

Tekoäly on täällä – mitä tekee kääntäjänkoulutus?

Raportti V jaoston syysseminaarista

Maarit Koponen & Minna Ruokonen
Itä-Suomen yliopisto

<https://doi.org/10.61200/mikael.138230>

Abstract

This article reports on the autumn seminar of the Teachers' and Researchers' Section of the Finnish Association of Translators and Interpreters (SKTL), organized in August 2023. The program consisted of presentations and workshop discussions on generative artificial intelligence in translator training, competence descriptors for specialized translation, and literary translators' technology use and attitudes. The seminar served as a starting point for translator trainers to discuss both the general principles and practical exercises for integrating AI use into translator training.

Keywords: translator training, translation levels, translation technology, artificial intelligence

Avainsanat: kääntäjänkoulutus, kääntämisen taitotasot, käännösteknologia, tekoäly

1 Johdanto

Suomen kääntäjien ja tulkkien liiton opettajien ja tutkijoiden jaoston (V jaosto) vuotuinen syysseminaari järjestettiin Joensuussa 17.8.2023 temalla Tekoäly on täällä – mitä tekee kääntäjänkoulutus? Tekoäly valittiin teemaksi, koska kevään 2023 aikana koneoppimiseen perustuvat kielimallit, etenkin OpenAI:n julkaisema ChatGPT-sovellus, olivat nousseet merkittäväksi puheenaiheeksi yliopisto-opetuksessa. ChatGPT perustuu suureen monikieliseen kielimalliin ja toimii siten, että käyttäjä antaa sovellukselle ns. kehotteen ja sovellus tuottaa (generoi) tekstiä ennustaen opetusaineistonsa perusteella todennäköisimpiä virkkeitä. Vastaavanlaisia sovelluksia ovat sittemmin julkaisseet myös esimerkiksi Google ja Microsoft. Vaikka niitä ei ole tarkoitettu kääntämiseen, niillä voidaan tuottaa konekäännöksiä antamalla sopivia kehoitteita (esim. Jiao ym. 2023).

Kielimallien tulo kaikkien ulottuville aiheutti tarpeen reagoida pikaisesti niiden käyttöön opetuksessa ja oppimisessa. Erityisesti generatiivisten sovellusten käyttö esseiden, tehtävien ja opinnäytteiden tekemiseen on herättänyt huolta. Tekoälysovelluksista voi toisaalta olla hyötyä esimerkiksi kielitaidon kehittämisessä (esim. Pekanheimo 2023a), mutta teknologiaan tukeutuminen saattaa haitata opiskelijan

omien kompetenssien kehittymistä (vrt. Carré ym. 2022). Yliopistot ovatkin laatineet periaatteita generatiivisten kielimallien käytölle. Esimerkiksi Itä-Suomen yliopiston ohjeistus (Itä-Suomen yliopisto 2023) mahdollistaa generatiivisten kielimallien käytön apuvälineenä, kunhan tämä ei haittaa oppimistavoitteiden saavuttamista ja käyttö ilmoitetaan asiaankuuluvalla tavalla. Opinnäytetyössä ”itsenäisen työn osuutta” on oltava riittävästi.

Yliopistotason ohjeistukset eivät kuitenkaan ota huomioon kääntämisen opetuksen erityispiirteitä tai vastaa juuri kääntämisen opettajia askarruttaviin kysymyksiin. Siksi syysseminaarin tavoitteena oli tarjota opettajille mahdollisuus keskustella näiden sovellusten vaikutuksista, jakaa kokemuksia ja ideoita käytänteitä esimerkiksi kurssien käännoستهävien ja muiden harjoitustöiden suunnitteluun.

Seminaarin aloitti Sampo Pyysalon alustus generatiivisista kielimalleista. Sen jälkeen seminaarin osallistujat pohtivat pienryhmissä kielimallien vaikutuksia opinnoissa. Seminaarin toisessa osassa Minna Kujamäki esitteli EFFORT-hankkeessa laadittuja erikoisalan kääntäjän taitotasokuvauksia, ja Kertova teksti, kääntäjä ja kone -hankkeen edustajat Kristiina Taivalkoski-Shilov, Minna Ruokonen ja Leena Salmi pitivät kirjallisuuden kääntämiseen ja teknologiaan liittyvät kommenttipuheenvuorot. Tässä raportissa kuvaamme seminaarin antia.

