmästä. Metodina tällainen lähestymistapa on erittäin tervetullut, koska varsinkaan kontrastiivisessa analyysissä ei saisi luottaa vain tutkijan intuitioihin: eihän tutkija pääsääntöisesti ole itse kaikkien tutkittavien kielten syntyperäinen puhuja. Vaikka tulosten analyysissä onkin useita ongelmia, joita kirjoittaja luettelee artikkelissaan, tämä tutkimus on kuitenkin hyvin olennainen edistysaskel kontrastiivisen kielentutkimuksen metodikassa.

Artikkeli (4) (»Koehenkilöiden taustan merkitys kokeellisessa leksikaalisessa semantiikassa ja sen sovelluksissa») käsittelee koehenkilöiden valintaa. Kirjoittaja on taas testannut syntyperäisten puhujien intuitioita eri sanaryhmien kohdalla (puheaktiverbit, *edullinen* – *halpa* ym.), mutta informantteina on tällä kertaa käytetty iältään ja muulta taustaltaan toisistaan eroavia ryhmiä. Informanttien mahdollisista taustatekijöistä mielenkiintoisin on kielellisen aktiivisuuden indeksi, jota ei tähän asti ole kovin paljon otettu huomioon kielenkäyttäjän sosiolingvistisessä kuvassa. Kyselytestimenetelmän käyttäjät ovat varmasti iloisia ar-

tikkelin tuloksista: koehenkilönä voi olla kuka tahansa äidinkielen kirjoitustaitojen¹ ihminen. Herää vain kysymys, miksei kuitenkaan lingvisti?

Artikkeleissa (5) (»Population testing: Extracting semantic information on near-synonyms from native speakers») ja (6) (»Finnish electronic dictionaries: Present and future challenges») käsitellään kyselytestimenetelmällä saatujen tulosten käyttöä elektronisessa sanakirjassa ja esitellään uutta sanakirjamallia (»uusi uljas sanakirja»). Sen jälkeen, kun lingvisti on saanut kyselytesteillä (tai muulla menetelmällä) uutta tietoa sanojen merkityksistä, nämä tiedot on saatava siihen muotoon, että niistä olisi hyötyä myös muille kielenkäyttäjille. Hyvin mielenkiintoinen on yritys esittää merkityksiä kolmiulotteisena kuvana, joka syntyy kyselytesteillä saatujen tulosten pohjalta. Valitettavasti painetussa muodossa ei ole ihan selvää, millaiselta tällainen kuva oikeasti näyttää.

Vanhatalo käsittelee siis eri näkökulmista ja eri metodien avulla eniten puheaktiverbejä mutta myös adjektiiviryhmiä *tärkeä*, *keskeinen*, *olennainen* ja *merkittävä*; *halpa* ja *edullinen*; *ihmeellinen*, *kummallinen*, *omituinen* ja *outo*. Vaikka kirjoittaja itse onkin pitänyt uutta leksikaalista tietoa sivutuotteena, se on silti arvokas osa väitöskirjaa. Hiukan epäselväksi kuitenkin jää, miksi juuri nämä sanaryhmät on valittu tutkimuksen kohteeksi. On selvää, että kyse on suomen (ja viron) monimutkaisista merkityskentistä, mutta miksi juuri niistä? Yhdeksi vastaukseksi kirjoittaja tarjoaa näiden sanojen konteksti-itsenäisyyttä, jolla tarkoitetaan ilmeisesti tulkinnan itsenäisyyttä, siis sitä, että informantit pystyvät arvioimaan näiden sanojen merkityksiä ilman laajempaa kontekstia.

¹ Ehkä tämäkään ei ole välttämätön kriteeri; voisihan testejä tehdä myös puhutussa muodossa.

Kokonaisuutena Ulla Vanhatalon väitöskirja antaa uutta tietoa semantiikan tutkimisen metodeista: teoreettiselle semantiikalle synonymiasta ja käytännönläheisemmälle sanastontutkimukselle usean sanaryhmän merkityksistä. Olisin odottanut enemmän ohjeita kyselytestin laatimiseen, kun kerran kyseessä on metodia edistävä tutkimus. Myös ilman tällaista osiota kirjasta on kuitenkin hyötyä jokaiselle, joka aikoo käyttää kyselytestejä omassa tutkimuksessaan.

Vanhatalon kirjasta näkyy selvästi kirjoittajan huolestuneisuus semantiikan ja leksikografian kohtalosta nykyisessä tietoyhteiskunnassa: uutta tietoa tulee paljon, mutta sen saaminen sellaiseen muotoon, että myös kielen syntyperäinen tai ei-syntyperäinen käyttäjä voisi päästä siihen käsiksi esimerkiksi sanakirjan avulla, on ihan eri asia. Laina on palautettava: kielenkäyttäjän intuiot, jotka on saatu kyselytestien

avulla, on saatava systemaattisessa, mutta ei liian monimutkaisessa muodossa takaisin maallikon käyttöön. Toivottavasti uusi uljas sanakirja on hyvä malli sellaiseen palautukseen. ■

RENATE PAJUSALU

Sähköposti: *renate.pajusalu@ut.ee*

LÄHTEET

- CRUSE, D. A. 1986: *Lexical semantics*. Cambridge: Cambridge University Press.
- JANTUNEN, JARMO H. 2001: »Tärkeä seikka» ja »keskeinen kysymys»: Mitä korpuslingvistinen analyysi paljastaa lähisyntyisistä? – *Virittäjä* 105 s. 170–192.
- RAUKKO, JARNO 2002: Pitämisen polysemia: miten koehenkilöt hahmottavat *pitää*-verbin merkitystyyppensä. – *Virittäjä* 106 s. 354–374.

VANHOJA ETYMOLOGIOITA UUSISSA KANSISSA

Kaisa Häkkinen *Nykysuomen etymologinen sanakirja*. WSOY 2004. 1633 s. ISBN 951-0-27108-X.

Kaisa Häkkinen *Nykysuomen etymologinen sanakirja* on kolmas mittava suomalainen etymologinen sanakirja, *Suomen kielen etymologisen sanakirjan* (SKES:n) ja *Suomen sanojen alkuperän* (SSA:n) seuraaja. Samalla kyseessä on ensimmäinen laaja populaari esitys suomen sanaston alkuperästä. Teos on paljon laajempi ja käyttökelpoisempi kuin Häkkinen aikaisemmin toimittama *Etymologinen sanakirja* (1987), mutta toisaalta tavoitteiltaan SSA:ää ja SKES:aa vaatimattomampi yksinenteinen teos.

Nykysuomen etymologinen sanakirja (tästä eteenpäin NSES) on myös ensimmäinen suomalainen yhden henkilön kirjoittama kattava etymologinen sanakirja. Vaikka Euroopassa etymologiset sanakirjat ovat yleisesti yhden tai muutaman kirjoittajan työtä, on uralistiikassa syystä tai toisesta muodostunut traditioksi, että etymologiset sanakirjat tehdään toimituskunnissa. Tämä ei kuitenkaan välttämättä ole taannut työn laatua. Siksi on huomattavaa, että yksittäinen tutkija on ottanut tehtäväkseen kokonaisen suomen etymologisen sanakirjan