

**LLM-teknologiaa hyödyntävä kääntäminen:
saksankielisten juridisten tekstien suomentaminen ja
jälkieditoinnin merkitys**

Laadullinen vertaileva käänösanalyysi

Eveliina Riihimäki

Pro gradu -tutkielma

Monikielisen käänösviestinnän tutkinto-ohjelma, saksan kieli

Kieli- ja käännöstieteiden laitos

Humanistinen tiedekunta

Turun yliopisto

Helmikuu 2026

Pro gradu -tutkielma

Monikielisen käännösviestinnän tutkinto-ohjelma, saksan kieli

Eveliina Riihimäki

LLM-teknologiaa hyödyntävä kääntäminen: saksankielisten juridisten tekstien suomentaminen ja jälkieditoinnin merkitys

Laadullinen vertaileva käännösanalyysi

Sivumäärät: 71 sivua + liitteet 24 sivua

Tässä pro gradu -tutkielmassa analysoidaan ChatGPT 5.2 -kielimallin ja DeepL Next-Gen -konekääntimen tuottamia suomennoksia saksankielisestä EU:n tuomioistuimen lehdistötiedotteesta (98/19) sekä tarkastellaan niiden jälkieditointitarvetta. Tutkielman tavoitteena oli selvittää, millaisia virheitä ja kielellisiä puutteita edistyneen käännösteknologian tuottamissa suomennoksissa esiintyy, kumpi järjestelmä tuottaa laadukkaamman käännöksen, miten suomen kielen erityispiirteet ja juridisen kääntämisen vaatimukset näkyvät käännösjäljessä sekä millaista asiantuntemusta käännösten jälkieditointi edellyttää.

Tutkimusmenetelminä käytettiin laadullista rinnakkaiskorpusanalyysiä ja tilastollista virheanalyysiä. Saksankielistä lähdetekstiä verrattiin ChatGPT-5.2:n ja DeepL Next-Genin tuottamiin käännöksiin sekä ihmiskääntäjän referenssikäännökseen. Virheet luokiteltiin MQM-virhemallilla ja jälkieditointitarve arvioitiin tutkimusta varten kehitetyllä viisiportaisella asteikolla.

Tulokset osoittivat, että ChatGPT-5.2 ja DeepL Next-Gen pystyvät usein tuottamaan sujuvaa suomea, mutta niiden käännöksissä esiintyy systemaattisia terminologia-, syntaksi- ja morfologisia virheitä, jotka useissa tapauksissa vaativat merkittävää jälkieditointia ja jopa vääristävät oikeudellista merkitystä; ChatGPT:n tuotokset olivat sujuvampia mutta vaativat enemmän terminologista ja substanssiin kajoavaa korjausta, kun taas DeepL:n käännökset olivat paikoin terminologisesti tarkempia mutta kärsivät enemmän saksan rakenteellisesta interferenssistä lauserakenteissa.

Tutkielman perusteella ammattikäntäjän juridinen ja kielellinen asiantuntemus on edelleen välttämätöntä kielimallien ja konekääntimien tuottamien juridisten käännösten viimeistelyssä. Jälkieditointi parantaa merkittävästi käännösten luettavuutta, kohderyhmäsopivuutta ja oikeudellista tarkkuutta. Tutkielma suosittaa, että konekäännöksiä hyödynnetään juridisessä kääntämisessä vain ammattimaisen jälkieditoinnin yhteydessä.

Avainsanat: konekääntäminen, kielimallit, ChatGPT, DeepL, jälkieditointi, saksa, suomi, käännöslaatu, korpusanalyysi, MQM, ISO 5060, virheluokittelu, juridinen kääntäminen, oikeuskieli

Sisällysluettelo

1	Johdanto	5
1.1	Tutkielman tausta, merkitys ja tavoitteet	5
1.2	Tutkimuskysymykset	6
1.3	Tutkielman menetelmät, aineisto ja rajaus	7
1.4	Tutkielman rakenne	8
2	Konekääntäminen ja juridinen kääntäminen LLM-tekniikan aikakaudella	9
2.1	Konekääntämisen historia ja siirtymä kohti suuria kielimalleja	9
2.2	ChatGPT-5.2 ja DeepL Next-Gen käännoystyökaluina	13
2.3	Juridisen suomentamisen erityishaasteet konekääntämisessä ja jälkieditoinnissa	15
2.3.1	<i>Juridisen tekstin normatiivinen luonne ja vaikutus jälkieditointiin</i>	16
2.3.2	<i>Suomen kielen morfologiset ja syntaktiset ominaisuudet jälkieditoinnin näkökulmasta</i>	18
3	Tutkimuksen toteutus	23
3.1	Tutkimusasetelma ja metodologiset valinnat	23
3.2	Aineiston valinta, keruu ja käsittely: Lehdistötiedote nro 98/19	25
3.2.1	<i>Aineiston valinta ja perustelu</i>	25
3.2.2	<i>Aineiston keruu</i>	26
3.2.3	<i>Aineiston käsittely ja rinnakkaiskorpuksen muodostaminen</i>	29
3.3	Analyysimenetelmät ja virheluokittelu	29
3.3.1	<i>MQM 2.0 -virhekategorioiden luokittelu</i>	30
3.3.2	<i>Virheiden vakavuusasteet ja pisteytys</i>	31
3.3.3	<i>Jälkieditointitarpeen viisiportainen asteikko</i>	32
3.3.4	<i>MQM-virhemallin ja jälkieditointiasteikon suhde</i>	33
3.4	Tutkimuksen rajoitteet ja metodologiset huomiot	34
4	Analyysi ja tulokset	36
4.1	Yleiskatsaus käännosten laatuun	36
4.2	Keskeiset virhekategorioiden luokittelu	38
4.2.1	<i>Leksikaalinen taso: Terminologinen täsmällisyys ja johdonmukaisuus</i>	38
4.2.2	<i>Morfosyntaktinen taso: Rakenteet ja interferenssi</i>	44
4.2.3	<i>Funktionaalinen taso: Juridisen rekisterin ja oikeusjärjestelmien heijastuminen</i>	51
4.3	Jälkieditoinnin tarve ja vaikutus käännoslaatuun	56

4.3.1	<i>Jälkiedointitarpeen jakautuminen viisiportaisella asteikolla</i>	56
4.3.2	<i>Jälkiedoinnin mahdollisuudet ja rajoitteet</i>	57
4.4	Konekääntimien väliset erot	59
4.4.1	<i>Virheiden painoarvo ja laatuerot</i>	59
4.4.2	<i>Jälkiedointitarpeen luonne</i>	60
4.4.3	<i>Yhteenveto ja suositukset</i>	61
5	Pohdinta	63
	Lähteet	65
	Liitteet	72
	Liite 1. Analysoitu rinnakkaiskorpusaineisto	72
	Liite 2. Excel-rinnakkaiskorpus ja MQM-pohjainen arviointi	80
	Liite 3. Tutkielman saksankielinen tiivistelmä – Deutsche Zusammenfassung der Abschlussarbeit	92

1 Johdanto

1.1 Tutkielman tausta, merkitys ja tavoitteet

Kääntäminen on kriittinen silta eri kielten ja kulttuurien välillä, ja sen rooli tietoyhteiskunnassa, jossa tieto liikkuu nopeasti ja vaivattomasti yli rajojen, on korostunut entisestään. Viimeisen kymmenen vuoden aikana käänösälalla on tapahtunut poikkeuksellisen nopea teknologinen murros. Neuraalisen konekääntämisen (Neural Machine Translation, NMT) ja sitäkin edistyneempien, syväoppimiseen pohjautuvien LLM-järjestelmien (Large Language Models) kyky tuottaa sujuvaa ja kontekstiherkkää tekstiä on saavuttanut tason, joka haastaa perinteiset käsitykset kääntäjän roolista käänösprosessissa.

Juridisten tekstien kääntäminen edellyttää omanlaistaan tarkkuutta ja asiantuntemusta. Juridisen kielen rekisteri on muodollinen, tiivis ja eksplisiittinen, ja sen kääntäminen vaatii kääntäjältä paitsi vahvaa kielitaitoa myös kontekstuaalista ymmärrystä (Lehto 1999: 120–121). Juridinen merkityssisältö on sidoksissa sekä lähde- että kohdekielen oikeusjärjestelmiin, minkä vuoksi virheellinen terminologia tai rakenteellinen epätarkkuus voi johtaa merkittäviin merkityksenmuutoksiin ja jopa kääntäjän oikeudelliseen vastuuseen (Galdia 2003: 11–13, 18). Tällaisessa ympäristössä käänöksen laatua ei voida taata pelkän yleiskielisen sujuvuuden tai näennäisen vastaavuuden perusteella, vaan käänöksen on säilytettävä tarkasti lähdetekstin sisällölliset, funktionaaliset ja oikeudelliset vaikutukset (Helin 1999: 43; Mattila 2017: 26, 116; ks. lisää luku 3).

Suomentamisen osalta haasteita koneelliselle kääntämiselle luovat esimerkiksi suomen kielen agglutinatiivinen luonne ja monipuolinen morfologia, jotka eroavat merkittävästi monista muista kielistä, kuten indoeurooppalaisista kielistä, joihin monet konekäännösjärjestelmät on ensisijaisesti optimoitu (Anhava 1998: 35; Lankinen ym. 2016: 1; Virpioja ym. 2007: 1). Vaikka LLM-järjestelmät ovat kehittyneet merkittävästi myös nuukemmin resursoitujen kielten osalta, niiden kyky käsitellä suomen kielen erityispiirteitä esimerkiksi juridisessa kontekstissa vaatii syvällisempää tutkimusta. Erityisesti eri kielikuntiin kuuluvien kieliparien, kuten saksa–suomi, kohdalla niiden merkittävät kielelliset ja rakenteelliset erot luovat lisähaasteita, jotka eivät välttämättä ratkea yksinomaan mallien laajalla koulutuksella tai datamäärällä (Lauscher 2020: 4491).

Suurten kielimallien edistyneimmät versiot, kuten esimerkiksi vuonna 2025 julkaistu OpenAI:n ChatGPT-5.2 ja Google Gemini 3.0 Pro, edustavat LLM-teknologian kärkeä ja tarjoavat uusia mahdollisuuksia ja haasteita käännoalalle. Niiden edistynyt kyky ymmärtää ja tuottaa tekstiä tekee niistä hyödyllisen työkalun myös juridisten tekstien kääntämiseen. Yleisten kielimallien rinnalla toimii myös kääntämiseen erikoistuneita ratkaisuja, kuten DeepL Next-Gen - konekääntin, joka hyödyntää LLM-teknologiaa mutta on koulutettu nimenomaan kääntämiseen (DeepL, ei pvm.). Kehittyneen käännösteknologian käyttö tehostaa työskentelyä ja nopeuttaa käännösprosessia merkittävästi. Samalla se siirtää yhä enemmän vastuuta virheiden tunnistamisesta ja korjaamisesta kääntäjälle, jonka työ painottuu entistä enemmän käännösten jälkieditointiin ja laadunvalvontaan perinteisen kielen tuottamisen rinnalla (Salmi 2015: 109).

Tämän tutkielman merkitys liittyy ajankohtaiseen tarpeeseen ymmärtää, miten LLM-teknologiaa hyödyntävillä käännöstyökaluilla tuotetut käännökset toimivat juridisessä kontekstissa ja millaisia vaatimuksia lopputulos asettaa jälkieditoinnille. Tutkielman tavoitteena on selvittää, millaisia virheitä ja puutteita kenties kehittyneimpinä pidettyjen käännösteknologioiden, ChatGPT-5.2:n ja DeepL Next-Genin, tuottamissa saksa-suomikäännöksissä esiintyy sekä millaista asiantuntemusta käännösten viimeistely edellyttää. Lisäksi tutkimus pyrkii lisäämään ymmärrystä näiden työkalujen vahvuuksista ja rajoituksista saksankielisten juridisten tekstien suomentamisessa. Vaikka LLM-teknologiaa hyödyntävä kääntäminen on jo laajasti käytössä, sen systemaattinen arviointi vaativilla erikoisaloilla on vielä vähäistä, varsinkin Suomen kaltaisten pienempien kielialueiden näkökulmasta.

1.2 Tutkimuskysymykset

Tutkielman tavoitteena on vastata seuraaviin kysymyksiin:

1. Millaisia virheitä ja puutteita ChatGPT-5.2 -kielimallin ja DeepL Next-Gen - konekääntimen tuottamissa saksankielisten juridisten tekstien suomennoksissa esiintyy? Tähän kysymykseen vastataan analysoimalla näiden käännöstyökalujen käännöksiä ja vertaamalla niitä ihmiskääntäjän tekemään referenssikäännökseen.
2. Toteutuvatko suomen kielen rakenteelliset erityispiirteet mainittujen työkalujen käännöksissä ja millaisia haasteita ne asettavat jälkieditoinnille? Tässä kysymyksessä paneudutaan kielikunnallisesti hyvin erilaisten kielten, saksan ja suomen, välisiin käännösongelmiin. Kysymyksessä tarkastellaan, kuinka hyvin käännöstyökalut

onnistuvat kääntämään suomen kielen monimutkaiset rakenteet, kuten agglutinatiivisuuden, monipuolisen morfologian ja sanajärjestyksen, ja millaisia korjaustoimenpiteitä jälkieditoijalta vaaditaan.

3. Toteutuvatko juridisen kääntämisen vaatimukset mainittujen käännöstyökalujen käännöksissä ja millaista asiantuntemusta niiden jälkieditointi vaatii? Tässä kysymyksessä keskitytään siihen, miten käännökset vastaavat oikeudellisen kielen vaatimukseen ja millainen oikeustieteellinen ja -kulttuurinen asiantuntemus on välttämätöntä konekäännöksen viimeistelyssä.

1.3 Tutkielman menetelmät, aineisto ja rajaus

Tutkielma on laadullinen korpuspohjainen tutkimus, jossa hyödynnetään rinnakkaiskorpusanalyysiä vertaamalla saksankielistä lähdetekstiä ChatGPT-5.2:lla ja DeepL Next-Genillä tuotettuihin konekäännöksiin sekä ihmiskääntäjän referenssikäännökseen. Virheet luokitellaan MQM-mallin avulla ja jälkieditointitarve arvioidaan viisiportaisella asteikolla (ks. luku 3.3). Painopiste on laadullisessa arvioinnissa, jota tukee määrällinen pisteytys.

Tutkimusaineisto käsittää yhden saksankielisen juridisen tekstin: Euroopan unionin tuomioistuimen lehdistötiedote nro 98/19 (Pressemitteilung Nr. 98/19). Lehdistötiedote tiivistää EU-tuomioistuimen ratkaisun monimutkaisessa tekijänoikeudellisessa asiassa, jossa käsitellään samplauksen¹ oikeudellisia rajoja ja poikkeuksia. Tästä lähdetekstistä on kolme rinnakkaista suomennosversiota, jotka on liitetty tutkielman liitteeseen 1:

1. ChatGPT 5.2 -kielimallilla tuotettu suomennos
2. DeepL Next-Gen -kääntimellä tuotettu suomennos
3. Ihmiskääntäjän laatima referenssisuomennos.

Tutkimus rajataan koskemaan yksinomaan saksankielisten juridisten tekstien suomentamista sekä kahden valitun käännösteknologian – ChatGPT-5.2:n ja DeepL Next-Genin – tuottamien käännösten vertailevaa analyysiä. Tutkimus keskittyy käännösten kielelliseen ja juridiseen laatuun sekä jälkieditoinnin asiantuntijavaatimukseen, jolloin prosessin ajalliset tai taloudelliset

¹ Samplaaminen on tekniikka, jossa äänitteestä otetaan elektronisten apuvälineiden avulla otteita, joita käytetään uuden sävellyksen osana toisessa äänitteessä. Euroopan unionin tuomioistuin, lehdistötiedote nro 98/19.

ulottuvuudet jäävät tämän tarkastelun ulkopuolelle. Aineisto on rajattu yhteen juridiseen tekstiin ja tekstilain, jotta analyysi pysyy syvällisenä ja hallittavana. Tutkimuksen tavoitteena ei ole tehdä yleistyksiä kaikkien LLM-teknologiaa hyödyntävien käännöstyökalujen toimivuudesta, vaan tuottaa tapaustutkimuksen avulla laadullista tietoa niiden mahdollisuuksista ja rajoituksista juridisten tekstien suomentamisessa.

1.4 Tutkielman rakenne

Tutkielma jakautuu viiteen pääosioon. Ensimmäinen luku, *Johdanto*, esittelee tutkimuksen taustan, tavoitteet ja tutkimuskysymykset sekä kuvaa käytetyt menetelmät, aineiston ja rajaukset. Toisessa luvussa, *Konekääntäminen ja juridinen kääntäminen LLM-teknologian aikakaudella*, tarkastellaan lyhyesti konekääntämisen kehitystä neuraalisista malleista suuriin kielimalleihin sekä kuvataan ChatGPT-5.2:n ja DeepL Next-Genin keskeisiä ominaisuuksia käännöstyökaluina. Luku käsittelee myös jälkieditoinnin roolia ja juridisen kääntämisen ominaispiirteitä. Kolmannessa luvussa, *Tutkimuksen toteutus*, esitellään tutkimuksen metodologia, tutkimusaineisto ja sen valintaperusteet sekä analyysimenetelmät.

Neljäs luku, *Analyysi ja tulokset*, muodostaa tutkielman ydinosan. Siinä esitetään tutkimuksen empiiriset havainnot ja vastataan tutkimuskysymyksiin tarkastelemalla käännösten virhetyyppejä, terminologista tarkkuutta, morfologisia ja syntaktisia rakenteita sekä jälkieditoinnin tarvetta ja vaikutusta käännöslaatuun. Viides luku, *Pohdinta*, kokoaa yhteen tutkimuksen keskeiset tulokset ja arvioi niiden merkitystä juridisen kääntämisen ja LLM-teknologian näkökulmista. Luvussa käsitellään myös tutkimuksen luotettavuutta ja rajoitteita sekä esitetään suosituksia ja jatkotutkimusehdotuksia.

2 Konekääntäminen ja juridinen kääntäminen LLM-tekniikan aikakaudella

2.1 Konekääntämisen historia ja siirtymä kohti suuria kielimalleja

Konekääntämisen kehitys on viime vuosikymmeninä ollut poikkeuksellisen nopeaa. Varhaiset sääntöpohjaiset järjestelmät (Rule-Based Machine Translation, RBMT) perustuivat manuaalisesti luotuihin kielellisiin sääntöihin (kielioppi, syntaksi, sanakirjat), joilla lähdetekstiä analysoitiin ja käännettiin sitten generoitiin kohdekielelle. 1990-luvulla kehittynyt tilastollinen konekääntäminen (Statistical Machine Translation, SMT) merkitsi merkittävää siirtymää kielitieteellisestä koodauksesta tilastolliseen estimointiin. Sen sijaan, että järjestelmälle olisi kerrottu säännöillä miten kääntää, järjestelmä opetettiin ennustamaan todennäköisimmät käännösvaihtoehdot suuren datamäärän havaintoihin perustuen. Tilastollisten mallien avulla SMT oppi kaiken tarvittavan tiedon automaattisesti kielidatasta, mikä poisti kokonaan tarpeen luoda käännessääntöjä manuaalisesti. Tilastolliset menetelmät jäivät taka-alalle neuraalisen konekääntämisen (Neural Machine Translation, NMT) yleistymisen myötä 2010-luvun puolivälissä. Muutos sai alkunsa tekoälyn ja erityisesti syväoppimisen (Deep Learning) läpimurroista, minkä ansiosta kielen mallintaminen muuttui huomattavasti tehokkaammaksi. (Chen 2024: 99–100; España-Bonet–Costa-jussa 2016: 5–7; Koehn 2010: xi; Koehn 2020: 11)

NMT-järjestelmät perustuvat monikerroksisiin neuroverkkoihin, jotka oppivat semanttisia ja kontekstuaalisia suhteita laajojen tekstiaineistojen pohjalta. Ne käsittelevät syötteenä annetun lähdekielisen lauseen ja tuottavat siitä suoraan kohdekielisen vastineen ilman käsin määriteltyjä sääntöjä tai erillisiä tilastollisia komponentteja. Koska malli oppii koulutusprosessin aikana ennustamaan kohdetekstiä lähdetekstin kontekstin perusteella (tyypillisesti lause- tai segmenttitasolla), se pystyy tuottamaan käännöksiä kokonaisina lauseina tai kappaleina, mikä mahdollistaa kontekstuaalisesti ja idiomien kannalta tarkemman kielen tuottamisen verrattuna perinteisiin menetelmiin. (Agrawal ym. 2018: 11; Stahlberg 2019: 343; Tan ym. 2020: 4)

Varsinainen läpimurto NMT:ssä tapahtui vuonna 2017, kun kahdeksan Googlen tutkijaa esitteli Transformer-arkkitehtuurin käännteentekevässä artikkelissaan ”Attention Is All You Need” (Vaswani ym. 2017). Transformerin ydin on itsehuomiointimekanismi (self-attention). Toisin kuin aiemmat neuraalimallit, jotka käsitelivät sanasekvenssejä yksi sana kerrallaan ja peräkkäin, Transformer kykenee prosessoimaan koko lauseen kaikki sanat rinnakkaisesti, mikä

tekee koulutuksesta ja käännöksen tuottamisesta huomattavasti nopeampaa ja tehokkaampaa. Itsehuomiointi antaa mallille kyvyn ymmärtää, kuinka kaukana toisistaan olevat sanat liittyvät toisiinsa lauseen sisällä. Se mallintaa, kuinka paljon painoa tai ”huomiota” yhden sanan tulisi antaa jokaiselle muulle sanalle lauseessa kontekstinsa määrittämiseksi. (mts. 1–3; Koehn 2020: 207–208; Tang ym. 2018: 4263) Kyky mallintaa etäisiä riippuvuuksia on huomattavan kriittinen niissä kielipareissa, joissa sanajärjestys ja lauserakenteet eroavat radikaalisti toisistaan. Huomiointimekanismin avulla malli pystyy yhdistämään lauseen alun ja lopun informaation tehokkaammin, mikä parantaa käännöksen kontekstuaalista tarkkuutta. (Chen ym. 2019: 1787–1788)

Transformer-arkkitehtuuri on sittemmin vakiintunut standardiksi paitsi konekääntämisessä, myös laajalti luonnollisen kielen käsittelyn sovelluksissa. Se muodostaa teknologisen perustan valtaosalle suurista kielimalleista, jotka hyödyntävät sen tehokasta itsehuomiointimekanismia globaalien riippuvuussuhteiden mallintamiseen. (Lai 2024: 60–65) Nykyaikaiset NMT-järjestelmät, kuten Googlen ja DeepL:n kaupalliset ratkaisut, rakentuvat pääosin Transformer-pohjaiselle enkooderi–dekooderi-arkkitehtuurille. Niiden koulutus perustuu edelleen massiivisiin rinnakkaiskorpuksiin, jotka tarjoavat lause-kohdistettuja käännösesimerkkejä. Syväoppimiseen ja sekvenssi–sekvenssi-malleihin perustuva lähestymistapa on tuottanut kiistattomia parannuksia käännösten sujuvuuteen, sanajärjestyksen luonnollisuuteen ja kieliopilliseen tarkkuuteen verrattuna aiempiin tilastollisiin menetelmiin. (Gain ym. 2025: 2–3) NMT:n suorituskyky on kuitenkin edelleen vahvasti riippuvainen koulutusdatan laadusta. Erikoisalojen kieli voi tuottaa mallille vaikeuksia, jos opetusdata ei kata alan terminologiaa ja rakenteita. Tämä niin sanottu domain mismatch lisää virheiden riskiä juuri niissä konteksteissa, joissa tarkkuus on kriittistä. (Briva-Iglesias ym. 2024: 1–2; Gezmu–Nürberger 2022: 459; Koehn 2020: 294)

NMT:n ja Transformer-arkkitehtuurin menestys on luonut teknologisen perustan siirtymälle kohti laajempia, yleiskäyttöisiä kielimalleja, joille kääntäminen ei ole enää vain erillinen erikoistehtävä, vaan yksi esiin nouseva kyky osana laajempaa kielellistä ymmärrystä ja generointikykyä. Suurten kielimallien taustalla on edelleen Transformer-arkkitehtuuri, mutta niiden koulutusdata ja -tavoite eroavat perustavanlaatuisesti NMT-malleista. Siinä missä NMT-mallit koulutetaan optimoimaan käännöstehtävää hyödyntämällä jäsenneltyjä, kaksikielisiä rinnakkaiskorpuksia, LLM:t koulutetaan massiivisilla, pääosin yksikielillä yleisaineistoilla. Niitä ei kouluteta ensisijaisesti kääntämään, vaan tuottamaan kieltä ennustamalla tekstiä

kontekstinsa perusteella. Käännöstoiminto rakentuu tämän yleisen kielenoppimisen varaan, ei erillisenä erikoistuneena tavoitteena. (Balashov 2025: 4–5, 12; Gain ym. 2025: 2–3)

Toisin kuin NMT-mallien tyypillinen enkooderi–dekooderi-arkkitehtuuri, useimmat suuret kielimallit, erityisesti generointiin suunnatut, perustuvat pelkkään dekooderiin. Tässä arkkitehtuurissa sama malli käsittelee syötteen ja tuottaa ulostulon yhden yhtenäisen prosessointipinon kautta: dekooderi generoi kohdekielisen tekstin autoregressiivisesti sananosan kerrallaan, hyödyntäen kaikkea aiemmin prosessoitua kontekstia, mukaan lukien kehote (prompting) ja lähdeteksti. Huolimatta tästä arkkitehtuurisesta erosta, LLM voi käytännössä simuloida enkooderi–dekooderi-toimintaa, kun kehoitteeseen syötetään lähdeteksti: kehoitus saattaa mallin implisiittisesti ”koodaamaan” lähteen merkityksen latentiin tilaansa, eli mallin sisäiseen, tiivistettyyn merkitysmalliin, ennen kohdekielen ulostulon generointia. (Balashov 2025: 11; Vaswani ym. 2017: 2–3)

LLM:t yhdistävät perinteisten NMT-järjestelmien tuottaman kieliopillisen sujuvuuden laajempaan kontekstintajuntaan, tyyllilliseen joustavuuteen ja kykyyn tuottaa tietoon perustuvaa tekstiä. Ne pystyvät hyödyntämään niin kutsuttua kontekstissaoppimista (In-Context Learning) ja noudattamaan käyttäjän antamia kehoitteita, minkä ansiosta tyyli ja terminologia voidaan mukauttaa nopeasti. Tämä on arvokasta erikoisaloilla, joilla LLM:ien on hyvällä kehoteohjauksella osoitettu pystyvän ylläpitämään terminologista johdonmukaisuutta tehokkaasti. (Ataman ym. 2025: 9; Balashov 2025: 17–18; Gain ym. 2025: 4–5)

Toisaalta LLM-mallit voivat tuottaa tyyliiltään tai muodollisuudeltaan epäyhtenäistä tekstiä, etenkin ellei niitä ohjata riittävän tarkasti, sillä niiden toimintaperiaate on tuottaa koulutusdatansa pohjalta uskottavalta vaikuttavaa kieltä. Perinteiset NMT-käännösjärjestelmät sen sijaan säilyttävät tyylin yleensä tasaisempana joko rajaamalla harjoitusaineiston tietyn tyyliseksi tai käyttämällä tyyli tageja eli merkintöjä, joiden avulla mallille määritellään esimerkiksi muodollinen tai epämuodollinen tyyli. Koska NMT-järjestelmät on suunniteltu ensisijaisesti siirtämään lähdetekstin sisältö mahdollisimman täsmällisesti kohdekielelle, ne kykenevät hyödyntämään tällaisia tyyliohjeita johdonmukaisesti. (Balashov 2025: 11, ks. myös 30–31)

LLM-mallien generatiivinen luonne aiheuttaa myös ajoittain hallusinaatioita: mallit voivat tuottaa sisällöllisesti virheellisiä, sepitettyjä tai asiayhteyden sopimattomia käännöksiä, jotka

kuitenkin vaikuttavat kielellisesti sujuvilta ja uskottavilta. Tämä on vakava haaste konteksteissa, joissa täsmällisyys on ehdotonta. Tutkimukset ovat osoittaneet hälyttävän korkeita hallusinointiasteita lakialan kyselyissä, ja jopa oikeudelliseen käyttöön tarkoitettut työkalut voivat tuottaa fiktiivisiä viittauksia tai virheellisiä merkitysisältöjä. (Dahl ym. 2024b: 2, 3, 6; Rahman ym. 2025: 1)

Konekäännöksen nykykehityksessä LLM- ja NMT-teknologioiden yhdistäminen on noussut yhdeksi ilmiöksi. Alan trendinä ovat hybridimallit, jotka pyrkivät yhdistämään NMT-mallien hyviä puolia, kuten nopeus ja tarkkuus, LLM-mallien vahvuuksiin, kuten laajempaan kontekstuaaliseen ymmärrykseen ja joustavuuteen. Esimerkiksi tuoreessa tutkimuksessa ”Combining the Best of Both Worlds: A Method for Hybrid NMT and LLM Translation” (Wu ym. 2024) tutkijat esittelevät järjestelmän, jossa NMT vastaa pääosin käännöstyöstä ja LLM aktivoidaan vain valikoivasti kun lause tai konteksti on monimutkainen tai vaatii laajempaa kielellistä ymmärrystä. Tällaisen opportunistisen aktivoinnin on tarkoitus mahdollistaa sekä korkealaatuinen lopputulos että kustannustehokas käännösprosessi. Lisäksi tutkimuksessa ”On-the-Fly Fusion of Large Language Models and Machine Translation” (Hoang ym. 2024) on tutkittu lähestymistapoja, joissa LLM ja NMT yhdistetään dynaamisesti siten, että NMT tuottaa peruskäännöksen, jonka LLM sitten hienosäätää. Tutkimuksen mukaan dynaamisesti yhdistetty NMT–LLM-malli oli laadullisesti ylivoimainen verrattuna kahden puhtaan NMT-mallin yhdistelmään. (Ataman ym. 2025: 20–21; Hoang ym. 2024: 520; Wu ym. 2024: 5140–5141, 5143–5144)

Kieli- ja hybridimallien muuttaessa käännösprosessia ja tuottaessa aiempaa monimuotoisempia käännöstyypppejä, myös käännösten laadunarvioinnin vaatimukset kasvavat. Perinteiset automaattiset metriikat, kuten n-gram-vastaavuuksiin perustuva BLEU (Bilingual Evaluation Understudy), joiden luotettavuus perustuu niiden korrelaatioon ihmisarvioiden kanssa, ovat osoittautuneet riittämättömiksi etenkin erikoisalojen käännöksissä. Ne on suunniteltu ensisijaisesti karkeaan laadunvarmistukseen eli tunnistamaan selvästi huonot käännökset, mutta niiltä puuttuu tarvittava tarkkuus. Koska ne mittaavat vain kielen pintatason samankaltaisuutta eivätkä ymmärrä semantiikkaa tai kontekstia, ne eivät kykene erottamaan hienovaraisia virheitä tai havaitsemaan LLM-käännöksille tyypillistä kontekstisidonnaista idiomaattisuutta ja tyyllillistä hienovaraisuutta, jotka ovat olennaisia asiantuntijatekstien kääntämisessä. Puute on olennainen, sillä LLM-mallien tuottama erittäin sujuva pintataso voi peittää alleen hienovaraisia mutta merkityksellisiä asiavirheitä. (Agrawal ym. 2024: 14481–14492; Kayano–

Sugawara 2025: 114, Koehn 2020: 59–60) Näin ollen ihmisarvioinnin merkitys korostuu LLM-käännösten todellista laatua arvioidessa.

2.2 ChatGPT-5.2 ja DeepL Next-Gen käännöstyökaluina

Tutkielman kohteena olevat työkalut edustavat edellä esiteltyjä lähestymistapoja: ChatGPT-5.2-kielimalli edustaa puhdasta LLM-ratkaisua ja DeepL Next-Gen -konekäännin hybridimallia, jossa LLM toimii yhdessä NMT:n kanssa.

ChatGPT-5 (tai GPT-5) on OpenAI:n kehittämä multimodaalinen suuri kielimalli, joka on suunniteltu olemaan edeltäjiään, kuten GPT-4:ää, merkittävästi älykkäämpi, nopeampi ja luotettavampi, erityisesti monimutkaisessa päättelyssä ja ongelmanratkaisussa. GPT-5:n paranneltuja vahvuuksia ovat eri sisältömuotojen, kuten tekstin, kuvien ja koodin, entistä analyttisempi käsittely. OpenAI:n toimitusjohtaja Sam Altmanin mukaan vuorovaikutus GPT-5:n kanssa vastaa keskustelua tohtoritason asiantuntijan kanssa; GPT-4:ää hän on pitänyt verrannollisena korkeakouluopiskelijaan. OpenAI:n sivut väittävät mallin kykenevän pilkkomaan monimutkaisia tehtäviä pienempiin osiin ja ratkaisemaan niitä järjestelmällisesti ja esimerkiksi jopa koodaamaan kokonaisia sovelluksia alusta loppuun. Mallin ”GPT-5.2 Thinking” -päättelytila varaa itselleen sisäisesti ajatteluaikaa ennen vastauksen tuottamista, vähentäen näin loogisia virheitä ja hallusinaatioita. Käytännössä tämä tarkoittaa, että malli pyrkii huomioimaan tehtävänannon kokonaisuuden, laajemman kontekstin sekä tekstin sisäiset suhteet ennen käännöksen generointia. Päättelytilan käyttö ei muuta mallin koulutusdataa tai kieliresursseja, vaan ainoastaan tapaa, jolla malli hyödyntää niitä vastauksen tuottamisessa. Vaikka mallit edelleen hallusinoivat, OpenAI kertoo GPT-5:n olevan selvästi täsmällisempi ja tarkempi kuin edeltäjänsä. GPT-5:n kerrotaan myös ilmaisevan selkeämmin omat rajoitteensa. (OpenAI 2025; OpenAI 2025b; OpenAI 2025c)

Mainituista parannuksista huolimatta kuitenkin GPT-5:n käännöskyky on muiden GPT-mallien tavoin emergentti ominaisuus, joka syntyy mallin koon, laajan koulutusdatan ja harjoitusmenetelmien yhteisvaikutuksesta. Mallit pystyvät käsittelemään laajoja tekstikokonaisuuksia ja tuottamaan luonnollisia, sujuvia käännöksiä, joskin ne edelleen saattavat myös hallusinoita ja tehdä faktuaalisia virheitä. Pitkissä dokumenteissa malli voi hukata osan kontekstista, tai se saattaa olla tietämätön ajankohtaisista tapahtumista tai päivityksistä. Tuotokset voivat vaihdella tyyliltään, eikä niiden paikkansapitävyyttä voida taata,

ja tarkkuutta vaativat numerot, kaavat tai yksityiskohdat voivat olla epäluotettavia. Lisäksi malli saattaa tehdä virheitä kohdekielen ilmauksissa ja sävyissä, etenkin milloin käyttäjä ei anna tarkkoja kehoitteita rekisterin, tyylin tai terminologian ylläpitämiseksi. (Delikoura ym. 2025: 3, 4; Hajikhani–Cole, ei pvm.: 1, 16; Huang–Liu 2024: 2)

Laadukkaista NMT-käännöskoneistaan tunnettu DeepL on ottanut askeleen kohti LLM-pohjaisuutta julkaistessaan Next-Gen-mallinsa. Next-Gen-malli hyödyntää LLM-infrastruktuuria ja sen on DeepL:n verkkosivujen mukaan tarkoitus parantaa käännöslaatua erityisesti pitkiä tekstejä sisältävissä yhteyksissä. Verkkosivujen mukaan uusi malli on koulutettu käännös- ja sisältödataa varten, hyödyntäen ”yli seitsemän vuoden ajan kertynyttä yksinoikeudellista dataa” sekä ”kieleen erikoistunutta LLM-teknologiaa”. Sen tavoitteena on tuottaa entistä ”ihmismäisempiä” käännöksiä samalla vähentäen yleismalleihin liittyviä hallusinaatioita ja väärän tiedon riskiä. Lisäksi DeepL esittää, että Next-Gen-malli vaatii kilpailijoita, kuten ChatGPT-4:ää, vähemmän jälkieditointia. (DeepL 2024; Deepl, ei pvm.) Epäselvää on, mitä DeepL Next-Gen käytännössä näillä kuvauksilla tarkoittaa, sillä se ei julkaise teknisiä mallejaan eikä täsmennä, yhdistääkö Next-Gen jollain tavalla NMT-komponentteja ja LLM-komponentteja, vai onko kyse jopa puhtaasta LLM-pohjaisesta käännösmallista, jota on vain koulutettu käännösaineistolla.

Viimeaikaiset laajat konekääntämisen vertailutkimukset osoittavat kielimallien vahvistuneen aseman perinteisiin käännöskoneisiin nähden. WMT25-konferenssin (Conference on Machine Translation) yleisen kääntämisen tehtävässä parhaiten suoriutunut järjestelmä oli Googlen Gemini 2.5 Pro, joka sijoittui ihmisarvioinnissa parhaaseen klusteriin 14:ssä arvioidusta 16:sta kieliparista (Kocmi ym. 2025: 374). Samassa yhteydessä tehty laajempi, 55 kieltä kattava WMT24+++-analyysi osoitti, että edistykselliset kielimallit (kuten OpenAI:n o1 ja Claude 3.5) suoriutuivat automaattisilla mittareilla mitattuna paremmin kuin perinteiset käännöspalveluntarjoajat, mukaan lukien DeepL, kaikissa testatuissa kielissä (Deutsch ym. 2025: 1). On kuitenkin huomattava, että automaattiset mittarit saattavat olla puolueellisia ihmiskääntäjiä vastaan ja suosia kielimalleja (mts. 8; Kocmi ym. 2025: 357). DeepL on tästä huolimatta säilyttänyt asemansa vahvana verrokkina; esimerkiksi Intenton vuoden 2024 raportissa DeepL ja GPT-4 suoriutuivat johdonmukaisesti muita malleja paremmin (Flückiger ym. 2025: 2).

Tuoreessa DeepL:n ja Supertextin vertailututkimuksessa havaittiin, että arvioijat suosivat DeepL:ää englanti–saksa-käännöksissä sekä segmentti- että dokumenttitasolla, todennäköisesti kilpailijan lausetason virheiden vuoksi. Muissa kielipareissa tilanne oli kuitenkin päinvastainen: arvioijat suosivat Supertextiä dokumenttitasolla kolmessa neljästä kielisuunnasta. Tutkimuksessa havaittiin, että DeepL:llä oli vaikeuksia ylläpitää terminologian johdonmukaisuutta pitkissä teksteissä – se saattoi kääntää saman termin eri tavoin eri kappaleissa. Vaikka lausetason laatuero oli vähäinen, dokumenttitason analyysi osoitti, että kielimallien laaja konteksti-ikkuna parantaa tekstin eheyttä ja terminologista johdonmukaisuutta. (mts. 1–4)

Yksi keskeinen ero NMT:n ja suurten kielimallien välillä liittyy siihen, miten ne suhtautuvat lähdetekstiin. NMT pyrkii säilyttämään alkuperäisen rakenteen ja terminologian usein hyvinkin konservatiivisesti, mikä voi tehdä käännöksestä jäykän tuntuksen. Toisaalta tämä lähestymistapa voi olla etu esimerkiksi oikeudellisissa teksteissä, joissa uskollisuus lähdetekstille on hyvin tärkeää. Sen sijaan kielimalli painottaa tekstin sujuvuutta ja yhtenäisyyttä, mikä johtaa usein luonnolliseen ja idiomaattiseen lopputulokseen. Samalla se saattaa kuitenkin ajautua liian vapaaseen tulkintaan tai ylitulkintaan, jolloin alkuperäinen merkitys vääristyy. Niin NMT kuin LLM:kin ovat herkkiä domain-siirtymille eli tilanteille, joissa aineisto poikkeaa koulutuksessa käytetystä datasta (ns. out-of-domain-aineisto), mikä voi heikentää terminologian johdonmukaisuutta ja semanttista tarkkuutta. Siksi keskeiset termit, oikeudelliset viittaukset ja kokonaismerkitys tulisi aina varmistaa manuaalisesti käytetystä työkalusta riippumatta. (Bestgen 2022: 1–2; Briva-Iglesias ym. 2024: 1–2; Sato ym. 2020: 4269; Zhang ym. 2025: 52–53)

2.3 Juridisen suomentamisen erityishaasteet konekääntämisessä ja jälkieditoinnissa

Perinteinen, täysin ilman teknologiaa tehtävä käännöstyö on käynyt harvinaiseksi. Käännöstoiminta on siirtynyt malliin, jota kutsutaan ”laajennetuksi kääntämiseksi” (augmented translation): käännösammattilaiset hyödyntävät rutiininomaisesti esimerkiksi konekäännöksiä ja käännösmuisteja, ja teknologia toimii kääntäjän kognitiivisen kapasiteetin jatkeena (Prieto Ramos 2024a: 2). Prieto Ramosin aineistossa (mts. 2) vain noin 12 prosenttia kääntäjistä työskentelee enää ilman teknisiä apuvälineitä.