2 Generatiivisten kielimallien perusteista

Ensimmäisen alustuksen piti yliopistotutkija Sampo Pyysalo Turun yliopiston TurkuNLP-tutkimusryhmästä, joka kehittää esimerkiksi suomenkielisiä kielimalleja (malleja voi testata osoitteessa <https://turkunlp.org/gpt3-finnish>). Pyysalo esitteli kielimallien toimintaperiaatteita sekä niiden käyttöä yliopisto-opetuksessa yleisesti.

Pyysalon mukaan kielimallien tuotokset ovat parantuneet erittäin nopeasti, koska niiden kouluttamiseen on käytetty entistä laajempia aineistoja ja enemmän laskenta-aikaa. Kielimallit pystyvät tuottamaan hyvää, sujuvaa tekstiä, jota ei käytännössä voi luotettavasti erottaa ihmisen kirjoittamasta. Kääntämisessä monikieliset kielimallit toimivat joissakin tapauksissa jo hyvin, vaikka varsinaiset konekääntimet saattavat pärjätä paremmin tietynlaisten erikoisalan tekstien kääntämisessä (Jiao ym. 2023). Luultavasti erot kaventuvat, ja kielimalleja integroidaan konekääntimiin.

Tätä kirjoittaessa kielimallisovellukset eivät kykene esimerkiksi ohjeiden seuraamiseen tai todelliseen vuorovaikutukseen. Todennäköisyyksiin perustuva tekstin tuottaminen voi myös tuottaa virheellistä tietoa tai toisintaa aineiston vinoumia. Asiasisällön oikeellisuuteen on siksi syytä suhtautua varauksella. Pyysalo mainitsi haasteena myös hyvälaatuisen aineiston saatavuuden: esimerkiksi suomenkielisiä tekstejä ei ole saatavilla riittävästi entistä suurempien kielimallien kouluttamiseen. Tulevaisuudessa tiettyihin aihealueisiin keskittyvät mallit todennäköisesti lisääntyvät ja mallien looginen päättelykyky kehittyy. Vaikka malleihin saattaakin kohdistua ylisuuria odotuksia, yliopisto-opinnoissa on siis harjoiteltava niiden kriittistä käyttöä.

3 Kielimallit kääntäjänkoulutuksessa

Pyysalon alustuksen jälkeen osallistujat pohtivat pienryhmissä kielimallien vaikutuksia kääntäjänkoulutuksessa. Yleisesti ryhmissä oltiin yhtä mieltä siitä, että kieliammatillaiset ja muut tietotyöläiset tarvitsevat tekoälylukutaitoa eli kykyä käyttää tekoälytyökaluja asiantuntevasti ja kriittisesti. Opiskelijan täytyy ymmärtää, miten kielimallit toimivat ja mitä tästä seuraa: Mihin tarkoituksiin tekoäly sopii ja mihin ei? Mitä ympäristö-, tietosuoja- ja tekijänoikeuskysymyksiä sen käyttöön liittyy? On myös entistä tärkeämpää perustella opiskelijoille, miksi oma oppiminen on arvokasta ja miksi on eettisesti olennaista raportoida, mitä työkaluja on käyttänyt ja miten. Samalla kriittisen luku- ja tiedonhakutaidon merkitys korostuu, kun tekoälygeneroitu tekstihäly lisääntyy.

Opiskelijoille on alusta lähtien tehtävä selväksi, miten kääntämisen taidot eroavat koneen tuottaman tekstin editoinnista ja miksi ”käsini” kääntämistä on edelleen tarkoituksenmukaista opetella: käännösprosessia, käännösongelmien tunnistamista ja ratkaisemista ei opi editoimalla. Näihin taitoihin perehtyminen auttaa myös hahmottamaan koneen heikkouksia. Tekoälyä on vaikea soveltaa esimerkiksi viestintätilanteisiin, joissa esiintyy kieli- ja kulttuurieroja, uusia ja kehittyviä aihepiirejä tai odotuksista poikkeavia ja monimerkityksisiä kielenkäyttötapoja.

