

Miksi tieteen kieli on epätarkkaa?

Arto Mustajoki

Esityksen tavoitteena on todistaa vääräksi seuraavat kaksi myyttiä:

- Tieteen kieli on täsmällistä ja tarkkaa
- Tiede pyrkii mahdollisimman täsmälliseen ja tarkkaan kielelliseen ilmaisuun

Aloittakaamme asian tarkastelu arkikieles-tä. Monitulkintaisuus (ambiguiteetti) on in-himillisen puheen ja tekstin perusominaisuuksia. Kun sanomme *Tien vieressä seisoi vanha Volvo ja Fiat* voimme tarkoittaa, että vain Volvo on vanha tai että sekä Volvo että Fiat ovat vanhoja. Lauseet (1-3) osoittavat, kuinka taustieto maailmasta ratkaisee tul-kinnan:

- 1 John studies Chinese language and literature.
- 2 John studies Chinese language and biology.
- 3 John studies Chinese language and philosophy.

Maailmankuvaamme nojaten tulkitsemme, että John opiskelee kiinalaista kirjallisuutta mutta ei kiinalaista biologiaa. Filosofian kohdalla käsitykset saattavat poiketa toisistaan, koska kiinalainen filosofia on mahdollinen käsite, toisaalta kieli ja filosofia eivät muodosta tavallista aineyhdistelmää.

Genetiivi on monessa kielessä monitulkintainen. Klassisia suomenkielisiä esi-

merkkejä ovat *metsästäjän ampuminen ja kanojen syöminen*, jotka mahdollistavat genetiivin viittaavan sekä objektiin että subjektiin. Lauseella *Missä se Liisan valokuva on?* on ainakin kolme tulkintaa: Liisan otta-ma kuva, Liisan omistama kuva tai kuva joka esittää Liisaa. Sama pätee vastaaviin lauseisiin muilla kielillä.

Monitulkintaisuutta käytetään laajasti hyväksi iskulauseissa ja mainoksissa: *Säästä ajoissa, Niemi muuttaa kaiken* (muuttofirman mainos). Joskus yllättävätkin sanat voivat mennä aidoissa puhetilanteissa sekaisin, kuten sanat *vaja* ja *vajaa*. Kun kesämökillä puoliso sanoo viinipullosta ”Mä vein jo sen vajaan”, väärintulkinta on täysin mahdollinen. Kielifilosofit näkevät mahdollisuuden yllättäville tulkinnoille silloinkin kun arkikeskustelussa tällainen ei tule mieleen. *Ursulalla on kolme lasta* tulkitaan yleensä yksiselitteisesti: tasan kolme lasta. Loogisesti tarkasteltuna lause voi tarkoittaa ”kolme tai enemmän”.

Kun tarkastellaan lausumien propositionaalisen sisällön ohella niiden pragmaattisia funktioita, päädytään helposti tulkintavyyyhteihin, joihin puhuja ja kuulija voivat sotkeutua pahemman kerran. Olettakaamme, että mies ja vaimo ovat sopineet lähtevänsä aamulla mökille, jos ei sada. Olettakaamme edelleen, että vaimo kovasti odot-

taa pääseväsä mökille, mies sen sijaan ei ole niin innostunut lähtemään. Kun sitten aamu koittaa, mies nousee sängystä, menee ikkunan ääreen, katsoo ulos ja sanoo "Ulkona sataa", lauseeseen liittyy suuret määrät mahdollisia tulkintoja ja tunnelatauksia. Ensinnäkään ei ole välttämättä selvää, miten sade määritellään: onko muutaman pisaran tihuttaminen jo sadetta. Toiseksi tärkeä merkitys on äänensävyllä: siitä ei saa välittyä tyytyväisyys tai vahingonilo. Väärä tulkinta tai äänensävy voi johtaa perhedraamaan.