Konekäännöksen jälkieditoinnilla tarkoitetaan prosessia, jossa kääntäjä tarkistaa ja korjaa konekääntimen tuottaman esiversion siten, että lopputulos on sisällöllisesti tarkka ja täyttää asetetut laatuvaatimukset. (Prieto Ramos 2024b: 155; Vigier-Moreno–Pérez-Macías 2022: 80) Jälkieditoinnin luonne ja laajuus riippuvat tekstilajista sekä käännöksen käyttötarkoituksesta. Juridisessa kääntämisessä tekstien normatiivinen luonne ja oikeudellisen täsmällisyyden vaatimus tekevät jälkieditoinnista poikkeuksellisen vaativaa; kompromisseille on vain vähän tilaa. Tämän vuoksi konekäännösten käyttö juridisessa kääntämisessä edellyttää ammattikäntäjän huomattavasti tiukempaa laadunvalvontaa kuin esimerkiksi journalistisissa tai markkinointiteksteissä, joissa semanttinen toleranssi on laajempi. (Prieto Ramos 2024a: 12–13; Vigier-Moreno–Pérez-Macías 2022: 75)

2.3.1 Juridisen tekstin normatiivinen luonne ja vaikutus jälkieditointiin

Juridisen tekstin ensisijainen funktio on välittää oikeusnormeja ja niihin liittyviä oikeudellisia vaikutuksia. Näin ollen käännöksen on säilytettävä täsmällisesti sekä sisällöllinen merkitys että oikeudellinen funktio (Mattila 2017: 51; Sajavaara 1999: 32). Tarkkuuden tarve liittyy olennaisesti siihen, että oikeuskieli on luonteeltaan vakiintunutta ja institutionaalista ja sen käsitteet ovat usein abstrakteja sekä oikeusjärjestelmässä historiallisesti muotoutuneita (Lehto 1999: 120). Koska termivalinnat, rakenteelliset ratkaisut ja tekstin koherenssi eivät voi vaihdella mielivaltaisesti asiakirjan eri kohdissa, jälkieditoinnille asetetut terminologiset ja rakenteelliset vaatimukset ovat tiukat: johdonmukaisuus on keskeinen oikeusvarmuuden edellytys (Prieto Ramos 2024a: 16).

Vaikka konekääntimet voivat tuottaa lauserakenteellisesti sujuvaa yleiskieltä, niiden kyky hallita oikeusjärjestelmiin sidottuja merkityksiä on huomattavasti heikompi (Vigier-Moreno–Pérez-Macías 2022: 81, 84–85). Tutkimukset osoittavat, että koneet saattavat käyttää samalle juridiselle käsitteelle eri kohdissa eri vastinetta tai valita yleiskielisen ilmauksen täsmällisen oikeustieteellisen termin sijaan (Briva-Iglesias ym. 2024: 19–20). Koska konekääntimet toimivat todennäköisyyslaskennan eivätkä juridisen ymmärryksen varassa, ne eivät kykene hahmottamaan käsitteellisiä epäsymmetrioita tai oikeuskulttuurisia eroja (Prieto Ramos 2024a: 3).

Nämä haasteet korostuvat erityisesti typologisesti etäisissä kielipareissa, kuten saksa ja suomi. Galdia (2003: 13–14) ja Lehto (1999: 122) huomauttavat, että kääntäjän on usein tehtävä valinta

kielen järjestelmän ja oikeusjärjestelmän välillä. Virheet syntyvät usein siitä, ettei käsitteiden historiallista taustaa huomioida. Esimerkiksi saksalainen termipari *natürliche Person* / *juristische Person* kääntyy suomeksi vakiintuneesti epäsymmetriseen muotoon *luonnollinen henkilö* / *oikeushenkilö*, vaikka *juridinen henkilö* olisi kielellisesti tarkempi vastine. Lisäksi suomen kielen morfologiset erityispiirteet voivat johtaa siihen, että konekäännös tuottaa ilmauksia, jotka poikkeavat vakiintuneesta oikeuskielestä (Koskeniemi ym. 2012: 47). Suoraa käännösvastinetta ei voi hyväksyä, mikäli se poikkeaa vakiintuneesta suomalaisesta oikeuskielestä. Myös harmonisoitu EU-oikeus tuo mukanaan termikerrostumia, jotka heijastavat useiden eri oikeusperinteiden ja kielialueiden välistä ylikansallista synteesiä ja vaativat siten erityistä tarkkuutta (Lehto 1999: 128–129).

Näistä lähtökohdista seuraa, että jälkieditoijan oikeudellinen asiantuntemus on keskeinen laadunvarmistuksen elementti. Tarvitaan kykyä tunnistaa sekä lähdekielen että kohdekielen oikeusjärjestelmien konventiot ja niiden väliset erot. Prieto Ramosin (2024b) päivitettyssä mallissa tämä jäsenyytys strategiseksi ja temaattiseksi kompetenssiksi. Strateginen kompetenssi on kääntämisen ”moottori”, se mekanismi, joka ohjaa käännösprosessia ja pitää sen käynnissä. Se analysoi konekäännöksen soveltuvuutta ja päättää, mitä muita taitoja, kuten kieli- tai substanssiosaamista, tarvitaan milloinkin. Sen tueksi välttämätön temaattinen ja kulttuurinen kompetenssi puolestaan tarkoittaa syvällistä tietoa lähde- ja kohdekielen oikeusjärjestelmistä ja käsitteiden hierarkiasta. Substanssiosaaminen on välttämätöntä koneellisen prosessoinnin riskejä hallittaessa, jotta kääntäjä pystyy tunnistamaan tilanteet, joissa koneen tuotos voi olla harhaanjohtava. (Prieto Ramos mts. 150, 155–157, 169)

Juridisen tekstin normatiivinen luonne ei aina tarkoita yksiselitteisyyttä. Lehto (1999: 128) nostaa esiin olennaisen periaatteen: jos lähdeteksti on tulkinnanvarainen, sama tulkinnanvaraisuus on säilytettävä myös käännöksessä. Tämä on huomattava haaste konekäännöksen jälkieditoinnille, sillä konekääntimet pyrkivät tyypillisesti ratkaisemaan moniselitteisyydet valitsemalla todennäköisimmän käännösvastineen, jolloin alkuperäinen tahallinen väljyys tai monitulkintaisuus saattaa kadota. Jälkieditoijan on kyettävä tunnistamaan, milloin epäselvyys on tekstin ominaisuus eikä virhe, ja palautettava käännökseen sen vaatima tulkinnanvara. (Rojas Plata–Castro Sánchez 2024: 2–3, 5; Mosavi Miangah–Zolfaghar Kondori 2015: 1–3)

Edellä kuvatut semanttiset, terminologiset ja tulkinnalliset riskit konkretisoituvat esimerkiksi institutionaalisessa juridisessa kääntämisessä, jonka osalta Prieto Ramos (2024a) tunnistaa kaksi keskeisestä laadun indikaattoria: terminologian intertekstuaalisen yhdenmukaisuuden ja instituutiokonventioiden noudattamisen. Intertekstuaalisella yhdenmukaisuudella tarkoitetaan sitä, että käännös on systemaattisesti linjassa organisaation aiemmin tuottamien ja hyväksytyjen tekstien kanssa, eikä ainoastaan yksittäisen lähdetekstin tai käännöstehtävän sisäisesti johdonmukainen. Prieto Ramosin mukaan institutionaaliset tekstit muodostavat normatiivisen jatkumon, jossa uudet käännökset suhteutuvat olemassa oleviin diskurssikonventioihin ja vakiintuneisiin terminologisiin ratkaisuihin oikeudellisen johdonmukaisuuden ja luotettavuuden säilyttämiseksi.

Nämä vaatimukset ovat merkityksellisiä konekäännöksen jälkieditoinnissa, sillä konekäännös ei kykene itsenäisesti tunnistamaan institutionaalista jatkumoa, aiempien käännösratkaisujen normatiivista painoarvoa eikä organisaatiokohtaisia diskurssikäytäntöjä. (mts. 16, 17, 22, 23) Vastuu terminologisten valintojen, käsitteellisen johdonmukaisuuden ja oikeudellisen funktion säilyttämisestä jää kääntäjälle.

2.3.2 Suomen kielen morfologiset ja syntaktiset ominaisuudet jälkieditoinnin näkökulmasta

Suomen kieli muodostaa konekäännösjärjestelmille merkittävän haasteen johtuen sen suuresta rakenteellisesta etäisyydestä dominoiviin lähde- ja koulutuskieliin, kuten englantiin ja saksaan. Tämä etäisyys ulottuu sanastoa syvemmälle kielen syvärakenteisiin: morfologiaan, syntaksiin ja informaatorakenteeseen. Nykyisten konekäännösjärjestelmien koulutusdata on painottunut vahvasti kielille, joissa merkitykset ilmaistaan usein erillisinä sanoina. Siksi mallit kohtaavat vaikeuksia suomen kaltaisissa agglutinatiivisissa kielissä, joissa merkitykset rakentuvat sanansisäisesti morfeemien välisen vuorovaikutuksen kautta. (Koskenniemi ym. 2012: 2, 3, 10–11; Löfberg ym. 2003: 458, 461; Singh ym. 2025: johdanto).

2.3.2.1 Morfologinen moniselitteisyys ja sijamuotojen funktionaalinen kuorma

Suomen sijamuotojen moniselitteisyys eli polysemia muodostaa erityisen haasteen konekääntimille. Indoeurooppalaisissa kielissä prepositiot rajaavat merkityksen usein tarkemmin kuin suomen laaja-alaiset sijapäätteet. Esimerkiksi inessiivi (-ssa/-ssä) ilmaisee tyypillisesti paikkaa (kuten englannin prepositio *in*), mutta verbin ”erehtyä” yhteydessä se

koodaa asiasisältöä: ”erehdyn pituudessa” ei viittaa sijaintiin, vaan virheeseen suhteessa pituuteen (Chen–Hulden 2022: 2610–2011).

Haasteellisuus kasvaa, kun yhdistetään morfologinen rikkaus ja kontekstiriippuvuus. Nomineilla voi olla jopa 2 000 ja verbeillä yli 12 000 eri taivutusmuotoa, kun huomioidaan sija-, luku- ja liitepartikkelitaivutus (Koskenniemi ym. 2012: 11). Neuroverkkokäännöksessä tämä johtaa datan hajaantumiseen: kone ei välttämättä ole nähnyt koulutusaineistossaan juuri sitä taivutusmuotoa, jota juridinen konteksti vaatisi, vaan arvaa muodon tilastollisen todennäköisyyden perusteella (Koponen 2016b: 47; Weller-Di Marco ym. 2022: 1).

Saksa–suomi-kielipari saattaa olla tässä suhteessa erityisen vaativa, sillä molemmat ovat morfologisesti rikkaita kieliä, vaikka kuuluvatkin eri kielikuntiin (Alho 2025: 278). Alhon tutkimuksessa toiseksi suurin virhekatgoria (suurimman ollessa ”väärä sana”, jonka osuus kaikista virheistä oli 34,9 %) oli mielenkiintoisesti ”oikea sana, väärä päätte” (32,9 %), joskaan nämä virheet eivät selittyneet pelkästään kielten morfologisilla eroilla, vaan johtuivat erityisesti haasteista muuntaa suomen imperatiivi- tai sinuttelumuotoja saksankielisille ohjeteksteille tyypillisiin infinitiivi- tai teitittelyrakenteisiin. (Alho 2025: 273, 283).

2.3.2.2 Informaatorakenne ja sanajärjestyksen haasteet

Suomen kielen sanajärjestys ei ole vapaa, vaan informaatorakenteen ohjaama ja ilmaisee systemaattisesti tutun ja uuden tiedon suhdetta – teeman ja reeman järjestystä (Vilkuna 1989: 9; Koskenniemi ym. 2012: 47). Juridisessa tekstissä tämä on kriittistä, sillä argumentaation looginen eteneminen ja painotusten täsmällisyys riippuvat siitä, missä järjestyksessä informaatio esitetään. Neuraalinen konekäännös tuottaa kuitenkin usein kirjaimellisia käännöksiä, jotka noudattavat lähdekielen syntaktista rakennetta. Tämä voi muuttaa alkuperäisen tekstin painotuksia ja heikentää koheesiota, erityisesti monimutkaisissa lauserakenteissa (Quinci–Pontrandolfo 2023: 178).

Saksa–suomi-kieliparia käsitteleviä tutkimuksia ei ole saatavilla, mutta Gerhalter (2024: 16–17) osoittaa, että DeepL:n espanja–portugali-käännöksissä informaatorakenne vääristyy systemaattisesti englannin kautta tapahtuvan kääntämisen seurauksena: esimerkiksi portugalin ”*Falar, falo eu*” (”Puhuja olen minä”, jossa korostetaan nimenomaan minän roolia puhujana) kääntyy espanjaksi muotoon ”*Yo hablaré*” (”Minä puhun”), minkä seurauksena verbin jälkeen

sijoitetun tekijän tuottama painotus häviää ja lauseen rakenne muuttuu tavanomaiseksi perussanajärjestykseksi.

2.3.2.3 Lauseenvastikkeet ja tiivis ilmaisutapa

Suomen kielen osalta tiiviys toteutuu usein lauseenvastikkeiden runsaana käyttönä. Vaikka lauseenvastikkeet säästävät tilaa, ne voivat hämärtää asioiden välisiä suhteita ja tehdä tekstistä monitulkintaista jopa äidinkielisille lukijoille. Esimerkiksi Rydzewska-Siemiątkowskan Kielikellossa julkaistussa artikkelissa (2024) on tarkasteltu lakitekstin ymmärrettävyyttä ja havaittu, että rakenteet kuten ”omaisuuden ositusta toimittaessa” ja ”avioliiton purkaututtua” osoittautuvat monille lukijoille vaikeiksi. Artikkelissa esitellään säädösteksti:

”Omaisuu den ositusta toimittaessa on omaisuutena, johon toisella puoliso lla ei ole avio-oikeutta, pidettävä: 1) omaisuutta, jonka puoliso on ansainnut taikka saanut perintönä, lahjana tai testamentin nojalla avioliiton purkaututtua toisen puolison kuoleman vuoksi; – –.”

Lauseenvastikkeiden tiiviys voi johtaa semanttiseen epäselvyyteen: edellä mainitussa esimerkissä lukijan on vaikea hahmottaa, mihin ajankohtaan tai ehtoon rakenne tarkalleen viittaa. Artikkelin mukaan lukijat saattaisivat ymmärtää kohdan helpommin, jos se purettaisiin sivulauseeksi (esim. ”sen jälkeen, kun [avioliitto] on purkautunut”). (Rydzewska-Siemiątkowska 2024, kohta ”Säädöstekstiä vaikeuttavat piirteet”)

Vaikka nimenomaan suomen kielen lauseenvastikkeiden ja konekääntämisen laadun välisestä suhteesta on toistaiseksi vähän suoraa empiiristä tutkimustietoa, neuraalisen konekääntämisen on yleisesti havaittu altistuvan virhetulkinnoille juuri monimutkaisten ja tiiviiden rakenteiden kohdalla (ks. Dahl ym. 2024b: 64, 69, 76).

2.3.2.4 Viittaussuhteiden katkeaminen ja artikkelittomuus

Suomen kieli eroaa monista muista kielistä siinä, ettei siinä ole artikkeleita määräisyyden ilmaisemiseksi. Artikkelien puute siirtää määräisyyden ilmaisun muihin keinoihin, kuten sanajärjestykseen, demonstratiivipronomineihin ja kontekstiin. Indoeurooppalaisissa kielissä sen sijaan artikkelit (kuten saksan *der/ein* tai englannin *the/a*) osoittavat eksplisiittisesti, viitataanko tiettyyn yksilöön vai mihin tahansa lajin edustajaan. (Chesterman 2005: 2–3, 90, 104) Koska tämä ero ei näy suomen pintarakenteessa, käännösprosessissa syntyy haasteita: esimerkiksi saksan *der Vertrag* (juuri tämä sopimus) ja *ein Vertrag* (mikä tahansa sopimus)

kääntyvät molemmat usein pelkäksi sopimukseksi. Tällöin jälkieditoijan on tarvittaessa täsmennettävä ilmaisua lisäämällä demonstratiivipronomineja tai tarkenteita, kuten *tämä sopimus* tai *kyseinen sopimus*.

Vaikka nykyaikaiset NMT- ja LLM-pohjaiset järjestelmät kykenevät hyödyntämään laajempaa kontekstia aiempaa paremmin, viittaussuhteiden hallinta ja diskurssitason johdonmukaisuus ovat edelleen konekäännöksen heikkoja kohtia. Perinteiset NMT-järjestelmät käsittelevät lauseita erillisinä yksiköinä, mikä vaikeuttaa mm. pronominien ja viittausten oikeaa tulkintaa ilman laajempaa tekstikontekstia. Tutkimukset osoittavat, että uudemmat context-aware-NMT-mallit tuottavat vain jossain määrin parannuksia perinteisiin lausekohtaisiin malleihin nähden, ja monet diskurssi-ilmiöt vaativat edelleen mittavia parannuksia mallien rakenteeseen ja arviointimenetelmiin (esim. Jin ym. 2023: 15246–15252). Erityisesti pitkissä ja monimutkaisissa teksteissä järjestelmät eivät pysty luotettavasti seuraamaan referenttejä, mikä voi johtaa viittaussuhteiden katkeamiseen. (Voita ym. 2018: 1264)

Voitan työryhmä käyttää NMT-järjestelmiä kuvaavana esimerkkinä englannin pronominia *it*, joka voi viitata mihin tahansa esineeseen, mutta esimerkiksi venäjässä tai ranskassa pronominin on vastattava viitattavan sanan kieliopillista sukua (maskuliini, feminiini tai neutri). Perinteinen NMT-malli valitsee usein yleisimmän suvun (kuten maskuliinin), vaikka konteksti vaatisi muuta. Esimerkiksi venäjän kielelle käännettäessä mallit tekevät merkittäviä virheitä, jos pronominin viittaussuhde vaatii feminiiniä tai monikkoa. Viittaussuhde siis katkeaa: jos järjestelmä ei ”näe” edellistä lausetta, se arvaa suvun väärin, rikkoen näin tekstin loogisen rakenteen. (Voita ym. 2018: 1269)

Karpinska ja Iyyer (2023: 427, Table 3) puolestaan osoittavat tutkimuksensa venäjä-puoläkäännösesimerkissä, miten venäjän feminiininen pronomini *ona* (hän/se) voi kääntyä virheellisesti puolan feminiinimuotoon *ona*, vaikka viittauskohde (paperi) on puolassa maskuliinisukuinen substantiivi ja vaatisi maskuliinimuotoisen pronominin *on*. Lausekohtainen käännös ei kykene päättelemään oikeaa kieliopillista sukua, koska viittauskohde mainitaan edellisessä lauseessa, mikä johtaa virheelliseen pronominivalintaan. Viittaussuhdeongelma on sovellettavissa myös saksa-suomi -kielipariin: saksan kieliopilliset suvut auttavat seuraamaan referenttejä (esim. *er/sie/es* viittaamassa edellisen lauseen substantiiviin), kun taas suomen sukuneutraali ”se” tai ”hän” kadottaa tämän tiedon samalla tavalla kuin venäjän esimerkeissä.

Suurten kielimallien kyky käsitellä pidempiä tekstikonteksteja on merkittävästi parempi kuin puhtaiden NMT-kääntäjien, mutta nekään eivät aina pysty varmistamaan loogisesti johdonmukaisten viittaussuhteiden säilymistä. Pitkät tekstit, joissa referentit rakentuvat monitasoisesti ja pronomini viittaukset kytkeytyvät kaukana aiemmassa diskurssissa esiintyviin elementteihin, edelleen usein johtavat epätarkkoihin tai epäselviin käännöksiin. Kuratovin työryhmän (2024: 1–5) mukaan suositut kielimallit käyttävät tehokkaasti vain noin 10–20 % käytettävissä olevasta kontekstista; esimerkiksi GPT-4 hyödyntää tehokkaasti vain noin 10 % 128 000 tokenin ikkunastaan, ja suorituskyky laskee jyrkästi tekstin pituuden ja tehtävän monimutkaisuuden kasvaessa.

Suomen kielen artikkelittomuuden ja konekääntämisen suhteesta tarvittaisiin tutkimusta. On mielenkiintoista nähdä, tulevatko viittaussuhteiden haasteet tai muut edellä esitetyt käännöshaasteet esiin tässä tutkielmassa hyödynnettyjen kääntimien tuotoksissa.

3 Tutkimuksen toteutus

Tutkimus toteutettiin laadullisena tapaustutkimuksena, joka sijoittuu deskriptiivisen käännöstutkimuksen kentälle; painopiste on käännösilmiöiden kuvailemisessa, havainnoinnissa ja selittämisessä. Toisin kuin perinteinen käännösarviointi, deskriptiivinen käännöstutkimus ei pyri ainoastaan arvostelemaan käännöksiä normatiivisesti ”oikeiksi” tai ”vääriksi”, vaan tarkastelee niitä todellisina kielellisinä ja kulttuurisina tuotoksina. (Tiittula–Kolehmainen 2020: 948) Deskriptiivinen viitekehys mahdollistaa konekäännösten ominaispiirteiden, kuten mahdollisten hallusinaatioiden tai indoeurooppalaisten rakenteiden heijastumien, yksityiskohtaisen arvioinnin suhteessa ihmiskääntäjän tuottamaan referenssiin. Tutkimuksessa hyödynnettiin metodologista pluralismia yhdistämällä systemaattinen laadullinen analyysi ja määrällinen virhefrekvenssien vertailu. Seuraavassa perustellaan valitut lähestymistavat ja niiden soveltuvuus tutkimusasetelmaan.

3.1 Tutkimusasetelma ja metodologiset valinnat

Tutkimusasetelma perustuu vertailevaan rinnakkaiskorpusanalyysiin. Saksankielistä juridista lähdetekstiä verrataan kolmeen suomenkieliseen vastineeseen: ChatGPT-5.2:lla ja DeepL Next-Genillä tuotettuihin raakakäännöksiin sekä ihmiskääntäjän laatimaan referenssikäännökseen. Analyysi etenee segmentoinnista virheiden tunnistamiseen ja MQM-pohjaiseen luokitteluun (ks. yksityiskohtainen kuvaus luvussa 3.2.3).

Tutkimuksen metodologiseksi lähtökohdaksi valittiin laadullinen tutkimus, sillä juridisen kielen kääntäminen edellyttää sellaista kontekstuaalista ja käsitteellistä tarkkuutta, jota perinteiset määrälliset menetelmät eivät kykene tavoittamaan. Laadullisille tutkimusstrategioille on ominaista pyrkimys ymmärtää tutkimuskohdetta ja sen merkityksiä kokonaisvaltaisesti. Laadullinen ote on käsillä olevalle tutkimukselle välttämätön, jotta voidaan tunnistaa ne hienovaraiset merkityserot ja rakenteelliset epätarkkuudet, jotka syntyvät juridisen kääntämisen prosessissa. (Luodonpää-Manni–Hamunen 2020: 13–14)

Laadulliselle tutkimukselle on ominaista analyysin syvällisyys ja monipuolisuus, joiden avulla voidaan lisätä ymmärrystä tutkittavan ilmiön luonteesta. Tutkimuksen kokonaisvaltaisen lähestymistavan vuoksi aineisto on usein määrällistä tutkimusta suppeampi, mikä mahdollistaa paneutumisen kielelliseen viestintään yksittäistapausten tasolla. Metodologisessa valinnassa keskeistä ei ole aineiston laajuus, vaan tutkimuskysymysten luonne. Tavoitteena ei ole niinkään

kielellisten ilmiöiden yleisyyden tai säännönmukaisuuksien mittaaminen, vaan ilmiön syvällinen ymmärtäminen ja sen tarkastelu, millaisia mahdollisuuksia kielellinen toiminta tarjoaa. (Luodonpää-Manni-Hamunen 2020: 13–14)

Konekääntämisen arvioinnissa pitkään käytetyt automaattiset metriikat, kuten BLEU (Bilingual Evaluation Understudy), perustuvat pääosin pintatason n-gram-vastaavuuksiin ihmislähdekäännösten kanssa. Esimerkiksi Seljan, Brkić ja Vičić (2012) osoittavat, että BLEU-metriikan antamat pistemäärät eivät juridisten tekstien kohdalla korreloi johdonmukaisesti ihmisten tekemien laatuarvioiden kanssa, mikä viittaa siihen, että n-gram-pohjainen automaattinen arviointi ei tavoita lakitekstien käännöksissä olennaisia merkitys- ja terminologiakysymyksiä. (Seljan ym. 2012: 2147) Nykyään konekäännösten laadun arvioinnissa käytetään yhä enemmän semanttisiin kielimalleihin perustuvia metriikoita, kuten BERTScorea ja COMETia. BERTScore vertailee käännöksen ja referenssin sanojen kontekstuaalisia upotuksia mahdollistaen näin synonyymien ja kielioppivariaatioiden tunnistamisen ja korreloi paremmin ihmisen arviointien kanssa kuin pinnalliset metriikat kuten BLEU. COMET puolestaan hyödyntää monikielisiä esikoulutettuja kielimalleja ja ennustaa käännösten laatua huomioiden sekä lähdetekstin että referenssin merkityssisällön, mikä tekee siitä erityisen tehokkaan semanttisesti tarkkojen arvioiden tekemiseen. (Glushkova ym. 2023: luvut 1 ja 3.3; Zhang ym. 2020: 1, 7, 15)

Kehittyneemmistä metriikoista huolimatta Agrawal, Farinhas, Rei ja Martins (2024) huomauttavat, että automaattiset arviointimenetelmät tavoittavat ainoastaan kyvyn erotella laadullisesti heikot ja vahvat käännösparit, mutta ne jäävät epäherkiksi vivahteikkaille laatueroille. Tämä metodologinen rajoite korostuu erityisesti tilanteissa, joissa käännösten yleinen laatutaso on korkea ja versioiden välinen varianssi vähäinen. (Agrawal ym. 2024: 14491–14492) Juridisessa viitekehyksessä vähäinenkin terminologinen epätarkkuus saattaa vääristää oikeudellisen argumentaation sisällön, vaikka käännöksen pintarakenteellinen muoto vaikuttaisi korrektilta. Laadullinen tutkimusote tarjoaa tässä tutkimuksessa tehokkaan tavan tarkastella systemaattisesti käännösten kriittisiä kohtia ja selvittää, millaisia vaikutuksia yksittäisillä käännösratkaisuilla on oikeudelliseen merkitykseen.

Vaikka tutkimuksen painopiste on laadullisessa tarkastelussa, laadullisen rinnakkaiskorpusanalyysin rinnalla käytetään määrällisenä elementtinä havaittujen virheiden ja puutteiden tilastollista vertailua. Analyysissä sovelletaan MQM-mallia, joka mahdollistaa

virheiden erittelyn sekä tyyppin että vakavuusasteen mukaan. (Lommel ym. 2024: luku 1) Määrällinen vertailu tuo tutkimukseen systemaattisuutta ja auttaa hahmottamaan ChatGPT-5.2:n ja DeepL Next-Genin välisiä suorituskykyeroja. Yhdistelmä laadullista analyysia ja sitä tukeva määrällinen vertailu vahvistaa tutkimuksen luotettavuutta ja tarjoaa kattavamman kuvan nykyisten kielimallien soveltuvuudesta juridiseen kääntämiseen.

3.2 Aineiston valinta, keruu ja käsittely: Lehdistötiedote nro 98/19

3.2.1 Aineiston valinta ja perustelu

Tutkimusaineisto perustuu yhteen saksankieliseen oikeudelliseen tekstiin: 29. heinäkuuta 2019 julkaistuun Euroopan unionin tuomioistuimen lehdistötiedotteeseen nro 98/19 (Pressemitteilung Nr. 98/19). Lehdistötiedote tiivistää Euroopan unionin tuomioistuimen ennakkoratkaisun asiassa C-476/17 (*Pelham GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian Schneider-Esleben*), joka koskee musiikin samplauksen oikeudellisia rajoja ja tekijänoikeudellisia poikkeuksia. Tapauksen taustalla on riita, jossa katsottiin, että noin kahden sekunnin pituinen rytmisekvenssi Kraftwerk-yhtyeen kappaleesta *Metall auf Metall* oli samplattu ilman suostumusta ja käytetty osana toista musiikkiteosta.

Ratkaisussa unionin tuomioistuin määrittelee, millä edellytyksillä äänitteeltä otetun katkelman käyttö kuuluu äänitteen valmistajan kappaleen valmistamista koskevan yksinoikeuden piiriin ja milloin samplaaminen voi jäädä tämän yksinoikeuden ulkopuolelle. Erityistä huomiota kiinnitetään tilanteisiin, joissa äänikatkelma on muunneltu siten, ettei se ole kuuntelemalla tunnistettavissa. Tapaus on juridisesti merkittävä, sillä se asettaa vastakkain tekijänoikeuden ja lähioikeuksien suojan sekä Euroopan unionin perusoikeuskirjassa turvatun taiteen vapauden ja edellyttää näiden perusoikeuksien välistä punnintaa.

Lehdistötiedote soveltuu tutkielman tutkimusasetelmaan useista syistä. Ensinnäkin aineiston oikeudellinen luonne ja haasteellisuus tekevät siitä otollisen tutkimuskohteen. Lehdistötiedote käsittelee monimutkaista tekijänoikeusoikeudellista kysymystä, jonka ymmärtäminen ja kääntäminen edellyttää täsmällistä terminologian hallintaa ja perehtyneisyyttä EU-oikeudelliseen viitekehykseen. Teksti on terminologisesti vaativa: se sisältää haastavia EU-oikeuden käsitteitä, kuten *verwandte Schutzrechte* / lähioikeudet, *Vervielfältigung* / kappaleenvalmistus ja *vollständige Harmonisierung* / täysimääräinen yhdenmukaistamistoimenpide, ja viittaa EU:n perusoikeuskirjaan.

Toiseksi lehdistötiedote yhdistää eri rekistereitä mielenkiintoisella tavalla. Vaikka se on virallinen oikeudellinen dokumentti, se on suunnattu sekä oikeusalan ammattilaisille että suurelle yleisölle. Tekstissä yhdistyvät lakitekninen täsmällisyys, kuten tarkat viittaukset direktiiveihin ja artikloihin, ja yleistajuiset selitykset, kuten alaviitteessä esitetty selvitys siitä, mitä samplaus tarkoittaa. Tämä antaa mahdollisuuden tutkia, kuinka joustavasti kääntimet kykenevät mukautumaan rekisterin vaihteluihin saman tekstikokonaisuuden sisällä.

Kolmanneksi virallisen suomenkielisen käännöksen puuttuminen tekee aineistosta otollisen käännösjäljen tarkastelulle ilman aiemman käännöksen ohjaavaa vaikutusta. Lisäksi aineiston avoin saatavuus EU-tuomioistuimen verkkosivuilla (<https://curia.europa.eu>) takaa sen, että tutkimuksen käännösratkaisuja voidaan tarkastella luotettavasti suhteessa alkuperäiseen lähteeseen.

Lisäksi aineiston laajuus on sopiva tutkielman tarkoituksiin. Saksankielinen lehdistötiedote käsittää noin 1200 sanaa ja 24 täyttä virkettä. Laajuus on riittävä vertailuanalyysin tekemiseen, mutta samalla hallittavissa syvälliseen laadulliseen tarkasteluun. Tekstin pituus mahdollistaa useiden erityyppisten käännösongelmien esiintymisen, mutta ei ole niin laaja, että yksityiskohtainen virhetarkastelu kävisi liian raskaaksi.

3.2.2 Aineiston keruu

Tutkimusaineisto kerättiin 11. tammikuuta 2026. Käännösten tuottamisessa käytettiin seuraavia menetelmiä ja ohjeistuksia:

- i) ChatGPT-5.2 -käännös: Käännös tuotettiin käyttämällä OpenAI:n ChatGPT-5.2 -kielimallia ChatGPT Plus -verkkokäyttöliittymän kautta (versio: GPT-5.2, malli julkaistu 11.12.2025). Mallille asetettiin ”Thinking” -päätelytila ja annettiin seuraava kehotus:

”Olet kokenut juridinen kääntäjä, joka on erikoistunut Euroopan unionin oikeuteen ja immateriaalioikeuteen. Tehtäväsi on kääntää liitteenä oleva Euroopan unionin tuomioistuimen saksankielinen lehdistötiedote suomeksi.

ANALYYSIVAIHE

Suorita ennen käännöstä sisäinen analyysi, jossa:

- *tunnistat tekstin oikeudellisen kontekstin ja keskeiset oikeudelliset käsitteet*

- varmistat keskeisten termien (esim. *Vervielfältigung, verwandte Schutzrechte, Tonträgerhersteller*) vakiintuneet EU-oikeudelliset suomenkieliset vastineet
- arvioit, missä kohdissa saksan ja suomen oikeusjärjestelmien sekä oikeuskielen erot edellyttävät erityistä huolellisuutta käännösratkaisuissa

KÄÄNNÖSPERIAATTEET

1. Terminologia

- Käytä johdonmukaisesti samaa suomenkielistä vastinetta kullekin juridiselle käsitteelle koko tekstissä
- Noudata EU:n virallista terminologiaa ja vakiintuneita suomalaisen oikeuskielen ilmauksia
- Priorisoi EU-oikeuden olemassa olevia suomenkielisiä käännöskonventioita myös silloin, kun yleiskielinen ilmaus vaikuttaisi luontevammalta
- Älä käännä direktiivien numeroita, otsikkotietoja tai virallisia viittauksia

2. Juridinen tarkkuus

- Säilytä alkuperäinen oikeusvaikutus täsmällisesti
- Varmista, että toimijuussuhteet, oikeudelliset vastuut ja normatiiviset suhteet käyvät yksiselitteisesti ilmi
- Älä lisää tekstiin tulkintoja, selityksiä tai täsmennyksiä, joita lähdetekstissä ei ole
- Jos et ole varma termin vakiintuneesta EU-oikeudellisesta vastineesta, käännä konservatiivisesti tai säilytä lähdetekstin termi

3. Suomen kieli

- Käytä oikeita sijamuotoja juridisten suhteiden ilmaisemiseen
- Huomioi objektin sijanvaihtelu (akkusatiivi / partitiivi) aspektin ja oikeusvaikutuksen ilmaisemisessa
- Rakenna lauseet suomen kielen luontaisen teema–reema-järjestyksen mukaisesti
- Vältä saksan syntaksin mekaanista jäljittelyä ja käännösuomea

4. Tyyli ja rekisteri

- Ylläpidä muodollista, mutta sujuvaa ja ymmärrettävää oikeuskieltä
- Säilytä EU-tuomioistuimen lehdistötiedotteen hybridityyli: lakitekninen täsmällisyys yhdistettynä yleistajuiseen esitystapaan
- Vältä arkikielisiä tai journalistisesti kärjistettyjä ilmaisuja

5. Ulkoasu ja muotoilu

- Säilytä lähdetekstin lihavoinnit ja kursivoinnit täsmällisesti vastaavissa kohdissa suomenkielistä käännöstä.
- Varmista, että tekstin visuaalinen hierarkia ja painotukset välittyvät lukijalle samalla tavalla kuin alkuperäisessä asiakirjassa.

TUOTANTO-OHJE

- *Tuota lopullinen suomenkielinen käännös yhtenäisenä tekstinä*
- *Älä kommentoi käännösratkaisuja*

Käännä nyt annettu teksti edellä mainittuja periaatteita noudattaen.”

”Thinking”-päätelytila valittiin hallusinaatioiden ehkäisemiseksi ja juridisen logiikan ylläpitämiseksi. Käytetty kehoitus kehitettiin iteratiivisesti hyödyntämällä ns. prompt engineering -tekniikoita. Lopullisen kehotteen muotoilussa käytettiin apuna useita eri suuria kielimalleja (ChatGPT, Claude, Gemini), joita pyydettiin optimoimaan kääntäjälle annettavia ohjeita juridisen tarkkuuden, EU-oikeudellisen terminologian ja suomen kielen vaatimusten näkökulmasta. Prosessilla pyrittiin varmistamaan, että kielimalli saa mahdollisimman tarkan kontekstin ja rooliohjeistuksen asiantuntijatasoon käännöksen tuottamiseksi. Lähdeteksti syötettiin mallille yhtenä kokonaisuutena, jotta se pystyisi hyödyntämään laajaa konteksti-ikkunansa ja säilyttämään terminologisen johdonmukaisuuden koko tekstin läpi. Käännös tuotettiin yhdellä kerralla ilman jälkikäteen tehtyjä korjauksia tai uudelleengenerointeja, jotta analyysi kohdistuisi nimenomaan kielimallin välittömään raakatuotokseen ilman tutkijan ohjaavaa vaikutusta prosessin aikana.

- ii) DeepL Next-Gen -käännös: Käännös tuotettiin maksullisen DeepL Pro -palvelun kautta käyttäen DeepL:n heinäkuussa 2024 käyttöön otettua viimeisintä Next-Gen -käännösteknologiaa. Käännöksen asetuksina käytettiin:
- Lähdekieli: saksa
 - Kohdekieli: suomi
 - Muut asetukset: oletusarvot

Koko lähdeteksti syötettiin järjestelmään yhtenä kokonaisuutena. DeepL-käännös tuotettiin yhdellä kerralla ilman manuaalisia väliintuloja tai muokkauksia.

- iii) Ihmiskääntäjän referenssikäännös: Vertailukohteena käytettiin ihmiskääntäjän tuottamaa referenssikäännöstä, jonka laatija on tämän tutkielman tekijä. Käännös on tuotettu osana kääntäjäkoulutuksen syventäviä opintoja, ja se on viimeistelty laadukkaaksi kokeneen ohjaajan opastuksella. On kuitenkin huomattava, että koska käännös on tehty koulutusympäristössä, se ei ole käynyt läpi institutionaalista tai kaupallista laadunvarmistusprosessia, joka olisi voinut tuottaa käännökselle vielä

organisaatiokohtaisia terminologisia tai tyyllisiä tarkennuksia. Prosessissa hyödynnettiin kielimallia (ChatGPT:n maksutonta versiota marraskuussa 2023) raakakäännöksen tuottamiseen, minkä jälkeen jokainen lause on analysoitu ja jälkieditoitu manuaalisesti vastaamaan juridiselta täsmällisyydeltään ja kielelliseltä laadultaan ammattimaista käännösjälkeä. Referenssikäännös edustaa sellaista laadullista tasoa, jossa kielellinen ja oikeudellinen asiantuntemus on korjannut teknologian mahdolliset virheet.