Periaatteena pidettiin sitä, että tekoälyn sovelluksia kokeillaan yhdessä opiskelijoiden kanssa. Tarvitaan erityyppisiä tehtäviä, joissa hyödynnetään sekä tekoälyä että omaa älyä. Käännöskurssilla tehtävät voivat olla esimerkiksi seuraavanlaisia:

- Analysoidaan esimerkkejä kielimallien ja konekäännösten virheistä (esim. Pekanheimo 2023b): Mitä on mennyt pieleen? Mikä siihen voi olla syynä?
- Vertaillaan eri kielimallien tai konekääntämien tuotoksia ja ihmiskäännöksiä: Mitä eroja teksteissä on ja mistä ne voivat johtua?
- Kasvatetaan rekisteritietoisuutta: Miten tekoäly sopii eri tekstilajeihin? Miten se sopii eri käyttötarkoituksiin?
- Lisätään käännöskommenttiin vakiokohta käytetyistä työkaluista ja pohdintaa niiden vaikutuksesta.

Tutkielmaopetuksessa voitaisiin testata tekoälyn käyttöä esimerkiksi oman tekstin jäsentämisessä ja tyylin, koherenssin tai metatekstin editoinnissa (vrt. Pekanheimo 2023a), tutkimuskysymysten ideoinnissa tai aiemman tutkimuksen hyödyntämisessä. Olennaista on myös arvioida lopputuloksia yhdessä: Jos kielimallia pyytää laatimaan tiivistelmän artikkelista, pitääkö tuotos paikkansa? Nostaako se esiin itseä kiinnostavia näkökulmia? Jos tekoälyä pyytää vertailemaan käsitteen määritelmiä, tunnistaako se relevantteja eroja ja yhtäläisyyksiä? Erityistä hyötyä saattaisi tuoda juuri käännöstieteellisellä aineistolla koulutettu kielimalli. Erikoistumattomat mallit voivat tuottaa ns. hallusinaatioita, kuten esimerkkinä mainittu ”käännöstieteen astraalikäänne”.

4 Kääntäjän kompetenssit, teknologia ja kirjallisuuden kääntäjät

Syysseminaarissa pohdittiin myös kääntäjien kompetensseja ja teknologisten apuvälineiden käyttöä erityisesti kirjallisuuden kääntämisessä. Tätä teemaa käsitelivät kahden hankkeen puheenvuorot, joita kuvaamme seuraavaksi.

4.1 Erikoisalojen kääntämisen kompetenssit

Ensin Minna Kujamäki (Itä-Suomen yliopisto) esitteli EU:n Erasmus+-ohjelman rahoittamaa EFFORT-hanketta (”Towards a European framework of reference for translation”, 2020–2023). Hankkeen tavoitteena on ollut määrittellä kääntäjänkoulutukselle vastaavanlaisia taitotasoja kuin kielten opetuksessa käytetyssä CEFR-viitekehyksessä. Kuvaukset tasoille A ja B, joilla ei vielä käsitellä erikoisaloja, on laadittu aiemmassa PACTE-tutkimusryhmän hankkeessa; EFFORT-hanke keskittyi tasoon C eli erikoisalojen kääntämiseen. Hanketta koordinoi Barcelonan autonominen yliopisto, ja siihen osallistui yhdeksän muuta yliopistoa, mukaan lukien syysseminaarin järjestänyt Itä-Suomen yliopisto.

Hankkeessa tarkastellut erikoisalot olivat laki, talous ja rahoitus, tekniikka, luonnontiede ja matematiikka sekä kaunokirjallisuus. Näille aloille laadittiin kuvaukset kunkin erikoisalan ominaispiirteistä ja kompetensseista. Kaunokirjallisuuden alan kuvauksista vastasivat Itä-Suomen yliopisto ja romanialainen Iaşin yliopisto. Kuvaukset, esimerkkejä tyypillisistä tekstilajeista sekä kääntämisen opiskelijoille ja opettajille suunnattu itsearviointityökalu on julkaistu hankkeen verkkosivuilla (<https://www.effortproject.eu/>).