Esimerkiksi sanojen epätarkkuudesta sopii hyvin kansallisuuksien nimitykset. Lauseen (4) suomalaisilla tarkoitetaan ilmeisesti Suomessa asuvia ihmisiä, joiden kahvinhimo näkyy tilastoissa. Näin ollen lauseen (5) suomalaiset ovat ihan eri ihmisiä. Seuraavan lauseen yleistys koskee lähinnä muutamaa Suomea edustavaa urheilijaa; suomalaisten Matti Meikäläisten hypytaidoista se ei kerro mitään. Vastaavasti lauseiden (7–10) takana on varsin erilaisia Suomi-tulkintoja.

- 4 Suomalaiset juovat kahvia eniten maailmassa.
- 5 Suomalaisia asuu Australiassa muutama tuhat.
- 6 Suomalaiset ovat hyviä mäkihyppääjiä.
- 7 Suomi kuuluu Skandinaviaan.
- 8 Suomi teki viisaasti kun siirtyi ajoissa euroon.
- 9 Suomi voitti vihdoinkin Ruotsin jalkapallossa.
- 10 Suomi taas kärjessä innovaatioissa / Pisa-testeissä

Otetaan vielä vaarattomalta vaikuttava esimerkki monitulkintaisuudesta. Lauseen *Työttömyys on Pohjois-Suomessa yleisempää kuin muualla maassa* kaikki tärkeät merkitystä kantavat osat voidaan ymmärtää useammalla kuin yhdellä tavalla: työt-

tömyydestä on erilaisia määritelmiä, Pohjois-Suomi voi tarkoittaa Lapin lääniä tai Lapin ja Oulun lääniä tai jollakin muulla tavalla rajattua aluetta, "yleisempää" voi merkitä 0,1 prosenttiyksikköä tai 5 prosenttiyksikköä.

Tarkkuusasteen ja määritelmän ongelmat tulevat hyvin esille myös vastauksessa *Kuinka pitkä matka on Tampereelle?* Tarkoitetaanko linnuntietä, rautatietä vai maantietä pitkin? Miten tarkkaa vastausta odotetaan: 176,33 kilometriä, 176 kilometriä, vähän alle 200 kilometriä vai parin tunnin ajomatka.

Edellä olevat esimerkit riittänevät todistamaan oikeaksi Ludwig Wittgensteinin (1921) toteamuksen siitä, että kielenkäyttö on aina enemmän tai vähemmän epämääräistä (vage) niin että se mitä me väitämme, ei voi koskaan olla täysin tarkkaa (völlig präzise). A. Cornelius Benjamin (*Science and Vagueness*, 1939) puki sen sanoiksi: "[---] our languages are not so precise as we should like to have them [---]".

Siirryn seuraavaksi tarkastelemaan tieteen kieltä. Päälähteenäni on Sabine Döninghausin kirja *Die Vagheit der Sprache* (Wiesbaden: Harrassowitz Verlag. 2005. 658 S.). Hänen mukaansa tieteen kielen epämääräisyys perustuu paljolti samanlaisiin ilmiöihin kuin arkikielessä, ne vain ilmenevät eri tavalla. Otan seuraavassa esille joi-takin tällaisia seikkoja.

Monitulkintaisuus (ambiguiteetti) on tietysti mahdollista myös tieteen kielessä. Esityksen alussa olleet esimerkit voidaan siirtää tieteelliseen tekstiin. Kun kirjoitamme *nopeat siittiöt ja munasolut* lukija ymmärtää että vain siittiö ovat nopeita, mutta kun kirjoitamme *harvasanaisten suomalaisten ja ruotsalaisten kohtaamisesta*, emme voi olla varmoja ovatko ruotsalaisetkin kirjoittajan mielestä harvasanaisia. Lauserakenteeseen

pohjautuva monitulkintaisuus on mahdollista siis myös tieteen kielessä, mutta se ei ole tärkein epätarkkuuden aiheuttaja.