3.2.3 Aineiston käsittely ja rinnakkaiskorpuksen muodostaminen

Kerätty aineisto käsiteltiin analyysia varten systemaattisesti Excel-pohjaiseksi rinnakkaiskorpukseksi. Ensimmäisessä vaiheessa saksankielinen lähdeteksti ja kolme suomennosvastinetta segmentoitiin manuaalisesti virkkeittäin. Analyysiyksiköksi valittiin virke, sillä se mahdollistaa tarkan virhetarkastelun säilyttäen silti kontekstuaalisen merkityksen ja syntaktiset suhteet.

Segmentoinnin jälkeen aineisto kohdistettiin Excel-taulukon niin, että kukin rivi muodostaa nelisarakeisen kokonaisuuden: lähdetekstin virke, ChatGPT-5.2 ”Thinking” -päätelytilan tuottama käännös, DeepL Next-Gen -käännös sekä ihmiskääntäjän laatima referenssikäännös. Virheanalyysia varten taulukkoon lisättiin sarakkeet MQM-mallin mukaiselle dokumentoinnille (virhetyyppi, vakavuusaste ja selitys). Koko rinnakkaiskorpus on esitetty tutkielman liitteessä 2.

3.3 Analyysimenetelmät ja virheluokittelu

Tutkielman metodologinen viitekehys yhdistää laadullisen sisällönanalyysin ja rinnakkaiskorpusanalyysin. Analyysin teoreettisena pohjana käytetään MQM (Multidimensional Quality Metrics) -virheluokittelua, joka mahdollistaa virhekategorioiden räätelöinnin tutkimusasetelman mukaan (Lommel ym. 2015: 456). MQM:n rinnalla sovelletaan viisiportaista jälkieditointitarpeen arviointiasteikkoa. Asteikko pohjautuu käännöstieteessä aiemmin hyödynnettyihin neli- ja viisiportaisiin malleihin (esim. Specia ym. 2010: 3376; Koponen 2016a: 37), mutta tasot on määritelty uudelleen siten, että ne tunnistavat juridisen tekstin normatiivisen luonteen ja saksankielisten rakenteiden suomentamiseen liittyvät kielelliset haasteet.

MQM-malli erottuu edukseen perinteisistä kaavamaisista arviointimalleista joustavuutensa ansiosta. Esimerkiksi teollisuuskäyttöön suunnattu DQF (Dynamic Quality Framework), nykyisin pitkälti MQM:ään sulautunut, tai vanhentunut lokalisointialan LISA QA, olivat usein vaikeasti mukautettavissa erilaisiin kääntämisen erityistarpeisiin. Myös teknisiin teksteihin painottuva SAE J2450 on soveltuvuudeltaan MQM-mallia kapeampi. MQM mahdollistaa virheiden tarkan erittelyn useisiin ulottuvuuksiin ja alaluokkiin sekä niiden painottamisen vakavuusasteen mukaan. Järjestelmä perustuu yli 100 virhetyypin pääluetteloon, josta tutkija voi valita kunkin tehtävän kannalta olennaiset kriteerit. MQM huomioi kattavasti kielelliset konventiot, tyylin ja kohderyhmäsopivuuden ja soveltuu näin hyvin ammattimaisen juridisen käännöksen arviointiin. (Lommel ym. 2024: luvut 1, 2, 5 ja 5.1; Lommel ym. 2015: 1, 2, 4)

Analyysiprosessi etenee segmentointivaiheesta virheiden tunnistamiseen ja luokitteluun. Analyysiyksikkönä käytetään virkettä, joka on riittävän laaja kontekstin huomioimiseksi mutta riittävän rajattu systemaattista tarkastelua varten. Vaikka painopiste on laadullisessa arvioinnissa, virheiden lukumäärän tilastollinen vertailu täydentää analyysiä. Huolimatta siitä, että analyysiyksikkönä käytetään yksittäistä virkettä, kokonaisuudessaan tarpeellisuutta ei sivuuteta. Tutkimuksessa kiinnitetään huomiota myös terminologiseen johdonmukaisuuteen koko tekstin laajuudelta, eikä havaintoja tarkastella ainoastaan irrallisina lausetason ilmiöinä.

3.3.1 MQM 2.0 -virhekatgoriat

Tutkimuksessa sovelletaan Lommelien tutkimusryhmän (2024) esittelemää MQM 2.0 -kehystä, joka on MQM-mallin päivitetty ja käytännön arviointiin yhtenäistetty versio. Tässä tutkimuksessa MQM 2.0:n keskeinen etu on sen tarjoama systemaattinen ja standardien mukainen tapa luokitella käännösvirheitä sekä arvioida niiden vakavuutta eri tekstikokonaisuuksissa. Erityisesti päivitetty MQM Core -virheluokitus vastaa alan vakiintuneita standardeja (kuten ISO 5060), mikä parantaa analyysin läpinäkyvyyttä ja vertailtavuutta. (mts. tiivistelmä ja luku 1)

Kehys on järjestetty hierarkkisesti seitsemään korkean tason virheluokkaan: terminologia, tarkkuus, kielelliset konventiot, tyyli, maakohtaiset konventiot, kohderyhmäsopivuus sekä ulkoasu ja merkinnät. (mts. luvut 5 ja 5.1) Havaitut virheet dokumentoidaan vastaavasti seuraavaan taksonomiaan:

Pääluokka	Alaluokat ja painopiste analyysissä
Terminologia (Terminology)	Johdonmukaisuus (Consistency). Oikeustieteellisen terminologian täsmällisyys sekä käsitteellinen yhtenäisyys läpi koko tekstin.
Tarkkuus (Accuracy)	Vääristymä (Mistranslation), Poisjätö (Omission), Hallusinaatio (Hallucination). Lähde- ja kohdekielen välisen semanttisen vastaavuuden ja juridisen asiasisällön säilyminen.
Kielelliset konventiot (Linguistic conventions)	Kielioppi (Grammar), Morfologia (Morphology), Syntaksi (Syntax), Sanajärjestys (Word order), Välimerkit (Punctuation). Suomen kielelle tyypillisten agglutinatiivisten rakenteiden, sijamuotojen hallinnan sekä teema–reema-rakenteen oikeellisuus.
Tyyli (Style)	Rekisteri, Muodollisuus (Register, Formality). Juridisen rekisterin, institutionaalisten konventioiden ja tyylillisen sujuvuuden toteutuminen.
Maakohtaiset konventiot (Locale conventions)	Standardit. Päivämäärien, lukujen ja muiden paikallisten muotoiluviittausten oikeellisuus suomalaisessa kontekstissa.
Kohderyhmäsopivuus (Audience appropriateness)	Ymmärrettävyys ja tarkoituksenmukaisuus. Arvio siitä, kuinka hyvin käännös palvelee lehdistötiedotteen kohderyhmää ja viestinnällistä tarkoitusta.
Ulkoasu ja merkinnät (Design and markup)	Asettelu. Tekstin visuaalisen rakenteen ja teknisten merkintöjen säilyminen.

Suomen kielen rakenteelliset ominaisuudet on integroitu analyysiin osana ”Kielellisten konventioiden” kategoriaa. Huomiota kiinnitetään esimerkiksi sijamuotojen hallintaan sekä informaatorakenteen vääristymiseen, mikä voi vaarantaa juridisen argumentaation eheyden ja tekstin tavoitteleman oikeusvaikutuksen täsmällisen välittymisen.

3.3.2 Virheiden vakavuusasteet ja pisteytys

Virheiden vakavuusaste on jaettu kolmeen tasoon MQM-standardin oletusmääritelmien mukaan:

Kriittiseksi virheeksi (critical) luokitellaan sellainen virhe, joka tekee koko tekstistä käyttötarkoitukseensa sopimattoman tai aiheuttaa vakavan riskin fyysisestä, taloudellisesta, oikeudellisesta tai mainehaitasta. MQM-pohjaisissa järjestelmissä yksikin kriittinen virhe johtaa usein hylättyyn arvosanaan (Fail Rating). Tämä painotus on tarpeellinen tutkielman kontekstissa, sillä juridisessa tekstissä kynnys virheen kriittisyydelle on usein matalampi kuin yleiskielessä, ja kääntäjä voi joutua oikeudelliseen vastuuseen epätarkkuuksista.

Suureksi virheeksi (major) luokitellaan sellainen virhe, joka vaikuttaa vakavasti sisällön ymmärrettävyyteen, luotettavuuteen tai käytettävyyteen. Se johtuu usein merkittävästä merkityksen muutoksesta tai virheestä tekstin näkyvässä tai tärkeässä osassa.

Pieneksi virheeksi (minor) luokitellaan sellainen virhe, joka ei merkittävästi haittaa sisällön käytettävyyttä, ymmärrettävyyttä tai luotettavuutta sen aiotussa käyttötarkoituksessa. Sillä on vain rajallinen vaikutus esimerkiksi tarkkuuteen, tyyliin tai sujuvuuteen. (MQMc, ei pvm.)

Käännösten laadun määrälliseksi arvioimiseksi tutkimuksessa sovelletaan MQM-malliin perustuvaa pisteytysjärjestelmää. Kullekin virheelle määritetään numeerinen virhepaino, tässä tapauksessa MQM:n oletuspainotus: pieni virhe 1 piste, suuri virhe 5 pistettä ja kriittinen virhe 25 pistettä. (MQMc, ei pvm.) Pisteytyksessä käytetään tavallisesti teoreettista maksimipistemäärää, joka asetetaan usein arvoon 100. Tämän valitun arvon tarkoituksena on skaalata laatuarvosana helpommin ymmärrettävään, prosentuaaliseen muotoon. Käännöksen lopullinen laatuarvosana saadaan vähentämällä painotettujen virhepisteiden summa tästä teoreettisesta maksimista, jolloin lopputulos edustaa arvioidun tekstin virheetöntä osuutta. (Lommel ym. 2024: luku A.2)

Laskennallinen kokonaisvirhepistemäärä mahdollistaa kääntimien systemaattisen tilastollisen vertailun, huomioiden virhemäärän lisäksi myös niiden merkitys viestinnän kannalta.

3.3.3 Jälkiedointitarpeen viisiportainen asteikko

Tutkielman tavoitteena on analysoida paitsi konekäännösten virheiden määrää myös sitä, millaista ja kuinka vaativaa asiantuntijatasen jälkiedointia juridinen teksti edellyttää. Pelkkä virheluokittelu ei tästä syystä riitä kuvaamaan kääntäjältä vaadittavaa työpanosta, sillä juridisessa kääntämisessä käännöksen hyväksyttävyyys määräytyy lopulta virkkeen oikeusvaikutuksen ja julkaisukelpoisuuden perusteella.

Jälkiedointitarpeen arviointiin on kääntämistieteessä kehitetty useita luokituksia. Tunnetuimpiin kuuluu Temnikovan (2010: 3487–3488) kymmenportainen asteikko, jossa virheet luokitellaan niiden aiheuttaman kognitiivisen kuormituksen mukaan. Alho (2025: 273, 285) soveltaa tätä mallia tutkiessaan suomi–saksa-konekäännösten laatua ja osoittaa, että mekaaniset virheet (esim. morfologiset poikkeamat) kuormittavat kääntäjää vähiten, kun taas lauserakenteelliset ja merkitystä vääristävät virheet vaativat huomattavasti enemmän asiantuntijatyötä.

Vaikka Temnikovan malli tarjoaa tarkan välineen yksittäisten virhetyyppien analysointiin, se ei sellaisenaan sovi tämän tutkielman tavoitteisiin. Juridisessa kääntämisessä kääntäjän työ ei rajoitu virheiden korjaamiseen, vaan edellyttää usein koko virkkeen uudelleenmuotoilua oikeusvaikutuksen säilyttämiseksi. Tämän vuoksi analyysissa sovelletaan MQM-mallin

rinnalla viisiportaista, holistista jälkiedointitarpeen asteikkoa, joka arvioi koko virkkeen saattamista julkaisukelpoiseksi.

Asteikko on muodostettu kääntäjän kognitiivisen ja ammatillisen kompetenssin näkökulmasta ja se jakautuu seuraaviin tasoihin:

- (i) **Taso 1:** Ei muutoksia. Koneen tuottama raakakäännös on juridisesti ja kielellisesti moitteeton, eikä vaadi kääntäjältä korjaustoimenpiteitä.
- (ii) **Taso 2:** Pinnallinen korjaus. Muutokset kohdistuvat mekaanisiin tai teknisiin seikkoihin, kuten oikeinkirjoitukseen, välimerkkeihin tai paikallisiin konventioihin (esim. päivämäärät), jotka eivät vaikuta asiasisältöön.
- (iii) **Taso 3:** Leksikaalinen ja morfologinen korjaus. Kääntäjä korjaa yksittäisiä sanoja tai päätteitä, mutta lauseen perusrakenne ja sanajärjestys pysyvät ennallaan.
- (iv) **Taso 4:** Rakenteellinen korjaus. Kääntäjän on muokattava lauseen syntaksia, sanavalintoja tai morfologiaa suomen kielen erityispiirteiden huomioimiseksi (esim. sanajärjestys tai sijamuotojen valinta), vaikka asiasisältö olisi tunnistettavissa.
- (v) **Taso 5:** Sisällöllinen uudelleenkirjoitus. Virke sisältää kriittisiä virheitä, hallusinaatioita tai merkityksen vääristymiä, jotka vaativat asiantuntijalta lauseen täydellistä uudelleenmuotoilua alkuperäisen oikeusvaikutuksen säilyttämiseksi.

Asteikko mahdollistaa kääntimien tuotosten vertailun sen perusteella, kuinka paljon asiantuntijatason jälkiedointia kunkin järjestelmän tuotos vaatii saavuttaakseen juridisen tekstin laatuvaatimukset.

Tasojen välillä on yhteneväisyyksiä aiempiin luokituksiin. Taso 2 kattaa Temnikovan asteikon matalamman kognitiivisen kuormituksen virheet, kun taas taso 3 vastaa vaativampia syntaktisia muutoksia. Alhon tunnistamien kriittisten selkeys- ja asiasisältövirheiden osalta viisiportainen asteikko tarjoaa kuitenkin tarkoituksenmukaisemman analyysivälineen, sillä sen taso 4 erottaa tilanteet, joissa konekäännös on juridisesti käyttökelvoton.

3.3.4 MQM-virhemallin ja jälkiedointiasteikon suhde

Tutkimuksessa sovelletaan kahta toisiaan täydentävää arviointijärjestelmää: MQM-virhepisteet mittaavat virheen vakavuutta käännöksen lopputuotteessa (alajakso 3.3.2.), kun taas jälkiedointiasteikko (alajakso 3.3.4.) arvioi vaadittavan työmäärän kääntäjän näkökulmasta.

Nämä järjestelmät eivät välttämättä kuitenkaan korreloi lineaarisesti. Esimerkiksi yksi kriittinen terminologinen virhe (25 MQM-pistettä) voi vaatia vain yksittäisen sanan korvaamista (Taso 2), mutta sen tunnistaminen edellyttää oikeustieteellistä asiantuntemusta. Toisaalta useampi pieni morfologinen virhe (yhteensä 5 MQM-pistettä) saattaa vaatia koko lauserakenteen uudelleenmuotoilua (Taso 3), mikä on ajallisesti ja kognitiivisesti vaativampaa.

Tästä syystä analyysissä raportoidaan molemmat metriikat erikseen ja niiden tuloksia tulkitaan eri tutkimuskysymysten valossa: MQM-pisteet vastaavat ensisijaisesti kysymykseen virheiden tyypeistä ja vakavuudesta (tutkimuskysymys 1) ja jälkieditointiasteikko valottaa vaadittavaa asiantuntemusta ja korjaustoimenpiteitä (tutkimuskysymykset 2 ja 3).

Valitulla lähestymistavalla huomioidaan virhepisteiden ja todellisen työmäärän välinen ero, mikä mahdollistaa juridisen kääntämisen moniulotteisuuden kattavamman tarkastelun.

3.4 Tutkimuksen rajoitteet ja metodologiset huomiot

On tärkeää tunnistaa tutkimusaineiston rajoitteet ja niiden vaikutus tulosten tulkintaan. Tutkimusta voisi luonnehtia kartoittavaksi tapaustutkimukseksi, joka keskittyy yhteen tekstilajiin (lehdistötiedote) ja yhteen oikeudenalaan (tekijänoikeus). Tulokset eivät ole suoraan yleistettävissä kaikkiin juridisiin tekstilajeihin, kuten sopimukseen, lainsäädäntöön tai tuomioistuinten ratkaisuihin, eivätkä muihin oikeudenaloihin. Tutkimuksen tavoitteena ei ole tuottaa tilastollisesti yleistettäviä tuloksia, vaan tarjota syvälinen ymmärrys yhdestä erityistapauksesta, tunnistaa tyypillisiä virhetyyppejä ja generoida hypoteeseja laajempaa tutkimusta varten.

Toiseksi tutkimus käsittelee nimenomaan saksa–suomi-kieliparia, joka on typologisesti haastava yhdistelmä konekääntämiselle. Tulokset eivät välttämättä päde muihin kielipareihin, joissa kielet ovat rakenteellisesti läheisempiä. Lisäksi yhden hieman yli tuhatsanaisen lähdetekstin analyysi on pieni korpus kieliteknologian tutkimuksessa.

On myös huomioitava, että ihmiskääntäjän referenssikäännös edustaa yhden kääntäjän tulkintaa ja terminologisia valintoja. Toinen kääntäjä saattaisi tehdä erilaisia, mutta yhtä päteviä valintoja. Tämä on kuitenkin väistämätöntä kaikessa käännöstutkimuksessa, ja referenssikäännöksen tarkoitus on ensisijaisesti tarjota vertailukohta eikä ainoa oikea ratkaisu.

MQM-pohjainen virheluokittelu sisältää väistämättä subjektiivisia elementtejä, kuten virheiden vakavuusasteiden määrittäminen (kriittinen, suuri, pieni) ja rajatapausten arviointi. Koska tässä tutkimuksessa sekä referenssikäännös että konekäännösten virheluokittelu on yhden arvioijan toteuttama, on olemassa riski, että poikkeamat tutkijan omista käännösratkaisuista arvioidaan herkemmin virheiksi, vaikka ne voisivat edustaa vaihtoehtoisia, yhtä päteviä ratkaisuja.

Tutkimuksen luotettavuutta vahvistaisi inter-rater reliability -analyysi, jossa riippumaton toinen arvioija luokittelisi osan aineistosta samoilla MQM-kriteereillä ja arvioijien välinen yhteneväisyys mitattaisiin tilastollisesti esimerkiksi Cohen's Kappa -kertoimella (ks. Agrawal ym. 2024: 14495; Schreier 2012: 5, 151, 171). Tässä tutkielmassa analyysi on kuitenkin rajattu yhden arvioijan suorittamaksi, mikä on huomioitava tulosten tulkinnassa.

Lisäksi on huomattava, että konekääntämistyökalut ovat luonteeltaan muuttuvia: sekä ChatGPT-5 että DeepL Next-Gen ovat jatkuvasti kehittyviä teknologioita. Tutkimus tarjoaa tilannekuvan niiden kyvyistä tammikuussa 2026, mutta tulokset ennen pitkää vanhenevat tulevien päivitysten myötä.

Mainituista rajoituksista huolimatta valittu tutkimusasetelma soveltuu hyvin aiheen laadulliseen tarkasteluun tapaustutkimuksen puitteissa. Yksittäisen tekstin syvälinen analyysi mahdollistaa sellaisten käännösongelmien havaitsemisen, jotka jäisivät helposti huomaamatta laajemmassa mutta pinnallisemmassa tarkastelussa. Juridisen kääntämisen kontekstissa tällaiset hienovaraiset ongelmat voivat olla kriittisiä oikeudellisen merkityksen kannalta.

4 Analyysi ja tulokset

4.1 Yleiskatsaus käännosten laatuun

Tässä luvussa vastataan tutkimuskysymyksiin analysoimalla ChatGPT-5.2:n ja DeepL Next-Genin tuottamien käännosten virheitä, niiden suhdetta suomen kielen erityispiirteisiin sekä juridisen kääntämisen vaatimuksiin. Analyysi rakentuu neljässä vaiheessa: tässä alajaksossa esitetään yleiskatsaus käännosten laatuun ja kokonaispisteisiin. Alajaksossa 4.2 tarkastellaan keskeisiä virhekatgorioita leksikaalisen, morfosyntaktisen ja funktionaalisen tason näkökulmista. Alajaksossa 4.3 käsitellään jälkieditoinnin tarvetta ja alajaksossa 4.4 konekääntimien välisiä eroja.

Analyysi perustuu 40 segmenttiin jaetun lehdistötiedotteen yksityiskohtaiseen tarkasteluun (täydelliset taulukot liitteessä 2). MQM-pisteytyksessä suurempi pistemäärä tarkoittaa enemmän virheitä. Virheetömän käännoksen teoreettinen pistemäärä olisi 0.

ChatGPT-5.2 saavutti kokonais-MQM-pisteiksi 231 ja DeepL Next-Gen 285 pistettä, mikä tarkoittaa, että ChatGPT tuotti kokonaisuutena laadukkaampia käännoksiä, mutta kumpikin järjestelmä oli selvästi kaukana virheetömästä lopputuloksesta. Virheiden vakavuusjakauma osoitti, että ChatGPT:llä oli 96 pientä, 17 suurta ja kaksi kriittistä virhettä (yhteensä 115 virhettä), kun taas DeepL:llä oli 85 pientä, 25 suurta ja 3 kriittistä virhettä (yhteensä 113 virhettä). DeepL tuotti siis hieman vähemmän virheitä kokonaismäärältään, mutta niistä suurempi osa oli suuria. Virhetiheys oli molemmilla kääntimillä korkea: keskimäärin 2,8–2,9 virhettä per segmentti, mikä tarkoittaa, että lähes jokaisessa virkkeessä oli useita korjattavia kohtia.

Taulukko 1. ChatGPT-5.2:n ja DeepL Next-Gen -käännosten MQM-virhepisteet ja virheiden vakavuusjakauma

Mittari	ChatGPT 5.2	DeepL Next-Gen
Kokonais MQM-pisteet	231	285
Pieniä virheitä	96	85
Suuria virheitä	17	25
Kriittisiä virheitä	2	3

Mittari	ChatGPT 5.2	DeepL Next-Gen
Yhteensä	115	113

MQM-vakavuusasteet konkretisoituvat aineistossa seuraavasti: pieni virhe ilmenee esimerkiksi segmentissä L6, jossa kääntimet käyttävät vakiintumatonta lainasanaa ”sampling” suomenkielisen ammattitermin ”samplaus” sijaan: tämä heikentää tekstin asiantuntevaa vaikutelmaa mutta ei estä asiasisällön ymmärtämistä (lisää aiheesta alajaksossa 4.2.1.1. ii.). Suuri virhe esiintyy esimerkiksi segmentissä L12, jossa viittaussuhdevirhe (”josta”) loi virheellisen käsityksen siitä, että fragmentti olisi peräisin oikeuksista eikä äänitteestä (ks. tarkempi analyysi alajaksossa 4.2.2.1 iii.). Kaikkein vakavimpia ovat kriittiset virheet, joissa oikeudellinen logiikka vääristyy täysin: tällainen on esimerkiksi segmentin L7 virhe, jossa fokuspartikkelin puuttuminen käänsi oikeudellisen säännön merkityksen päinvastaiseksi (ks. tarkempi analyysi alajaksossa 4.2.3.2 ii.).

Yleisenä huomiona voidaan todeta, että molemmat kääntimet tuottivat jokseenkin sujuvaa ja ymmärrettävää suomea, mutta kumpikaan ei tuottanut juridisesti julkaisukelpoista käännöstä ilman jälkieditointia. Molemmat kääntimet kamppailevat samojen perusongelmien kanssa, joskin painotukset eroavat hieman. Molempien kääntimien käännösten toistuvimpia virhekatteorioita olivat terminologia (DeepL 30 kpl, ChatGPT 24 kpl), kielelliset konventiot (DeepL 31 kpl, ChatGPT 30 kpl) ja tarkkuus (DeepL 27 kpl, ChatGPT 27 kpl). Laatu oli useimmiten riittävä viestinnälliseen ymmärrettävyyteen, mutta ei yltänyt vaadittavalle tasolle oikeudellisen tarkkuuden tai institutionaalisen tyylin osalta.

Eriyisen silmiinpistävä piirre molempien kääntimien käännöksissä on saksankielisen lähdetekstin interferenssi eli lähdekielen rakenteiden mekaaninen heijastuminen kohdekieleen lähes kaikissa (37/40) analysoiduissa segmenteissä vähintään toisen kääntimen kohdalla. Tässä yhteydessä interferenssi tarkoittaa sitä, että kääntimet tuottavat rakenteeltaan tai ilmaisuiltaan saksan mukaisesti muotoiltua suomea, joka ei vastaa vakiintunutta suomenkielistä ilmaisutapaa tai juridisen tekstin rekisteritasoa. Interferenssi on ongelmallista juridisessa tekstissä, sillä se voi hämärtää tekstin merkityksiä, heikentää luettavuutta ja jopa johtaa väärin oikeudellisiin tulkintoihin. Interferenssi ilmenee aineistossa muun muassa siinä, että kääntimet segmentissä L3 noudattavat saksan mukaista paikkaa ja päiväystä koskevaa rakennetta (”Luxemburg, 29.

heinäkuuta”), vaikka unionin tuomioistuimen lehdistötiedotteissa vakiintunut muoto edellyttäisi inessiivin käyttöä (”Luxenburgissa 29.7.”). Näitä virheitä analysoidaan tarkemmin seuraavissa alajaksoissa.

Koneiden ohella myös vertailukohtana toimineen ihmiskääntäjän havaittiin tehneen referenssikäännöksessään virheitä, joista kaksi luokiteltiin suuriksi (L15 joka-relatiivipronominiin liittyvä viittaussuhdevirhe ja L23 terminologinen epätarkkuus tyhjentävästi-termin kohdalla). Havainto korostaa juridisen tekstilajin haastavuutta ja osoittaa, ettei ihmisvetoinen kääntäminenäkään takaa virheetöntä lopputulosta ilman huolellista oikolukua. Referenssikäännöksen virheet olivat kuitenkin luonteeltaan erilaisia kuin konekääntämien virheet: ne koskivat lähinnä yksittäisiä terminologisia valintoja tai ulkoasuseikkoja, kun taas konekääntämien virheet sisälsivät systemaattisempia ongelmia.

4.2 Keskeiset virhekategorioid

Seuraavissa alajaksoissa analysoidaan aineistosta havaittuja virheitä ja kielellisiä puutteita kolmella eri tasolla: leksikaalisella, morfosyntaktisella ja funktionaalilla. Valittu jako mahdollistaa erottelun yksittäisten termien ja sanavalintojen, lauseiden sisäisten rakenteiden sekä tekstin yleisen pragmaattisen vaikuttavuuden välillä. Leksikaalisella tasolla keskitytään juridisen terminologian hallintaan ja johdonmukaisuuteen. Morfosyntaktisessa osiossa analysoidaan kieliopillisia rakenteita ja erityisesti saksan kielen interferenssiä suomen kielen erityispiirteiden näkökulmasta. Lopuksi funktionaalilla tasolla tarkastellaan käännösten rekisteriä ja niiden onnistumista juridisen oikeusvaikutuksen välittäjinä suomalaiselle kohdeyleisölle.

4.2.1 Leksikaalinen taso: Terminologinen täsmällisyys ja johdonmukaisuus

Leksikaalisen tason virheet koskevat yksittäisten sanojen ja termien valintaa. Ne jakautuvat kolmeen pääryhmään: juridisten erikoistermien hallinta, institutionaalisten nimeämiskäytäntöjen tunnistaminen ja terminologinen johdonmukaisuus sekä tekstin sisällä että suhteessa laajempaan oikeudelliseen diskurssiin. Toisin kuin rakenteelliset virheet (käsitellään alajaksossa 4.2.2), leksikaalisen tason ongelmat ovat usein korjattavissa yksittäisten sanojen vaihtamisella, mutta niiden juridinen painoarvo voi olla merkittävä.

4.2.1.1 Juridisten erikoistermien hallinta

Konekääntimet kamppailevat erikoistermien kanssa kolmella tavalla: ne valitsevat liian yleiskielisiä sanoja täsmällisten juridisten termien sijasta, käyttävät vierasperäisiä lainasanoja vakiintuneiden suomenkielisten vastineiden sijaan tai sekoittavat keskenään lähellä toisiaan olevia juridisia käsitteitä.

i. Yleiskieliset sanat täsmällisten juridisten termien sijaan

Tyypillinen esimerkki on termi ”kielto”, jota DeepL käyttää segmentissä L11 saksan ”Unterlassung” vastineena. ”Kielto” on yleiskielinen sana, joka viittaa johonkin, mikä on kiellettyä. Oikeudellisessa kontekstissa kyse on kuitenkin tuomioistuimen antamasta prosessuaalisesta toimenpiteestä, jolla vastaaja määrätään lopettamaan oikeudenloukkaus. Täsmällinen suomenkielinen termi on kieltomääräys, joka on vakiintunut prosessioikeudellinen käsite ja vastaa saksan ”Unterlassungsverfügung”-käsitettä. Termivalinta ei ole vain tyylillinen kysymys: ”kieltomääräys” ilmaisee yksiselitteisesti sekä toimenpiteen luonteen (tuomioistuimen määräys) että sen oikeusvaikutuksen (loukkauksen kieltäminen).

Samankaltainen ongelma ilmeni termissä ”puuttuminen” (L6, L7), jota ChatGPT käytti saksan ”Eingriff” vastineena. Puuttuminen on arkikielen sana, joka ei kerro lukijalle, minkälaisesta oikeudellisesta toimesta on kyse. Tekijänoikeudellisessa kontekstissa vakiintunut termi on DeepL:n ja ihmiskääntäjän käyttämä loukkaus tai loukkaaminen, joka viittaa nimenomaan immateriaalioikeuden loukkaamiseen.

Kolmas esimerkki on DeepL:n käyttämä termi ”riita-asia” (L29, L30) saksan ”Rechtsstreit” vastineena. Suomen oikeuskielessä ”riita-asia” on kapea termi, joka viittaa vain siviilioikeudellisiin riita-asioihin ja sulkee pois esimerkiksi rikosasiat ja hallinto-oikeudelliset asiat. Saksan ”Rechtsstreit” on kuitenkin yleistermi, joka kattaa kaikki oikeudenkäyntityypit. Ennakkoratkaisupyynnöissä ja EU-tuomioistuimen yhteydessä sopiva termi olisi asia (esim. ”asia C-476/17”), joka on neutraali ja kattaa kaikki oikeusprosessit.

ii. Vierasperäiset lainasanat vakiintuneiden suomenkielisten vastineiden sijaan

Toinen merkittävä ongelmatyyppi on vierasperäisten lainasanojen käyttö tilanteissa, joissa suomenkielinen vakiintunut termi on olemassa ja laajasti käytössä.

Molemmat kääntimet käyttivät sanaa ”sampling” (L6, L10) koko tekstin ajan, vaikka suomenkielinen alalla vakiintunut termi on samplaus tai samplaaminen. Lainasanan käyttö voi olla ymmärrettävää yleiskielisessä tekstissä tai tilanteessa, jossa suomenkielistä vastinetta ei ole, mutta juridisessa asiategistissä, joka on suunnattu ammattilaisyleisölle, vakiintuneen suomenkielisen termin käyttö on ammattimaisempaa.

Vastaava ongelma ilmeni termien ”audiofragmentti” (ChatGPT, L7) ja ”äänifragmentti” (DeepL, L7) kanssa. Molemmat ovat tarpeettoman teknisiä ja vierasperäisiä sanoja tilanteessa, jossa luontevat suomenkieliset vastineet äänikatkelma tai ääninäyte olisivat käytettävissä. Nämä termit ovat yleiskielessä tunnettuja ja tekevät tekstistä helpommin lähestyttävän myös lukijalle, joka ei ole tekijänoikeuden tai musiikintuotannon asiantuntija.

iii. Lähikäsitteiden sekoittaminen

Kolmas ongelmatyyppi liittyy juridisiin käsitteisiin, jotka ovat lähellä toisiaan mutta joiden välillä ei vallitse täyttä käsitteellistä vastaavuutta.

Esimerkki tästä on ChatGPT:n käyttämä termi ”esittävät” (L10) tilanteessa, jossa asianosainen tekee väitteen oikeudenkäynnissä (saksan ”machen geltend”). Juridisessa kontekstissa ”esittävät” on neutraali termi, joka sopii moneen tilanteeseen, mutta kun kyse on nimenomaan osapuolen oikeudellisesta väitteestä, täsmällisempi termi olisi väittävät tai vetoavat siihen, että. Ero on hienovarainen mutta olennainen: ”väittävät” korostaa, että kyse on vasta osapuolen näkemyksestä, jonka todenperäisyys on arvioitavana, kun taas ”esittävät” neutraalimpana valintana häivyttää sitä tosiasiaa, että kyseessä on nimenomaan kiistanalainen väite.

Toisena esimerkkinä segmentissä L36 DeepL käyttää verbiä ”myöntää” lainsäädäntökontekstissa (poikkeuksien myöntäminen), vaikka oikeuskielellisesti lainsäätäjät ”säättää” poikkeuksia ja viranomaiset ”myöntää” lupia. Tämä sekoittaa oikeuslähteiden hierarkian ja antaa virheellisen kuvan normien synnystä. Vastaava ongelma esiintyy myös aiemmin mainitun termin ”Rechtsstreit” suomennoksissa segmenteissä L29 ja L30 (riita-asia vs. asia).

4.2.1.2 Institutionaaliset nimeämiskäytännöt ja intertekstuaalisuus

Juridinen kääntäminen edellyttää kääntäjältä kykyä tunnistaa, milloin kyse on virallisesta nimestä, jolla on vakiintunut muoto kohdekielellä. Konekääntimet eivät tässä onnistuneet, vaan

tuottivat usein mekaanisia sanasta sanaan -käännöksiä, jotka eivät vastaa virallisia nimeämiskäytäntöjä.

i. EU-direktiivien viralliset nimet

Kaikkein selvin esimerkki on direktiivien käsittely. EU-direktiiveillä on viralliset, vakiintuneet nimet kaikilla unionin kielillä. Nämä nimet on julkaistu EUR-Lex-tietokannassa, ja niitä käytetään johdonmukaisesti kaikessa virallisessa viestinnässä.

Segmentissä L35 lähdetekstissä viitataan kahteen direktiiviin niiden virallisilla saksankielisillä nimillä: ”Richtlinie 2001/29/EG zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft” ja ” Richtlinie 2006/115/EG zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums”.

ChatGPT jätti säädösnimet kokonaan kääntämättä ja kopioi saksan tekstin sellaisenaan käännökseen. Tämä on vakava virhe, sillä direktiiveillä on viralliset suomenkieliset nimet: ”Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY tekijänoikeuden ja lähioikeuksien tiettyjen piirteiden yhdenmukaistamisesta tietoyhteiskunnassa” ja ” Euroopan parlamentin ja neuvoston direktiivi 2006/115/EY vuokraoikeudesta ja lainausoikeudesta sekä tietyistä tekijänoikeuden lähioikeuksista immateriaalioikeuden alalla”.

DeepL puolestaan yritti kääntää direktiivien nimet suoraan saksasta, mutta ei noudattanut virallista suomenkielistä muotoa. Kumpikaan konekäännin ei tunnista, että kyse on viittauksesta virallisten säädösnimien verkostoon.

ii. Tuomioistuinten nimeäminen

Vastaava ongelma ilmeni tuomioistuinviittauksissa. Segmentissä L12 lähdetekstissä mainitaan ”Bundesgerichtshof” eli Saksan liittotasavallan ylin yleinen tuomioistuin. Tämä on yksiselitteinen virallinen nimi, joka viittaa tiettyyn tuomioistuimeen Saksan oikeusjärjestelmässä.

DeepL käänsi tämän muotoon ”liittovaltion tuomioistuin” (L12) ja toisessa kohdassa ”liittovaltion korkein oikeus” (L13). Molemmat ovat epätarkkoja yleistyksiä. ”Liittovaltion tuomioistuin” on liian yleinen ilmaisu, sillä Saksassa on useita liittovaltion tuomioistuimia.

”Liittovaltion korkein oikeus” taas antaa väärän kuvan, sillä Bundesgerichtshof ei ole korkein tuomioistuin kaikissa asioissa, sillä perustuslakiasioissa korkein on Bundesverfassungsgericht.

ChatGPT jätti kysymyksen auki toisella tavalla: se ei maininnut lainkaan, keneltä ennakkoratkaisua pyydetään (vom Gerichtshof = unionin tuomioistuimelta). Ihmiskääntäjän referenssikäännös esittää toimivan ratkaisun: ”Bundesgerichtshof (liittovaltion ylin yleinen tuomioistuin, Saksa)”. Ratkaisu säilyttää virallisen erisimen (Bundesgerichtshof), tarjoaa selittävän suomennoksen suluissa, sekä lisää maakontekstin (Saksa) selkeyden takaamiseksi.

Valittu muotoilu osoittaa ihmiskääntäjän kyvyn huomioida kohdeyleisön tarpeet: juridinen teksti on ymmärrettävä itsenäisesti ilman, että lukijalla on syvällistä tietoa lähdekulttuurin oikeusjärjestelmästä.

iii. Lehdistötiedotteiden standardiosiot

Kolmas esimerkki institutionaalisista käytännöistä on EU-tuomioistuimen lehdistötiedotteiden standardiosiot. Saksan ”Pressekontakt” (L38) on vakiintunut muotoilu, joka vastaa suomeksi ”Lisätietoja”. Tämä ei ole sanasta sanaan -käännös, vaan instituutiokohtainen konventio.

Molemmat kääntimet tuottivat mekaanisen käännöksen ”Lehdistökontakti” (L38), joka on ymmärrettävä mutta paljastaa välittömästi tekstin konekäännetyksi. EU:n tuomioistuimen kaikissa suomenkielisissä lehdistötiedotteissa käytetään otsikkoa ”Lisätietoja”, ja tästä poikkeaminen rikkoo tekstin institutionaalista yhtenäisyyttä. Vastaavia käännösvirheitä esiintyi myös segmenteissä L3, L4, L36, L37 ja L39.

4.2.1.3 Terminologinen johdonmukaisuus

Terminologinen johdonmukaisuus jakautuu kahteen ulottuvuuteen: sisäiseen johdonmukaisuuteen eli siihen, että samaa termiä käytetään samasta asiasta läpi tekstin, ja ulkoiseen johdonmukaisuuteen eli siihen, että termit ovat linjassa laajemman alan ja institutionaalisten käytänteiden kanssa.

i. Tekstin sisäinen johdonmukaisuus

Myönteinen havainto oli, että molemmat järjestelmät kykenivät säilyttämään sisäisen terminologisen johdonmukaisuuden useimpien keskeisten käsitteiden osalta.

ChatGPT käytti johdonmukaisesti termiä ”äänitetuottaja” läpi tekstin viittaamaan samaan oikeudelliseen käsitteeseen (saksan Tonträgerhersteller). DeepL puolestaan käytti järjestelmällisesti termiä ”äänitallenteen valmistaja”. Vaikka nämä termit poikkeavat toisistaan, kumpikin kääntäjä pysyi omassa valinnassaan, mikä on tärkeää tekstin sisäisen koherenssin kannalta.