EFFORT-kuvauksissa keskeisiä erikoisalojen kompetensseja ovat kulttuurinen tieto, tiedonhaku- ja dokumentaatiotaidot sekä työkalujen käyttötaito (*instrumental competence*). Kaunokirjallisuuden kääntämisenkin kuvauksessa nousee esiin kyky arvioida kieliteknologiatyökalujen hyötyä sekä taito käyttää niitä tarpeen mukaan. Vaikka generatiivista tekoälyä sellaisenaan ei ole erikseen mainittu, teknologian rooli on siis tunnistettu.

4.2 Teknologian käyttö kaunokirjallisuuden kääntämisessä

Kommenttipuheenvuorona EFFORT-hankkeen kompetenssikuvauksiin kuultiin tuloksia Suomen Akatemian rahoittamasta hankkeesta Kertova teksti, kääntäjä ja kone. Kristiina Taivalkoski-Shilov (Turun yliopisto) esitteli yleisesti tätä hanketta, joka tutkii kaunokirjallisen proosan luonnetta ja käännösprosessia sekä kehittää käyttäjäystävällisiä ja eettisesti kestäviä työkaluja. Sen jälkeen Minna Ruokonen (Turun yliopisto ja Itä-Suomen yliopisto) ja Leena Salmi (Turun yliopisto) kertoivat keväällä 2023 tekemästään kyselystä, jossa selvitettiin suomeen tai suomesta kääntävien kirjallisuuden kääntäjien työkalujen käyttöä ja teknologianäkemyksiä. Kysely perustui vastaavaan tutkimukseen, joka oli tehty Hollannissa (Daems 2022).

Suurin osa vastaajista (n = 72) oli kuullut käännösmuisteista ja sanastotyökaluista sekä konekääntimistä yleensä; neuroverkkokääntimet ja vuorovaikutteiset konekääntimet olivat vieraampia. Vain kymmenkunta vastaajaa käytti käännösmuistia tai konekääntimiä ”usein” tai ”aina”, mutta suurin osa arvioi, että niistä voisi ”joskus” olla apua kaunokirjallisuuden kääntämisessä. Käännösteknologian hyvinä puolina mainittiin esimerkiksi monipuoliset hakutoiminnot sekä termien ja nimien kääntämisen helpottaminen. Toisteisten ilmausten tunnistaminen työkaluilla auttaa sekä huomaamaan tyylikeinoja että välttämään toistoa. Huonojakin puolia nähtiin useita. Teknologian käyttäminen oli monen mielestä ristiriidassa kaunokirjallisuuden kääntämisen luonteen kanssa: luovuus ja taiteellisuus vaativat hitautta ja käsityötä. Kustantajien nykyiset prosessit eivät tue käännösteknologian käyttöä, ja työkalut ovat kalliita. Vastaajia mietitytti myös, kuka päättää milloin ja miten työkaluja käytetään. Tosin jotkut kaunokirjallisuuden kääntäjät käyttävät käännösteknologiaa jo nyt ja näkevät käytössä enemmän hyviä kuin huonoja puolia. Olisikin tärkeää herättää avointa keskustelua teknologian käytöstä, jotta alalle ei synny sisäisiä jakolinjoja.

5 Lopuksi

Syysseminaarin tavoitteena oli tarjota kääntämisen opettajille yleiskuva kielimallien toimintaperiaatteista, potentiaalista ja rajoituksista sekä avata keskustelua generatiivisen tekoälyn käytöstä kääntäjänkoulutuksessa. Seminaarin jälkimmäinen osa puolestaan korosti työkaluihin liittyvien kompetenssien merkitystä: edes kaunokirjallisuuden kääntämistä ei voi pitää teknologiavapaana vyöhykkeenä. Pienryhmäkeskusteluissa hahmottui varsin yhtenäinen näkemys siitä, että tekoälysovellusten kriittistä käyttöä on syytä sisällyttää kääntäjänkoulutukseen mutta käsin kääntäminen on edelleen olennainen taito. Lisäksi ideoitiin erilaisia käytännön harjoituksia. Lisää pohdintaa ja keskustelua aiheesta kuitenkin tarvitaan.

