

## KESKUSTELUA

### Tesaurus- ja guidelinesproblematiikasta

Tore Ahlbäck kirjoittaa tämän lehden numerossa 2/1985 indeksikielistä ja tesaurusproblematiikasta. Hän tarkastelee sekä tesaurusten termien välisten suhteiden että myös koko indeksointityön tulevaisuuden ongelmia. Nämä kysymykset ovat tärkeitä etenkin, kun ajatellaan elektronisen julkaisemisen läpimurron tuomia mahdollisuuksia sekä tiedon (viitteiden) että varsinaisten dokumenttien hakuun. Kirjoittajan näkemykset jälkimmäisestä ongelmasta ansaitsevat tulla huomatuksi. Edellistä ongelmaa käsitellessään kirjoittaja kuitenkin arvostelee Austinin ja Dalen opasta 'Guidelines for the establishment and development of multilingual thesauri' (1981) tavalla, joka antaa aiheen joihinkin reunahuomautuksiin.

Kirjoittaja lainaa 'guidelinesin' esimerkkiä eihierarkkisesta termien välisestä suhteesta, joka kuitenkin on 'mentally associated': BIRDS ja ORNITOLOGY. Hän pitää esimerkkiä virheellisenä — — siis sellaisena, joka ei oikealla tavalla edusta tarkoitettua RT-suhdetta — — ja pikemminkin tyypillisenä esimerkkinä hierarkkisesta suhteesta. Guidelinesin esimerkki on kuitenkin käsittääkseni moitteeton. Hierarkkinen suhde nimittäin määritellään (myös guidelinesissa) joko geneeriseksi tai partitiiviseksi. Edellinen tarkoittaa laji—alalaji -suhdetta ja jälkimmäinen kokonaisuus—osa -suhdetta. Termien BIRDS ja ORNITOLOGY suhde ei ole kumpaakaan lajia. BIRDS ei ole ORNITOLOGYn alalaji l. lintutieteen joku haara vaan sen kohde. Lintutieteen alalajin tulee olla myös lintutiede, kuten esim. 'petolintutiede' tai 'vesilintutiede'. Selvää on myös, ettei BIRDS ole ORNITOLOGYn jokin osa samassa mielessä kuin pyrstö on osa lintua.

Sama ongelma toistuu kirjoittajan kritisoidussa toisessa esimerkissä hevosten, aasien ja muulien suhteesta. 'Guidelinesin' mukaan muulit ovat RT-suhteessa hevosiin ja aaseihin kun taas kirjoittajan mukaan suhde on *par preference* geneerinen suhde. Käsittääkseni suhteessa on geneeristä vain se, että muuli saa geeninsä hevoselta ja aasilta. Muulit eivät kuitenkaan ole sen paremmin hevosten kuin aasienkaan alalaji, jota geneerinen suhde edellyttäisi.

Mielestäni 'guidelinesin' vihjeet RT-termien valintaan eri kategorioista ovat luonnollisia ja

tarpeellisia. Olen kirjoittajan kanssa kuitenkin samaa mieltä siitä, että ne mahdollisesti sopivat hyvin vain luonnontieteen ja tekniikan tesaurusten rakentamiseen. Sen sijaan en oikein ymmärrä kirjoittajan kritiikkiä sitä vastaan, että PLANTS RT HERBICIDES -suhde kirjataan tesaurukseen. Saan kirjoituksesta vaikutelman, että kirjoittaja ajattelee kirjauksen vaativan kasvimyrkkyjä käsittelevän artikkelin indeksointia myös termillä PLANTS (ja päinvastoin), mikä heikentää tiedonhaun relevanssia ja siirtää suotta työtä tiedonhakijalle, mikäli kasvimyrkkyä käsittelevä artikkeli ei varsinaisesti käsitelisi myös kasveja. Saatan tässä syüllistyä ylitulkintaan ja lukea sellaista, mitä ei ole tarkoitettu, josta minua oikaistakoon. Mikäli vaikutelmani on oikea, on seuraava huomautus tarpeen. Tesaurus on vain väline sopivien termien etsintään sekä dokumenttien että kyselyjen sisällön kuvailussa. Se 'ehdottaa' mahdollisesti hyödyllisiä termejä indeksoijan ja tiedonhakijan *harkittavaksi*, mutta ei pakota tiettyjen termien käyttöön (lukuunottamatta USE-viittauksia). Siis kirjaus HERBICIDES RT PLANTS ei määrää, että kasvimyrkkyjä käsittelevä artikkeli on myös indeksoitava termillä PLANTS, ellei itse dokumentti anna tähän aiheita. Kirjaus kuitenkin vihjaa tesauruksen käyttäjälle, että kasvimyrkkyjä tarkasteltaessa myös termi PLANTS on usein harkinnan arvoinen. (RT-viittaus vihjaa myös mahdollisesti *parempiin* termeihin, mikäli tesauruksen käyttäjän tarve ei ole vielä täsmentynyt.) RT-termit ovat aina arvokkaita myös tiedonhakijalle, mikä ei suinkaan ole ilmaus indeksoinnin puutteellisuuksista. Samalla tavalla kuin indeksoijaa niin myös tiedon hakijaa ne ohjaavat parhaimpien mahdollisten termien käyttöön kysymyksen kuvailussa. En usko, että ainakaan luonnontieteen ja tekniikan tiedonhakija hämmästelee sitä, että prosessista viitataan sen kohteeseen, välineisiin tuloksiin tai vaikutuksesta (voimasta) sen vasta-vaikutukseen (vastavoimaan). Näin siitä huolimatta, ettei sääntöjä tyhjentävästi olekaan voitu esittää.

Olen kirjoittajan kanssa samaa mieltä siitä, että indeksoinnilla perinteisessä mielessä ei välttämättä ole pitkää tulevaisuutta. Vapaatekstihaku mahdollisesti tekee indeksoinnin tarpeettomaksi. Tesauruksia ja niiden rakentamista tällainen kehitys ei kuitenkaan tee tarpeettomaksi. Vapaatekstihaussa, jos missä, tiedonhakija tarvitsee apua mielikuvitukselleen hän ajattelemiinsa termeihin mentaalisesti liittyvistä muista termeistä. Tesaurukset ovat tässä hyviä apuvälineitä.

Kalervo Järvelin