Vastaavasti molemmat kääntimet osoittivat johdonmukaisuutta termien ”tekijänoikeus” (Urheberrecht), ”lähioikeus” (verwandtes Schutzrecht) ja ”ennakkoratkaisupyyntö” (Vorabentscheidungsersuchen) kohdalla. Nämä termit toistuivat aineistossa useita kertoja, ja kääntimet kykenivät palauttamaan saman suomenkielisen vastineen jokaisessa esiintymiskohdassa.

Johdonmukaisuudessa esiintyi kuitenkin myös poikkeuksia. Segmentissä L7 DeepL käänsi saksan termin ”Audiofragment” muotoon ”äänifragmentti”, mutta segmentissä L12 samasta termistä muotoa ”ääninäyte”. Lisäksi DeepL käänsi termin ”Bundesgerichtshof” segmentissä L12 muotoon ”liittovaltion tuomioistuin” ja segmentissä L13 muotoon ”liittovaltion korkein oikeus”.

ChatGPT kärsi niin ikään epäjohdonmukaisuudesta käyttäessään segmenteissä L29, L31 ja L32 pelkkää epätarkkaa ilmaisua ”tuomioistuin” ilman määritettä ”unionin”. Segmenteissä L1 ja L30 se puolestaan käytti täsmällisempää muotoa ”unionin tuomioistuin”. Segmentissä L29 ChatGPT käänsi saksan termin Rechtsstreit epäidiomaattisella ilmaisulla ”oikeusriita”, mutta heti seuraavassa segmentissä L30 se käänsi saman sanan epätarkkaan muotoon ”oikeudenkäynti”.

ii. Tekstin johdonmukaisuus suhteessa laajempaan diskurssiin

Vaikka kääntimet onnistuivat sisäisessä terminologisessa johdonmukaisuudessa useimpien käsitteiden osalta, ulkoinen johdonmukaisuus, eli linjaus EU:n vakiintuneiden termivalintojen ja alan yleisten käytänteiden kanssa, osoittautui ongelmallisemmaksi.

Edellä mainittu ”sampling”-termin käyttö koko tekstin ajan on esimerkki sisäisestä johdonmukaisuudesta, joka kuitenkin epäonnistuu ulkoisen johdonmukaisuuden tasolla: termi on johdonmukainen tekstin sisällä, mutta ei linjassa alan suomenkielisen terminologian kanssa.

Toinen vastaava esimerkki on saksan termin ”Verbreitungsrecht” kääntäminen. Molemmat konekääntimet tuottivat verbaaliteydyliä noudattavan muodon ”yksinoikeus levittää”, kun luontevampaa olisi ollut käyttää substantiiviteydylistä termiä ”yksinomainen levitysoikeus”. Vaikka verbaalinen ilmaus on kielipillisesti oikein ja välittää merkityksen, se ei vastaa suomalaisessa oikeuskielessä vallitsevaa tapaa, jossa oikeudet nimetään substantiivimuotoisina käsitteinä (esim. ”tekijänoikeus”, ”levitysoikeus”).

4.2.1.4 Yhteenveto

Terminologiset virheet muodostivat yleisen ongelmakategorian molempien kääntimien tuotoksissa: ChatGPT:n käännöksissä esiintyi 24 terminologista virhettä 19 segmentissä, kun taas DeepL:n käännöksissä virheitä oli 30 kappaletta 23 segmentissä.

Virheanalyysi paljastaa, että konekääntimet hallitsevat kohtuullisen hyvin sisäisen terminologisen johdonmukaisuuden yksittäisen tekstin sisällä, mutta ne eivät kykene yhtä luotettavasti tunnistamaan ja noudattamaan laajempia, tekstin ulkopuolisia terminologisia standardeja tai institutionaalisia käytänteitä. Rajoite on erityisen merkittävä juridisessa kääntämisessä, jossa terminologian on oltava linjassa sekä oikeudellisen diskurssin että institutionaalisten konventioiden kanssa.

Konekääntä tuottamat terminologiset virheet eivät johdu niinkään kielellisestä kyvyttömyydestä vaan pikemminkin siitä, etteivät ne kykene hallitsemaan laajempaa oikeudellista ja institutionaalista kontekstia tai tunnistamaan tekstien välisiä suhteita, joissa termit toimivat. Johtopäätöksenä voidaan todeta, että asiantuntijan tekemä jälkieditointi säilyy välttämättömyytenä näiden ongelmien välttämiseksi.

4.2.2 *Morfosyntaktinen taso: Rakenteet ja interferenssi*

Morfosyntaktisen tason virheet koskevat kielen rakenteellisia piirteitä: miten sanat taipuvat ja miten ne järjestäytyvät lauseiksi. Toisin kuin leksikaalisen tason virheet, jotka ovat korjattavissa yksittäisten sanojen muokkaamisella tai vaihtamisella, erityisesti syntaktiset virheet ja rakenteellinen interferenssi edellyttävät usein koko lauseen informaatorakenteen uudelleenarviointia.

Kuten luvussa 4.1 todettiin, saksankielisen lähdetekstin rakenteellinen interferenssi on aineistossa hallitseva ilmiö. Tarkempi analyysi osoittaa, että morfosyntaktiset virheet

jakautuivat epätasaisesti morfologisten ja syntaktisten virheiden välillä: syntaktiset ongelmat olivat selvästi yleisempiä kuin puhtaat morfologiset virheet. Verbin sijoitteluun ja sanajärjestykseen liittyviä ongelmia havaittiin 16 segmentissä (40 %) (segmentit L7, L10-L12, L14, L18, L19, L21, L23-L28, L33, L34), kun taas sijamuotovirheitä esiintyi 9 segmentissä (22,5 %) (segmentit L3, L7, L11, L13, L15, L17, L18, L24 ja L33) ja modusvirheitä vain 2 segmentissä (5,0 %) (L10, L11). Havainto viittaa siihen, että vaikka konekääntimet hallitsevat nykyisin suomen morfologian perustason kohtuullisen hyvin, syntaktinen interferenssi saksan kielestä on systemaattisempi ja vaikeammin ratkaistavissa oleva ongelma.

4.2.2.1 Syntaktinen interferenssi

Syntaktinen interferenssi muodostaa aineiston merkittävimmän rakenteellisen ongelman. Siinä missä morfologiset virheet olivat suhteellisen harvinaisia ja rajautuivat yksittäisiin sijamuotoihin tai modusvalintoihin, syntaktinen interferenssi oli läsnä lähes jokaisessa analysoidussa segmentissä. Verbin sijoitteluun ja sanajärjestykseen liittyviä ongelmia havaittiin 16 segmentissä (40 % aineistosta), mikä tekee niistä selvästi yleisimmän morfosyntaktisen virhekatteorian.

i. Verbin sijoittelu ja saksan kehysrakenne

Saksan kielen tunnusomainen kehysrakenne (Satzklammer), jossa finiittiverbi sijaitsee lauseen toisella paikalla ja verbikompleksin toinen osa lauseen lopussa, siirtyy systemaattisesti molempien kääntimien tuotoksiin. Tämä johtaa suomen kielelle vieraaseen verbin etäisyyteen subjektistaan ja tekee tekstistä vaikeasti luettavaa.

Selvä esimerkki tästä on segmentin L12 ChatGPT-käännös, jossa saksan kysymysrakenne ”Möchte der Bundesgerichtshof vom Gerichtshof unter anderem wissen, ob...” kääntyy muotoon ”Bundesgerichtshof [...] haluaa [...] muun muassa tietää, merkitseekö...”. Verbi-kompleksi ”möchte... wissen” on jaettu kahtia saksan mallin mukaisesti, ja apuverbi ”haluaa” erotetaan pääverbistä ”tietää” huomattavalla määrällä väliin sijoitettavia lauseenjäseniä. Suomen kielessä verbit tulisi sijoittaa peräkkäin heti subjektin jälkeen: ”Bundesgerichtshof haluaa tietää muun muassa, merkitseekö...”.

Vastaava ongelma ilmenee segmentissä L14, jossa kääntimet tuottivat erilaiset mutta yhtä lailla epäluonnolliset sanajärjestykset. ChatGPT tuotti rakenteen ”ovatko unionin oikeuden kanssa

yhteensopivia Saksan oikeussäännökset”, joka sijoittaa subjektin virheellisesti lauseen loppuun. DeepL puolestaan noudatti saksan ob-lauseen kehysrakennetta, jossa verbi (sind) edeltää subjektia (die deutschen Rechtsvorschriften) lauseen lopussa: ”ovatko Saksan säännökset [pitkä sivulause] unionin lainsäädännön mukaisia”. Suomessa subjekti tulisi sijoittaa lauseen alkuun heti kysymyssanan jälkeen, koska se edustaa tunnettua tietoa (teemaa): ”ovatko Saksan oikeussäännökset sopusoinnussa unionin oikeuden kanssa”.

Verbin sijoittelun ongelmia esiintyy tutkimusaineistossa myös segmenteissä L11, L19 ja L28. Segmentissä L19 ChatGPT sijoitti velvoitetta ilmaisevan verbirakenteen ”on turvattava” lauseen loppuun, mikä on suora kopio saksan infinitiivirakenteen sijoittelusta ja heikentää asiatyylisen virkkeen luettavuutta. Segmentissä L28 DeepL kääntää saksan ilmaisun ”auszuschließen ist” mekaanisena käännöslainana muotoon ”on... suljettava pois”, jolloin verbi päättyy jälleen lauseen loppuun rikkoen suomen kielen luonnollisen sanajärjestyksen. Myös segmentissä L11 havaittu kankeus juontaa juurensa saksan mallia noudattaviin rakenteisiin, joissa virke aloitetaan mekaanisesti pitkällä syylauseella (koska-lause), mikä siirtää päälauseen finiittiverbin kauas virkkeen alusta.

ii. Sanajärjestys ja informaatorakenne

Verbin sijoittelun ohella konekääntimet kamppailevat suomen kielen teema-reema-rakenteen kanssa, sillä ne seuraavat kuitenkin mekaanisesti saksan sanajärjestyksiä, mikä vääristää tekstin informaatorakennetta ja heikentää luettavuutta.

Juuri edellä kehysrakenteen yhteydessä mainittu segmentti L14 on kuvaava esimerkki myös informaatorakenteen vääristymisestä: siinä jo tunnettu asia (teema) sijoittuu virheellisesti lauseen loppuun, mikä tekee virkkeestä kognitiivisesti raskaan.

Segmentissä L19 ChatGPT, kuten jo aiemmin mainittiin, sijoitti velvoitetta ilmaisevan verbirakenteen ”on turvattava” lauseen loppuun saksankielisen rakenteen mukaisesti.

iii. Relatiivilauseiden ja kiilalauseiden käsittely

Saksan kielelle tyypilliset pitkät, sisäkkäiset relatiivilauseet ja muut kiilalauseet aiheuttavat systemaattisia ongelmia molemmille kääntimille. Koneet pyrkivät säilyttämään saksan lauserakenteen, mikä johtaa raskaslukuisiin kokonaisuuksiin.

Segmentissä L25 DeepL säilyttää saksan monimutkaisen relatiivilauseketjun lähes sellaisenaan: ”...äänitallenteesta otetun äänifragmentin käyttö, josta voidaan tunnistaa teos, josta se on otettu...”. Relatiivipronominien ketjutus tekee tekstistä puuromaista: sanojen väliset suhteet hämärtyvät. Suomen asiatyylissä tällainen ketjutus tulisi purkaa partisiippirakenteella tai erillisillä lauseilla.

Mielenkiintoinen tapaus on L21, jossa DeepL upottaa relatiivilauseen vertailurakenteen sisään: ”Esine, joka – kuten asiassa, josta on annettu viittauspäättös, kyseessä oleva esine – sisältää...”. Suomelle on vierasta sijoittaa pitkä kiilalause määritteen sisään ennen pääsanaa, ja tämä vaikeuttaa virkkeen hahmottamista.

Segmentissä L23 DeepL tuottaa raskaan ”jotta”-rakenteen (“...jotta varmistetaan sisämarkkinoiden toimivuus...”) ja sijoittaa fokuspartikkelin ”myös” epäluonnollisesti (“...on myös säädetty tyhjentävästi...”). Nämä ratkaisut niin ikään noudattavat saksan lauserakennetta dynaamisuuden kustannuksella.

Muita relatiivilauseisiin liittyviä ongelmia esiintyy segmenteissä L11, L12, L15, L21 ja L25. Segmentissä L11 ChatGPT käyttää kömpelöä relatiivirakennetta ”äänitteiden, joiden nimi on ’Nur mir’, luovuttamista”, vaikka informaation tiivistäminen sujuvampaan yhdyssanarakenteeseen (esim. ”’Nur mir’ -nimisten äänitteiden”) olisi ollut luontevampaa. Segmentissä L12 ChatGPT aloittaa virkkeen raskaalla relatiivilauseella ”Bundesgerichtshof, jonka käsiteltävänä asia on...”, joka katkaisee dynaamisuuden heti lauseen alussa. Samassa segmentissä ChatGPT:n käyttämä relatiivipronomi ”josta” viittaa loogisesti väärään edeltävään sanaan. Relatiivilauseen huolimaton sijoittelu luo virheellisen käsityksen siitä, että äänifragmentti olisi peräisin oikeuksista eikä äänitallenteesta.

Segmentissä L15 DeepL sortuu virheelliseen viittaussuhteeseen käyttämällä pronominia ”joka”, joka viittaa kieliopillisesti sanaan ”piiriin” loogisen kohteen ”poikkeus” sijaan. Vastaava viittaussuhdevirhe toistuu segmentissä L21 ChatGPT-käännöksessä, jossa pronominin ”jotka” ja sen pääsanana ”musiikkifragmentteja” väliin sijoitettu saksankielinen kiilalausemäärite ”tarvittaessa muutetussa muodossa” katkaisee loogisen yhteyden. Segmentissä L25 ChatGPT epäonnistuu sivulauseen integroinnissa, erottaen sivulauseen muusta virkkeestä kömpelösti ajatusviivoina (“– jotka koskevat kappaleen valmistamiseen ja yleisölle välittämiseen liittyviä yksinoikeuksia –”).

4.2.2.2 Morfologiset haasteet

Morfologiset ongelmat liittyivät usein siihen, että konekääntimet pyrkivät mekaaniseen vastaavuuteen saksan ja suomen välillä. Saksan prepositioita vastaavat suomessa sijapäätteet, mutta koneet eivät aina osaa muuntaa prepositiolausekkeita idiomaattisiksi sijamuodoiksi. Kääntimet eivät vaikuta hahmottavan funktionaalaisia vastaavuuksia eli pysty arvioimaan sitä, että täyttävätkö saksan prepositio + perusmuoto ja suomen sijapäätte saman kielellisen funktion vai tarvitaanko toisenlaisia rakenteellisia keinoja.

Vastaava mekaanisuus heijastui myös moduksen ja aikamuotojen hallintaan. Konekääntimillä oli haasteita tunnistaa saksan konjunktiivin ja erilaisten tempusrakenteiden pragmaattisia funktioita, mikä johti suomennoksissa juridisesti epätarkkoihin modusvalintoihin ja kerronnallisen koheesion kannalta epäloogisiin aikamuotosiirtoihin.

i. Sijapäätteiden interferenssi

Klassinen esimerkki konventioiden törmäyksestä on jo aiemmin mainittu päiväyksen muotoilu (L3). Lähdetekstissä on muoto ”Luxemburg, den 29. Juli 2019”. Molemmat kääntimet tuottivat käännöksen ”Luxemburg, 29. heinäkuuta 2019”, jossa paikannimi on nominatiivissa. Suomen kielessä ja unionin tuomioistuimen lehdistötiedotteissa vakiintuneen käytännön mukaan prepositioiden funktio hoidetaan kuitenkin sijapäätteillä: ”Luxemburgissa 29.7.2019”. Nominatiivin käyttö (”Luxemburg”) on suora heijastuma saksan rakenteesta, jossa paikannimet pysyvät perusmuodossa prepositioiden (esim. ”in Luxemburg”) yhteydessä.

Segmentin L33 DeepL:n käännöksessä ilmeni adessiivi-inessiivi-sekaannus. DeepL tuotti muodon ”toisella äänitallenteella” ja ChatGPT ”toisella äänitteellä” (adessiivi), kun idiomaattinen suomen muoto olisi ”toisessa äänitallenteessa” (inessiivi). Suomen kielessä sijaintia ilmaistavissa yhteyksissä käytetään kuitenkin inessiiviä, kun kyse on abstraktista tai konkreettisesta sisällymisestä (ääninäyte sijaitsee äänitallenteessa).

DeepL:n osalta sijapäätteongelmia esiintyi lisäksi segmenteissä L7 ja L18 epäidiomaattinen peräkkäisten inessiivien käyttö (kuuntelussa tunnistamattomassa), mikä seuraa saksan ”beim Hören” -rakennetta. Segmentissä L13 DeepL käytti partitiivia (”esittää myös kysymyksiä”), vaikka konteksti viittaa nimenomaisiin, tiettyihin kysymyksiin. Juridisessa tekstissä tällöin tulisi käyttää (”esittää myös kysymykset”), joka ilmaisee, että kyse on juuri näistä mainituista

kysymyksistä, ei mistä tahansa kysymyksistä. Segmentissä L15 se jätti genetiivin pois lauseenvastikkeen loogiselta subjektilta: ”voidaanko sampling katsoa” tulisi olla ”voidaanko samplingin katsoa”. Segmentissä L24 DeepL käytti adessiivia (”vapaalla käytöllä”), kun asiayhteys vaatisi adverbiaalista ilmaisua (”vapaasti käyttäen”). Koneen versio tekee vapaasta käytöstä ikään kuin välineen tai keinon, kun kyse on oikeudellisesta toimintatavasta tai säännöksestä.

ChatGPT-5.2:n osalta sijapäätevirheitä oli havaittavissa useissa eri kohdissa. Segmentissä L11 ilmeni haasteita partitiivin käytössä verbin ”loukata” yhteydessä, jolloin kone käytti muotoa oikeus, vaikka oikea muoto olisi ollut partitiivi oikeutta. Segmentissä L17 kääntimellä oli vaikeuksia hallita genetiivirakenteita, mikä johti liialliseen genetiiviketjuun ilmaisussa ”audiofragmentin kappaleen valmistaminen”; tämä on suora heijastuma saksan kielen genetiivirakenteista. Segmentissä L18 havaittiin väärä sijamuoto verbin ”olla kyse” yhteydessä, sillä kone käytti muotoa ”valmistamista”, kun asiasisältö vaatisi elatiivia ”valmistamisesta”. Lisäksi segmentissä L33 esiintyi sijamuotovirhe, jossa kone käytti adessiivimuotoa ”toisella äänitteellä”, vaikka merkitys vaatisi inessiiviä ”toisessa äänitteessä”.

ii. Moduksen ja aikamuotojen hallinta

Kriittisimmät morfologiset ongelmat liittyivät modusvalintoihin. Segmentissä L10 molemmat kääntimet tuottivat indikatiivin ”on kopioinut”, kun konteksti olisi vaatinut konditionaalin ”olisi kopioinut”.

Saksan lähdeteksti käyttää konjunktiiivia I:tä: ”Pelham habe... kopiert”, joka ilmaisee epäsuoraa esitystä. Suomen kielessä epäsuora esitys merkitään tyypillisesti konditionaalilla, joka säilyttää ilmauksen väitteen luonteen. Indikatiivin käyttö antaa ymmärtää, että kopiointi on jo todettu faktaksi, mikä on juridisesti virheellinen tulkinta: kyse on vasta osapuolen väitteestä, jonka todenperäisyys on arvioitavana.

Vaikka tämä ongelma voidaan teknisesti luokitella morfologiseksi (väärä modus), sen vaikutus ulottuu pragmaattiselle tasolle muuttaen tekstin juridista merkitystä. Morfologiset valinnat eivät näin ole pelkästään kieliopillisia kysymyksiä vaan kantavat juridista painoarvoa.

Segmentissä L11 ilmeni aikamuotojen epäjohdonmukaisuutta. ChatGPT tuotti virheellisen moduksen ”katsovat... on loukattu”, jossa preesensin ja perfektin yhdistelmä ei vastaa tekstin

kerronnallista kokonaisuutta. Saksan lähdetekstissä käytetään rakennetta ”der Auffassung sind, dass... verletzt worden sei”. Tässä sind-verbi on preesensissä, mutta sivulauseen sei-apuverbi on konjuktiivin perfekti, joka ylläpitää saksalle tyypillisen epäsuoran esityksen moduksen.

Suomen kielessä aikamuotojen johdonmukaisuus narratiivisessa kontekstissa on tärkeää. Kun raportoidaan menneistä tapahtumista ja osapuolten esittämistä väitteistä, imperfekti on tyypillisesti luontevampi kuin preesensin ja perfektin sekaannus. Oikea muoto olisi ollut käyttää partisiippirakennetta, kuten ”katsoivat... oikeuttaan loukatun” (tai muodollisemmin ”katsoivat... oikeuttaan olleen loukattu”). Jos taas kerronnassa haluttaisiin jostain syystä säilyttää preesens, muoto olisi ”katsovat... oikeuttaan loukattavan” tai ”olevan loukattu”.

Tämä virhe on suora käännöslaina saksan kieliopista, jossa uskomusta ilmaiseva verbi on preesensissä (sind) ja passiivinen mennyt tapahtuma ilmaistaan konjuktiivin perfektillä (verletzt worden sei). Konekääntimet eivät hahmottaneet, että suomen kielen tempusjärjestelmä ja kerronnan koheesio vaativat tässä yhteydessä aikamuodon siirtoa tai partisiippirakennetta.

4.2.2.3 Yhteenveto

Syntaktinen interferenssi ei rajoitu yksittäisiin virheisiin, vaan se muodostaa systemaattisen kuvion, joka läpäisee lähes koko aineiston. Molempien kääntimien kohdalla saksan kielen lauserakenteet, kuten kehysrakenne, kiilalauseet ja relatiivilauseketjut, siirtyvät suomeen mekaanisesti ilman kielten välisten rakenne-erojen huomioimista.

Morfologiset virheet olivat aineistossa vähemmän yleisiä kuin syntaktiset virheet, mutta niiden juridinen painoarvo on yhtä merkittävä. Erityisesti modusvalinnat vaikuttavat suoraan tekstin oikeusvaikutukseen, kun taas sijapäätteiden virheet heikentävät ennen kaikkea tekstin idiomaattisuutta ja luettavuutta. Syntaktisten virheiden painoarvo on usein pienempi kuin morfologisten virheiden juridisessa tekstissä, sillä ne näyttävät vaikuttavan pääasiassa tekstin luettavuuteen, sujuvuuteen ja tyylilliseen ammattimaisuuteen.

Konekääntimet näyttävät hallitsevan suomen perusmorfologian, kuten tavalliset taivutusmuodot ja genetiivin muodostuksen, kohtuullisen hyvin, mutta kamppailevat niissä kohdissa, joissa saksan prepositiorakenteet on muunnettava suomen sijajärjestelmään. Vaikuttaa siltä, että konekääntimien ongelma ei ole suomen morfologian tuntemuksessa vaan kielten välisten rakenne-erojen hahmottamisessa.

4.2.3 Funktionaalinen taso: Juridisen rekisterin ja oikeusjärjestelmien heijastuminen

Siinä missä terminologinen ja morfosyntaktinen analyysi paljastaa kääntimien teknisiä rajoitteita, juridisen rekisterin tarkastelu siirtää painopisteen tekstin funktionaaliseen laatuun: siihen, toimiiko käänнос uskottavana institutionaalisenä viestinä suomalaiselle kohdeyleisölle ja säilyttääkö se alkuperäisen tekstin oikeusvaikutuksen. Vaikka morfosyntaktiset virheet on edellä analysoitu kieliopillisesta näkökulmasta, ne ovat merkityksellisiä myös funktionaalisella tasolla, sillä ne voivat heikentää tekstin juridista funktiota ja oikeusvaikutusta.

Funktionaalisen tason virheet ovat usein vaikeimpia tunnistaa, sillä ne eivät välttämättä ole kieliopillisesti väärin tai semanttisesti virheellisiä. Kyse on siitä, että kielellisesti oikea ilmaus on kontekstissaan sopimaton, virheellisen vaikutelman luova tai juridisesti merkitykseltään väärä. Tällaiset virheet vaativat jälkieditoijalta ymmärrystä juridisen tekstilajin konventioista ja oikeusjärjestelmien eroista.

Pragmaattisen tason virheet jakautuvat kolmeen pääryhmään: tekstilajin konventioiden rikkominen, juridisten oikeusvaikutusten vääristyminen ja oikeusjärjestelmien välisen vertailun puutteet.

4.2.3.1 Tekstilajin konventiot

Analyysin perusteella konekääntimet kärsivät rekisterin epävakaudesta, joka ilmenee taipumuksena liukua kohti yleiskielisyyttä tai arkisuutta. Kun juridisen kielen keskeisiä piirteitä ovat muodollisuus ja tietty etäisyys, kääntimet murentavat tätä rekisteriä epäonnistuneilla verbivalinnoilla.

Segmenteissä L16 ja L17 ChatGPT käyttää verbiä ”muistuttaa” saksan hinweisen-verbin vastineena. Kontekstina ovat EU-tuomioistuimen tuomion perustelut, joissa viitataan vakiintuneeseen oikeuskäytäntöön. Valinta on tyylillisesti ongelmallinen, sillä ”muistuttaa” on hieman liian yleiskielinen, ja se antaa vaikutelman siitä, että tuomioistuin opastaa lukijaa. Tämä laskee tekstin muodollisuusastetta ja muuttaa virallisen kannanoton epäviralliseksi huomautukseksi. DeepL:n valinta ”toteaa” noudattaa suomalaisen oikeuskielen konventioita, mutta ihmiskääntäjän ”kiinnittää huomiota” on vielä täsmällisempi: se säilyttää tuomioistuimen auktoriteetin ja saksan alkuperäisen painotuksen, joka on hieman pehmeämpi kuin kategorinen toteaminen.

Vastaava rekisterin arkistuminen ilmenee segmenteissä L12 ja L14, joissa ChatGPT käyttää ilmaisua ”haluaa tietää” kuvaamaan kansallisen tuomioistuimen ennakkoratkaisupyyntöä (ks. alajakso 4.2.2.1. i.). Ilmaus on tyyliltään liian arkkinen ja subjektiivinen juridisessa kontekstissa. Vakiintuneemmat ilmaisut olisivat ”tiedustele”, ”pyytää selvitystä” tai ”esittää kysymyksen”, jotka säilyttävät muodollisen virallisuuden.

iii. Kohteliaisuuskäytänteet ja tittelien käyttö

Kulttuurisidonnaiset kohteliaisuuskäytänteet aiheuttivat ongelmia ChatGPT:lle. Segmenteissä L9, L10 ja L11 ChatGPT toistaa tarpeettomasti ”herra” (Herr) nimien edessä: ”herra Ralf Hütter ja herra Florian Schneider-Esleben”. Saksan lähdetekstissä tittelin käyttö on pakollista muodollisessa kontekstissa: ”Herr Ralf Hütter und Herr Florian Schneider-Esleben”. Suomalaisessa asiatyylissä titteleitä ei kuitenkaan toisteta tällä tavoin, vaan pelkkä ”Ralf Hütter ja Florian Schneider-Esleben” on asianmukainen muoto. Tittelien toistaminen luo vaikutelman liioittelusta muodollisuudesta, joka on vierasta suomalaiselle hallinnolliselle kielelle.

iv. Vakiintuneet institutionaaliset fraasit ja muotoilut

EU-tuomioistuimen lehdistötiedotteilla on vakiintunut rakenne ja tiettyjä standardifraaseja, joiden käyttö on osa institutionaalista viestintää.

Segmentissä L38 molemmat kääntimet tuottivat mekaanisen käännöksen ”Lehdistökontakti” saksan ”Pressekontakt” vastineena. EU:n tuomioistuimen kaikissa suomenkielisissä lehdistötiedotteissa käytetään kuitenkin otsikkoa ”Lisätietoja”. Tämä ei ole sanasta sanaan käännös, vaan vakiintunut instituutiokohtainen konventio. Sama kyvyttömyys noudattaa vakiintuneita fraaseja toistuu segmentissä L37, jossa käsitellään tuomion julkaisemista Curia-sivustolla.

Saksan lähdetekstin lause ”Der Volltext des Urteils wird am Tag der Verkündung auf der Curia-Website veröffentlicht” on käännetty molempien kääntimien toimesta sinänsä ymmärrettävästi: ChatGPT: ”Tuomion täysi teksti...”, DeepL: ”Tuomion koko teksti...”. Kumpikaan eivät kuitenkaan käytä EU-tuomioistuimen virallista vakiomuotoista ilmaisua: ”Tuomion koko teksti on luettavissa CURIA-sivustolla julistamispäivänä.”

Vastaavasti segmentissä L3 molemmat kääntimet tuottivat aiemmin mainitun päiväysmuodon ”Luxemburg, 29. heinäkuuta 2019”, kun EU:n tuomioistuimen suomenkielisisä käännöksissä vakiintunut muoto on ”Luxemburgissa 29.7.2019”.

Tällaiset pienet mutta näkyvät poikkeamat vakiintuneista fraaseista paljastavat välittömästi tekstin konekäännetyksi. Ne eivät estä ymmärtämistä, mutta ne rikkovat tekstin institutionaalista yhtenäisyyttä ja heikentävät ammattimaisen vaikutelman syntymistä.

v. Passiivin käyttö

Virallisissa yhteyksissä passiivin avulla vältytään turhalta subjektin mainitsemiselta ja korostetaan itse toimintaa tai tapahtumaa.

Segmentissä L6 molemmat kääntimet syyllistyivät mekaaniseen vastaavuuteen kääntämällä saksan funktionaalisen *erfolgen-verb*in suomen *tapahtua-verb*iksi. Koska kyseessä on samplaamisen kaltainen tietoinen toiminta, ilmauksen kontekstuaalinen merkitys vääristyy: suomessa *tapahtua-verb*i viittaa tyypillisesti prosessiin, jolla ei ole aktiivista toimijaa. Juridisessa asiatyylissä rakenne vaatiikin toimijaan viittaavan passiivin: ”jos se tehdään ilman suostumusta”.

Vastaavia passiiviongelmia ilmeni segmenteissä L11, L17 ja L21. Näissä tapauksissa kääntimet turvautuivat usein kankeisiin kiertoilmauksiin, kuten segmentin L17 ”käyttäjän toimesta” -agenttirakenteeseen (*durch einen Nutzer*), joka on suora käännöslaina saksasta. Suomen juridinen asiatyyli suosii dynaamisempaa kieltä ilman turhia byrokraattisia apurakenteita: ”se, että käyttäjä jäljentää”.

vi. Lauseiden välisten loogisten suhteiden ilmaiseminen

Lauseiden välisten loogisten suhteiden ilmaiseminen paljastaa myös tekstilajisidonnaisia konventioita. Molemmat konekäännökset käyttivät segmentissä L11 ”koska”-alkuisia lauseita käännettäessä saksan ”*Da*” -rakennetta. Vaikka ”koska” on kieliopillisesti oikein, suomen asiatyylissä ja erityisesti juridisessa tekstissä suositaan muita, dynaamisempia rakenteita kausaalisuhteiden ilmaisemiseen.

Samoin seuraussuhteiden ilmaisemisessa molemmat koneet käyttivät arkista ”joten”-sanaa (segmentit L17, L28), kun juridisessa tekstissä painavammat ilmaukset, kuten ”minkä vuoksi” tai ”näin ollen”, olisivat tyyllillisesti sopivampia ja vastaisivat paremmin juridisen argumentaation vakavuutta.

Konekääntimien valinnat eivät ole virheitä semanttisessa mielessä, mutta ne heijastavat kykenemättömyyttä erottaa yleiskieli ja erikoisalan rekisteri toisistaan.

4.2.3.2 Kielelliset valinnat oikeusvaikutusten rakentajina

Juridiset tekstit eivät ole pelkästään informatiivinen, vaan ne rakentavat oikeudellista todellisuutta, määrittelevät osapuolten asemia ja ilmaisevat tiedon varmuusastetta. Konekääntimet eivät näytä hahmottavan tätä performatiivista ulottuvuutta, mikä johtaa oikeusvaikutuksen vääristymiin. Tämä tiivistyy segmentin L10 modusvirheeseen epäsuorassa esityksessä.

i. Modaalisuuden hallinta: väite vs. fakta

Kaikkein kriittisin pragmaattisen tason virhe liittyy siihen, miten kieli ilmaisee tiedon varmuusastetta ja totuusarvoa. Tämä tiivistyy aiemmin alajaksossa 4.2.2.2 kohdassa ii. analysoituun segmentin L10 modusvirheeseen. Vaikka virheen tekninen tausta on saksan konjunkttiivin epätarkassa kääntämisessä, sen pragmaattinen seuraus on juridisesti kohtalokas: molemmat konekääntimet tuottivat indikatiivisen ilmauksen ”on kopioinut”, joka esittää tapahtumienkulun toteutuneena ja riidattomana asiantilana, rikkoen näin tuomioistuimen kielellisen neutraaliuden.

ii. Oikeusvaikutuksen vääristymä: logiikan kääntyminen

Juridisen logiikan vääristyminen näkyy kriittisellä tavalla segmentissä L7, jossa DeepL-käännin käänsi lauseen merkityksen täysin päinvastaiseksi. Saksan lähdetekstin ilmaus ”stellt jedoch auch ohne Zustimmung keinen Eingriff dar” tarkoittaa, ettei loukkausta tapahdu edes ilman suostumusta. DeepL:n tuottama käännös ”loukkaukseksi ilman suostumusta” jätti kuitenkin huomiotta negaation ”keinen” ja fokuspartikkelin ”auch” välisen vuorovaikutuksen, jolloin käännöksen merkitykseksi tuli, että loukkaus muodostuisi nimenomaan suostumuksen puuttuessa.

Tämä mekaaninen virhe muuttaa tuomion ydinsisällön ja vääristää sen oikeusvaikutuksen täysin. Ongelma ei ole siinä, että kääntäjä ei tuntisi sanoja ”loukkaus” tai ”suostumus”, vaan siinä, että se ei hahmota negaation ja fokuspartikkelin vuorovaikutusta kokonaisuudessa, vaan kääntimen kyvyttömyydestä hahmottaa monimutkaisten kieliopillisten osien muodostamaa loogista kokonaisuutta. Tapaus on varoittava esimerkki siitä, miten rakenteellisesti monimutkaisen lauseen mekaaninen kääntäminen voi johtaa juridisesti kohtalokkaaseen lopputulokseen.

iii. Normatiivisen voiman säilyttäminen

Juridinen teksti luo normatiivista todellisuutta: se ei vain kuvaa vaan määrittää, mikä on oikein ja väärin, sallittua ja kiellettyä. Normatiivisen voiman säilyttäminen käännöksessä edellyttää tarkkuutta kielellisten valintojen juridisessa painoarvossa.

Segmentissä L11 kääntimet käyttivät ilmausta ”kielto”, kun juridisesti täsmällisempi termi olisi ”kieltomääräys” (käsitelty terminologiana 4.2.1.1:ssä, mutta sen oikeudellinen vaikutus on tässä olennainen). ”Kielto” on yleiskielinen sana ilman juridista viitekehystä, kun taas ”kieltomääräys” viittaa nimenomaan tuomioistuimen antamaan päätökseen, jolla on suora oikeusvaikutus.

Konekäännös voi myös muuttaa lauseen juridista luonnetta hienovaraisesti. Esimerkiksi segmentissä L26 DeepL käytti muotoa ”Jos teosta ei voida tunnistaa”, kun ihmiskääntäjä käytti muotoa ”jos teos ei ole tunnistettavissa”. Kielitieteellisesti tarkasteltuna konekäännös muuttaa saksan alkuperäisen adjektiivipredikaatin (erkennbar ist) toimintaverbiin perustuvaksi rakenteeksi. Juridisesta näkökulmasta tämä on ongelmallista: se siirtää painopisteen kohteen objektiivisesta ominaisuudesta (tunnistettavuus) subjektiiviseen havaintoon tai kykyyn (tunnistaminen). Tällainen rekisterivirhe voi hämärtää oikeudellista määritelmää, vaikka lause on yleiskielellisesti ymmärrettävä.

4.2.3.3 Yhteenveto

Pragmaattisen tason virheet ovat juridisessa kääntämisessä kaikkein vakavimpia, sillä ne voivat vääristää tekstin oikeusvaikutusta, heikentää sen institutionaalista uskottavuutta ja johtaa virheellisiin tulkintoihin. Vaikka nämä virheet eivät aina ole kieliopillisesti väärin tai semanttisesti virheellisiä, ne rikkovat tekstin funktionaalista tarkoitusta. Juridinen

asiantuntemus on siksi välttämätöntä jälkieditoinnissa. Jälkieditoijan on kyettävä tunnistamaan juridisen rekisterin normit ja korjaamaan poikkeamat, arvioimaan kielellisten valintojen juridista painoarvoa ja oikeusvaikutusta, ja toimimaan tulkkina eri oikeusjärjestelmien välillä ja huomioimaan kohdeyleisön tarpeet.

4.3 Jälkieditoinnin tarve ja vaikutus käänöslaatuun

Analyysin perusteella juridisen konekäännöksen saattaminen julkaisukelpoiseksi edellyttää monitasoista asiantuntijatyötä, jossa korostuvat kääntäjän kielellinen ja kognitiivinen kompetenssi ja oikeudellinen asiantuntemus. Tutkimuksessa hyödynnetty viisiportainen jälkieditointiasteikko osoittaa, miten tarvittava asiantuntijatyö jakautuu ChatGPT-5.2:n ja DeepL Next-Genin välillä (ks. taulukko 2).

4.3.1 Jälkieditointitarpeen jakautuminen viisiportaisella asteikolla

Jälkieditointitarvetta arvioitiin luvussa 3.3.3 kuvatulla viisiportaisella asteikolla, jossa taso 1 tarkoittaa moitteetonta käänöstä ja taso 5 täydellistä uudelleenkirjoitusta. Taulukko 2 esittää molempien kääntimien jakautumisen eri tasoille.

Taulukko 2: Jälkieditointitarpeen jakautuminen ChatGPT-5.2 ja DeepL Next-Gen -käännöksissä

Taso	Kuvaus	ChatGPT	DeepL
Taso 1	Ei muutoksia (julkaisukelpoinen)	5 seg. (12,5 %)	3 seg. (7,5 %)
Taso 2	Pinnallinen korjaus (oikeinkirjoitus jne.)	3 seg. (7,5 %)	4 seg. (10,0 %)
Taso 3	Leksikaalinen korjaus (sanat, termit)	18 seg. (45,0 %)	15 seg. (37,5 %)
Taso 4	Rakenteellinen korjaus (syntaksi)	12 seg. (30,0 %)	15 seg. (37,5 %)
Taso 5	Täydellinen uudelleenkirjoitus	2 seg. (5,0 %)	3 seg. (7,5 %)
Yhteensä		40 seg.	40 seg.

Jakautuminen paljastaa neljä keskeistä havaintoa. Ensinnäkin, vain harva segmentti oli julkaisukelpoinen sellaisenaan: ChatGPT 12,5 % ja DeepL 7,5 %. Tämä osoittaa selkeästi, että konekääntimet eivät tuota juridisesti julkaisukelpoista tekstiä ilman jälkieditointia. Toiseksi, leksikaalinen korjaus (taso 3) oli yleisin tarpeen ChatGPT:llä (45,0 % segmenteistä). Nämä

virheet koskevat yksittäisiä sanoja, termejä ja päätteitä, ja niiden korjaaminen edellyttää terminologista osaamista sekä juridisen rekisterin tuntemusta. Kolmanneksi, rakenteellinen korjaus (taso 4) oli yleisin tarpeen DeepL:llä (35,7 % segmenteistä). Näissä tapauksissa koko lauserakenne on muotoiltava uudelleen, mikä edellyttää suomen kielen syvällistä hallintaa ja kykyä hahmottaa informaatorakennetta. Neljänneksi, kriittisiä, täydellistä uudelleenkirjoitusta vaativia virheitä (taso 5) oli vähän mutta ne olivat merkittäviä: ChatGPT 5,0 % ja DeepL 7,5 %. Nämä liittyivät yleensä juridisen merkityksen vääristymiseen tai logiikan muuttumiseen.