Kääntäjänkoulutuksessa on hyvät valmiudet hyödyntää kielimalleja, koska käännösteknologiaa käytetään jo opetuksessa aktiivisesti. ”Perinteisen” konekäännöksen ja generatiivisen kielimallin tuottaman tekstin käytössä on joitakin yhtäläisyyksiä, mutta myös eroja, jotka saattavat vaikuttaa monin tavoin sekä käytännön toimintaan että kognitiiviseen prosessointiin. Tutkimusta on kuitenkin toistaiseksi vähän, ja kielimallien käyttö käännösosalalla kehittynee lähivuosina nopeasti. Kehityksen seuraamisessa olennainen tekijä on kääntäjänkoulutuksen tiiviit yhteydet alan toimijoihin.

Lähdeluettelo

Carré, Alice, Dorothy Kenny, Caroline Rossi, Pilar Sánchez-Gijón & Olga Torres-Hostench 2022. Machine Translation for Language Learners. Teoksessa: Dorothy Kenny (toim.) *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*. Berlin: Language Science Press, 187–207. <https://doi.org/10.5281/zenodo.6760024>

Daems, Joke 2022. Dutch Literary Translators' Use and Perceived Usefulness of Technology: The Role of Awareness and Attitude. Teoksessa: James Luke Hadley, Kristiina Taivalkoski-Shilov, Carlos S.C. Teixeira & Antonio Toral (toim.) *Using Technologies for Creative-Text Translation*. New York: Routledge, 40–65. <https://doi.org/10.4324/9781003094159-3>

Itä-Suomen yliopisto 2023. Tekoälyn käyttäminen opetuksessa ja tutkimuksessa. Saatavissa: <https://kamu.uef.fi/tietopankki/opiskelijan-oikeudet-ja-velvollisuudet/tekoalyn-kayttaminen-opetuksessa-ja-tutkimuksessa/> [viitattu 23.10.2023].

Jiao, Wenxiang, Wenxuan Wang, Jen-tse Huang, Xing Wang & Zhaopeng Tu 2023. Is ChatGPT A Good Translator? Yes With GPT-4 As The Engine. <https://doi.org/10.48550/arXiv.2301.08745>

Pekanheimo, Eeva 2023a. Lost in language acquisition – Tekoäly ja vieraiden kielten oppiminen. Saatavissa: <https://www.humak.fi/blogit/tekoaly-ja-kieli-osa-iii-lost-in-language-acquisition-tekoaly-ja-vieraiden-kielten-oppiminen/> [viitattu 23.10.2023].

Pekanheimo, Eeva 2023b. Lost in translation – Tekoäly, käännöstyöt ja vieraskielisen tekstin tuottaminen. Saatavissa: <https://www.humak.fi/blogit/tekoaly-ja-kieli-osa-ii-lost-in-translation-tekoaly-kaannostyot-ja-vieraskielisen-tekstin-tuottaminen/> [viitattu 23.10.2023].

Kirjoittajat

Maarit Koponen toimii Itä-Suomen yliopistossa käännöstieteen professorina. Hänen tutkimuksensa keskittyy konekääntämisen ja muun käännösteknologian käyttöön erilaisissa tilanteissa ammattikäntäjän työstä arkipäivän tiedonhakuun ja viestintään. Sähköpostiosoite: [maarit.koponen\(at\)uef.fi](mailto:maarit.koponen(at)uef.fi)

Minna Ruokonen on englannin kielen ja kääntämisen yliopistonlehtori Itä-Suomen yliopistossa. Hän toimii Kääntämisen ja tulkkauksen opetusmenetelmävarannon ylläpitäjänä ja on perehtynyt etenkin käntäjän ammatin arvostuksen ja työolojen tutkimukseen. Sähköpostiosoite: [minna.ruokonen\(at\)uef.fi](mailto:minna.ruokonen(at)uef.fi)