



KAITSTUD DOKTORITÖÖD

11. mail kaitses Tallinna Ülikooli humanitaarteaduste instituudi doktorant Marju Taukar Zoomi keskkonnas doktoritöö „Algupäraste ja tõlgitud eestikeelsete tekstide sarnasused ja erinevused“. Juhendajad olid Arvi Tavast ja Heiki-Jaan Kaalep (TÜ), oponendid Heete Sahkai (EKI) ja Terje Loogus (TÜ).

Marju Taukari uurimismaterjaliks olid eestikeelsed ilukirjanduslikud tekstid. Enne tekstide analüüsimist uuris autor, kuidas inimesed üldse tõlgitud tekste tajuvad ning kuidas muudab tõlgitud teksti toimetamine. Peamiseks uurimismaterjaliks oli doktoritöös 13 miljoni sõnaline tõlgitud tekstide alamkorpusest ja algupäraste tekstide alamkorpusest koosnev keelekorpus. Tõlgitud tekstide korpus

on koostatud Marju Taukari doktoritöö tarbeks. Algupärased tekstid on uurimistöös sellised, mis on kirjutatud kohe eesti keeles, ilma võõrkeelse lähtetekstita. Eesti keelde tõlgitud tekstid pärinesid erinevatest lähtekeeltest.

Väitekirja eesmärk oli leida tõlkelisele keelele omaseid jooni. Enamasti öeldakse sihttekstis selgemalt välja seda, mida lähtetekstis nii ilmselgelt kirja ei panda. Sellele viitab sagedaste ja üldisema tähendusega sõnade suurem osakaal, samuti unikaalsete keelendite harvem kasutamine, kuna nendel puudub lähtekeeles vormiliselt sarnane vaste.

Doktoritöös uuriti, kas algupäraste ja tõlgitud eestikeelsete tekstide keelekasutus erineb nii palju, et neid tekste saaks automaatselt klassifitseerida. Klassifitseerimine õnnestus hästi kõigi erineva pikkusega sõnaloendite põhjal. See tähendab, et sõnade kasutussagedused erinevad tõlgetes ja algupärastes tekstides sellisel määral, et arvuti suudab ära arvata, kas tegemist on tõlkega või algupäraste tekstiga.

Kõige üllatavam tulemus töös oli see, et sõnaloendite põhjal saab nii täpselt tekste klassifitseerida. Lisaks tavalistele sõnaloenditele katsetas autor oma uurimuses

veel tähemärkide, sõnade, lemmade ja sõnaliikide pikemate jadadega. Vastupidiselt mitmetele teistes keeltes tehtud samalaadsetele töödele ei eristunud täpsuse poolest sõnade sagedusloendi kõige ülemine ots, kuhu kuuluvad pigem funktsioonisõnad (erinevad ase- ja sidesõnad).

Väitekirjast selgus veel, et tõlgitud tekst ei muutu toimetamise tõttu nii palju, et avaldatud tõlked ei sobiks tõlkimise uurimiseks. Teisalt oli toimetamata ja toimetatud tekste uuritud väga piiratud mahus võrreldes klassifitseerimisülesandega. Toimetamine ja teksti muutumine toimetamise käigus vajab kindlasti lähemat uurimist. Seni on nii Eestis kui mujal peamiseks piiranguks selle teema arendamises tekstide kättesaadavus.

Marju Taukari töö on esimene, mis käsitleb eesti keelde tõlgitud tekste nii suure ulatuses. Kuna eesti keeles ei ole tõlgitud tekstide keelekasutust kvantitatiivselt uuritud, võiks töös leitu anda edasisi suundi nii kvantitatiivsete kui ka kvalitatiivsete uurimisküsimustega tegelevatele keeleteadlastele. (<https://www.tlu.ee/teadus/uudised/doktoritoo-uuris-kuidas-tolgitud-eesti-keel-erineb-alguparasest-eesti-keelest>)