Konkreettisesti eri tasot ilmenevät aineistossa seuraavasti. Hyvä esimerkki tasosta 3 on segmentti L7, jossa ChatGPT käyttää yleiskielistä termiä ”puuttuminen” täsmällisen juridisen termin ”loukkaaminen” sijaan. Lauserakenne on sinänsä oikea, mutta sanavalinnat paljastavat puutteita juridisen rekisterin hallinnassa. Tason 4 esimerkki on segmentti L12, jossa ChatGPT sijoittaa verbin ”tietää” lauseen loppuun saksan kehysrakenteen mukaisesti, kun suomessa verbi tulisi heti apuverbin ”haluaa” jälkeen. Tämä vaatii koko lauserakenteen uudelleenjärjestelyä. Kriittisin tason 5 virhe esiintyy esimerkiksi segmentissä L28, jossa DeepL jättää keskeisen termin ”kappaleenvalmistus” kokonaan kääntämättä, mikä muuttaa juridista merkitystä kriittisesti ja vaatii paitsi virheen havaitsemista myös lähdetekstin uudelleenarviointia ja koko lauseen merkityssisällön varmistamista.

4.3.2 Jälkieditoinnin mahdollisuudet ja rajoitteet

Konekääntimien tuottamat käännökset tarjoavat käytännön hyötyjä mutta sisältävät myös merkittäviä riskejä, jotka jälkieditoijan on tunnistettava. Positiivisena puolena konekäännös tarjoaa nopean raakaversio, joka säästää aikaa erityisesti pitkissä teksteissä. Terminologinen johdonmukaisuus on usein hyvä, sillä molemmat kääntimet käyttivät samaa termiä läpi tekstin (esim. ChatGPT ”äänitetuottaja”, DeepL ”äänitallenteen valmistaja”). Morfologia on pääosin oikein, sillä suomen taivutus hallitaan molemmilla kääntimillä kohtuullisen hyvin. Konekäännös voi siis toimia tehokkaana apuvälineenä terminologisen pohjan luomisessa ja kokonaiskuvan hahmottamisessa.

Kriittinen näkökulma on kuitenkin se, että jälkieditointi on yhtä vaativaa kuin alusta asti kääntäminen. Jälkieditoijan on luettava koko lähdeteksti huolellisesti varmistaakseen, ettei konekäännös ole vääristänyt merkitystä. Koska 35,7-45,0 % segmenteistä vaatii tason 3 tai korkeampaa korjausta, kyse ei ole pelkästä viilailusta vaan merkittävästä uudelleenmuotoilusta.

Kognitiivinen kuormitus voi olla jopa korkeampi kuin suoraan lähtötekstistä kääntämisessä, koska jälkieditoijan on ikään kuin koodattava auki konekäännöksen logiikka ja pääteltävä, miksi järjestelmä on valinnut tietyn ilmaisun.

Tutkimuksessa havaittiin useita tilanteita, joissa konekäännös voi johtaa jälkieditoijaa harhaan. Ensinnäkin, sujuva mutta virheellinen teksti: konekäännös voi olla kieliopillisesti oikein ja sujuvaa, mutta sisältää hienovaraisen terminologisen virheen. Esimerkiksi ”yksinoikeus levittää” on ymmärrettävä, mutta ei juridisesti täsmällinen – oikea termi on ”yksinomainen levitysoikeus”. Toiseksi, interferenssin sokeuttava vaikutus: kun saksan lauserakenne näkyy suomen käännöksessä toistuvasti, jälkieditoija voi ”tottua” siihen eikä enää huomaa interferenssiä. Kolmanneksi, luottamus koneen ”älykkyyttä” kohtaan: jälkieditoija voi olettaa, että LLM ”ymmärtää” tekstin ja luottaa sen valintoihin, vaikka todellisuudessa malli vain tuottaa todennäköisimmältä vaikuttavaa tekstiä ilman todellista ymmärrystä.

On myös huomattava, että ihmiskääntäjäkään ei tuota virheetöntä tekstiä ilman huolellista oikolukua. Tutkimuksessa käytetyssä referenssikäännöksessä havaittiin kaksi virhettä: relatiivipronominin viittaussuhdevirhe segmentissä L15 (viittaus sanaan ”piiriin” sanan ”poikkeuksen” sijasta) ja terminologinen epätarkkuus segmentissä L23, jossa käytettiin termiä ”kattavasti” kun oikea termi olisi ollut ”tyhjentävästi”. Tämä osoittaa, että juridisessa tekstissä jokainen käännös, oli se sitten kokonaan ihmisen tekemä tai jälkieditoitu konekäännös, vaatii laadunvarmistuksen.

Näiden havaintojen pohjalta voidaan määritellä, milloin konekäännös on hyödyllinen ja milloin ei. Konekäännös soveltuu raakaversioiden tuottamiseen terminologisen pohjan luomiseksi, nopean kokonaiskuvan saamiseen vieraskielisestä tekstistä, terminologisen johdonmukaisuuden ylläpitämiseen pitkissä teksteissä, sekä tilanteisiin, joissa on käytettävissä pätevä juridinen jälkieditoija, joka hallitsee sekä lähdekielen että kohdekielen. Konekäännös ei sovellu julkaisukelpoisiksi sellaisenaan (vain 7,5–12,5 % segmenteistä oli julkaisukelpoisia), tilanteisiin, joissa jälkieditoijalla ei ole juridista asiantuntemusta, tilanteisiin, joissa ei ole mahdollista tarkistaa lähdekieltä (esimerkiksi jos jälkieditoija ei hallitse saksaa), eikä kriittisiin oikeudellisiin asiakirjoihin, joissa semanttinen täsmällisyys on ehdoton vaatimus (esimerkiksi sopimukset, tuomiot).

Tutkimuksen tulokset osoittavat selkeästi, että konekäännökset eivät korvaa ammattikäntäjää juridisessa kääntämisessä, mutta ne voivat toimia tehokkaana apuvälineenä oikein käytettyinä. Konekäännös on työkalu, joka vaatii asiantuntijaa käyttämään sitä. Se ei nykyisellään poista kääntäjän tarvetta, vaan muuttaa kääntäjän roolia alkuperäisen tekstin tuottajasta laadunvarmistajaksi ja viimeistelijäksi. Tämä rooli edellyttää juridista asiantuntemusta, syvällistä kohdekielen hallintaa, lähdekielen hallintaa, teknologista ymmärrystä sekä Prieto Ramosin mainitsemaa strategista kompetenssia, eli kaikkia niitä taitoja, joita perinteinen kääntäjäkin tarvitsee (Prieto Ramos 2024b: 155).

4.4 Konekääntimien väliset erot

Vaikka molemmat järjestelmät tuottavat systemaattisia virheitä, ChatGPT-5.2:n ja DeepL Next-Genin väliset erot korostuvat virheiden painotuksessa ja jälkieditointitarpeen luonteessa. Vertailussa nousevat esiin erityisesti virhepisteiden merkitys suhteessa oikeudelliseen logiikkaan sekä kääntäjältä vaadittava asiantuntemus.

4.4.1 Virheiden painoarvo ja laatuerot

Virheanalyysi osoittaa, että ChatGPT-5.2 saavutti kokonaispistemääräksi 231 pistettä ja DeepL Next-Gen 285 pistettä, mikä tarkoittaa ChatGPT:n tuottaneen laadukkaampia käännöksiä. Tämä tulos saattaa ensi silmäyksellä vaikuttaa ristiriitaiselta, sillä DeepL:llä oli hieman vähemmän virheitä kokonaismäärältään (113 vs. 115). Ero selittyy kuitenkin virheiden painoarvolla: ChatGPT:llä oli 96 pientä, 17 suurta ja kaksi kriittistä virhettä, kun taas DeepL:llä oli 85 pientä, 25 suurta ja kolme kriittistä virhettä. DeepL:n korkeampi pistemäärä johtuu siis siitä, että sen virheistä suurempi osa kuului suuri tai kriittinen -luokkiin, jotka painottuvat pisteityksessä voimakkaammin.

Juridisessa kääntämisessä virheiden painoarvolla on erityinen merkitys. Pienet virheet, kuten satunnaiset vierasperäiset lainasanat (esim. ”sampling” termin ”samplaus” sijaan, L6) tai epätarkka päiväysformaatti (L3 Luxemburg-esimerkki), heikentävät tekstin institutionaalista uskottavuutta ja asiantuntevaa vaikutelmaa, mutta eivät estä asiasisällön ymmärtämistä. Suuret virheet sen sijaan saattavat hämärtää oikeudellista tulkintaa: esimerkiksi segmentissä L12 ChatGPT:n tuottama viittausuhdevirhe (”josta”) loi virheellisen käsityksen siitä, että fragmentti olisi peräisin oikeuksista eikä äänitteestä. Kaikkein vakavimpia ovat kriittiset virheet, joissa oikeudellinen logiikka vääristyy täysin. Tällainen on esimerkiksi segmentin L7

virhe DeepL:n käännöksessä, jossa fokuspartikkelin (”auch”) huomiotta jättäminen käänsi oikeusvaikutuksen päinvastaiseksi: sen sijaan että fragmentin käyttö nimenomaan ei loukkaa oikeuksia, partikkelin puuttuminen suomennoksesta antoi ymmärtää, että suostumuksen puuttuminen johtaisi oikeuden loukkaamiseen.

Virheiden painoarvo heijastuu suoraan jälkieditointityön vaativuuteen. Pienet virheet ovat tyypillisesti korjattavissa nopeasti, kun taas suuret ja kriittiset virheet edellyttävät syvällistä oikeudellista ymmärrystä ja huomattavaa kognitiivista panosta. Jälkieditoijan on kyettävä tunnistamaan, milloin virhe muuttaa juridista merkitystä, milloin se vain heikentää tyyliä, ja milloin se voidaan jättää huomiotta. Tämä edellyttää paitsi kielellistä ja terminologista osaamista, myös kykyä hahmottaa oikeusjärjestelmien välisiä eroja ja tekstin oikeusvaikutusta kokonaisuutena.

4.4.2 Jälkiedoitointitarpeen luonne

Viisiportaisen jälkieditointiasteikon perusteella ChatGPT:n käännöksissä yleisin korjaustarve on selvästi taso 3 (leksikaaliset korjaukset, 45,0 % segmenteistä). DeepL:n käännöksissä puolestaan tasot 3 (leksikaaliset korjaukset, 37,5 %) ja 4 (rakenteelliset korjaukset, 37,5 %) ovat yhtä yleisiä. Tämä ero korostaa virheiden luonnetta: ChatGPT vaatii enemmän terminologista ja sanatasoista tarkistusta, kun taas DeepL edellyttää tasapainoisemmin sekä sanatasoista että rakenteellista uudelleenmuotoilua.

ChatGPT:n tason 3 virheet keskittyvät yksittäisiin sanavalintoihin ja terminologiseen tarkkuuteen. Tyypillinen esimerkki on segmentti L7, jossa ChatGPT käyttää yleiskielistä termiä ”puuttuminen” täsmällisen juridisen termin ”loukkaaminen” sijaan. Lauserakenne on sinänsä oikea, mutta sanavalinnat paljastavat puutteita juridisen rekisterin hallinnassa. Vastaavasti segmentissä L10 ChatGPT tuotti virheellisen moduksen (”on kopioinut”) konditionaalnin (”olisi kopioinut”) sijaan, mikä muutti osapuolen väitteen todetuksi faktaksi. ChatGPT:n tuottama petollinen sujuvuus vaatii asiantuntijaa, joka näkee sujuvan pinnan alle ja tunnistaa ne kohdat, joissa oikeudellinen logiikka ontuu, vaikka kieli olisi moitteetonta.

DeepL:n tason 4 virheet puolestaan painottuvat rakenteellisiin ongelmiin, jotka edellyttävät koko lauseen uudelleenmuotoilua. Nämä virheet liittyvät usein saksan kielen rakenteelliseen interferenssiin: DeepL kopioi saksalaisia lauserakenteita suoraan suomeen tavalla, joka tekee tekstistä kankeaa ja vaikeasti luettavaa. Esimerkiksi segmentissä L14 DeepL säilyttää

saksankielisen kehysrakenteen, jossa finiittiverbi ja sen ohjaama loppuosa erotetaan toisistaan huomattavalla määrällä muita lauseenjäseniä. Tämä tekee suomenkielisestä tekstistä raskaslukuista ja häiritsee ajatuksen kulkua. Vastaavasti segmentissä L7 DeepL tuottaa kankean rakenteen ”kuuntelussa tunnistamattomaksi”, kun sujuvampi muoto olisi ”kuuntelemalla tunnistamattomassa muodossa”. Näiden virheiden korjaaminen edellyttää jälkieditoijalta syvällistä suomen kielen syntaksin hallintaa sekä kykyä hahmottaa informaatorakennetta ja teema-reema-jakoa. Hän ei voi tyytyä yksittäisten sanojen vaihtamiseen, vaan hänen on purettava saksan mallin mukaiset lauseet ja rakennettava ne uudelleen suomen kielen logiikan mukaisesti.

Käytännössä tämä ero tarkoittaa, että ChatGPT:n käännökset ovat hieman sujuvampia ja luonnollisemmin muotoiltuja, mutta ne vaativat jälkieditoijalta DeepL:ää enemmän terminologista ja oikeudellista tarkkuutta. DeepL:n käännökset ovat terminologisesti tarkempia yksittäisten sanojen tasolla, mutta käännöslauseet edellyttävät laajempaa rakenteellista muokkausta, jotta teksti saavuttaa suomen kielelle luontevan ilmaisun ja juridisen tekstin vaatiman selkeyden. Kääntimestä riippumatta vastuu juridisen merkityksen ja oikeusvaikutusten säilymisestä jää viime kädessä jälkieditoijalle, jonka on tunnistettava ja korjattava näihin vaikuttavat virheet.

4.4.3 Yhteenveto ja suositukset

Vertaileva analyysi osoittaa, että ChatGPT-5.2:n ja DeepL Next-Genin väliset erot eivät ole niinkään jälkieditointitarpeen kokonaismäärässä vaan sen luonteessa. Molemmat järjestelmät tuottavat käännöksiä, jotka vaativat merkittävää asiantuntevaa jälkieditointia ennen kuin ne ovat juridisesti julkaisukelpoisia. Vain harva segmentti oli sellaisenaan julkaisukelpoinen (ChatGPT 12,5 %, DeepL 7,5 %), ja valtaosa käännöksistä vaati vähintään tason 3 tai 4 korjauksia.

Keskeinen ero on siinä, että ChatGPT vaatii jälkieditoijalta erityisesti substanssiosaamista eli kykyä tunnistaa ja korjata terminologisia epätarkkuuksia, modusvirheitä ja juridisen rekisterin horjahduksia. DeepL puolestaan vaatii erityisesti kieliosaamista eli kykyä purkaa ja uudelleenmuotoilla rakenteellisesti kankeita lauseita sekä poistaa saksan kielen interferenssiä. Molemmissa tapauksissa ammattikäntäjän juridinen ja kielellinen asiantuntemus on

välttämätöntä, eikä konekääntimien tuottamia käännöksiä voida käyttää sellaisenaan ilman huolellista laadunvarmistusta.

Tulosten perusteella voidaan suositella, että LLM-teknologiaa hyödyntävien käännöstyökalujen käyttöä juridisessa kääntämisessä harkittaessa kiinnitetään erityistä huomiota jälkieditointiin ja varmistetaan, että jälkieditoijilla on riittävä sekä oikeudellinen että kielellinen osaaminen.

5 Pohdinta

Tämän tutkielman tulokset tammikuulta 2026 piirtävät kuvan murrosvaiheesta, jossa kääntämisen teknologinen kehitys on saavuttanut tason, joka pakottaa määrittelemään ammattikäntäjän roolin uudelleen. Vaikka ChatGPT-5.2 ja DeepL Next-Gen kykenevät tuottamaan pintapuolisesti sujuvaa ja usein vakuuttavaltakin vaikuttavaa oikeuskieltä, analyysi osoitti, että juridisesti julkaisukelpoisen tekstin saavuttaminen vaatii edelleen osaavan ihmiskääntäjän tekemää asiantuntevaa laadunvarmistusta.

Tutkimus vahvistaa näkemystä siitä, että kääntäjän työ on siirtynyt peruuttamattomasti kohti Prieto Ramosin esittelemää laajennettua kääntämistä, jossa teknologia on kääntäjän kognitiivisen kapasiteetin jatke. Tulevaisuudessa juridisten tekstien parissa työskentelevän kääntäjän keskeisin ammattitaito ei ole enää tekstin tuottaminen, vaan kyky toimia kielellisenä riskienhallitsijana.

Kun konekäännös on kieliopillisesti virheetöntä, mutta sisältää hienovaraisia oikeusvaikutuksen kääntäviä virheitä, kääntäjän vastuu korostuu. Ammatilaisen on kyettävä näkemään sujuvan pinnan alle ja tunnistamaan ne kohdat, joissa oikeudellinen logiikka ontuu. Tämä edellyttää siirtymää jälkieditoinnista strategiseen laadunvarmistukseen, jossa kääntäjä kantaa lopullisen vastuun tekstin juridisesta integriteetistä.

ChatGPT-5.2:n ”Thinking”-päätelytila ja DeepL:n hybridimalli viittaavat siihen, että tulevaisuuden käänösprosessi on dynaaminen yhdistelmä eri teknologioita. Kääntäjän on osattava valita oikea työkalu kunkin tekstityypin mukaan: ChatGPT vaatii jälkieditoijalta erityisesti substanssiosaamista terminologisten epätarkkuuksien korjaamiseen, kun taas DeepL vaatii vahvaa kielellistä otetta saksan rakenteellisen interferenssin purkamiseen. Kehotemuotoilun merkitys kääntäjän työkalupakissa kasvaa jatkossa osana ammatillista ydinosaamista.

Tutkimuksen tulokset antavat selvän viestin kääntäjäkoulutukselle: erikoisalojen kääntämiseen vaadittava asiantuntemus ei hälvene teknologian myötä, vaan siitä tulee välttämätön edellytys tekoälyavusteisen kääntämisen laadunvarmistukselle. Koska vain murto-osa koneen tuottamista segmenteistä oli sellaisenaan julkaisukelpoisia, kääntäjän on hallittava oikeusjärjestelmien ja oikeuskielien erot sekä juridinen terminologia voidakseen toimia koneen tuloksia arvioivana asiantuntijana. Kääntäjän on lisäksi ymmärrettävä, miksi malli hallusinoi

tai miksi se suosii tiettyjä rakenteita, eli hallita teknologia mutta säilyttää kriittinen etäisyys sen tuotoksiin.

Koska suomi on pieni kielialue ja rakenteellisesti etäällä valtaryhmistä, on olemassa riski, että saksan tai englannin rakenteellinen interferenssi alkaa muokata suomalaista juridista kielenkäyttöä, jos konekäännöksiä hyödynnetään ilman riittävää kritiikkiä. Konekääntäminen voi parantaa oikeudellisen tiedon saavutettavuutta, mutta se ei saa tapahtua oikeusvarmuuden kustannuksella.

Tulevaisuudessa on keskeistä löytää ja ylläpitää tasapaino teknologisen tehokkuuden ja inhimillisen asiantuntijuuden välillä, jotta suomenkielinen oikeuskieli säilyy täsmällisenä ja itsenäisenä myös tekoälyn aikakaudella.

Lähteet

- Agrawal, Sweta – Farinhas, António – Rei, Ricardo – Martins, André Filipe Torres 2024. Can Automatic Metrics Assess High-Quality Translations? [Verkkodokumentti.] – *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, s. 14491–14502. Toim. Yaser Al-Onaizan, Mohit Bansal ja Yun-Nung Chen. Association for Computational Linguistics, Miami. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.18653/v1/2024.emnlp-main.810>.
- Alho, Marjut 2025. Open AI and translation quality in translations from Finnish into German. A case study. – *trans-kom* 18 (1), s. 273–287. [Verkkolehti]. Viitattu 12.2.2026. Saatavissa: <http://hdl.handle.net/10138/602425>.
- Anhava, Jaakko 1998. *Maaailman kielet ja kielikunnat*. Gaudeamus, Helsinki.
- Ataman, Duygu – Birch, Alexandra – Habash, Nizar ym. 2025. Machine Translation in the Era of Large Language Models: A Survey. – *Information* 16 (9), s. 723. [Verkkolehti]. MDPI, Basel. Viitattu 10.12.2025. Saatavissa: <https://doi.org/10.3390/info16090723>.
- Balashov, Yuri 2025. Translation in the Wild. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2505.23548v2 [cs.CL]. Viitattu 5.1.2026. Saatavissa: <https://doi.org/10.3390/info16121077>.
- Bestgen, Yves 2022. Comparing Formulaic Language in Human and Machine Translation: Insight from a Parliamentary Corpus. – *Proceedings of the Workshop ParlaCLARIN III within the 13th Language Resources and Evaluation Conference*, s. 101–106. Toim. Darja Fišer, Maria Eskevich, Jakob Lenardič ja Franciska de Jong. European Language Resources Association, Marseille.
- Briva-Iglesias, Vicent – Cavalheiro Camargo, João Lucas – Dogru, Gokhan 2024. Large Language Models ”Ad Referendum”: How Good Are They at Machine Translation in the Legal Domain? [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2402.07681 [cs.CL]. Viitattu 15.12.2025. Saatavissa: <https://doi.org/10.48550/arXiv.2402.07681>.
- Chen, Daniel – Hulden, Mans 2022. My Case, For an Adposition: Lexical Polysemy of Adpositions and Case Markers in Finnish and Latin. – *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, s. 2610–2616. Toim. Nicoletta Calzolari, Frédéric Béchet, Philippe Blache ym. European Language Resources Association, Marseille.
- Chen, Kehai – Wang, Rui – Utiyama, Masao – Sumita, Eiichiro 2019. Neural Machine Translation with Reordering Embeddings. [Verkkodokumentti.] – *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, s. 1787–1799. Toim. Anna Korhonen, David Traum ja Lluís Màrquez. Association for Computational Linguistics, Firenze. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.18653/v1/P19-1174>.
- Chen, Yuduo 2024. The metamorphosis of machine translation: The rise of neural machine translation and its challenges. – *Applied and Computational Engineering* 43 (1), s. 99–106. Sichuan University, Chongqing. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.54254/2755-2721/43/20230815>.

- Chesterman, Andrew 2005. *On Definiteness: A Study with Special Reference to English and Finnish*. Cambridge University Press, Cambridge.
- Costa-jussà, Marta Ruiz – Rapp, Reinhard – Lambert, Patrik ym. (toim.) 2016. *Hybrid Approaches to Machine Translation*. Theory and Applications of Natural Language Processing. Springer, Cham. España-Bonet, Cristina – Costa-jussà, Marta Ruiz: Hybrid Machine Translation Overview, s. 1–24.
- Dahl, Matthew – Magesh, Varun – Suzgun, Mirac – Ho, Daniel 2024. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. – *Journal of Legal Analysis* 16 (1), s. 64–93. [Verkkolehti]. Oxford University Press, Oxford/Stanford. Viitattu 29.11.2025. Saatavissa: <https://doi.org/10.1093/jla/lae003>.
- DeepL 2024. DeepL’s next-gen LLM outperforms ChatGPT-4, Google, and Microsoft for translation quality. Päivitetty 16.7.2024. [Verkkodokumentti.] Viitattu 30.11.2025. Saatavissa: <https://www.deepl.com/en/blog/next-gen-language-model>.
- ei pvm. About DeepL language models. [Verkkodokumentti.] Viitattu 16.10.2025. Saatavissa: <https://support.deepl.com/hc/en-us/articles/14241705319580-About-DeepL-language-models>).
- Delikoura, Iris – Fung, Yi – Hui, Pan 2025. From Superficial Outputs to Superficial Learning: Risks of Large Language Models in Education. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2509.21972. Viitattu 1.1.2026. Saatavissa: <https://arxiv.org/abs/2509.21972>.
- Deutsch, Daniel – Briakou, Eleftheria – Caswell, Isaac ym. 2025. WMT24++: Expanding the Language Coverage of WMT24 to 55 Languages and Dialects. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2502.12404. Viitattu 7.12.2025. Saatavissa: <https://arxiv.org/abs/2502.12404>.
- Flückiger, Alex – Amrhein, Chantal – Graf, Tim ym. 2025. A comparison of translation performance between DeepL and Supertext. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2502.02577. Viitattu 7.12.2025. Saatavissa: <https://arxiv.org/abs/2502.02577>.
- Gain, Baban – Bandyopadhyay, Dibyanayan – Ekbal, Asif 2025. Bridging the Linguistic Divide: A Survey on Leveraging Large Language Models for Machine Translation. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2504.01919. Viitattu 20.12.2025. Saatavissa: <https://doi.org/10.48550/arxiv.2504.01919>.
- Galdia, Marcus 2003. Juridisten tekstien kääntäminen erityisesti suomalais-saksalaisen vertailun näkökulmasta. – *Lakimies* 101 (1), s. 3–22. Suomalainen Lakimiesyhdistys, Helsinki.
- Glushkova, Taisiya – Zerva, Chrysoula – Martins, André Filipe Torres 2023. BLEU Meets COMET: Combining Lexical and Neural Metrics Towards Robust Machine Translation Evaluation. [Verkkodokumentti]. – *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, s. 47–58. Toim. Mary Nurminen, Judith Koppeimann ym. European Association for Machine Translation, Tampere. Viitattu 3.1.2026. Saatavissa: <https://aclanthology.org/2023.eamt-1.6>.

- Green, Spence – Schwartz, Lane (toim.) 2016. *Conferences of the Association for Machine Translation in the Americas: MT Researchers' Track*. The Association for Machine Translation in the Americas, Austin. Lankinen, Matti – Takala, Pyry – Heikinheimo, Hannes – Raiko, Tapani – Karhunen, Juha: A Character-Word Compositional Neural Language Model for Finnish, s. 1–10.
- Hajikhani, Arash – Cole, Carolyn 2024. A critical review of large language models: Sensitivity, bias, and the path toward specialized AI. – *Quantitative Science Studies* 5 (3), s. 736–756. [Verkkolehti]. Viitattu 12.2.2026. Saatavissa: https://doi.org/10.1162/qss_a_00310.
- Helin, Markku 1999. Kieli oikeustodellisuuden rakentajana. – *Oikeuden kielet: oikeus ja oikeudellinen ajattelu monikielisessä maailmassa*, s. 43–55. Toim. Antero Jyränki. WSOYpro, Helsinki.
- Hoang, Hieu – Khayrallah, Huda – Junczys-Dowmunt, Marcin 2024. On-the-Fly Fusion of Large Language Models and Machine Translation. [Verkkodokumentti]. – *Findings of the Association for Computational Linguistics: NAACL 2024*, s. 520–532. Toim. Kevin Duh, Helena Gomez ja Steven Bethard. Association for Computational Linguistics, Mexico City. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.18653/v1/2024.findings-naacl.33>.
- Huang, Yan – Liu, Wei 2024. Evaluating the translation performance of large language models based on EUAS-20. [Verkkodokumentti]. arXiv-ennakkopainos arXiv:2408.03119. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.48550/arXiv.2408.03119>.
- Jamali, Lily – McMahon, Liv 2025. OpenAI claims GPT-5 model boosts ChatGPT to 'PhD level'. [Verkkodokumentti.] *BBC News*. Päivitetty 7.8.2025. Viitattu 4.12.2025. Saatavissa: <https://www.bbc.com/news/articles/cy5prvgw0r1o>.
- Jin, Linghao – He, Jacqueline – May, Jonathan – Ma, Xuezhong 2023. Challenges in Context-Aware Neural Machine Translation. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2411.00244. Viitattu 10.1.2026. Saatavissa: <https://arxiv.org/abs/2411.00244>.
- Jyränki, Antero (toim.) 1999. *Oikeuden kielet: oikeus ja oikeudellinen ajattelu monikielisessä maailmassa*. WSOYpro, Helsinki.
- Kayano, Yoko – Sugawara, Saku 2025. Specification-Aware Machine Translation and Evaluation for Purpose Alignment. – *Proceedings of the Tenth Conference on Machine Translation (WMT)*, s. 113–141. [Verkkodokumentti] Toim. Barry Haddow, Tom Kocmi, Philipp Koehn ja Christof Monz. Association for Computational Linguistics, Suzhou. Viitattu 7.12.2025. Saatavissa: <https://aclanthology.org/2025.wmt-1.22>.
- Kocmi, Tom – Avramidis, Eleftherios – Bawden, Rachel ym. 2025. Findings of the WMT25 General Machine Translation Shared Task: Time to Stop Evaluating on Easy Test Sets. – *Proceedings of the Tenth Conference on Machine Translation (WMT)*, s. 355–413. Toim. Barry Haddow, Tom Kocmi, Philipp Koehn ja Christof Monz. [Verkkodokumentti.] Association for Computational Linguistics, Suzhou. Viitattu 7.12.2025. Saatavissa: <https://aclanthology.org/2025.wmt-1.22>.

- Koehn, Philipp 2010. *Statistical Machine Translation*. Cambridge University Press, Cambridge.
- 2020. *Neural Machine Translation*. Cambridge University Press, Cambridge.
- Koehn, Philipp – Haddow, Barry – Kocmi, Tom – Monz, Christof (toim.) 2023. *Proceedings of the Eighth Conference on Machine Translation (WMT)*. Association for Computational Linguistics, Singapore. Karpinska, Marzena – Iyyer, Mohit: Large language models effectively leverage document-level context for literary translation, but critical errors persist, s. 419–451.
- Koponen, Maarit 2016a. Is Machine Translation Post-Editing Worth the Effort? A Survey of Research into Post-Editing and Effort. – *The Journal of Specialised Translation* 25, s. 131–153.
- 2016b. *Machine Translation Post-editing and Effort: Empirical Studies on the Post-editing Process*. Väitöskirja. Helsingin yliopisto, Helsinki.
- Koskenniemi, Kimmo – Lindén, Krister – Carlson, Lauri – Vainio, Martti – Arppe, Antti ym. 2012. *Suomen kieli digitaalisella aikakaudella – The Finnish Language in the Digital Age*. META-NET White Paper Series. Springer, Heidelberg.
- Kuratov, Yuri – Bulatov, Aydar – Anokhin, Petr ym. 2024. BABILong: Testing the Limits of LLMs with Long Context Reasoning-in-a-Haystack. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2406.10149 [cs.CL]. Viitattu 10.1.2026. Saatavissa: <https://arxiv.org/abs/2406.10149>.
- Lai, Kaijie 2024. Research on Key Technologies of Neural Machine Translation Based on Deep Learning. – *Proceedings of the 1st International Conference on Modern Logistics and Supply Chain Management – Volume 1: MLSCM*, s. 60–65. Toim. Colin W. K. Chen, Tariq H. Malik, Qiufang Fu ja Haiyan Xuan. SciTePress, Suzhou.
- Lauscher, Anne – Ravishankar, Vinit – Vulić, Ivan – Glavaš, Goran 2020. From Zero to Hero: On the Limitations of Zero-Shot Language Transfer with Multilingual Transformers. [Verkkodokumentti]. – *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, s. 4483–4499. Toim. Bonnie Webber, Trevor Cohn, Yulan He ja Yang Liu. Association for Computational Linguistics.
- Lehto, Leena 1999. Juridisten tekstien kääntämisen yleisistä ongelmista. – *Oikeuden kielet: oikeus ja oikeudellinen ajattelu monikielisessä maailmassa*, s. 119–129. Toim. Antero Jyränki. WSOYpro, Helsinki.
- Lommel, Arle – Burchardt, Aljoscha – Uszkoreit, Hans 2015. Multidimensional Quality Metrics: A Flexible System for Assessing Translation Quality. *The Journal of Specialised Translation* 23, s. 33–52.
- Lommel, Arle 2018. Augmented Translation: A New Approach to Combining Human and Machine Capabilities. – *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)*, s. 5–12. Toim. Janice Campbell,

- Alex Yanishevsky, Jennifer Doyon ja Doug Jones. Association for Machine Translation in the Americas, Boston.
- Lommel, Arle – Gladkoff, Serge – Melby, Alan – Wright, Sue Ellen – Strandvik, Ingemar ym. 2024. The Multi-Range Theory of Translation Quality Measurement: MQM scoring models and Statistical Quality Control. [Verkkójulkaisu.] Viitattu 3.1.2026. Saatavissa: <https://arxiv.org/html/2405.16969v5>.
- Luodonpää-Manni, Milla – Hamunen, Markus – Konstenius, Reetta 2020. Tutkimuksen käytäntö. – *Kielentutkimuksen menetelmiä I–IV*, s. 1–31. Toim. Milla Luodonpää-Manni, Markus Hamunen ja Reetta Konstenius. Suomalaisen Kirjallisuuden Seuran toimituksia 1457. Suomalaisen Kirjallisuuden Seura, Helsinki.
- Mattila, Heikki E. S. 2017. *Vertaileva oikeuslingvistiikka: juridinen kielenkäyttö, lakimieslatina, kansainväliset oikeuskielet*. Alma Talent, Helsinki.
- Mosavi Miangah, Tayebeh – Zolfaghar Kondori, Zohreh 2015. Lexical Disambiguation of Polysynthetic Adjectives in Machine Translation: A Corpus Based Study. *Iranian Journal of Information Processing and Management* 30 (3), s. 719–735. [Verkkójulkaisu.] Viitattu 21.12.2025. Saatavissa: <http://jipm.irandoc.ac.ir>.
- MQM (Multidimensional Quality Metrics) ei pvm a. [Verkkodokumentti.] Viitattu 6.11.2025. Saatavissa: <https://themqm.org/>.
- ei pvm b. [Verkkodokumentti.] Viitattu 4.1.2026. Saatavissa: <https://themqm.org/uses-of-mqm/>.
- ei pvm c. [Verkkodokumentti.] Viitattu 6.1.2026. Saatavissa: <https://themqm.org/error-types-2/the-mqm-scoring-models/>.
- OpenAI 2025a. Introducing GPT-5. [Verkkodokumentti.] Päivitetty 7.8.2025. Viitattu 4.12.2025. Saatavissa: <https://openai.com/index/introducing-gpt-5/>.
- 2025b. GPT-5.1: A smarter, more conversational ChatGPT. [Verkkodokumentti.] Päivitetty 12.11.2025. Viitattu 4.12.2025. Saatavissa: <https://openai.com/index/gpt-5-1/>.
- 2025c. GPT-5.1 Instant and GPT-5.1 Thinking System Card Addendum. [Verkkodokumentti.] Päivitetty 12.11.2025. Viitattu 4.12.2025. Saatavissa: <https://openai.com/index/gpt-5-system-card-addendum-gpt-5-1/>.
- Pérez-Ortiz, Juan Antonio – Sánchez-Martínez, Felipe – Esplà-Gomis, Miquel ym. (toim.) 2018. *Proceedings of the 21st Annual Conference of the European Association for Machine Translation*. European Association for Machine Translation, Alicante. Agrawal, Ruchit – Turchi, Marco – Negri, Matteo: Contextual Handling in Neural Machine Translation: Look Behind, Ahead and on Both Sides, s. 31–40.
- Prieto Ramos, Fernando 2024a. Patterns of Human-Machine Interaction in Legal and Institutional Translation: From Hype to Fact. – *Polissema* (Erikoisnumero), s. 1–27. ISCAP, Porto.

- 2024b. Revisiting translator competence in the age of artificial intelligence: the case of legal and institutional translation. – *The Interpreter and Translator Trainer* 18 (2), s. 148–173. [Verkkolehti]. Viitattu 12.1.2026. Saatavissa: <https://doi.org/10.1080/1750399X.2024.2344942>.
- Quinci, Carla – Pontrandolfo, Gianluca 2023. Testing neural machine translation against different levels of specialisation. An exploratory investigation across legal genres and languages. – *trans-kom* 16 (1), s. 174–209. [Verkkolehti]. Viitattu 12.2.2026. Saatavissa: <https://hdl.handle.net/11577/3483820>.
- Rahman, Subhey Sadi – Islam, Mohammad Adnanul – Alam, Mohammad Mahbub – Zeba, Musarrat – Rahman, Mohammad Abdur ym. 2025. Hallucination to Truth: A Review of Fact-Checking and Factuality Evaluation in Large Language Models. [Verkkodokumentti.] arXiv-ennakkopainos arXiv:2508.03860. Viitattu 4.12.2025. Saatavissa: <https://arxiv.org/abs/2508.03860>.
- Rocha, Ana Paula – Steels, Luc – van den Herik, Jaap (toim.) 2022. *Proceedings of the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022) – Volume 1*. SciTePress, Setúbal. Gezmu, Andargachew Mekonnen – Nürnberger, Andreas: Transformers for Low-resource Neural Machine Translation, s. 459–466.
- Rojas Plata, Daniel – Castro Sánchez, Noé Alejandro 2024. Neural and Statistical Machine Translation: Confronting the State of the Art. *International Journal on Natural Language Computing (IJNLC)* 13 (5/6), s. 1–13. [Verkkojulkaisu.] Viitattu 21.12.2025. Saatavissa: <https://doi.org/10.5121/ijnlc.2024.13601>.
- Rydzewska-Siemiątkowska, Joanna 2024. Miten eritaustaiset lukijat ymmärtävät lakitekstiä? – *Kielikello*. [Verkkodokumentti.] Julkaistu 20.9.2024. Viitattu 9.1.2026. Saatavissa: <https://kielikello.fi/miten-eritaustaiset-lukijat-ymmartavat-lakitekstia/>.
- Sajavaara, Anu 1999. Kieli oikeustodellisuuden rakentajana. – *Oikeuden kielet: oikeus ja oikeudellinen ajattelu monikielisessä maailmassa*, s. 131–144. Toim. Antero Jyränki. WSOYpro, Helsinki.
- Schreier, Margrit 2012. *Qualitative Content Analysis in Practice*. SAGE Publications, London.
- Specia, Lucia – Cancedda, Nicola – Dymetman, Marc 2010. A Dataset for Assessing Machine Translation Evaluation Metrics. – *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, s. 3374–3379. Toim. Nicoletta Calzolari ym. European Language Resources Association (ELRA), Valletta.
- Tan, Zhixing – Wang, Shuo – Yang, Zonghan – Chen, Gang – Huang, Xuancheng ym. 2020. Neural machine translation: A review of methods, resources, and tools. *AI Open* 1, s. 5–21. Elsevier B.V., Amsterdam. [Verkkojulkaisu.] Viitattu 13.12.2025. Saatavissa: <https://doi.org/10.1016/j.aiopen.2020.11.001>.
- Tiittula, Liisa – Kolehmainen, Leena 2020. Käännöstiede ja sen menetelmiä. – *Kielentutkimuksen menetelmiä I–IV*, s. 942–983. Toim. Milla Luodonpää-Manni, Markus Hamunen ja Reetta Konstenius. Suomalaisen Kirjallisuuden Seuran toimituksia 1457. Suomalaisen Kirjallisuuden Seura, Helsinki.