Käsitteiden epätarkkuus on merkittävä ongelma monella tieteenalalla. Ihmistieteissä se erityinen ymmärtämistä haittaava tekijä (esimerkiksi kulttuuri-käsitteelle on ainakin sata erilaista määritelmää), mutta ei se ole vierasta luonnontieteissäkään. Diodoros Kronos (350 e.a.a.) pohti jo aikanaan harmaatukkaisuuden käsitettä: montako harmaata hiusta tällaiseen määrittelyyn tarvitaan prosenteissa tai absoluuttisesti. Aristoteles askarteli hiekanjyvien kanssa: montako niitä tarvitaan, että voidaan puhua kasasta. Lehtikasaan tarvitaan paljon vähemmän lehtiä, lautakasaan vielä vähemmän lautoja. Näin ollen osasten määrä ei yksin määrää sitä rajaa kuinka monta niitä tarvitaan kasaan. Käsitteiden ”kieli” ja ”murre” lingvistinen määritelmä on itsessään epätarkka: eri murretta puhuvat ymmärtävät toisiaan; jos he eivät ymmärrä, kyseessä on eri kielet. Määritelmä selvästikin vuotaa ainakin parista kohtaa: mitä tarkoittaa ”ymmärtää” – koskaan ei voi ymmärtää kaikkea, aina voi ymmärtää jotakin; mitä jos osa murteenpuhujista ymmärtää hyvin, osa taas ei. Kaiken lisäksi lingvistinen määritelmä jää toiseksi silloin, kun nuori valtakunta haluaa määritellä identiteettiään oman kielen kautta. Siksi meillä on nykyään erilliset kielet kroaatti, serbia ja bosnia. Eliölajien määritelmien luulisi olevan selviä, mutta niissäkin on prototyyppisiä ja perifeerisiä tapauksia, niinpä varpunen on ”enemmän” lintu kuin pingviini.

Määrittelyn tarkkuusaste voi vaihdella. Kun puhumme maantieteellisestä paikasta, voimme antaa eksaktit koordinaatit, paikakunnan, seutukunnan, läänin tai valtion. Ihmisten luokittelussa voimme tyytyä yh-

teen kategoriaan (suomalainen, venäläinen jne.) tai voimme selvittää tarkemmin jokaisen tausta, sen kuinka ”moniprosenttisesti” he kuuluvat vanhempien ja isovanhempien kautta tiettyyn kategoriaan. Puutalot, muurarahaiset ja maalit – sanoja jotka voivat esiintyä tieteellisessä tekstissä – voidaan jakaa tai olla jakamatta alaryhmiin.

Kvantifioinnin epätarkkuus on myös tieteellisen kirjoittamisen ominaisuus. Kyse on sekä tiedon epätarkkuudesta tai halusta välttää liiasta tarkkuudesta koituvaa virheen riskiä. Fossiililöydön iänmäärittämisessä riittää sadan tai tuhannen vuoden tarkkuus. Tätä kuvaa vanha vitsi museonhoitajasta joka kertoi, että mammutin luu on 9 000 vuotta ja kaksi kuukautta vanha. Kun museon vierailija ihmetteli näin tarkkaa tietoa, tämä totesi: ”Täällä kävi kaksi kuukautta sitten asiantuntija, joka totesi luun olevan 9 000 vanha”.

Laatuluonnehdinnat ovat jo lähtökohtaisesti epätarkkoja. Sellaiset sanat kuin *lyhyt*, *suuri*, *kylmä*, *matala* ja *vanha* ovat vahvasti kulttuuri- ja tilannesidonnaisia. Vanhoille puille, ihmisille, hiirille, autoille, tavoille, taloille, käsityksille ja kallioille on jokaiselle oma ikäskaalansa; kaiken lisäksi ihmisillä, myös tutkijoilla, voi olla oma individuaalinen mielipiteensä asiasta. Tästä huolimatta tieteellisessäkin tekstissä esiintyy laatuaadjektiveja.

Kuvakieli ja metaforat eivät kuulu tieteellisen kieleen – näin yleensä ajatellaan. Kuitenkin tiukan luonnontieteellisissä julkaisuissa voidaan käyttää sellaisia ilmauksia kuin ”Cell-to-Cell communication is absolutely essential for multicellular organisms”. Kommunikaatio ymmärretään tässä metaforisesti. Solut eivät läheta tietoisesti viestettä, vaikka tieto siirtyykin solusta toiseen.

