

“Ihmiskääntämisen viimeisellä linnakkeella”

# Katsaus kaunokirjallisuuden koneavusteisen kääntämisen tutkimukseen

Julkisuudessa liikkuu välillä huhuja sensaatiomaisista läpimurroista kaunokirjallisuuden konekääntämisessä. Aihetta koskeva tieteellinen tutkimus on kuitenkin toistaiseksi ollut vähäistä. Tutkijoita on kiinnostanut erityisesti runouden kääntäminen, jälkieditointi sekä kääntäminen sukulaiskielten välillä. Asialla ovat olleet lähinnä tietojenkäsittelytieteen tutkijat, jotka eivät ole kaunokirjallisuuden tai kääntämisen asiantuntijoita. Professoriluentoani varten tutustuin kahteen kertovien tekstien koneavusteista kääntämistä koskevaan tutkimukseen.

Dublin City Universityn tietojenkäsittelytieteen laitoksessa meneillään olevassa tutkimuksessa **Andy Way** ja **Antonio Toral** ovat tehneet kokeiluja kaunokirjallisuuden konekääntämisessä erityisesti sukulaiskielten välillä. Wayn ja Toralin mielestä on aika kokeilla tilastollista konekääntämistä ”ihmiskääntämisen viimeiseen linnak-

keeseen”, koska kaunokirjallisuuden kääntäjien palkkiot ovat niin matalat, ettei niillä tule toimeen. He uskovat, että kone voisi nopeuttaa kääntäjien työtä ja kasvattaa heidän ansaintomahdollisuuksiaan. Toral ja Way kertovat saaneensa lupaavia tuloksia kokonaisen romaanin (Carlos Ruiz Zafónin *El prisionero del cielo*) konekään-



Kristiina Taivalkoski-Shilov kuvattuna Vaasan yliopistolla KäTü-symposiumin yhteydessä.  
Kuva: Maarit Laitinen

## Tekstit rakentuvat monille äänille, joista tavallisimpia ovat kertojan ja henkilöhahmojen puhetta, ajatuksia ja näkökulmaa kuvastavat äänet.

tämisessä espanjasta katalaaniin. Heidän vuonna 2015 ilmestyneen konferenssijulkaisunsa tarkka lukeminen kuitenkin osoittaa, että kyseistä konekäännöstä testattiin vain kahdella henkilöllä, jotka lukivat tekstiä virkkeittäin Appraise Tool -nimisen työkalun avulla. Toralin ja Wayn toisen, *Translation Spaces* -lehdessä ilmestyneen artikkelin tiivistelmässä taas huomion kiinnittää väite, että Google Translate on kehittynyt vuoden 2013 ja kesän 2015 välillä *Camus'n Sivullisen* kääntämisessä ranskasta englantiin. Todellisuudessa artikkelissa vertaillaan Google-kääntäjien versioita 109-sanaisesta tekstinkappaleesta.

Toinen esittelemäni tutkimus tehtiin Grenoblen yliopistossa vuonna 2014. Professori **Laurent Besacier** testasi konekäännöksen jälkieditointia 20-sivuisen esseen kääntämisessä. **Richard Powersin** "The Book of Me" ranskannettiin ensin lehtitekstien kääntämistä varten räätälöidyllä konekääntimellä ja jälkieditointiin sitten käyttäen työkalua, jolla oli aiemmin muokattu tietosanakirja- ja internet-tekstien käännöksiä. Lähtöteksti käännettiin ja jälkieditoitiin kolmessa osassa, ja konekäännintä muokattiin jokaisessa vaiheessa jälkieditoinnista saadun datan avulla. Lopputulos todettiin kuitenkin hyvin epätydyttäväksi. Kaksi henkilöä muokasi tekstiä vielä kymmenisen tuntia, minkä jälkeen se annettiin yhdeksän testihenkilön ja Powersin nimikoranskantajan, **Jean-Yves Pellegrinin**, arvioitavaksi. Enemmistö testihenkilöistä piti tekstin laatua varsin hyväksyttävänä. Pellegrinin näkemyksen mukaan tekstin asiasisältö välittyi hyvin, mutta tyyli- ja kulttuurivii- tauksissa oli ongelmia. Hän totesi myös, ettei omassa työssään voisi hyödyntää koneen tekemää raakakäännöstä. Besacier kirjoittaa artikkelinsa johtopäätöksissä, että konekääntämisen ja jälkieditoinnin yhdistelmällä voidaan tuottaa nopeasti halpoja käännöksiä, joiden laatu kuitenkin jää puutteelliseksi.

Kummankin esittelemäni tutkimuksen ongelmana on perusolettamus, että kaunokirjallisen tekstin sisältö ja muoto voitaisiin erottaa toisistaan, kuten asiatekstien kohdalla voidaan osittain tehdä, vaikka niissäkin tekstin tyyli on osa merkitysrakennetta. Kaunokirjallisuudessa tekstin sisällön ja muodon yhteys on kuitenkin vielä tiiviimpi. Niiden erottamattomuus on ihmiskääntäjillekin keskeinen ongelma kielten ja kulttuurien eroavaisuuksien takia. Kaunokirjallisen tekstin sanat eivät välitä tietoa pelkästään teoksen juonesta ja teemoista vaan ne myös kannattelevat koko tekstin kerronnallista rakennetta. Tekstit rakentuvat monille äänille, joista tavallisimpia ovat kertojan ja henkilöhahmojen puhetta, ajatuksia ja näkökulmaa kuvastavat äänet. Äänien tunnusmerkit voivat olla pronomineja, verbien tempuksia, deiktisiä ajan ja paikan määreitä, kielen eri tyylirekistereihin liittyviä elementtejä tai typografisia korostuskeinoja. Kyseiset kielen elementit saavat tekstissä erilaisia funktiota suhteessa paikalliseen kontekstiinsa ja teokseen kokonaisuutena. Tekstin pienelläkin yksityiskohdalla voi olla kerronnallinen merkitys. Tästä syystä kaunokirjallisuuden kääntäjä tarvitsee tulkintansa pohjaksi koko alkuperäisen tekstin eikä koneen tekemää raakakäännöstä sellaisenaan tai jälkieditoituna. Tästä menetelmästä ei ole kaunokirjallisuuden kääntämiselle mitään hyötyä vaan suoranaista haittaa.

Konekääntämisen ja käännösteknologian kehittäminen kaunokirjallisuuden kääntäjien tarpeisiin ei sinällään ole huono ajatus, ja oletettavaa on, että jotkut kääntäjät hyödyntävät jo nyt esim. käännösmuistiohjelmia. Kehittämistyössä tarvittaisiin kuitenkin monitieteistä tutkimusta, jossa olisi mukana myös kielen, kääntämisen ja kirjallisuuden asiantuntijoita.

**Kristiina Taivalkoski-Shilov**

Kirjoittaja on monikielisen käännösviestinnän professori Turun yliopiston kieli- ja käännöstieteiden laitoksella. Taivalkoski-Shilov, joka on erikoistunut käännöskirjallisuuden tutkimukseen, piti professoriluentonsa Turun yliopistossa 11. tammikuuta 2017. Professoriluento löytyy kokonaisuudessaan SKTL:n internet-sivulta osoitteesta: <https://www.sktl.fi/?x1239065=1239069>

## Tekstin pienelläkin yksityiskohdalla voi olla kerronnallinen merkitys.