- Vaswani, Ashish – Shazeer, Noam – Parmar, Niki ym. 2017. Attention Is All You Need. [Verkkodokumentti.] – *Advances in Neural Information Processing Systems 30*, s. 5998–6008. Toim. Isabelle Guyon, Ulrike Von Luxburg, Samy Bengio ym. Curran Associates, Long Beach. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.5555/3295222.3295349>.
- Webber, Bonnie – Cohn, Trevor – He, Yulan – Liu, Yang (toim.) 2020. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics.
- Wu, Taiqiang – Hou, Cheng – Lao, Shanshan ym. 2024. Combining the Best of Both Worlds: A Method for Hybrid NMT and LLM Translation. [Verkkodokumentti.] – *Findings of the Association for Computational Linguistics: ACL 2024*, s. 5138–5146. Toim. Lun-Wei Ku, André Filipe Torres Martins ja Vivek Srikumar. Association for Computational Linguistics, Bangkok. Viitattu 12.2.2026. Saatavissa: <https://doi.org/10.18653/v1/2024.findings-acl.305>.
- Zhang, Tianyi – Kishore, Varsha – Wu, Felix – Weinberger, Kilian Q. – Artzi, Yoav 2020. BERTScore: Evaluating Text Generation with BERT. *International Conference on Learning Representations (ICLR)*. [Verkkojulkaisu.] Viitattu 11.1.2026. Saatavissa: <https://openreview.net/forum?id=SkeHuCVFDr>.

Liitteet

Liite 1. Analysoitu rinnakkaiskorpusaineisto

- 1a. ChatGPT 5.2 -kielimallilla tuotettu suomennos
- 1b. DeepL Next-Gen -kääntimellä tuotettu suomennos
- 1c. Ihmiskääntäjän laatima referenssisuomennos

1a. ChatGPT 5.2 -kielimallilla tuotettu suomennos

Euroopan unionin tuomioistuim

LEHDISTÖTIEDOTE nro 98/19

Luxemburg, 29. heinäkuuta 2019

Lehdistö ja tiedotus

Tuomio asiassa C-476/17 Pelham GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian Schneider-Esleben

Sampling voi merkitä puuttumista äänitetuottajan oikeuksiin, jos se tapahtuu ilman tämän suostumusta

Äänitteestä otetun audiofragmentin käyttäminen muutetussa muodossa, joka ei kuunneltaessa ole tunnistettavissa, ei kuitenkaan merkitse näihin oikeuksiin puuttumista edes ilman suostumusta

Musiikkiryhmä Kraftwerk julkaisi vuonna 1977 äänitteen, jolla on musiikkikappale *Metall auf Metall*. Herra Moses Pelham ja herra Martin Haas ovat säveltäneet musiikkikappaleen *Nur mir*, joka julkaistiin vuonna 1997 Pelham GmbH:n äänitteillä. Kraftwerk-yhtyeen kaksi jäsentä, herra Ralf Hütter ja herra Florian Schneider-Esleben, väittävät, että Pelham on kopioinut noin kaksi sekuntia rytmijaksosta kappaleesta *Metall auf Metall sampling*¹-tekniikan avulla ja asettanut sen kappaleen *Nur mir* taustalle jatkuvana toistona. Koska herra Hütter ja herra Schneider-Esleben katsovat, että heidän kyseisen äänitteen tuottajina² omistama lähioikeus on loukattu, he vaativat muun muassa kiellon määräämistä, vahingonkorvausta ja *Nur mir* -nimistä äänitettä sisältävien tallenteiden luovuttamista hävitettäviksi.

Bundesgerichtshof, jonka käsiteltävänä asia on, haluaa unionin tekijänoikeuden ja lähioikeuksien³ sekä Euroopan unionin perusoikeuskirjassa taattujen perusoikeuksien kannalta muun muassa tietää, merkitseekö puuttumista sen äänitteen tuottajan oikeuksiin, josta audiofragmentti (*Sample*) on otettu, jos tämä audiofragmentti liitetään ilman tuottajan suostumusta *samplingin* avulla toiseen äänitteeseen. Bundesgerichtshof esittää myös kysymyksiä unionin oikeudessa säädetyistä oikeudenhaltijoiden oikeuksiin kohdistuvista poikkeuksista ja rajoituksista. Se haluaa tältä osin tietää, ovatko unionin oikeuden kanssa yhteensopivia Saksan oikeussäännökset, joiden mukaan itsenäinen teos, joka on luotu suojatun teoksen vapaata käyttöä hyödyntäen, voidaan lähtökohtaisesti julkaista ja sitä voidaan hyödyntää ilman oikeudenhaltijoiden suostumusta. Se haluaa lisäksi tietää, voiko *sampling* kuulua sitaattia koskevan poikkeuksen piiriin; tämä poikkeus vapauttaisi käyttäjän velvollisuudesta hankkia äänitetuottajan suostumus suojatun äänitteen käyttöön.

Tuomioistuim muistuttaa tämänpäiväisessä tuomiossaan ensiksi, että äänitetuottajilla on yksinoikeus sallia tai kieltää äänitteidensä kappaleen valmistaminen kokonaan tai osittain. Näin ollen **tällaisen – myös hyvin lyhyen – äänitteestä otetun audiofragmentin kappaleen valmistaminen käyttäjän toimesta on lähtökohtaisesti tämän äänitteen osittaista kappaleen valmistamista**, joten tällainen kappaleen valmistaminen kuuluu äänitetuottajan yksinoikeuden piiriin.

¹ *Sampling* on tekniikka, jossa elektronisten laitteiden avulla otetaan äänitteestä otteita, jotta niitä voidaan käyttää uuden sävellyksen osina toisella äänitteellä.

² Äänitteiden tuottajat ovat luonnollisia henkilöitä tai oikeushenkilöitä, jotka rahoittavat äänitteiden valmistamisen.

3 Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft (ABl. 2001, L 167, S. 10) und Richtlinie 2006/115/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums (ABl. 2006, L 376, S. 28).

”Kappaleen valmistamista” ei kuitenkaan ole kyseessä, jos käyttäjä taiteen vapauttaan käyttäen ottaa äänitteestä audiofragmentin liittääkseen sen muutettuna ja kuunneltaessa tunnistamattomana muotona uuteen teokseen. Sen toteaminen, että tällainen audiofragmentin käyttö on kappaleen valmistamista, joka edellyttää äänitetuottajan suostumusta, olisi muun muassa ristiriidassa sen vaatimuksen kanssa, että toisaalta tekijänoikeuksien ja lähioikeuksien haltijoiden etujen – heidän perusoikeuskirjaan kirjatun immateriaalioikeutensa suojan – ja toisaalta suojattujen kohteiden käyttäjien etujen ja perusoikeuksien, mukaan lukien perusoikeuskirjassa niin ikään taatun taiteen vapauden, sekä yleisen edun välillä on turvattava asianmukainen tasapaino.

Tuomioistuin toteaa seuraavaksi, että esine, joka omaksuu kaikki tai olennaisen osan äänitteeseen tallennetuista äänistä, on kyseisen äänitteen kappale, johon äänitetuottajalla on yksinoikeus levittää. **Tällainen kappale ei kuitenkaan ole esine, joka** – kuten pääasiassa kyseessä oleva – **omaksuu vain musiikkifragmentteja**, tarvittaessa muutetussa muodossa, **jotka siirretään tästä äänitteestä, jotta luodaan uusi ja siitä riippumaton teos.**

Lisäksi unionin oikeudessa säädetyt oikeudenhaltijoiden oikeuksia koskevat poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjä on ottanut huomioon suojattujen kohteiden tuottajien ja käyttäjien edut sekä yleisen edun. Nämä poikkeukset ja rajoitukset on myös säännelty tyhjentävästi, jotta tekijänoikeutta ja lähioikeuksia koskevien sisämarkkinoiden toimivuus voidaan varmistaa. Näin ollen **Saksan oikeussäännökset**, joissa poikkeusten ja rajoitusten tyhjentävästä luonteesta huolimatta säädetään poikkeuksesta tai rajoituksesta, jota unionin oikeudessa ei tunneta ja jonka mukaan toisen teoksen vapaata käyttöä hyödyntäen luotu itsenäinen teos voidaan lähtökohtaisesti julkaista ja sitä voidaan hyödyntää ilman käytetyn teoksen tekijän suostumusta, **eivät ole unionin oikeuden mukaisia.**

Sitaatteja koskevien, jäsenvaltioiden unionin oikeuden nojalla mahdollisesti säätämien poikkeusten ja rajoitusten osalta – jotka koskevat kappaleen valmistamiseen ja yleisölle välittämiseen liittyviä yksinoikeuksia – tuomioistuin toteaa, että **äänitteestä otetun audiofragmentin käyttö, joka tekee siitä teoksesta, josta fragmentti on otettu, tunnistettavan, voi tietyin edellytyksin olla sitaatti**, erityisesti silloin, kun käytön tavoitteena on olla vuorovaikutuksessa kyseisen teoksen kanssa. Jos teos ei ole tunnistettavissa, **fragmentin käyttö ei sitä vastoin ole sitaatti.**

Lopuksi tuomioistuin toteaa, että jäsenvaltiot voivat, silloin kun unionin oikeus ei määritä niiden toimintaa täysin, soveltaa unionin oikeutta täytäntöön pannessaan kansallisia perusoikeuksien suojan tasoja edellyttäen muun muassa, ettei perusoikeuskirjan suojan taso heikkene. Äänitetuottajan yksinomaisen kappaleen valmistamisoikeuden aineellinen sisältö on kuitenkin täydellisesti yhdenmukaistettu, joten tällainen käyttö on tältä osin poissuljettu.

HUOMAUTUS: Ennakkoratkaisupyynnön perusteella jäsenvaltioiden tuomioistuimet voivat niiden käsiteltävänä olevassa oikeusriidassa esittää tuomioistuimelle kysymyksiä unionin oikeuden tulkinnasta tai unionin toimen pätevydestä. Tuomioistuin ei ratkaise kansallista oikeusriitaa. Kansallisen tuomioistuimen tehtävänä on ratkaista asia tuomioistuimen ratkaisun

mukaisesti. Tuomioistuimen ratkaisu sitoo vastaavalla tavalla muita kansallisia tuomioistuimia, joiden käsiteltäväksi vastaava ongelma tulee.

Välineille tarkoitettu epävirallinen asiakirja, joka ei sido tuomioistuinta.

Tuomion täysi teksti julkaistaan julistamispäivänä Curia-sivustolla.

Lehdistöyhteyshenkilö: Hartmut Ost (+352) 4303 3255 Tuomion julistamistilaisuudesta on saatavilla videomateriaalia palvelun kautta

„Europe by Satellite“ (+32) 2 2964106

www.curia.europa.eu

www.curia.europa.eu

1b. DeepL Next-Gen -kääntimellä tuotettu suomennos



Lehdistö ja tiedotus

Euroopan unionin tuomioistuin

LEHDISTÖTIEDOTE Nro 98/19Luxemburg, 29. heinäkuuta
2019Tuomio asiassa C-476/17 Pelham
GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian
Schneider-Esleben**Sampling voi loukata äänitallenteen valmistajan oikeuksia, jos se tapahtuu ilman tämän suostumusta***Äänitallenteesta otetun äänifragmentin käyttö muutetussa ja kuuntelussa tunnistamattomassa muodossa ei kuitenkaan loukkaa näitä oikeuksia ilman suostumusta*

Musiikkiyhtye Kraftwerk julkaisi vuonna 1977 äänitallenteen, jolla on kappale *Metall auf Metall*. Moses Pelham ja Martin Haas ovat säveltäneet kappaleen *Nur mir*, joka julkaistiin vuonna 1997 Pelham GmbH:n äänitallenteilla. Kaksi Kraftwerk-yhtyeen jäsentä, Ralf Hütter ja Florian Schneider-Esleben, väittävät, että Pelham on kopioinut noin kahden sekunnin pituisen rytmisekvenssin kappaleesta *Metall auf Metall* sampling-tekniikan avulla ja lisännyt sen toistuvasti kappaleeseen *Nur mir*. Koska Hütter ja Schneider-Esleben katsovat, että heidän oikeuksiaan kyseisen äänitallenteen² valmistajina on loukattu, he vaativat muun muassa lopettamista, vahingonkorvausta ja *Nur mir* -kappaleen sisältävien äänitallenteiden luovuttamista tuhoamista varten.

Asiaa käsittelevä liittovaltion tuomioistuin pyytää unionin tuomioistuimelta muun muassa selvitystä siitä, voidaanko tekijänoikeuslainsäädännön ja unionin lähioikeuksien³ sekä Euroopan unionin perusoikeuskirjassa taattujen perusoikeuksien nojalla loukkaa äänitallenteen valmistajan oikeuksia, jos äänitallenteesta otettu ääninäyte (*sample*) lisätään ilman valmistajan suostumusta ääninäytteiden ottamisen avulla toiseen äänitallenteeseen. Liittovaltion korkein oikeus esittää myös kysymyksiä unionin lainsäädännössä säädetyistä poikkeuksista ja rajoituksista oikeudenhaltijoiden oikeuksiin. Se haluaa tietää, ovatko Saksan säännökset, joiden mukaan itsenäinen teos, joka on luotu vapaasti käyttämällä suojattua teosta, voidaan periaatteessa julkaista ja hyödyntää ilman oikeudenhaltijoiden suostumusta, unionin lainsäädännön mukaisia. Se haluaa lisäksi tietää, voidaanko *sampling* katsoa kuuluvan lainauspoikkeuksen piiriin, joka vapauttaa käyttäjän velvollisuudesta hankkia äänitallenteen valmistajan suostumus suojatun äänitallenteen käyttöön.

Tämänpäiväisessä tuomiossaan tuomioistuin toteaa ensinnäkin, että äänitallenteiden valmistajilla on yksinoikeus sallia tai kieltää äänitallenteidensa kopiointi kokonaan tai osittain. Näin ollen käyttäjän tekemä äänitallenteesta otetun, jopa hyvin lyhyen äänifragmentin kopiointi on periaatteessa kyseisen äänitallenteen osittainen kopiointi, joten tällainen kopiointi kuuluu äänitallenteen valmistajan yksinomaiseen oikeuteen.

¹ *Sampling* on tekniikka, jossa äänitallenteesta otetaan elektronisten laitteiden avulla otteita, joita käytetään osana uutta sävellystä toisella äänitallenteella.

² Äänitallenteiden valmistajia ovat luonnolliset tai oikeushenkilöt, jotka rahoittavat äänitallenteiden valmistuksen.

³ Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY, annettu 22 päivänä toukokuuta 2001, tekijänoikeuden ja lähioikeuksien tiettyjen tietoyhteiskunnan alalla sovellettavien näkökohtien yhdenmukaistamisesta (EYVL 2001, L 187, s. 10) ja Euroopan parlamentin ja neuvoston direktiivi 2006/115/EY, annettu 12 päivänä joulukuuta 2006, vuokrausoikeudesta ja lainausoikeudesta sekä tietyistä tekijänoikeuden alaisista immateriaalioikeuksista (EUVL 2006, L 376, s. 28).

Kopiointia ei kuitenkaan ole kyseessä, jos käyttäjä käyttää taiteellista vapauttaan ja ottaa äänitallenteesta äänifragmentin, jonka hän lisää uuteen teokseen muutetussa ja kuuntelussa tunnistamattomassa muodossa. Oletus, että tällainen äänifragmentin käyttö on kopiointia, joka edellyttää äänitallenteen valmistajan suostumusta, olisi ristiriidassa muun muassa vaatimuksen kanssa, jonka mukaan on varmistettava asianmukainen tasapaino yhtäältä tekijänoikeuksien ja lähioikeuksien haltijoiden etujen, jotka liittyvät perusoikeuskirjassa vahvistettuun immateriaalioikeuksien suojaan, ja toisaalta suojattavien kohteiden käyttäjien etujen ja perusoikeuksien, mukaan lukien perusoikeuskirjassa taattu taiteellinen vapaus, sekä yleisen edun välillä.

Tuomioistuin toteaa edelleen, että esine, joka sisältää kaikki tai olennaisen osan äänitallenteeseen tallennetuista äänistä, on kyseisen äänitallenteen kopio, johon äänitallenteen valmistajalla on yksinoikeus levittää. Esine, joka – kuten asiassa, josta on annettu viittauspäätös, kyseessä oleva esine – sisältää vain äänitallenteesta siirrettyjä, mahdollisesti muunnettuja musiikkikatkelmia uuden ja siitä riippumattoman teoksen luomiseksi, ei kuitenkaan ole tällainen kopio.

Lisäksi unionin lainsäädännössä säädetty oikeudenhaltijoiden oikeuksien poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjät on ottanut huomioon suojattujen esineiden valmistajien ja käyttäjien edut sekä yleisen edun. Nämä poikkeukset ja rajoitukset on myös säädetty tyhjentävästi, jotta varmistetaan sisämarkkinoiden toimivuus tekijänoikeuden ja lähioikeuksien alalla. Siksi **Saksan lainsäädäntö**, jossa säädetään unionin lainsäädännössä sääntelemättömästä poikkeuksesta tai rajoituksesta huolimatta poikkeusten ja rajoitusten lopullisesta luonteesta, jonka mukaan itsenäinen teos, joka on luotu toisen teoksen vapaalla käytöllä, voidaan periaatteessa julkaista ja hyödyntää ilman käytetyn teoksen tekijän suostumusta, ei ole unionin lainsäädännön mukainen.

Jäsenvaltioiden unionin lainsäädännön nojalla suojatun teoksen lainauksille mahdollisesti myöntämien jäljentämis- ja esittämisoikeuksien poikkeusten ja rajoitusten osalta tuomioistuin toteaa, että äänitallenteesta otetun äänifragmentin käyttö, josta voidaan tunnistaa teos, josta se on otettu, voi tietyissä olosuhteissa olla lainaus, erityisesti jos käytön tarkoituksena on olla vuorovaikutuksessa kyseisen teoksen kanssa. Jos teosta ei voida tunnistaa, fragmentin käyttö ei sen sijaan ole lainaus.

Lopuksi tuomioistuin toteaa, että jos jäsenvaltioiden toimintaa ei säännellä kokonaan unionin lainsäädännöllä, ne voivat soveltaa unionin lainsäädännön täytäntöönpanossa kansallisia perusoikeuksien suojanormeja, edellyttäen että tämä ei heikennä muun muassa perusoikeuskirjassa vahvistettua suojan tasoa. Äänitallenteiden valmistajan yksinoikeuden aineellinen sisältö on kuitenkin täysin yhdenmukaistamisen kohteena, joten tällainen käyttö on tässä suhteessa suljettava pois.

HUOMAUTUS: Jäsenvaltioiden tuomioistuimet voivat esittää tuomioistuimelle ennakkoratkaisupyyntö, joka koskee unionin oikeuden tulkintaa tai unionin oikeudellisen toimenpiteen pätevyyttä niiden käsiteltävänä olevassa riita-asiassa. Tuomioistuin ei ratkaise kansallista riita-asiata. Kansallisen tuomioistuimen tehtävänä on ratkaista asia tuomioistuimen päätöksen mukaisesti. Tuomioistuimen päätös sitoo samalla tavalla muita kansallisia tuomioistuimia, jotka käsittelevät samanlaista asiaa.

Tämä ei-virallinen asiakirja on tarkoitettu tiedotusvälineiden käyttöön, eikä se sido tuomioistuinta.

Tuomion [koko teksti](#) julkaistaan Curia-verkkosivustolla sen julistamispäivänä.

Lehdistöyhteyshenkilö: Hartmut Ost ☎ (+352) 4303 3255

Videotallenteita tuomion julistamisesta on saatavilla osoitteessa [Europe by Satellite](#) -palvelusta ☎ (+32) 2 2964106

1c. Ihmiskääntäjän laatima referenssisuomennos



Media ja lehdistö

Euroopan unionin tuomioistuin
Lehdistötiedote nro 98/19
 Luxemburgissa 29.7.2019

Tuomio asiassa C-476/17
 Pelham GmbH, Moses Pelham ja
 Martin Haas / Ralf Hütter ja Florian
 Schneider-Esleben

Samplaaminen voi merkitä äänitteen valmistajan oikeuksien loukkaamista, jos se tehdään ilman tämän suostumusta

Äänitteeltä otetun äänikatkelman käyttö muunnellussa ja kuuntelemalla tunnistamattomassa muodossa ei kuitenkaan loukkaa valmistajan oikeuksia edes tilanteessa, jossa suostumusta ei ole pyydetty

Vuonna 1977 musiikkiryhmä Kraftwerk julkaisi äänitteen, joka sisältää musiikkikappaleen *Metall auf Metall*. Moses Pelham ja Martin Haas ovat säveltäneet musiikkikappaleen *Nur mir*, joka julkaistiin vuonna 1997 Pelham GmbH:n tuottamana. Kaksi Kraftwerk-ryhmän jäsentä, Ralf Hütter ja Florian Schneider-Esleben, esittävät, että Pelham olisi kopioinut noin kaksi sekuntia rytmisekvenssiä *Metall auf Metall* -kappaleesta samplaustekniikan¹ avulla ja lisännyt sen kappaleeseen *Nur mir* peräjälkeen toistettavaksi. Hütter ja Schneider-Esleben katsoivat, että tekijänoikeutta, joka heille kyseisen äänitteen valmistajina² kuului, oli loukattu, ja vaativat muun muassa kieltomääräystä, vahingonkorvausta ja *Nur mir* -kappaleen äänitteiden luovuttamista tuhoamista varten.

Asiaa käsitellyt Bundesgerichtshof (liittovaltion ylin yleinen tuomioistuin, Saksa) tiedusteleo unionin tuomioistuimelta muun muassa sitä, onko unionin tekijänoikeuslainsäädännön ja lähioikeuksien³ sekä unionin perusoikeuskirjassa taattujen perusoikeuksien mukaan kyseessä äänitteen valmistajan oikeuksien loukkaus, jos valmistajan äänitteestä on otettu äänikatkelma (näyte) liitettäväksi samplaamalla toiseen äänitteeseen ilman valmistajan suostumusta. Bundesgerichtshof nostaa esiin myös kysymykset unionin oikeudessa säädetyistä oikeudenhaltijoiden oikeuksia koskevista poikkeuksista ja rajoituksista. Tältä osin se haluaa tietää, ovatko Saksan lain säännökset, joiden mukaan tekijänoikeudella suojattua teosta vapaasti käyttäen luotu itsenäinen teos voidaan lähtökohtaisesti julkaista ja hyödyntää ilman oikeudenhaltijoiden suostumusta, sopusoinnussa unionin oikeuden kanssa. Se haluaa myös tietää, voiko samplaaminen kuulua lainauksia koskevan poikkeuksen piiriin, joka vapauttaa käyttäjän velvollisuudesta hankkia äänitteen valmistajan suostumus suojatun äänitteen käyttöön.

Tänään antamassaan tuomiossa unionin tuomioistuin kiinnittää huomiota ensinnäkin siihen, että äänitteen valmistajalla on yksinoikeus sallia tai kieltää kokonaan tai osittain äänitteensä valmistaminen. Näin ollen se, että käyttäjä jäljentää toisen valmistamasta äänitteestä otetun äänikatkelman – jopa hyvin lyhyen – on lähtökohtaisesti kyseisen äänitteen osittaista jäljentämistä, joka kuuluu äänitteen valmistajan yksinoikeuden piiriin.

¹ *Samplaaminen* on tekniikka, jossa äänitteestä otetaan elektronisten apuvälineiden avulla otteita, joita käytetään uuden sävellyksen osana toisessa äänitteessä.

² Äänitteen valmistajalla tarkoitetaan luonnollista henkilöä tai oikeushenkilöä, joka ensimmäisenä tallentaa esityksen äänen tai muut äänet.

³ Tekijänoikeuden ja lähioikeuksien tiettyjen piirteiden yhdenmukaistamisesta tietoyhteiskunnassa 22.5.2001 annettu Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY (EYVL L 167, s. 10) ja vuokraus- ja lainausoikeuksista sekä tietyistä tekijänoikeuden lähioikeuksista henkisen omaisuuden alalla 12.12.2006 annettu Euroopan parlamentin ja neuvoston direktiivi 2006/115/EY (EUVL L 376, s. 28).

"Jäljentämisestä" ei kuitenkaan ole kyse, jos käyttäjä taiteen vapauttaan harjoittaessaan ottaa äänitteeltä äänikatkelman ja käyttää sitä muunnellussa muodossa uudessa teoksessa, jossa se ei ole kuuntelemalla tunnistettavissa. Oletus, että sellainen äänikatkelman käyttö olisi kopiaamista, joka edellyttää äänitteen valmistajan suostumusta, olisi ristiriidassa muun muassa sen vaatimuksen kanssa, että on tarpeen pitää tasapainossa yhtäältä tekijänoikeuksien ja niihin liittyvien suoja-oikeuksien haltijoiden intressi suojata peruskirjassa turvattua oikeuttaan henkiseen omaisuuteensa, toisaalta suojattujen kohteiden käyttäjien intressi ja perusoikeudet, muun muassa peruskirjassa taattu taiteen vapaus, ja lopulta vielä yleinen etu.

Tuomioistuin toteaa edelleen, että tuotos, joka käyttää kaikkia tai huomattavaa osaa äänitteeseen tallennetuista äänistä, on kyseisen äänitteen jäljennös, jonka osalta äänitteen valmistajalla on yksinomainen levitysoikeus. Kuitenkaan tällaisia jäljennöksiä eivät ole – kuten nyt kyseessä olleessa menettelyssä – tuokset, jotka käyttävät musiikin katkelmia mahdollisesti muunnelluissa muodoissa, siirrettyinä alkuperäisestä äänitteestä uuden ja itsenäisen teoksen osaksi.

Lisäksi unionin oikeudessa säädetyt oikeudenhaltijoiden oikeuksia koskevat poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjä on ottanut huomioon suojan piirissä olevien teosten valmistajien ja käyttäjien edut sekä myös yleisen edun. Näitä poikkeuksia ja rajoituksia säännellään kattavasti myös sisämarkkinoiden toiminnan varmistamiseksi tekijänoikeuden ja sen lähioikeuksien alalla. Näin ollen Saksan lain säännökset, joissa unionin oikeuden poikkeusten ja rajoitusten tyhjentävästä luonteesta huolimatta säädetään poikkeuksesta tai rajoituksesta, josta ei ole säädetty unionin oikeudessa ja joiden mukaan toisen teosta vapaasti käyttäen luotu itsenäinen teos voidaan lähtökohtaisesti julkaista ja hyödyntää ilman käytetyn teoksen luojaan suostumusta, eivät ole sopusoinnussa unionin oikeuden kanssa.

Mitä tulee yksinomaiseen valmistus- ja levitysoikeuteen liittyviin poikkeuksiin ja rajoituksiin, joita jäsenvaltiot voivat unionin oikeuden mukaisesti säätää suojatun teoksen lainaamisen osalta, tuomioistuin toteaa, että äänitteeltä otetun äänikatkelman käyttö, joka mahdollistaa alkuperäisen teoksen tunnistamisen, voi tietyin edellytyksin olla lainaus, erityisesti silloin, kun käytön tarkoituksena on vuorovaikutus alkuperäisen teoksen kanssa. Jos alkuperäinen teos ei sitä vastoin ole tunnistettavissa, katkelman käyttö ei ole lainaus.

Tuomioistuin huomauttaa lopuksi, että milloin unionin oikeus ei tyhjentävästi määrää jäsenvaltioiden toimintaa, jäsenvaltiot voivat soveltaa perusoikeuksien suojelua koskevia kansallisia normeja pannaan täytäntöön unionin oikeutta, edellyttäen, että muun muassa perusoikeuskirjan suojan taso ei vaarannu. Äänitteen valmistajan yksinomaisen kappaleenvalmistusoikeuden aineellinen sisältö on kuitenkin täysimääräisen yhdenmukaistamistoimenpiteen kohteena, minkä vuoksi kansallisia normeja ei tältä osin voida soveltaa.

HUOMAUTUS: Ennakkoratkaisupyyntö antaa jäsenvaltioiden tuomioistuimille mahdollisuuden tiedustella niiden käsiteltävänä olevan yksittäisen asian puitteissa unionin tuomioistuimelta unionin oikeuden tulkintaa tai unionin toimen pätevyyttä. Unionin tuomioistuin ei ratkaise kansallisen tuomioistuimen käsiteltävänä olevaa asiaa. Kansallisen tuomioistuimen tehtävänä on ratkaista asia unionin tuomioistuimen ratkaisun mukaisesti. Unionin tuomioistuimen ratkaisu sitoo vastaavalla tavoin muita kansallisia tuomioistuimia, jotka käsittelevät samanlaista ongelmaa.

Tämä on tiedotusvälineiden käyttöön tarkoitettu epävirallinen asiakirja, joka ei sido Euroopan unionin tuomioistuinta.

Tuomion [koko teksti](#) julkaistaan CURIA-sivustolla julistamispäivänä.

Lisätietoja: Hartmut Ost (+352) 4303 3255

Kuvia tuomion julistamisesta on saatavilla sivustolla "[Europe by Satellite](#)"

(+32) 2 2964106. Yleisarvio

www.curia.europa.eu

Liite 2. Excel-rinnakkaiskorpus ja MQM-pohjainen arviointi

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT	PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL	PE / DeepL
L1	Gerichtshof der Europäischen Union	Euroopan unionin tuomioistuin	Euroopan unionin tuomioistuin	Ei virheitä.	0	1	Euroopan unionin tuomioistuin	Ei virheitä.	0	1
L2	PRESSEMITTEILUNG Nr. 98/19	LEHDISTÖTIEDOTE nro 98/19	LEHDISTÖTIEDOTE nro 98/19	Ei virheitä.	0	1	LEHDISTÖTIEDOTE Nro 98/19	Iso "N"-alkukirjain lyhenteessä: saksan kielen interferenssi (saksassa lyhenne "Nr." kirjoitetaan aina isolla alkukirjaimella, kun taas suomenkielisessä lehdistöedotteessa standardina lyhenne "nro" kirjoitetaan pienellä) (Locale convention: Minor).	1	2
L3	Luxemburg, den 29. Juli 2019	Luxemburgissa 29.7.2019	Luxemburg, 29. heinäkuuta 2019	Pääte "-ssa" puuttuu: saksan kielen interferenssi (saksassa paikkakunta ilmoitetaan päiväyksen edellä perusmuodossa eli nominatiivissa) (Linguistic conventions/Morphology: Minor, 1). Pilkun lisäys: saksan kielen interferenssi (saksassa paikan ja päiväyksen välissä on pilkku) (Linguistic conventions, Punctuation: Minor, 1). Kuukautta ei kirjoiteta auki EU:n tuomioistuimen lehdistöedotteissa: käännöskoneen kyvyttömyys noudattaa asiakaskohtaisia paikalliskonventioita (Locale convention: Minor, 1).	3	3	Luxemburg, 29. heinäkuuta 2019	Pääte "-ssa" puuttuu: saksan kielen interferenssi (saksassa paikkakunta ilmoitetaan päiväyksen edellä perusmuodossa eli nominatiivissa) (Linguistic conventions/Morphology: Minor, 1). Pilkun lisäys: saksan kielen interferenssi (saksassa paikan ja päiväyksen välissä on pilkku) (Linguistic conventions, Punctuation: Minor, 1). Kuukautta ei kirjoiteta auki EU:n tuomioistuimen lehdistöedotteissa: käännöskoneen kyvyttömyys noudattaa asiakaskohtaisia paikalliskonventioita (Locale convention: Minor, 1).	3	3
L4	Presse und Information	Media ja lehdistö	Lehdistö ja tiedotus	Muoto "Lehdistö ja tiedotus" poikkeaa institutionaalisesta nimeämisikäytännöstä (Terminology, Style/Locale convention: Minor, 1)	1	2	Lehdistö ja tiedotus	Muoto "Lehdistö ja tiedotus" poikkeaa institutionaalisesta nimeämisikäytännöstä (Terminology, Style/Locale convention: Minor, 1)	1	2
L5	Urteil in der Rechtssache C-476/17 Pelham GmbH, Moses Pelham und Martin Haas / Ralf Hütter und Florian Schneider-Esleben	Tuomio asiassa C-476/17 Pelham GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian Schneider-Esleben	Tuomio asiassa C-476/17 Pelham GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian Schneider-Esleben	Ei virheitä.	0	1	Tuomio asiassa C-476/17 Pelham GmbH, Moses Pelham ja Martin Haas / Ralf Hütter ja Florian Schneider-Esleben	Ei virheitä.	0	1
L6	Das Sampling kann einen Eingriff in die Rechte des Tonträgerherstellers darstellen, wenn es ohne dessen Zustimmung erfolgt	Samplaaminen voi merkitä äänitteen valmistajan oikeuksien loukkaamista, jos se tehdään ilman tämän suostumusta	Sampling voi merkitä puuttumista äänitetuottajan oikeuksiin, jos se tapahtuu ilman tämän suostumusta	Lauseen ilhavointi puuttuu (Design & Markup: Minor, 1). Puhtaasti visuaalinen/tekniinen poikkeama lähdetekstistä. Se ei muuta asiasisältöä, mutta rikkoee tekstin yhtenäisyyttä. "puuttuminen" on yleisempi sana. Oikeudellisessa tekstissä, jossa käsitellään immateriaalioikeuksien suojausta, vakiintunut termi on loukkaus tai loukkaaminen. Konekäännin ei tunne alan ammattikieltä. (Terminology: Minor, 1). "tapahtuu" on epädiomaattinen valinta - aktiivinen toiminta (samplaaminen) vaatii tyyllillisesti passiivin ("tehdään"). "tapahtuu"-verbin ylikäyttö on klassinen saksan kielen interferenssin merkki, jossa kääntävyys on muodollisesti oikein mutta tyyllillisesti vierasta. (Style/Linguistic conventions: Minor, 1). Vakiintunut termi "sampling": Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1)	4	4	Sampling voi loukata äänitteen valmistajan oikeuksia, jos se tapahtuu ilman tämän suostumusta	Kursiivi vain "sampling"-sanan ensimmäisessä kolmessa kirjaimessa (Design & Markup: Minor, 1). "tapahtuu" on epädiomaattinen valinta - aktiivinen toiminta (samplaaminen) vaatii tyyllillisesti passiivin ("tehdään"). "tapahtuu"-verbin ylikäyttö on klassinen saksan kielen interferenssin merkki, jossa kääntävyys on muodollisesti oikein mutta tyyllillisesti vierasta. (Style/Linguistic conventions: Minor, 1). Vakiintunut termi "sampling": Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1).	3	4

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT	PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL	PE / DeepL
L7	Die Nutzung eines Audiofragments, das einem Tonträger entnommen wurde, in geänderter und beim Hören nicht wiedererkennbarer Form stellt jedoch auch ohne Zustimmung keinen Eingriff in diese Rechte dar	Ääniteeltä otetun äänikatkelman käyttö muunnellussa ja kuuntelemalla tunnistamattomassa muodossa ei kuitenkaan loukkaa valmistajan oikeuksia edes tilanteessa, jossa suostumusta ei ole pyydetty	Ääniteestä otetun audiofragmentin käyttäminen muutetussa muodossa, joka ei kuunneltaessa ole tunnistettavissa, ei kuitenkaan merkitse ilman suostumusta	"audiofragmentin" on tarpeettoman tekninen vierassana vakiintuneemman suomenkielisen sanan (äänikatkelma/ääninäyte) sijaan (Terminology: Minor, 1) "puuttumista" toistuva juridinen epätarkkuus (Terminology: Minor, 1).	2	3	Äänitalenteesta otetun äänifragmentin käyttö muutetussa ja kuuntelussa tunnistamattomassa muodossa ei kuitenkaan loukkaa näitä oikeuksia ilman suostumusta	"äänifragmentin" tarpeettoman tekninen vierassana vakiintuneemman suomenkielisen sanan ("äänikatkelma/ääninäyte") sijaan (Terminology: Minor, 1) "kuuntelussa tunnistamattomassa" on väärä muoto - luontevampaa olisi käyttää modaalirakennetta (kuuntelemalla) tai adjektiivia (korvakuuloita tunnistamaton). Heijastaa liian suoraan saksan kielen beim Hören -rakennetta (käännöslaina). (Linguistic conventions / Style, Minor, 1) Logiikan päinvastaisuus: "ilman suostumusta" antaa kuvan että loukkaus tapahtuu jos suostumus on saatu, jättää huomiotta lähdekstän fokuspartikkelin auch (myös/edes). (Accuracy/Mistranslation: Critical, 25)	27	4
L8	Die Musikgruppe Kraftwerk veröffentlichte im Jahr 1977 einen Tonträger, auf dem sich das Musikstück Metall auf Metall befindet.	Vuonna 1977 musiikkiryhmä Kraftwerk julkaisi äänitteen, joka sisältää musiikkikappaleen Metall auf Metall.	Musiikkiryhmä Kraftwerk julkaisi vuonna 1977 äänitteen, jolla on musiikkikappale Metall auf Metall.	Ilmaisu "jolla on" on turhan arkinen ja mekaaninen vastine rakenteelle auf dem sich ... befindet (Style/Register: Minor, 1).	1	2	Musiikkiryhmä Kraftwerk julkaisi vuonna 1977 äänitteen, jolla on kappale Metall auf Metall.	Ilmaisu "jolla on" on turhan arkinen ja mekaaninen vastine rakenteelle auf dem sich ... befindet (Style/Register: Minor, 1).	1	2
L9	Herr Moses Pelham und Herr Martin Haas sind die Komponisten des Musikstücks Nur mir, das im Jahr 1997 auf Tonträgern der Pelham GmbH erschienen ist.	Moses Pelham ja Martin Haas ovat säveltäneet musiikkikappaleen Nur mir, joka julkaistiin vuonna 1997 Pelham GmbH:n tuottamana.	Herra Moses Pelham ja herra Martin Haas ovat säveltäneet musiikkikappaleen Nur mir, joka julkaistiin vuonna 1997 Pelham GmbH:n äänitteillä.	Tarpeeton herroittelu (Herr Moses Pelham -> herra Moses Pelham); suomalaisessa asiatyylissä titeletään hyvin harvoin toistetaan nimen edellä. (Style / Register, Formality: Minor, 1) "Julkaiseminen äänitteillä" epädiomaattinen ilmaisu, mekaaninen vastine "auf Tonträgern... erschienen ist" (Style / Register, Formality: Minor, 1).	2	2	Moses Pelham ja Martin Haas ovat säveltäneet kappaleen Nur mir, joka julkaistiin vuonna 1997 Pelham GmbH:n äänitteenä.	"Julkaistiin äänitteenä" epädiomaattinen ilmaisu, mekaaninen vastine "auf Tonträgern... erschienen ist". (Style / Register, Formality: Minor, 1)	1	2
L10	Zwei Mitglieder der Gruppe Kraftwerk, Herr Ralf Hütter und Herr Florian Schneider-Esleben, machen geltend, Pelham habe etwa zwei Sekunden einer Rhythmussequenz aus dem Titel Metall auf Metall mit Hilfe der Sampling1-Technik kopiert und dem Titel Nur mir in fortlaufender Wiederholung unterlegt.	Kaksi Kraftwerk-ryhmän jäsentä, Ralf Hütter ja Florian Schneider-Esleben, esittävät, että Pelham olisi kopioinut noin kaksi sekuntia rytmisekvenssiä Metall auf Metall -kappaleesta samplausmekaniikalla avulla ja lisännyt sen kappaleeseen Nur mir peräjäkeen toistettavaksi.	Kraftwerk-yhtyeen kaksi jäsentä, herra Ralf Hütter ja herra Florian Schneider-Esleben, väittävät, että Pelham on kopioinut noin kaksi sekuntia rytmijaksosta kappaleesta Metall auf Metall samplausmekaniikalla avulla ja asettanut sen kappaleen Nur mir taustalle jatkuvana toistona.	Tarpeeton herroittelu (Herr Moses Pelham -> herra Moses Pelham); suomalaisessa asiatyylissä titeletään hyvin harvoin toistetaan nimen edellä. (Style / Register, Formality: Minor, 1) Väärä modus "on kopioinut" kun oikea olisi konditionaali "olisi kopioinut" - indikatiivi kuulostaa siltä, että tuomioistuimien jo todennut kopioinnin tapahtuneeksi (Accuracy, linguistic conventions, Mistranslation, Critical, 25) "esittävät" (machen geltend) on juridisessa kontekstissa usein neutraalimpi ja vakiintuneempi termi kuin "väittävät" (Style / Register, Formality: Minor, 1). "asettanut taustalle" ei vastaa lähdekstää, jossa merkityssisältö on se, että näyte on sisällytetty osaksi teoksen rakennetta (Accuracy/Hallucination, Mistranslation: Major, 5). "rytmijakso" on epämääräinen, kun taas "rytmisekvenssi" olisi täsmällinen tekninen termi (Style/Register, Minor, 1). Vakiintunut termi "sampling"; Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1)	34	4	Kaksi Kraftwerk-yhtyeen jäsentä, Ralf Hütter ja Florian Schneider-Esleben, väittävät, että Pelham on kopioinut noin kahden sekunnin pituisen rytmisekvenssin kappaleesta Metall auf Metall samplausmekaniikalla avulla ja lisännyt sen toistuvasti kappaleeseen Nur mir.	Puuttuva alaviitenumero "1" (Design & Markup, Minor 1). väärä modus "on kopioinut" kun oikea olisi konditionaali "olisi kopioinut" - indikatiivi kuulostaa siltä, että tuomioistuimien jo todennut kopioinnin tapahtuneeksi (Accuracy, Mistranslation: Critical, 25) "esittävät" (machen geltend) on juridisessa kontekstissa usein neutraalimpi ja vakiintuneempi termi kuin "väittävät" (Style / Register, Formality: Minor, 1). Virheellinen määrittäminen "lisännyt sen toistuvasti": jos sanotaan "lisännyt sen toistuvasti", adverbiaali toistuvasti määrittää itse tekoa (lisäämistä) - jos taas puhutaan "loopeista", kyse on lopputuloksesta (rytmi toistuu kappaleen sisällä) (Accuracy, Mistranslation: Major, 5). Prepositiolause "zum Zweck..." käännetty turhan mekaanisesti, suomen kielen asiatyylissä translatiivi (hävitettäväksi) on dynamisempi ja luonnollisempi (Style / Register, Formality: Minor, 1). Vakiintunut termi "sampling"; Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1).	34	4

