

Laskennallisen kielitieteen nykynäkymiä

Minulla oli tilaisuus osallistua vuoden 1978 aikana huhtikuussa Birminghamissa pidettyyn ALLC:n Computers in Literary and Linguistic Research -symposiumiin ja elokuussa Bergenissä pidettyyn seitsemänten Conference on Computational Linguistics -konferenssiin. Yleisenä linjana molemmissa tilaisuuksissa oli laskennallisen kielitieteen perustana olevan lingvistisen teorian merkityksen korostaminen. Tässä vaiheessa katsottiin edistymisen riippuvan ensisijaisesti teorian kehittämistä; teoriattomia, matemaattisen mallin raakaan voimaan perustuvia yrityksiä sen sijaan kavahdettiin.

Tietokoneen käytön kielitieteilijän ja filologin apuvälineenä saattoi havaita arkipäiväistyneen sekä monipuolistuneen nyt, kun pioneeriaika on tullut sivuuteksi. Tietokonetta käytetään perinteisissä kielitieteissä nimenomaan apuvälineenä, joka ei sinänsä vaikuta varsinaisen tutkimuksen teoriaan vaan pyrkii esim. nopeuttamaan työtä. Laajinta käyttö lienee mm. melko suppeisiin tekstikorpuksiin perustuvissa leksikografisissa hankkeissa ja filologisia tarkoituksia varten laadittavissa konkordansseissa, mutta myös monella muulla alueella.

Varsinaisella laskennallisella kielitieteellä tarkoitetaan kuitenkin laajasti ottaen kaikkea sellaista kielitiedettä, joka saa innoituksensa tietojenkäsittely-

opin menetelmistä (esim. tietorakenteista tai algoritmeista). Suppeimman tulkinnan mukaan taas kaikki laskennallinen kielitiede tulisi lukea joko teoreettiseksi kielitieteeksi tai tekoälyn tutkimukseksi. Tulkintojen pääasiallinen ero koskee kuitenkin vain nimitystä, ei sisältöä.

Laskennallisen kielitieteen kohteena on luonnollisen kielen analysointi ja tuottaminen. Morfologian, syntaksin ja semantiikan kuvaamisessa käytetään hyvin moninaisia malleja ja kielioppeja, joista mainittakoon sijakielioppi, dependenssikielioppi, ns. systeemikielioppi, ns. junktiokielioppi, ATN-kieliopit (augmented transition network), GSP (general syntactic processor) jne. Merkille pantavaa on puhtaan transformaatioteorian vähäisen käyttö, mikä johtuu TG-mallin yksisuuntaisuudesta.

Tekoälyllä on keskeinen asema pyritessä kohti luonnollisten kielten auto-

maattista käsittelyä, koska tietynlainen »ymmärtäminen» näyttää olevan välttämätön ja riittävä ehto erityisesti kielen kääntämiselle. Tekoälyn viimeaikainen kehitys on johtanut siihen, että suppeita osa-alueita käsittelevää tekstiä voidaan tällä tavoin mekaanisesti »ymmärtää», ts. niistä voidaan tehdä ne johtopäätökset, jotka ihminenkin huomaamattaan tekee. Ainakaan lähivuosina ei kuitenkaan ole odotettavissa, että nämä kokeilut johtavat tekoälyn soveltamiseen laaja-aiheisiin alueisiin.

Kiinnostus automaattista kielenkääntämistäkin kohtaan on 1960-luvun epäonnistumisten jälkeen taas viriämässä tältä uudelta pohjalta. Käytännön toteutuksissa näytävät realistisimmilta järjestelmät, joissa ihminen avustaa konetta ratkaisevassa ymmärrystä edellyttävässä vaiheessa.

Kimmo Koskenniemi