

Tilastollisen menetelmien perusteita

Kaius Sinnemäki

1. Johdanto tilastolliseen kielentutkimukseen

Tilastollisessa kielentutkimuksessa keskeinen tavoite on tehdä yleistyksiä jostain kielellisestä ilmiöstä. Tutkimus on empiiristä ja perustuu kieliaineiston analyysiin. Kieliaineiston kerääminen on usein työlästä ja vie aikaa, joten varsinkin perusopintovaiheessa on suositeltavaa käyttää valmiita saatavilla olevia kieliaineistoja. Hyviä vinkkejä esimerkiksi saatavilla oleviin korpusaineistoihin tarjoaa Laippala & Palander-Collin (2020).

Määrällisessä tutkimuksessa käytetään nykyään tietokoneohjelmia niiden nopeuden ja helppokäyttöisyyden vuoksi. Usein käytetään avoimesti saatavilla olevaa R-ohjelmointiympäristöä, josta on tullut keskeinen työkalu. Vauhtiin pääseminen sen kanssa vaatii kuitenkin opettelua ja käytännössä oman kurssinsa. Tässä oppimateriaalissa käytetään MS EXCEL:ä, koska sen kanssa pääsee vauhtiin nopeasti ja koska useimmilla opiskelijoilla on jotain kokemusta siitä tai vastaavasta taulukkolaskentaohjelmasta.

Ensimmäisessä videossa on erittäin lyhyt johdanto tilastollisesta lähestymistavasta kielentutkimuksessa sekä lisätietoa tästä metodipankin osiosta.

Viitteet:

Laippala, Veronika & Minna Palander-Collin 2020. Korpusaineistot. Teoksessa Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne & Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä II: Aineistonkeruusta analyysiin*, 460–486. Helsinki: Suomalaisen Kirjallisuuden Seura. <https://doi.org/10.21435/skst.1457>