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGP ^{PE} / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL ^{PE} / DeepL		
L11	Da Herr Hütter und Herr Schneider-Esleben der Auffassung sind, dass das ihnen als Hersteller des betroffenen Tonträgers zustehende verwandte Schutzrecht verletzt worden sei, beantragten sie u. a. Unterlassung, Schadensersatz und Herausgabe der Tonträger mit dem Titel <i>Aur mir</i> zum Zweck ihrer Vernichtung.	Hütter ja Schneider-Esleben katsoivat, että tekijänoikeutta, joka heille kyseisen äänitteen valmistajana2 kuului, oli loukattu, ja vaativat muun muassa kielomääräystä, vahingonkorvausta ja <i>Aur mir</i> -kappaleen äänitteiden luovuttamista tuhoamista varten. "Huom. ihmiskääntäjä on tässä hieman epätarkka käyttäessään sanaa "tekijänoikeus", vaikka se usein maallikoille tarkoittaa samaa asiaa. (Minor: 1)	Koska herra Hütter ja herra Schneider-Esleben katsovat, että heidän kyseisen äänitteen tuottajina2 omistama lähiöikeus on loukattu, he vaativat muun muassa kiellon määräämistä, vahingonkorvausta ja <i>Aur mir</i> -nimistä äänitettä sisältävien tallenteiden luovuttamista hävitettäväksi.	Tarpeeton herroittelu: suomalaisessa asiatyylissä titteleitä hyvin harvoin toistetaan nimien edellä. (Style / Register, Formality: Minor, 1). "koska"-sanalla alkava lauserakenne on turhan mekaaninen vastine saksankielisen lauseen rakenteille (Style, Linguistic conventions: Minor, 1). "betroffenen": "asianomainen" viittaa juridisessa kielessä yleensä henkilöön (osapuoleen). Esineistä tai asioista (kuten äänitteistä) puhuttaessa "kyseinen" tai "kyseessä oleva" on oikea valinta. (Accuracy/Mistranslation: Minor, 1) "äänitteiden, joiden nimi on "Nur mir", luovuttamista" on kömpelö relatiivirakenne verrattuna DeepL:n ja Fefin sujuvampaan adjektiivimääräitteen "Nur mir"-nimisten äänitallenteiden". Kone ei osaa tiivistää informaatiota suomalaisen yhdyssanarakenteeseen. (Linguistic conventions/Syntax: Minor, 1) Virhe moduksissa: aikamuotojen epäohdonmukaisuus ("katsovat", "on loukattu") suhteessa tekstin kerronnalliseen kokonaisuuteen (narratiivinen imperfekti vs. presens). Aikamuoto on siis suora käännöslaina saksan kielipolista. (Linguistic conventions/Grammar, Minor, 1) "kielto" on terminä liian yleiskielinen ja epämääräinen - "kieltomääräys" on juridisesti täsmällisempi ja vastaa paremmin suomalaista prosessioikeudellista terministöä. Se viittaa	7	4	Koska Hütter ja Schneider-Esleben katsovat, että heidän oikeuksiaan kyseisen äänitallenteen2 valmistajina on loukattu, he vaativat muun muassa lopettamista, vahingonkorvausta ja <i>Aur mir</i> -kappaleen sisältävien äänitallenteiden luovuttamista tuhoamista varten. Virhe moduksissa: aikamuotojen epäohdonmukaisuus ("katsovat", "on loukattu") suhteessa tekstin kerronnalliseen kokonaisuuteen (narratiivinen imperfekti vs. presens). Aikamuoto on siis suora käännöslaina saksan kielipolista. (Linguistic conventions/Grammar, Minor, 1) "lopettamista" on terminä liian yleiskielinen ja epämääräinen - "kieltomääräys" on juridisesti täsmällisempi ja vastaa paremmin suomalaista prosessioikeudellista terminstöä. Se viittaa nimenomaan tuomioistuimen antamaan määräykseen, jolla loukkaus kielletään (Terminology). "lopettaminen" vaatisi tässä myös määritteen: käännöksessä ei ngt kerrota minkä lopettamista vaaditaan? (Accuracy/Omission) Major, 5.	8	4	
L12	Der mit der Sache befasste Bundesgerichtshof möchte vom Gerichtshof u. a. wissen, ob es nach dem Urheberrecht und dem Recht verwandter Schutzrechte der Union3 sowie nach den durch die Charta der Grundrechte der Europäischen Union garantierten Grundrechten einen Eingriff in die Rechte des Herstellers eines Tonträgers, dem ein Audiofragment (<i>Sample</i>) entnommen wurde, darstellt, wenn dieses Audiofragment ohne dessen Zustimmung mittels <i>Sampling</i> in einen anderen Tonträger eingefügt wird.	Asiaa käsittelevä Bundesgerichtshof (liittovaltion ylin yleinen tuomioistuin, Saksa) tiedustelee unionin tuomioistuimelta muun muassa sitä, onko unionin tekijänoikeuslainsäädännön ja lähiöoikeuksien3 sekä unionin perusoikeuskirjassa taattujen perusoikeuksien mukaan kyseessä äänitteen valmistajan oikeuksien loukkaus, jos valmistajan äänitteestä on otettu äänikatkelmia (näyte) liitettäväksi samplaamalla toiseen äänitteeseen ilman valmistajan suostumusta. "Huom. ihmiskääntäjä on laiminlyönyt merkittävät "Sample" ja "Sampling" kursivoilla, kuten lähdeteksti on tehnyt. (Design & Markup: 1)	Bundesgerichtshof, jonka käsiteltävänä asia on, haluaa unionin tekijänoikeuden ja lähiöoikeuksien3 sekä Euroopan unionin perusoikeuskirjassa taattujen perusoikeuksien kannalta muun muassa tietää, merkitseekö puuttumista sen äänitteen tuottajan oikeuksiin, josta audiofragmentti (<i>Sample</i>) on otettu, jos tämä audiofragmentti liitetään ilman tuottajan suostumusta <i>samplingin</i> avulla toiseen äänitteeseen.	Vakintunton termi "sampling": Käytetään lainaa, vaikka alalla on vakintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1) Verbi "merkitä" vaatisi subjektin ("tämä", "se"), mutta ngt se puuttuu (Accuracy, Omission: Minor, 1). "kö"-liite: epäsuora kysymys on muotoiltu kömpelösti. Asiantuntijatyylissä kuuluisi käyttää muotoa "onko" tai "sitä, onko" (Linguistic conventions/Grammar, Minor, 1). "Puuttuminen" yleiskielinen ja epätarkka (Terminology: Minor, 1). Raskas relatiivilause "Bundesgerichtshof, jonka käsiteltävänä asia on..." katkaisee pitkän lauseen tarpeettomasti jo heti alussa, suomenle tyypillinen partisiippi määrä "Asiaa käsittelevä Bundesgerichtshof..." olisi tiiviimpi ja dynamisempi (Linguistic conventions/Word order: Minor, 1). "haluaa tietää" on liian arkinen sanavalinta viralliseen oikeudelliseen tekstityhteyteen. Poikkeaa lajityypin konventioista. (Style/ Register, Formality: Minor, 1) Kone on sijoittanut pääverbin "tietää" lauseen loppuun, suomessa se on syntaktinen virhe (Linguistic conventions/Word order: Minor, 1).	24	4	Asiaa käsittelevä liittovaltion tuomioistuin pyytää unionin tuomioistuimelta muun muassa selvitystä siitä, voidaanko tekijänoikeuslainsäädännön ja unionin perusoikeuskirjassa taattujen perusoikeuksien nojalla loukkaa äänitallenteen valmistajan oikeuksia, jos äänitallenteesta otettu ääninäyte (<i>sample</i>) lisätään ilman valmistajan suostumusta <i>ääninäytteiden ottamisen</i> avulla toiseen äänitallenteeseen.	"voidaanko... loukkaa" on kieliooppivirhe: Passiivisen apuverbin voidaanko kanssa tulisi käyttää infinitiiviä loukata (Linguistic conventions, Grammar: Minor, 1). "Ääninäytteiden ottamisen avulla": Kone ylisellittää termin, jolle on jo vakintunut suomenkielinen vastine (sampling) (Terminology: Minor, 1). "ääninäyte" on sujuva käännös sanalle "Audiofragment", huom. L7 DeepL käänsi saman sanan "äänifragmentti" (Terminology/Consistency: Minor, 1). Bundesgerichtshof: Kone korvaa erisnimen epätarkalla yleisnimityksellä, lukija ei tiedä, mitä tuomioistuinta tarkoitetaan. (Accuracy/Terminology, Omission: Major, 5) "Muun muassa" sanajärjestys on epäluonnollinen suomessa, "muun muassa" pitäisi sijoittaa määrittämään kysymyksen sisältöä, ei pyyntöverbiä. Lähdekielen sanajärjestys (u. a. wissen) on siirretty suomeen sellaisenaan. (Linguistic conventions/Word order: Minor, 1) "unionin" määrittäminen sijoitella aiheuttaa, että määritteen vaikutusala on väärä. Käännös antaa ymmärtää, että vain toinen mainituista oikeudenaloista on unionin tasoinen (Accuracy/Mistranslation: Major, 5).	16	4

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L13	Der Bundesgerichtshof wirt auch Fragen zu den im Unionsrecht vorgesehenen Ausnahmen und Beschränkungen für die Rechte der Rechtsinhaber auf.	Bundesgerichtshof nostaa esiin myös kysymykset unionin oikeudessa säädettyistä oikeudenhaltijoiden oikeuksista koskevista poikkeuksista ja rajoituksista.	Bundesgerichtshof nostaa esiin myös kysymykset unionin oikeudessa säädettyistä oikeudenhaltijoiden oikeuksista koskevista poikkeuksista ja rajoituksista.	Ei virheitä.	0	1	Liittovaltion korkein oikeus esittää myös kysymyksiä unionin lainsäädännössä säädettyistä poikkeuksista ja rajoituksista oikeudenhaltijoiden oikeuksiin.	"liittovaltion korkein oikeus" on epätarkka käänös, "liittovaltion ylin yleinen tuomioistuin" on juridisesti oikea vastine, mutta käänntimen tulisi silti pikemminkin säilyttää erisnimi "Bundesgerichtshof" tarkan viitteen takaamiseksi. (Terminology: Minor, 1) "Unionin lainsäädäntö" liian kapea termi joka viittaa vain säädettyihin normeihin, "Unionsrecht" kattaa muutakin kuin lainsäädännön. (Terminology: Minor, 1) "esittää (myös kysymyksiä)" on neutraali ja korrekti, mutta asiantuntijakielessä "nostaa esiin" tai "tiedustelee" olisi dynaamisempi ja tekstilajiin paremmin sovelia. (Style/Register, Formality: Minor, 1) "kysymyksiä" epätarkkuus: käyttämällä partitiivia (kysymyksiä) kone antaa ymmärtää, että kysymyksiä on joitakin, kun taas akkusatiivi (kysymykset) viittaisi juuri niihin nimellisiksi kysymyksiin, jotka on mainittu tekstissä. (Accuracy/Mistranslation: Minor, 1)	4	3
L14	Er möchte insoweit wissen, ob die deutschen Rechtsvorschriften, wonach ein selbständiges Werk, das in freier Benutzung eines geschützten Werks geschaffen worden ist, grundsätzlich ohne die Zustimmung der Rechtsinhaber veröffentlicht und verwertet werden darf, mit dem Unionsrecht vereinbar sind.	Tältä osin se haluaa tietää, ovatko Saksan lain säännökset, joiden mukaan tekijänoikeudella suojattua teosta vapaasti käyttäen luotu itsenäinen teos voidaan lähtökohtaisesti julkaisua ja hyödyntää ilman oikeudenhaltijoiden suostumusta, sopusoinnussa unionin oikeuden kanssa.	Se haluaa tältä osin tietää, ovatko unionin oikeuden kanssa yhteensopivia Saksan oikeussäännökset, joiden mukaan itsenäinen teos, joka on luotu suojatun teoksen vapaata käyttöä hyödyntäen, voidaan lähtökohtaisesti julkaisua ja sitä voidaan hyödyntää ilman oikeudenhaltijoiden suostumusta.	"ovatko unionin oikeuden kanssa yhteensopivia Saksan oikeussäännökset" sijoittaa subjektin epäluonnollisesti lauseen loppuun. Lauseen painotus on väärä, mikä tekee tekstistä vaikeaselkoisen ja epäidiomaattisen. (Linguistic conventions/Word order: Minor, 1) "vapaata käyttöä hyödyntäen" antaa ymmärtää, että kyseessä on toimijan tietoinen valinta käyttää jotakin "vapautta", vaikka kyse on juridisen kriteerin täyttymisestä. Termi vapaa käyttö pitäisi liittää säännökseen, ei teoksen ominaisuuteen. (Accuracy/Mistranslation: Major, 5) "hyödyntäen": kone on lisännyt merkityksen (aktiivinen hyödyntäminen), jota alkutekstissä ei ole. (Accuracy/Hallucination: Major, 5)	11	4	Se haluaa tietää, ovatko Saksan säännökset, joiden mukaan itsenäinen teos, joka on luotu vapaasti käyttämällä suojattua teosta, voidaan periaatteessa julkaisua ja hyödyntää ilman oikeudenhaltijoiden suostumusta, unionin lainsäädännön mukaisia.	"säännökset" on yleiskielinen ja jättää täsmennyksen, että kyse on nimenomaan oikeudellisista normeista. (Terminology: Minor, 1) Liian pitkä välilyönti subjektin "Saksan säännökset" ja predikaatiivin "unionin lainsäädännön mukaisia" välillä katkaisee ajatuksen ja tekee lauseesta vaikeaselkoisen. Lähdekielen kehysrakenteen mekaaninen siirto suomeen. (Linguistic conventions, Word order: Minor, 1) "voidaan" verbi on sijoitettu liian kauas subjektista saksan kielen mallin mukaisesti. Piikoo suomen kielen luonnollisen sanajärjestyksen. (Linguistic conventions/Word order: Minor, 1) Kone on jättänyt kääntämättä sanan <i>insoweit</i> , joka on keskeinen sidossana lauseen loogisen rajauksen kannalta. (Accuracy/Omission: Major, 5)	8	4
L15	Er möchte außerdem wissen, ob das <i>Sampling</i> unter die Ausnahme für Zitate fallen kann, die den Nutzer von der Pflicht befreit, für die Nutzung des geschützten Tonträgers die Zustimmung des Tonträgerherstellers einzuholen.	Se haluaa myös tietää, voiko samplaaminen kuulua lainauksia koskevan poikkeuksen piiriin, joka vapauttaa käyttäjän velvollisuudesta hankkia äänitteiden valmistajan suostumus suojatun äänitteen käyttöön. "Huom. ihmiskääntäjä on unohtanut merkitä "samplaaminen" kursiivilla, kuten lähdeteksti on tehnyt. (Design & Markup: 1) Ihmiskääntäjä käyttää relatiivipronominia "joka", joka viittaa kielilopillisesti sanaan piiriin, vaikka loogisesti sen pitäisi viitata sanaan poikkeus. (Linguistic conventions/Grammar: Major, 5)	Se haluaa lisäksi tietää, voiko <i>sampling</i> kuulua sitaattia koskevan poikkeuksen piiriin; tämä poikkeus vapauttaisi käyttäjän velvollisuudesta hankkia äänitetuottajan suostumus suojatun äänitteen käyttöön.	Vakiintunut termi "sampling": Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1)	1	3	Se haluaa lisäksi tietää, voidaanko <i>sampling</i> katsoa kuuluvan lainauspoikkeuksen piiriin, joka vapauttaa käyttäjän velvollisuudesta hankkia äänitteen valmistajan suostumus suojatun äänitteen käyttöön.	Vakiintunut termi "sampling": Käytetään lainaa, vaikka alalla on vakiintunut suomenkielinen vastine "samplaus". (Terminology: Minor, 1) Kone äänin käyttää relatiivipronominia "joka", joka viittaa kielilopillisesti sanaan piiriin, vaikka loogisesti sen pitäisi viitata sanaan poikkeus. (Linguistic conventions/Word order, Grammar: Major, 5) "voidaanko sampling": lauseenvastikkeen looginen subjekti on nominatiivissa, vaikka sen pitäisi olla genetiivissä "samplingin". (Linguistic conventions/Grammar, Morphology: Minor, 1) "lainauspoikkeuksen" arkisuus: "lainauksia koskevan poikkeuksen" olisi muodollisempi ja analyttisempi vironomaiskielen ja säädöskielen perinteitä noudattava rakenne. (Style/register, Formality: Minor, 1) "kopiointi" ei tavoita oikeudellisen tekstilajin vaatimaa teknistä tarkkuutta. "kappaleen valmistaminen" kattaa kaikenlaisen tallentamisen, monistamisen ja toistamisen, oli kyseessä sitten digitaalinen tiedosto tai fyysinen levy. (Terminology: Minor, 1)	8	4
L16	In seinem heutigen Urteil weist der Gerichtshof zunächst darauf hin, dass die Tonträgerhersteller das ausschließliche Recht haben, die Vervielfältigung ihrer Tonträger ganz oder teilweise zu erlauben oder zu verbieten.	Tänään antamassaan tuomiossa unionin tuomioistuin kiinnittää huomiota ensinnäkin siihen, että äänitetuottajilla on yksinoikeus sallia tai kieltää äänitteidensä kappaleen valmistaminen kokonaan tai osittain.	Tuomioistuin muistuttaa tämänpäiväisessä tuomiossaan ensiksi, että äänitetuottajilla on yksinoikeus sallia tai kieltää äänitteidensä kappaleen valmistaminen kokonaan tai osittain.	"muistuttaa": hieman liian yleiskielinen, vaikuttaa siltä, että tuomioistuin "opastaa" lukijaa. Laskee tekstin muodollisuusastetta, "toteaa" noudattaa suomalaisen oikeuskielen konventioita ja säilyttää tuomioistuimen auktoriteetin (Style/Register, Formality: Minor, 1).	1	3	Tämänpäiväisessä tuomiossaan tuomioistuin toteaa ensinnäkin, että äänitallentajien valmistajilla on yksinoikeus sallia tai kieltää äänitallentajien kopiointi kokonaan tai osittain.	"kopiointi" ei tavoita oikeudellisen tekstilajin vaatimaa teknistä tarkkuutta. "kappaleen valmistaminen" kattaa kaikenlaisen tallentamisen, monistamisen ja toistamisen, oli kyseessä sitten digitaalinen tiedosto tai fyysinen levy. (Terminology: Minor, 1)	1	3

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL	
L17	Folglich ist die Vervielfältigung eines – auch nur sehr kurzen – Audiofragments, das einem Tonträger entnommen wurde, durch einen Nutzer grundsätzlich eine teilweise Vervielfältigung dieses Tonträgers , so dass eine solche Vervielfältigung unter das ausschließliche Recht des Tonträgerherstellers fällt.	Näin ollen se, että käyttäjä jäljentää toisen valmistamasta äänitteestä otetun äänikatkelman – jopa hyvin lyhyen – on lähtökohtaisesti kyseisen äänitteen osittaista jäljentämistä, joka kuuluu äänitteen valmistajan yksinoikeuden piiriin. *Huom. ihmiskääntäjä on laiminlyönyt lihavoida tekstin lähdetekstin mukaisesti. (Design & Markup: Minor, 1)	Näin ollen tällaisen – myös hyvin lyhyen – äänitteestä otetun audiofragmentin kappaleen valmistaminen käyttäjän toimesta on lähtökohtaisesti tämän äänitteen osittaista kappaleen valmistamista , joten tällainen kappaleen valmistaminen kuuluu äänitetuottajan yksinoikeuden piiriin.	"joten" yhdessä lauseen aloittavan "näin ollen" - ilmauksen kanssa luo tyyllisesti kömpelön kaksinkertaisen seurausrakenteen. Ammattimaisessa juridisessa suomessa tällaiset rakenteet korvattaisiin usein sujuvammalla jolloinrakenteella tai muuttamalla passiivimainen agenttirakenne aktiiviseksi. Tämä osoittaa kääntimen kyvyttömyyden muokata lauseen informaatorakennetta kohdekielen konventioiden mukaiseksi. (Style/Register, Formality: Minor, 1) Genetiiviketju "audiofragmentin kappaleen valmistaminen"; peräkkäiset genetiivimääräet malkivat lähdetekstiä ja tekevät tekstistä kankeaa ja vaikeasti hahmotettavaa. (Linguistic conventions/Morphology, Word order: Minor, 1) "audiofragmentin" on tarpeettoman tekninen vierassana (k.s. edell. kommentti). (Terminology: Minor, 1)	3	4	Näin ollen käyttäjän tekemä äänitalenteesta otetun, jopa hyvin lyhyen äänifragmentin kopiointi on periaatteessa kyseisen äänitalteen osittainen kopiointi , joten tällainen kopiointi kuuluu äänitalteen valmistajan yksinomaiseen oikeuteen. "periaatteessa" sisältää arkkielisen konnotaation teoreettisuudesta tai epävarmuudesta ("periaatteessa joo, periaatteessa ei"), mikä voi heikentää tekstin juridista auktoriteettia. Termi "lähtökohtaisesti" viesti lukijalle vahvemmin siitä, että kyseessä on oikeudellinen pääsääntö. (Style/Register, Formality: Minor, 1) "tekemä osittainen kopiointi" -rakenne on kankea ja epäluonnollinen; se suosii turhan raskaata substantiivirakennetta aktiivisen ilmaisuuden sijasta. (Style/Register, Formality: Minor, 1) "äänifragmentin" on tarpeettoman tekninen vierassana, k.s. aiempi kommentti (Terminology: Minor, 1)	4	4
L18	Keine „Vervielfältigung“ liegt jedoch vor, wenn ein Nutzer in Ausübung seiner Kunstfreiheit einem Tonträger ein Audiofragment entnimmt, um es in geänderter und beim Hören nicht wiedererkennbarer Form in ein neues Werk einzufügen.	"Jäljentämisestä" ei kuitenkaan ole kyse, jos käyttäjä taiteen vapauttaan harjoittaessaan ottaa äänitteeltä äänikatkelman ja käyttää sitä muunnellussa muodossa uudessa teoksessa, jossa se ei ole kuuntelemalla tunnistettavissa.	"Kappaleen valmistamista" ei kuitenkaan ole kyseessä , jos käyttäjä taiteen vapauttaan käyttäen liittyy äänitteestä audiofragmentin muotona uuteen teokseen.	Osittainen lihavointi: Kone ei noudata lähdetekstin lihavoitteja (Design & Markup: Minor, 1). "valmistamista... kyseessä": verbi "olla kyseessä" ei sovi yhteen elatiivimuotoisen määritteen kanssa. "valmistamisesta... kyse" noudattaa suomen kielen vakiintunutta ja oikeaa rakennetta. (Linguistic conventions/Morphology, Grammar, Minor, 1) "vapauttaan käyttäen" arkisuus: "harjoittaessaan" olisi muodollinen ja tarkka vastine saksan Ausübung-termille ja vahvistaisi tekstin juridista asiantuntemusta. (Style/Register, Formality: Minor, 1) "audiofragmentin" on tarpeettoman tekninen vierassana (k.s. edell. kommentti). (Terminology: Minor, 1) "muutettuna" yleiskielisyys: asiallinen, mutta ei tavoita oikeudenalan teknistä spesifisyyttä. "muunnellussa muodossa" on vivahteikkaampi ja tarkempi merkityksen välittäjä. (Style/Terminology, register/Accuracy: Minor, 1) "tunnistamattomana muotona" epädiomaattisuus: kankea essiiviketju, joka ei vastaa suomen luonnollista virtaa. (Linguistic conventions/Syntax, Word order, Minor, 1).	6	4	Kopiointia ei kuitenkaan ole kyseessä, jos käyttäjä käyttää taiteellista vapauttaan ja ottaa äänitalenteesta äänifragmentin, jonka hän lisää uuteen teokseen muutettuna ja kuuntelussa tunnistamattomassa muodossa. "hän" on liian personoitu pronomini juridisessa asiantuntijatekstissä. (Style/Register, Morphology: Minor, 1) "käyttää... ja ottaa" heikko sidosteisuus: tekee tekstistä tarinamaisen ja vähentää juridista argumentatiivisuutta. Saksan kielen ilmaisu "in Ausübung" vaatii suomessa tiiviimpää, alisteista rakennetta. Ammattimaisessa juridisessa kielessä suosittava muoto "vapauttaan harjoittaessaan" (lauseenvastike) osoittaa, että toiminta tapahtuu nimenomaan perusoikeuden toteuttamisena. (Style/Linguistic conventions, Register, Word order: minor, 1) "kuuntelussa tunnistamattomassa" peräkkäiset inessiivirakenne on staattinen ja epäselvä. Se kuulostaa siltä, että tunnistamattomuus tapahtuu "jonkin sisällä", mikä on abstraktia ja vaikeaselkoista. (Linguistic conventions/Word order, Grammar: Minor, 1) "äänifragmentin" on tarpeettoman tekninen vierassana, k.s. aiempi kommentti (Terminology: Minor, 1).	5	4

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L19	Die Annahme, dass eine solche Nutzung eines Audiofragments eine Vervielfältigung darstellt, die der Zustimmung des Tonträgerherstellers bedarf, widerspräche u. a. dem Erfordernis, einen angemessenen Ausgleich zu sichern zwischen der Inhaber von Urheber- und verwandten Schutzrechten am Schutz ihres in der Charta verankerten Rechts am geistigen Eigentum und auf der anderen Seite dem Schutz der Interessen und Grundrechte der Nutzer von Schutzgegenständen, darunter der ebenfalls durch die Charta gewährleisteteten Kunstfreiheit, sowie dem Allgemeininteresse.	Oletus, että sellainen äänikatkelman käyttö olisi kopiointia, joka edellyttää ääniteen valmistajan suostumusta, olisi ristiriidassa muun muassa sen vaatimuksen kanssa, että on tarpeen pitää tasapainossa yhtäältä tekijänoikeuksien ja niihin liittyvien suoja- ja oikeuksien välillä ja toisaalta suoja- ja oikeuksien välillä. Lisäksi "niihin liittyvien suoja- ja oikeuksien" on epädiomaattinen, kun taas "lähioikeuksien" noudattaa suomalaisen tekijänoikeuslain vakiintunutta terminologiaa. (Terminology: Minor, 1)	Sen toteaminen, että tällainen audiofragmentin käyttö on kappaleen valmistamista, joka edellyttää äänitetuottajan suostumusta, olisi muun muassa ristiriidassa sen vaatimuksen kanssa, että toisaalta tekijänoikeuksien ja lähioikeuksien haltijoiden etujen – heidän perusoikeuskirjaan kirjattun immateriaalioikeutensa suojan – ja toisaalta suoja- ja oikeuksien välillä on turvattava asianmukainen tasapaino.	"audiofragmentin" on tarpeettoman tekninen vierassana. (Ks. edell. kommentti). (Terminology: Minor, 1) "Sen toteaminen" muuttaa oletuksen (Annahme) faktaksi (toteaminen) (Accuracy/Mistranslation: Major, 5). "muun muassa ristiriidassa": adverbii siirtyy verbin viereen, mikä hämää painotusta. "ristiriidassa muun muassa" kohdistaisi painotuksen tarkasti oikeudelliseen vaatimukseen. (Accuracy/Mistranslation: Minor, 1) "on turvattava asianmukainen tasapaino" sijoittaminen lauseen loppuun seuraa saksan syntaksia ja on liian kaukana siitä ohjaavasta sanasta (vaatimus) [vrt. ihmiskääntäjän käännös]. (Style/Register, Word order: Minor, 1) "kirjatun immateriaalioikeutensa" on mekaaninen käännös saksan yksiköstä. Suomen juridisessa kielenkäytössä monikko on vakiintuneempi viittaessa haltijan oikeuksiin. (Linguistic conventions/Morphology: Minor, 1) "edut" on asiallinen ja oikein, mutta hieman yleiskielinen. Ei tavoita intressipunninnan teknistä sävyä yhtä hyvin kuin "intressi". (Style/Register, Formality: Minor, 1) "turvattava" viittaa enemmänkin suojan lopputulokseen (oikeus on turvassa), kun taas "vahvistettu" vastaisi hyvin saksan sanaa verankert (ankkuroitu/vahvistettu/kirjattu) - se viittaa siihen, että oikeus on säädetty asiakirjassa. (Terminology: Minor, 1) "toisaalta... ja toisaalta" on väärä konnektiivipari; "yhtäältä" puuttuu, mikä rikkoo lauseen sisäisen logiikan. (Accuracy/Mistranslation: Major, 5)	16	4	Oletus, että tällainen äänifragmentin käyttö on kopiointia, joka edellyttää äänitallenteen valmistajan suostumusta, olisi ristiriidassa muun muassa vaatimuksen kanssa, jonka mukaan on varmistettava asianmukainen tasapaino yhtäältä tekijänoikeuksien ja lähioikeuksien haltijoiden etujen, jotka liittyvät perusoikeuskirjassa vahvistettuun immateriaalioikeuksien suojaan, ja toisaalta suoja- ja oikeuksien välillä on turvattava asianmukainen tasapaino.	"äänifragmentin" on tarpeettoman tekninen vierassana, k.s. aiempi kommentti (Terminology: Minor, 1). "edut" on asiallinen ja oikein, mutta hieman yleiskielinen. Ei tavoita intressipunninnan teknistä sävyä yhtä hyvin kuin "intressi". (Style/Register: Minor, 1)	2	3
L20	Der Gerichtshof stellt sodann fest, dass ein Gegenstand, der alle oder einen wesentlichen Teil der in einem Tonträger festgelegten Töne übernimmt, eine Kopie dieses Tonträgers ist, für die der Tonträgerhersteller über ein ausschließliches Verbreitungsrecht verfügt.	Tuomioistuimien toteaa edelleen, että tuotos, joka käyttää kaikkia tai huomattavaa osaa äänitteeseen tallennetuista äänistä, on kyseisen ääniteen jäljennös, jonka osalta ääniteen valmistajalla on yksinomainen levitysoikeus.	Tuomioistuimien toteaa seuraavaksi, että esine, joka on maksu kaikki tai olennaisen osan äänitteeseen tallennetuista äänistä, on kyseisen ääniteen kappale, johon äänitetuottajalla on yksinoikeus levittää.	"esine" on rekisteriltään liian fyysinen verrattuna asiantuntijakieleen. (Style/Register, Formality: Minor, 1) "omaksuu" on epädiomaattinen: tuotos ei voi "omaksua" ääniä samalla tavalla kuin ihminen omaksuu ajatuksia. (Style/Accuracy/Linguistic conventions: Major, 5) "yksinoikeus levittää" on sujuvaa suomea (kuvaava sitä, mitä haltija saa tehdä), mutta verbaalilylyi jää voimemmaksi kuin oikeuskielessä suosittu substantiivilyly "yksinomainen levitysoikeus" (vakiintunut termi sille oikeudelle, joka haltijalla on), joka vastaa täsmällisesti saksan termiä "Verbreitungsrecht". (Style/Register, Formality: Minor, 1)	7	3	Tuomioistuimien toteaa edelleen, että esine, joka sisältää kaikki tai olennaisen osan äänitallenteeseen tallennetuista äänistä, on kyseisen äänitallenteen kopio, johon äänitallenteen valmistajalla on yksinoikeus levittää.	"yksinoikeus levittää" on sujuvaa suomea (kuvaava sitä, mitä haltija saa tehdä), mutta verbaalilylyi jää voimemmaksi kuin oikeuskielessä suosittu substantiivilyly "yksinomainen levitysoikeus" (vakiintunut termi sille oikeudelle, joka haltijalla on), joka vastaa täsmällisesti saksan termiä "Verbreitungsrecht". (Style/Register, Formality: Minor, 1)	1	3

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L21	Keine solche Kopie ist jedoch ein Gegenstand, der – wie der im Ausgangsverfahren fragliche –, nur Musikfragmente, gegebenenfalls in geänderter Form, übernimmt, die von diesem Tonträger übertragen werden, um ein neues und davon unabhängiges Werk zu schaffen.	Kuitenkaan tällaisia jäljennöksiä eivät ole – kuten nyt kyseessä olleessa menettelyssä – tuotokset, jotka käyttävät musiikin katkelmia mahdollisesti muunnelluissa muodoissa, siirretynä alkuperäisestä äänitteestä uuden ja itsenäisen teoksen osaksi. *Huom. Ihmiskääntäjän käännöksessä lihavoinnit eivät täsmää lähdetekstin kanssa: lihavoinnina pitäisi olla "tuotokset, jotka" ja "mahdollisesti muunnelluissa muodoissa" pitäisi olla normaalla tekstillä. (Design & Markup: Minor, 1)	Tällainen kappale ei kuitenkaan ole esine, joka – kuten pääasiassa kyseessä oleva – omaksuu vain musiikkifragmentteja, tarvittaessa muutettuna muodossa, jotka siirretään tästä äänitteestä, jotta luodaan uusi ja siitä riippumaton teos.	"pääasiassa" sekoittuu yleiskieliseen adverbiin "enimmäkseen", mikä hämää oikeudellista viittausta. (Accuracy: Major, 5) "kuten pääasiassa kyseessä oleva" on mekaaninen käännös saksan partisiipista ilman suomen vaatimaa täsmennyttä. Siitä puuttuu pääsana: saksa sallii adjektiivin tai partisiipin käytön substantiivina (wie das ... fragliche), kun taas suomessa partisiippi "oleva" vaatii lähes poikkeuksetta pääsanon (substantiivin), jota se määrittää. Lisäksi lauseesta puuttuu osoittava pronomini (kuten tässä tai kyseisessä). (Accuracy, Omission / Syntax: Minor, 1) "esine" on rekisteriltään liian fyysinen verrattuna asiantuntijakieleen. (Style/Register: Minor, 1) "omaksuu" on epäidiomaattinen: tuotos ei voi "omaksua" ääniä samalla tavalla kuin ihminen omaksuu ajatuksia. (Style/Register: Minor, 1) "tarvittaessa" on juridisesti epätarkka, sillä se viittaa tavoitteelliseen toimintaan tai tarkoituksenmukaisuusharkintaan. Kyseisessä oikeudellisessa kontekstissa ei kuitenkaan ole kyse toimijan tarpeesta, vaan objektin (ääninäytteen) faktisesta tilasta. Näyte on joko muutettu tai muuttamaton; kyse on asiantilasta, joka on oikeudellisen arvioinnin edellytys, ei toimintaohjeesta. (Accuracy, Mistranslation: Major, 5) "jotka" viittaa loogisesti kauempana olevaan monikolliseen sanaan "musiikkifragmentteja", väliin tullut määrite ("tarvittaessa muutettuna muodossa") katkaisee viittaussuhteen. (Accuracy, Linguistic conventions: Minor, 1) Pronomini "siitä" luo heikon viittaussuhteen ja kärsii pronomini-inflaatiosta (liikaa pronomineja samassa virkkeessä: jotka, tästä, siitä). Luo konekäännösmäisyyttä. (Style: Minor, 1)	16	4	Esine, joka – kuten asiassa, josta on annettu viittauspäätös, kyseessä oleva esine – sisältää vain äänitallenteesta siirrettäviä, mahdollisesti muunnettuja musiikkikatkelmia uuden ja siitä riippumattoman teoksen luomiseksi, ei kuitenkaan ole tällainen kopio.	"esine" on rekisteriltään liian fyysinen verrattuna asiantuntijakieleen. (Style/Register: Minor, 1) Suomen kielelle on vierasta sijoittaa pitkä relativause ("josta on annettu viittauspäätös") vertailurakenteen sisään ennen pääsanaa. Kyseessä saksalle ominaisen kiilalauseen säilyttäminen: lauseen sisään sijoitettu toinen lause tai määrite. (Linguistic conventions/word order: Minor, 1) Verbi "sisältää" kuvaa staattista tilaa, kun taas esim. "käyttää" kuvaisi paremmin tekijänoikeudellista prosessia eli näytteen (sampling) ottamista. (Accuracy/Mistranslation: Minor, 1)	9	4
L22	Außerdem spiegeln die im Unionsrecht vorgesehenen Ausnahmen und Beschränkungen für die Rechte der Rechtsinhaber bereits wider, dass der Unionsgesetzgeber die Interessen der Hersteller und der Nutzer von geschützten Gegenständen sowie das Allgemeininteresse berücksichtigt hat.	Lisäksi unionin oikeudessa säädetty oikeudenhaltijoiden oikeuksia koskevat poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjä on ottanut huomioon suojan piirissä olevien teosten valmistajien ja käyttäjien edut sekä myös yleisen edun.	Lisäksi unionin oikeudessa säädetty oikeudenhaltijoiden oikeuksia koskevat poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjä on ottanut huomioon suojujien kohteiden tuottajien ja käyttäjien edut sekä yleisen edun.	"Suojattu" on yleiskielisempi ja voi kuulostaa mekaaniselta käännökseltä (saks. geschützt). "Suojan piirissä oleva" tai "suojan kohteiden" nostaa tekstin rekisteriä vastaamaan alkuperäisen oikeudellisen tekstin juhlallisuutta ja tarkkuutta. (Style/Register: Minor, 1)	1	3	Lisäksi unionin lainsäädännössä säädetty oikeudenhaltijoiden oikeuksien poikkeukset ja rajoitukset heijastavat jo sitä, että unionin lainsäätäjä on ottanut huomioon suojujien esineiden valmistajien ja käyttäjien edut sekä yleisen edun.	"lainsäädännössä säädetty" on liian kapea: ei kata oikeuskäytäntöä tai perussopimuksia. Vrt. "unionin oikeudessa säädetty". (Accuracy, Terminology: Major, 5) "esineiden" on rekisteriltään liian fyysinen verrattuna asiantuntijakieleen. (Style/Register: Minor, 1) "Suojattu" on yleiskielisempi ja voi kuulostaa mekaaniselta käännökseltä (saks. geschützt). "Suojan piirissä oleva" tai "suojan kohteiden" nostaa tekstin rekisteriä vastaamaan alkuperäisen oikeudellisen tekstin juhlallisuutta ja tarkkuutta. (Style/Register: Minor, 1)	7	3