Vastaavalla tavalla sana *kommunikaatio* on otettu käyttöön myös tietokoneiden välisessä ”yhteydenpidossa”.

Edellä kuvatut epätasallisuuden muotojen esiintymät painottuvat eri tavoin tieteellisten julkaisujen eri osiin. Johdannossa ja tutkimuksen taustan kuvauksessa käytetään usein melko epämääräisiä viittauksia muiden tutkijoiden mielipiteisiin ja tutkimustuloksiin: *practically all; some; is often*. Ilkka Hanski kirjoittaa Nature-artikkelissa (Nature, 396, 41–49, 1998) ”Theoretical ecologists **typically** assume ...”, ”Landscape ecologists **tend** to analyse ...” (lihavointi minun). Näin jätetään pieni varaus sille, että väite ei koskekaan kaikkia kyseisen suuntauksen edustajia.

Analyysiosassa pyritään suurimpaan tarkkuuteen, mutta siinäkin voidaan varoa liiallista tarkkuutta eri syistä. Ensinnäkin voi johtaa **tekstin vaikeaselkoisuuteen**.

Siksi tavoitteena on löytää balanssi tarkkuuden ja luettavuuden välillä.

John Keynes (1936) totesi suunnilleen näin: ”Jos yritetään välttää kaikkea epätarkkuutta ja olla mahdollisimman tarkkoja, määrittelyissä joudutaan hiusten halkomiseen, joka ei kuitenkaan johda asian tarkkaan ymmärtämiseen.”

Liika tarkkuus voi lisätä myös **virheellisyyden** (valheen) riskiä. Tieteellisissäkin teksteissä esiintyy sanoja *muutama, melkein, yli, ainakin, enemmistö*. Jos ilmoitamme, että jotakin on 97,34 %, luvun on oltava tarkka, muuten tieto on väärä. Kun sanomme ”selvä enemmistö” olemme turvallisilla vesillä, vaikka tästä on seurauksena, että lukijoiden tulkinnat voivat vaihdella 75–99,9 prosentin välillä. Tarkkaa lukumäärää voidaan välttää myös sen vuoksi, että se antaa **väärän mielikuvan tiedon luotettavuudesta**. Jos sanomme, että Suomessa

asuu 38 657 venäläistä, syntyy harhainen kuva tiedonkeruun tarkkuudesta.

Vaikka tieteen kieli on lähtökohtaisesti asiaproosaa, josta on karsittu pois kaikki kruusailut, tutkijakin voi kokea tarpeelliseksi lisätä tekstiinsä sellaisia **pehmentäjiä ja varaumia** (hedges) kuten *I think / find / feel; I am not sure; sounds, seems, may, might, could, would; probably, perhaps, possibly*. Ne antavat takaportin siktä varalta, jos tieto ei osoittaudukaan oikeaksi.

Kaiken edellä kuvatun vakuudeksi sopivat Bertrand Russellin (1923) sanat:

”All knowledge is more or less uncertain and more or less vague. There are, in a sense, opposing characters: vague knowledge has more likelihood of truth than precise knowledge, but is less useful. One of the aims of science is to increase precision without diminishing certainty. But we cannot confine the word ‘knowledge’ to what has the biggest degree of both these qualities; we must include some propositions that are rather vague and some that are only rather probable. It is important, however, to indicate vagueness and uncertainty where they are present, and, if possible, to estimate their degree. [...] in most cases precision in this respect is impossible.”

Yleissääntö tieteellisessä kielessä on siis täysin sama kuin arkipuheessa: tavoitteena **kommunikaatiotilanteen kannalta riittävä tarkkuus**. On epätarkoituksenmukaista ja epäviisasta pyrkiä liialliseen tarkkuuteen, koska siitä on enemmän haittaa kuin hyötyä.

Johtopäätökset:

- Tieteen kieli ei ole aina tarkkaa ja täsmällistä.
- Tieteen kieli ei pyrikään olemaan mahdollisimman tarkkaa ja täsmällistä.

M.O.T.