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L23	Diese Ausnahmen und Beschränkungen sind auch erschöpfend geregelt, um das Funktionieren des Binnenmarkts im Bereich des Urheberrechts und der verwandten Schutzrechte zu sichern.	Näitä poikkeuksia ja rajoituksia säännellään kattavasti myös sisämarkkinoiden toiminnan varmistamiseksi tekijänoikeuden ja sen lähioikeuksien alalla. "Huom. lähdetekstin termi "erschöpfend" on käännetty virheellisesti muotoon "kattavasti". Kyseessä olevan direktiivin 2001/29/EY) kontekstissa termi tulee kääntää muotoon "tähjntävästi". "Tähjntävy" tarkoittaa oikeudellisesti suljettua järjestelmää, jossa luettelon ulkopuoliset poikkeukset eivät ole mahdollisia, kun taas "kattava" on merkitykseltään yleiskielisempi ja jättää sääntelyn avoimeksi muille tulkinnoille. (Terminology: Major, 5)	Nämä poikkeukset ja rajoitukset on myös säännelty tähjntävästi, jotta tekijänoikeutta ja lähioikeuksia koskevien sisämarkkinoiden toimivuus voidaan varmistaa.	"jotta"-rakenne on suora käänntösalina ja raskaampi kuin lauseenvastike. (Style/Register, Minor 1) Rakenne "tekijänoikeutta ja lähioikeuksia koskevien sisämarkkinoiden" muuttaa sisämarkkinat ikään kuin nimenomaan näitä oikeuksia varten oleviksi markkinoiksi. Saksan lähdetekstin <i>im Bereich des ...</i> viittaa pikemminkin toimintaan tietyllä alalla. (Accuracy/Mistranlation: Major 5)	6	4	Nämä poikkeukset ja rajoitukset on myös säädetty tähjntävästi, jotta varmistetaan sisämarkkinoiden toimivuus tekijänoikeuden ja lähioikeuksien alalla.	"Jotta"-rakenne tekee lauseesta tarpeettoman pitkän ja hieman kerronnallisen. "...sisämarkkinoiden toimivuuden varmistamiseksi". Se on dynamisempi ja vastaa paremmin suomalaisen viiakielien konventioita. (Style/Register, Formality: Minor, 1) "on myös säädetty tähjntävästi" - sanan "myös" sijoittelu muuttaa lauseen painotusta: muoto, jossa painopiste on verbissä tai ominaisuudessa "...on säädetty myös tähjntävästi" tai "...on säädetty tähjntävästi myös sisämarkkinoiden..." olisi luontevampi. (Linguistic conventions/Word order: Minor, 1)	2	4
L24	Daher sind die deutschen Rechtsvorschriften, die trotz des abschließenden Charakters der Ausnahmen und Beschränkungen eine nicht im Unionsrecht geregelte Ausnahme oder Beschränkung vorsehen, nach der ein selbständiges Werk, das in freier Benutzung des Werks eines anderen geschaffen wurde, grundsätzlich ohne Zustimmung des Urhebers des benutzten Werks veröffentlicht und verwertet werden darf, nicht mit dem Unionsrecht vereinbar.	Näin ollen Saksan lain säännökset, joissa unionin oikeuden poikkeusten ja rajoitusten tähjntävästä luonteesta huolimatta säädetään poikkeuksesta tai rajoituksesta, josta ei ole säädetty unionin oikeudessa ja joiden mukaan toisen teosta vapaasti käyttäen luotu itsenäinen teos voidaan lähioikeuksien alalla julkistaa ja hyödyntää ilman käytetyn teoksen luojan suostumusta, eivät ole sopuinnussa unionin oikeuden kanssa.	Näin ollen Saksan oikeussäännökset, joissa poikkeusten ja rajoitusten tähjntävästä luonteesta huolimatta säädetään poikkeuksesta tai rajoituksesta, jota unionin oikeudessa ei tunneta ja jonka mukaan toisen teoksen vapaata käyttöä hyödyntäen luotu itsenäinen teos voidaan lähioikeuksien alalla julkistaa ja sitä voidaan hyödyntää ilman käytetyn teoksen tekijän suostumusta, eivät ole unionin oikeuden mukaisia.	Lähdetekstissä viitataan nimenomaan säännösten puuttumiseen, "ei tunneta" on epätasällinen ja avoin tulkinnalle. (Accuracy/Mistranlation: Minor, 1) "oikeussäännökset" on mekaaninen käänntösvastine saksan sanalle <i>Rechtsvorschriften</i> . "lain säännökset" on ihmiskääntäjän valinta, ja se on suomalaisessa oikeuskielessä vakiintuneempi ja idiomatiikaltaan vahvempi. (Style/Register: Minor, 1)	2	3	Siksi Saksan lainsäädäntö, jossa säädetään unionin lainsäädännössä sääntelemättömästä poikkeuksesta tai rajoituksesta huolimatta poikkeusten ja rajoitusten lopullisesta luonteesta, jonka mukaan itsenäinen teos, joka on luotu toisen teoksen vapaalla käytöllä, voidaan periaatteessa julkistaa ja hyödyntää ilman käytetyn teoksen tekijän suostumusta, ei ole unionin lainsäädännön mukainen.	"siksi" on hieman liian yleiskielinen ja keventää tekstiä tavalla, joka ei sovi oikeuden tuomion tyyliin (Style/Register: Minor, 1). "periaatteessa" on liian yleiskielinen. Se voi viitata epävarmuuteen tai teoreettisuuteen, eikä se ole juridisesti yhtä painokas kuin "lähioikeuksien alalla". (Accuracy: Minor, 1) "Saksan lainsäädäntö" on hieman epätarkka. Muuttaa yksittäiset säännökset kollektiiviseksi lainsäädännöksi. Sujuvaa, mutta kadottaa saksan tekstin fokuksen tiettyihin normeihin. (Style/Register: Minor, 1) "Lopullisesta luonteesta" on harhaanjohtava. "Lopullinen" viittaa suomessa usein ajalliseen päättymiseen tai muuttamattomuuteen, ei luettelon kattavuuteen tai poissulkevuteen. (Accuracy/Mistranlation: Major, 5) Ilmaisu "sääntelemättömästä poikkeuksesta" on kömpelö. Juridisessa suomessa on vakiintuneempaa sanoa, että poikkeuksesta "ei ole säädetty". (Style/Register: Minor, 1) Lauseessa esiintyy sana "lainsäädäntö" kolme kertaa, mikä tekee tekstistä raskaslukuista ja epäammattimaisista. Ammattimainen kääntäjä käyttäisi synonyymeja tai rakenteita, joissa toisto vältetään (esim. "unionin oikeus", "sen säännökset"). (Style: Minor, 1) "Huolimatta"-rakenne seuraa saksan lausejärjestystä tavalla, joka katkaisee ajatuksen. (Linguistic conventions/Syntax: Minor, 1) "Vapaalla käytöllä" ei vastaa juridisia konventioita toiminnan kuvaamisessa. Koneen versio tekee "vapaasta käytöstä" ikään kuin työkalun (vapaalla käytöllä), kun kyse on oikeudellisesta toimintatavasta tai säännöksestä. (Linguistic conventions/Morphology: Minor, 1)	12	4

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L25	Hinsichtlich der Ausnahmen und Beschränkungen für die ausschließlichen Rechte zur Vervielfältigung und Wiedergabe, die von den Mitgliedstaaten nach dem Unionsrecht für Zitate aus einem geschützten Werk vorgesehen werden können, stellt der Gerichtshof fest, dass die Nutzung eines Audiofragments, das einem Tonträger entnommen wurde und das Werk, dem es entnommen ist, erkennen lässt, unter bestimmten Voraussetzungen ein Zitat sein kann, insbesondere dann, wenn die Nutzung zum Ziel hat, mit diesem Werk zu interagieren.	Mitä tulee yksinomaiseen valmistus- ja levitysoikeuteen liittyviin poikkeuksiin ja rajoituksiin, joita jäsenvaltiot voivat unionin oikeuden mukaisesti säätää suojatun teoksen lainaamisen osalta, tuomioistuin toteaa, että äänitteeltä otetun äänikatkelman käyttö, joka mahdollistaa alkuperäisen teoksen tunnistamisen, voi tietjin edellytksin olla lainaus, erityisesti silloin, kun käytön tarkoituksena on vuorovaikutus alkuperäisen teoksen kanssa.	Sitaatteja koskevien, jäsenvaltioiden unionin oikeuden nojalla mahdollisesti säätämien poikkeusten ja rajoitusten osalta – jotka koskevat kappaleen valmistamisen ja yleisölle välittämiseen liittyviä yksinoikeuksia – tuomioistuin toteaa, että äänitteestä otetun audiofragmentin käyttö, joka tekee siitä teoksesta, josta fragmentti on otettu, tunnistettavan, voi tietjin edellytksin olla sitaatti, erityisesti silloin, kun käytön tavoitteena on olla vuorovaikutuksessa kyseisen teoksen kanssa.	Lihavointi puuttuu lauseen loppuosasta: "erityisesti silloin, kun käytön tavoitteena on olla vuorovaikutuksessa kyseisen teoksen kanssa." (Design & Markup: Minor, 1)	10	4	Jäsenvaltioiden unionin lainsäädännön nojalla suojatun teoksen lainauksille mahdollisesti myöntämien jäljentämis- ja esittämisoikeuksien poikkeusten ja rajoitusten osalta tuomioistuin toteaa, että äänitalenteesta otetun äänifragmentin käyttö, josta voidaan tunnistaa teos, josta se on otettu, voi tietjissä olosuhteissa olla lainaus, erityisesti jos käytön tarkoituksena on olla vuorovaikutuksessa kyseisen teoksen kanssa.	"Jäljentämis- ja esittäminen" ovat väärät termit tekijänoikeuden taloudellisille oikeuksille. (Terminology: Major, 5) "Mahdollisesti" muuttaa toimivallan epävarmuudeksi. (Accuracy/Mistranslation: Major, 5) "Myöntämien" (lainsäädäntö). Lainsäätäjät säätää poikkeuksia, viranomainen myöntää lupia. Sekoittaa oikeuslähteiden hierarkian. (Accuracy, Mistranslation: Major, 5) "Olosuhteissa" on epätarkka vastine juridiselle vaatimukselle (edellytys). (Accuracy: Minor, 1) "myöntämien" on väärä verbi kuvaamaan lainsäädäntöprosessia. Sekoittaa lainsäätämisen ja hallinnollisen lupapäätöksen. (Accuracy, Mistranslation: Major, 5) relatiivipronominien ketju "...äänifragmentin käyttö (A), josta (B) voidaan tunnistaa teos (C), josta (D) se on otettu..." tekee tekstistä "puuromaista", koska sanojen väliset suhteet hämärtyvät. Onko kyseessä "lainsäädännön nojalla suojattu teos" vai "lainsäädännön nojalla myönnetty poikkeus"? Rakenne on kielipolisesti mahdollinen mutta tyylillisesti viiallinen. (Style, Linguistic conventions / Syntax: Minor, 1)	22	4
L26	Ist das Werk nicht zu erkennen, stellt die Nutzung des Fragments hingegen kein Zitat dar.	Jos alkuperäinen teos ei sitä vastoin ole tunnistettavissa, katkelman käyttö ei ole lainaus.	Jos teos ei ole tunnistettavissa, fragmentin käyttö ei sitä vastoin ole sitaatti.	"ei sitä vastoin ole" kömpelö sanajärjestys ja raskas konjunktiovalinta. Teksti tuntuu mekaaniselta käännökseltä. Kysymykseksi jää, mitä vastoin? (Style/Word order: Minor, 1)	2	4	Jos teosta ei voida tunnistaa, fragmentin käyttö ei sen sijaan ole lainaus.	"fragmentin" on tarpeeton lainasana. Suositeltavampaa käyttää vakiintunutta termiä (näyte/atkkelma). (Terminology: Minor, 1) Käyttämällä partitiiviä (teosta) ja toimintaverbiä (voida tunnistaa) lause muuttuu subjektiiviseksi havainnoksi. Juridisesti pitäisi määritellä kohteen ominaisuus ("teos ei ole tunnistettavissa"). (Accuracy/Mistranslation: Major, 5)	6	4
L27	Schließlich führt der Gerichtshof aus, dass die Mitgliedstaaten, wenn ihr Handeln nicht vollständig durch das Unionsrecht bestimmt wird, bei der Durchführung des Unionsrechts nationale Schutzstandards für die Grundrechte anwenden dürfen, sofern dadurch u. a. nicht das Schutzniveau der Charta beeinträchtigt wird.	Tuomioistuin huomauttaa lopuksi, että milloin unionin oikeus ei tyhjentävästi määrää jäsenvaltioiden toimintaa, jäsenvaltiot voivat soveltaa perusoikeuksien suojelua koskevia kansallisia normeja pannesaan täytäntöön unionin oikeutta, edellyttäen, että muun muassa perusoikeuskirjan suojan taso ei vaarannu.	Lopuksi tuomioistuin toteaa, että jäsenvaltiot voivat, silloin kun unionin oikeus ei määritä niiden toimintaa täysin, soveltaa unionin oikeutta täytäntöön pannesaan kansallisia perusoikeuksien suojan tasoja edellyttäen muun muassa, ettei perusoikeuskirjan suojan taso heikkene.	"ei määrittä... täysin" käsitteellinen epätarkkuus. EU-oikeudessa vakiintunut termi on "tyhjentävästi". Se viittaa siihen, onko sääntely suljettua [exhaustive]. "Täysin" on yleiskielinen ja juridisesti epätarkka. (Terminology: Minor, 1) "niiden" (vs. jäsenvaltioiden) viittausuuteen epäselvyys. Pirkissä juridisissa virkkeissä pronominit ovat riskialtita. Substantiivien toisto [jäsenvaltioiden] varmistaa ehdottoman selkeyden, mikä on tekstilajin vaatimus. (Style: Minor, 1) "määrittää" vs. "määrätä" säyvero. Unionin oikeus määrää (prescribes/dictates) jäsenvaltioiden toimivallasta. "Määrittää" on luonteeltaan kuvaileva. Se viittaa rajojen asettamiseen tai parametrien määrittelyyn (vt. "määritellä termi" tai "määrittää koordinaatti"). Se kuulostaa siltä, että unionin oikeus on passiivinen viitekehys. (Accuracy/Mistranslation: Minor, 1) "täytäntöön pannesaan... suojan tasoja": Täytäntöönpanon kohde* ei voi olla "suojan taso". Jäsenvaltiot panevat täytäntöön "unionin oikeutta". (Accuracy/Mistranslation: Major, 5)	8	5	Lopuksi tuomioistuin toteaa, että jos jäsenvaltioiden toimintaa ei säännellä kokonaan unionin lainsäädännöllä, ne voivat soveltaa unionin lainsäädännön täytäntöönpanossa kansallisia perusoikeuksien suojanormeja, edellyttäen, että tämä ei heikennä muun muassa perusoikeuskirjassa vahvistettua suojan taso.	"unionin lainsäädännöllä" asiasisältövirhe: <i>Unionsrecht</i> on unionin oikeus. Lainsäädäntö on vain osa oikeutta. (Terminology/Accuracy: Major, 5) "säännellä kokonaan" yleiskielisyys ja epätarkkuus. Juridinen avaintermi "tyhjentävästi" puuttuu. (Terminology: Minor, 1) "suojanormeja" epädiomaattisuus: Sana on teknisesti oikein muodostettu, mutta se ei kuulu suomalaisen juridiseen sanastoon. Teksti paljastuu käännöksiksi. Termi ei herätä juridisissa selkeää mielikuvaa siitä, tarkoitetaanko säädöstä vai tavoitetasoa. (Terminology/Style: Minor, 1) "vahvistettua" sanalle ei ole vastinetta lähdetekstissä. Se on kääntäjän keksimä tyylillinen lisäys. (Accuracy/Addition: Minor, 1)	8	3

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL	
L28	Der materielle Gehalt des ausschließlichen Vervielfältigungsrechts des Tonträgerherstellers ist jedoch Gegenstand einer Maßnahme zur vollständigen Harmonisierung, so dass eine solche Nutzung insofern auszuschließen ist.	Äänitteen valmistajan yksinomaisen kappaleenvalmistusoikeuden aineellinen sisältö on kuitenkin täysimääräisen yhdenmukaistamistoimenpiteen kohteena, minkä vuoksi kansallisia normeja ei tältä osin voida soveltaa.	Äänitetuottajan yksinomaisen kappaleen valmistamisoikeuden aineellinen sisältö on kuitenkin täydellisesti yhdenmukaistettu, joten tällainen käyttö on tältä osin poissuljettu.	"täydellisesti yhdenmukaistettu": "Täydellinen" on yleiskielinen ja juridisesti epätarkka. (Style/Register: Minor, 1) "Gegenstand einer Maßnahme" on jätetty kääntämättä. Kyse on nimenomaan toimenpiteestä (direktiivistä tms.), joka on yhdenmukaistanut oikeuden. (Accuracy/Omission: Major, 5) "kappaleen valmistamisoikeus" on yleiskielinen rakenne, joka seuraa tarkasti saksan kielen lauserakennetta. Käännösmäinen vaikutelma, koska yhdessä nimitys puuttuu. (Style: Minor, 1) "tällainen käyttö" väärä viittausuhde. Antaa ymmärtää, että kyse on teoksen käytöstä, vaikka kyse on kansallisten normien soveltamisesta. Muuttaa tuomion logiikan. (Accuracy/Mistranslation: Critical, 25) "poissuljettu" epäidiomaattisuus. Ei vastaa suomalaisen oikeuskielen tapaa ilmaista normin soveltumattomuutta. (Style/Register: Minor, 1) "joten" on tyylillisesti heikko: Pitkässä ja painavassa juridisessa lauseessa se ei kannattele edeltävän väitteen (täysimääräinen harmonisointi) ja sitä seuraavan johtopäätöksen (harkintavallan poissulkeminen) välistä loogista painoarvoa. (Style: Minor, 1)	34	3	Äänitallenteiden valmistajan yksinoikeuden aineellinen sisältö on kuitenkin täysin yhdenmukaistamisen kohteena, joten tällainen käyttö on tässä suhteessa suljettava pois. "täysimääräisen" on käännöslainaa. Suora käännös saksan verbistä auszuschließen. Juridisesti luontevampaa olisi sanoa, ettei kansallisia normeja "voida soveltaa". (Style: Minor, 1) "on [tässä suhteessa] suljettava pois" Analyysi: Syntaktinen interferenssi ja sanajärjestysvirhe: käännös noudattaa mekaanisesti saksan kehysrakennetta (Satzklammer), jossa finiittiiverbi ja infinitiivi on erotettu toisistaan ja verbi sijoittuu lauseen loppuun. Suomen kielessä rakenne on epäidiomaattinen ja tekee lauseesta takapainoisen. Kategoria ja pisteet: Style/Syntax: Minor, 1 "tässä suhteessa": "Tältä osin" on vakiintuneempi juridinen ilmaisu kuin yleiskielinen "tässä suhteessa". (Style/Register: Minor, 1) "joten" on tyylillisesti heikko: Pitkässä ja painavassa juridisessa lauseessa se ei kannattele edeltävän väitteen (täysimääräinen harmonisointi) ja sitä seuraavan johtopäätöksen (harkintavallan poissulkeminen) välistä loogista painoarvoa. (Style: Minor, 1)	43	5
L29	HINWEIS: Im Wege eines Vorabentscheidungsersuchens können die Gerichte der Mitgliedstaaten in einem bei ihnen anhängigen Rechtsstreit dem Gerichtshof Fragen nach der Auslegung des Unionsrechts oder nach der Gültigkeit einer Handlung der Union vorlegen.	HUOMAUTUS: Ennakkoratkaisupyyntö antaa jäsenvaltioiden tuomioistuimille mahdollisuuden tiedustella niiden käsiteltävänä olevan yksittäisen asian puitteissa unionin tuomioistuimelta unionin oikeuden tulkintaa tai unionin toimen pätevyyttä.	HUOMAUTUS: Ennakkoratkaisupyyntö perusteella jäsenvaltioiden tuomioistuimet voivat niiden käsiteltävänä olevassa oikeusriidassa esittää tuomioistuimelle kysymyksiä unionin oikeuden tulkinnasta tai unionin toimen pätevyydestä.	"tuomioistuimelle" Epätasavertaisuus. Ei yksilöi unionin tuomioistuinta, mikä on juridisesti tärkeää sekannusten välttämiseksi. (Terminology: Minor, 1) "oikeusriidassa" epäidiomaattisuus. Suomen oikeuskielessä "asia" tai "yksittäinen asia" on vakiintuneempi ilmaisu tässä kontekstissa. (Linguistic conventions: Minor, 1) "perusteella" on mekaaninen käännös. Seuraa saksan prepositiotalpimusta. Ref-version verbi rakenne on sujuvampaa suomea. (Style: Minor, 1) "kysymyksiä... tulkinnasta" selkeä, mutta ihmiskääntäjän version "tiedustella... tulkintaa" on kielellisesti hienostuneempi. (Style/Register: Minor, 1)	4	3	HUOMAUTUS: Jäsenvaltioiden tuomioistuimet voivat esittää tuomioistuimelle ennakkoratkaisupyyntö, joka koskee unionin oikeuden tulkintaa tai unionin oikeudellisen toimenpiteen pätevyyttä niiden käsiteltävänä olevassa riita-asiaa. "oikeudellisen toimenpiteen" on epästANDARDI termi. Handlung der Union on vakiintuneesti "unionin toimi". "Oikeudellinen toimenpide" on yleiskielisempi ja kömpelö tässä kontekstissa. (Terminology: Major, 5) Sana "oikeudellinen" on tarpeeton lisäys, jota lähdetekstistä ei löydy. (Accuracy/Addition: Minor, 1) "esittää... ennakkoratkaisupyyntö, joka koskee..." Rakennevirhe. Lähdetekstissä esitetään kysymyksiä (Fragen vorlegen) ennakkoratkaisupyyntö puitteissa, mutta kone kääntää, että esitetään pyyntö. Nuanssiero on pieni, mutta asiantuntijakielestä tarkkuus on tärkeää. (Style: Minor, 1) "joka koskee" on kuvaileva ja passiivinen. Tyypillistä yleiskielelle tai heikolle virkakiellelle. "koskevan" partisiippi rakenne on suositeltava asiantuntijatekstissä. (Style: Minor, 1)	14	5

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L30	Der Gerichtshof entscheidet nicht über den nationalen Rechtsstreit.	Unionin tuomioistuin ei ratkaise kansallisen tuomioistuimen käsiteltävänä olevaa asiaa.	Unionin tuomioistuin ei ratkaise kansallista oikeudenkäyntiä.	"ratkaise... oikeudenkäyntiä": epädiomaattisuus: Oikeuskielessä ratkaistaan asia, ei prosessia (oikeudenkäynti). (Style: Minor, 1) "kansallista oikeudenkäyntiä" yleiskielisyys. <i>Rechtsstreit</i> on tässä yhteydessä nimenomaan se viireillä oleva oikeusasia. (Terminology: Minor, 1)	2	3	Tuomioistuin ei ratkaise kansallista riita-asiaa.	"Tuomioistuin" epätarkkuus. Ei yksilöi Unionin tuomioistuinta, mikä on tärkeää viittaussuhteiden selkeyden vuoksi. (Terminology: Major, 5) "riita-asiaa" liian kapea termi. Rajaa periaatteen vain sivullioikeuteen, vaikka se koskee kaikkia oikeudenkäyntilajeja. (Terminology: Major, 5)	10	3
L31	Es ist Sache des nationalen Gerichts, über die Rechtssache im Einklang mit der Entscheidung des Gerichtshofs zu entscheiden.	Kansallisen tuomioistuimen tehtävänä on ratkaista asia unionin tuomioistuimen ratkaisun mukaisesti.	Kansallisen tuomioistuimen tehtävänä on ratkaista asia tuomioistuimen ratkaisun mukaisesti.	"Tuomioistuin" epätarkkuus. Ei yksilöi Unionin tuomioistuinta, mikä on tärkeää viittaussuhteiden selkeyden vuoksi. (Terminology: Major, 5)	5	3	Kansallisen tuomioistuimen tehtävänä on ratkaista asia tuomioistuimen päätöksen mukaisesti.	"Tuomioistuin" epätarkkuus. Ei yksilöi Unionin tuomioistuinta, mikä on tärkeää viittaussuhteiden selkeyden vuoksi. (Terminology: Major, 5)	5	3
L32	Diese Entscheidung des Gerichtshofs bindet in gleicher Weise andere nationale Gerichte, die mit einem ähnlichen Problem befasst werden.	Unionin tuomioistuimen ratkaisu sitoo vastaavalla tavoin muita kansallisia tuomioistuinta, jotka käsittelevät samanlaista ongelmaa.	Tuomioistuimen ratkaisu sitoo vastaavalla tavalla muita kansallisia tuomioistuinta, joiden käsiteltäväksi vastaava ongelma tulee.	"Tuomioistuin" epätarkkuus. Ei yksilöi Unionin tuomioistuinta, mikä on tärkeää viittaussuhteiden selkeyden vuoksi. (Terminology: Major, 5)	5	3	Tuomioistuimen päätös sitoo samalla tavalla muita kansallisia tuomioistuinta, jotka käsittelevät samanlaista asiaa.	"Tuomioistuin" epätarkkuus. Ei yksilöi Unionin tuomioistuinta, mikä on tärkeää viittaussuhteiden selkeyden vuoksi. (Terminology: Major, 5)	5	3
L33	Das <i>Sampling</i> ist eine Technik, bei der mit Hilfe elektronischer Geräte einem Tonträger Auszüge entnommen werden, um sie als Bestandteile einer neuen Komposition auf einem anderen Tonträger zu nutzen.	1 <i>Sampling</i> on tekniikka, jossa äänitteestä otetaan elektronisten apuvälineiden avulla otteita, joita käytetään uuden sävellyksen osana toisessa äänitteessä.	1 <i>Sampling</i> on tekniikka, jossa elektronisten laitteiden avulla otetaan äänitteestä otteita, joita voidaan käyttää uuden sävellyksen osana toisella äänitteellä.	"Sampling" on vierassana. Suomenkielisessä asiantuntijatekstissä "sampling" on vakiintuneempi ja paremmin kieleen mukautuva. (Style: Minor, 1) "toisella äänitteellä" Sijamuotovirhe. Äänite on abstrakti tallenne, jonka sisällä otteita käytetään. Pitäisi olla inessiivi "äänitteessä". Adessiivi (-lla) viittaisi fyysisen esineeseen (kuten levyyn). (Accuracy: Minor, 1)	2	3	1 <i>Sampling</i> on tekniikka, jossa äänitallenteesta otetaan elektronisten laitteiden avulla otteita, joita käytetään osana uutta sävellystä toisella äänitallenteella.	"Sampling" on vierassana. Suomenkielisessä asiantuntijatekstissä "sampling" on vakiintuneempi ja paremmin kieleen mukautuva. (Style: Minor, 1) "toisella äänitallenteella" lähdekieleen vaikutus: sijamuoto on valittu saksan preposition mukaan, ei suomen merkitysopin mukaan. (Accuracy: Minor, 1)	2	3
L34	2 Hersteller von Tonträgern sind natürliche oder juristische Personen, die die Herstellung von Tonträgern finanzieren.	2 Äänitteen valmistajalla tarkoitetaan luonnollista henkilöä tai oikeushenkilöä, joka ensimmäisenä tallentaa esityksen äänen tai muut äänet.	2 Äänitteiden tuottajat ovat luonnollisia henkilöitä tai oikeushenkilöitä, jotka rahoittavat äänitteiden valmistamisen.	"ovat... henkilöitä": Epätypillinen muoto. Lakikielessä käytetään "tarkoitetaan"-rakennetta. (Linguistic conventions: Minor, 1) "äänitteiden tuottajat" (mon.) (vrt. ihmiskääntäjän "äänitteen valmistajalla" (yks.)): suomessa käsitteet määritellään lähes poikkeuksetta yksikössä. (Linguistic conventions: Minor, 1)	2	3	2 Äänitallenteiden valmistajat ovat luonnolliset tai oikeushenkilöt, jotka rahoittavat äänitallenteiden valmistuksen.	"äänitallenteiden valmistajat ovat" (mon.) (vrt. ihmiskääntäjän "äänitteen valmistajalla" (yks.)): suomessa käsitteet määritellään lähes poikkeuksetta yksikössä. (Linguistic conventions: Minor, 1) "luonnolliset tai oikeushenkilöt" rakennevirhe. Suomen kieli vaatii sanan "henkilöt" molempien adjektiivien yhteyteen tai selkeämmän rakenteen. (Linguistic conventions/Syntax: Minor, 1) "valmistajat ovat" epätypillinen muoto. Määritelmät pitäisi antaa yksikössä "X:llä tarkoitetaan". (Linguistic conventions: Minor, 1)	3	4
L35	3 Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft (ABl. 2001, L 167, S. 10) und Richtlinie 2006/115/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums (ABl. 2006, L 376, S. 28).	3 Tekijänoikeuden ja lähioikeuksien tiettyjen piirteiden yhdenmukaistamisesta tietoyhteiskunnassa 22.5.2001 annettu Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY (EYVL L 167, s. 10) ja vuokraus- ja lainausoikeuksista sekä tietystä tekijänoikeuden lähioikeuksista henkisen omaisuuden alalla 12.12.2006 annettu Euroopan parlamentin ja neuvoston direktiivi 2006/115/EY (EUVL L 376, s. 28).	3 Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft (ABl. 2001, L 167, S. 10) und Richtlinie 2006/115/EG des Europäischen Parlaments und des Rates vom 12. Dezember 2006 zum Vermietrecht und Verleihrecht sowie zu bestimmten dem Urheberrecht verwandten Schutzrechten im Bereich des geistigen Eigentums (ABl. 2006, L 376, S. 28).	Kääntämättä jättäminen. Tekoäly on olettanut, että kyseessä on viite, jota ei tule kääntää, vaikka kyse on direktiivin virallisesta nimestä, jolla on vakiintunut suomennos. (Accuracy/Omission: Major, 5)	5	5	3 Euroopan parlamentin ja neuvoston direktiivi 2001/29/EY, annettu 22 päivänä toukokuuta 2001, tekijänoikeuden ja lähioikeuksien tiettyjen tietoyhteiskunnan alalla sovellettavien näkökohtien yhdenmukaistamisesta (EYVL 2001, L 167, s. 10) ja Euroopan parlamentin ja neuvoston direktiivi 2006/115/EY, annettu 12 päivänä joulukuuta 2006, vuokraus- ja lainausoikeudesta sekä tietystä tekijänoikeuden alaisista immateriaalioikeuksista (EUVL 2006, L 376, s. 28).	Vaikka koneen suomennos on kielipöytäsiivoksi ymmärrettävä, se on virheellinen, koska se ei käytä EU:n virallisesti vahvistettuja suomenkielisiä nimiä. (Linguistic conventions: Major, 5) Esimerkiksi: Suomalainen säädöskieli noudattaa direktiivien nimeämisessä täysin erilaista loogista ketjua kuin saksan ja englanti. Kone (Saksan malli): Direktiivi X, annettu [pvm], [aiheesta]. Ihmiskääntäjä (Suomen virallinen malli): [Aiheesta] [pvm] annettu direktiivi X.	5	5
L36	Zur Verwendung durch die Medien bestimmtes nichtamtliches Dokument, das den Gerichtshof nicht bindet.	Tämä on tiedotusvälineiden käyttöön tarkoitettu epävirallinen asiakirja, joka ei sido tuomioistuinta.	Välineille tarkoitettu epävirallinen asiakirja, joka ei sido tuomioistuinta.	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta. Käännöksessä on myös huonoja käännösvalintoja, esim. sana "välineet" ei tarkoita "mediaa" suomeksi ilman tarkentintaa. (Linguistic conventions: Minor, 1)	1	3	Tämä ei-virallinen asiakirja on tarkoitettu tiedotusvälineiden käyttöön, eikä se sido tuomioistuinta.	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta. Käännöksessä on myös huonoja käännösvalintoja, esim. "ei-virallinen" epädiomaattisuus: suomen hallintokielessä "epävirallinen" on vakiintuneempi. (Linguistic conventions: Minor, 1)	1	3

ID	Lähdeteksti (ST)	Ihmiskääntäjä (Ref)	ChatGPT 5.2	MQM / ChatGPT	MQM Score / ChatGPT PE / ChatGPT	DeepL Next-Gen	MQM / DeepL	MQM Score / DeepL PE / DeepL		
L37	Der Volltext des Urteils wird am Tag der Verkündung auf der Curia-Website veröffentlicht.	Tuomion koko teksti julkaistaan CURIA-sivustolla julistamispäivänä.	Tuomion täysi teksti julkaistaan julistamispäivänä Curia-sivustolla.	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta. (Linguistic conventions: Minor, 1)	1	3	Tuomion koko teksti julkaistaan Curia-verkkosivustolla sen julistamispäivänä.	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta. (Linguistic conventions: Minor, 1)	1	3
L38	Pressekontakt: Hartmut Ost (+352) 4303 3255	Tiedotusyhteyshenkilö: Hartmut Ost (+352) 4303 3255 *Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, jossa "Pressekontakt" käännetään "Lisätietoja". (Linguistic conventions: Minor, 1)	Lehdistökontakti: Hartmut Ost (+352) 4303 3255	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, jossa "Pressekontakt" käännetään "Lisätietoja". (Terminology, Linguistic conventions: Minor, 1)	1	3	Lehdistökontakti: Hartmut Ost (+352) 4303 3255	Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, jossa "Pressekontakt" käännetään "Lisätietoja". (Terminology, Linguistic conventions: Minor, 1)	1	3
L39	Filmaufnahmen von der Verkündung des Urteils sind verfügbar über „Europe by Satellite“...	Tuomion julistamisesta tehtyjä videotallenteita on saatavilla Europe by Satelliten kautta... *Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, joka on muotoa "Kuvia tuomion julistamisesta on saatavilla "Europe by Satellite" (+32) 2 2964106". (Linguistic conventions: Minor, 1)	Videotallenteita tuomion julistamisesta on saatavilla "Europe by Satellite" -palvelun kautta...	*Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, joka on muotoa "Kuvia tuomion julistamisesta on saatavilla "Europe by Satellite" (+32) 2 2964106". (Terminology, Linguistic conventions: Minor, 1)	1	3	Tuomion julistamisesta tehtyt filmaukset ovat saatavilla "Europe by Satellite" -palvelun kautta...	*Suomennos on virheellinen, koska se ei käytä lehdistötiedotteiden suomenkielistä standardiausetta, joka on muotoa "Kuvia tuomion julistamisesta on saatavilla "Europe by Satellite" (+32) 2 2964106". (Terminology, Linguistic conventions: Minor, 1)	1	3
L40	www.curia.europa.eu	www.curia.europa.eu	www.curia.europa.eu	Ei virheitä.	0	1	www.curia.europa.eu	Ei virheitä.	0	1
		25			231	Yhteensä: 1: 5 kpl 2: 3 kpl 3: 18 kpl 4: 12 kpl 5: 2 kpl		285		Yhteensä: 1: 3 kpl 2: 4 kpl 3: 15 kpl 4: 15 kpl 5: 3 kpl
		Minor 15, Major 2, Critical 0			Minor 96, Major 17, Critical 2			Minor 85, Major 25, Critical 3		
					Yhteensä 115			Yhteensä 113		

Liite 3. Tutkielman saksankielinen tiivistelmä – Deutsche Zusammenfassung der Abschlussarbeit

LLM-basierte maschinelle Übersetzung: Deutsch-finnische Übersetzung juristischer Texte und die Bedeutung des Post-Editings – Eine qualitative vergleichende Übersetzungsanalyse

Die vorliegende Masterarbeit untersucht die Qualität maschinell erzeugter Übersetzungen juristischer Texte vom Deutschen ins Finnische unter Einsatz moderner Large-Language-Models (LLM). Im Zentrum der Analyse stehen die Übersetzungsleistungen des Sprachmodells ChatGPT-5.2 und des auf neuronaler maschineller Übersetzung basierenden Systems DeepL Next-Gen sowie der jeweils erforderliche Post-Editing-Aufwand. Die Studie basiert auf der Übersetzung einer deutschen Pressemitteilung des Gerichtshofs der Europäischen Union (Nr. 98/19) ins Finnische und analysiert systematisch Fehler auf lexikalischer, morphosyntaktischer und funktionaler Ebene anhand eines einzelnen, nicht offiziell übersetzten juristischen Ausgangstextes.

Rechtssprachliche Texte stellen besondere Anforderungen an die Übersetzungsqualität. Juristische Terminologie ist präzise und normativ, und bereits geringfügige Ungenauigkeiten können zu erheblichen Bedeutungsverschiebungen führen. Gleichzeitig weist das Finnische als agglutinierende Sprache mit komplexer Morphologie strukturelle Besonderheiten auf, die sich erheblich von indoeuropäischen Sprachen wie dem Deutschen unterscheiden. Diese linguistischen Divergenzen zwischen Ausgangs- und Zielsprache schaffen zusätzliche Herausforderungen für maschinelle Übersetzungssysteme, die nicht allein durch große Datenmengen bewältigt werden können.

Methodisch basiert die Arbeit auf einer qualitativen Parallelkorporusanalyse. Der deutsche Ausgangstext wird mit drei finnischen Versionen verglichen: den maschinellen Übersetzungen von ChatGPT-5.2 und DeepL Next-Gen sowie einer von einem professionellen Übersetzer angefertigten Referenzübersetzung. Zur systematischen Fehleridentifikation und -klassifikation wird das MQM-Modell (Multidimensional Quality Metrics) angewandt, das eine differenzierte Analyse nach Fehlertyp und Schweregrad ermöglicht. Ergänzend wurde für diese Studie eine fünfstufige Post-Editing-Skala entwickelt, mit der das für jeden Satz erforderliche Niveau an Fachexpertise bewertet wird, um eine publikationsreife juristische Übersetzung zu erreichen.

Terminologische Fehler und Ungenauigkeiten

Die Analyse auf lexikalischer Ebene zeigt systematische terminologische Probleme in beiden Systemen. ChatGPT-5.2 wies 24 terminologische Fehler in 19 Segmenten auf, während DeepL Next-Gen 30 Fehler in 23 Segmenten verzeichnete. Während beide Systeme bei der internen terminologischen Konsistenz innerhalb

eines Textes relativ erfolgreich waren, zeigten sich erhebliche Schwächen bei der Einhaltung etablierter Fachterminologie und institutioneller Konventionen.

Ein besonders gravierendes Problem war die Verwendung des englischen Lehnworts “sampling” anstelle des etablierten finnischen Fachbegriffs “samplaaminen” tai “samplaus”. ChatGPT verwendete durchgehend “sampling”, während DeepL zwischen verschiedenen deutschen Begriffen schwankte, ohne die etablierte finnische Terminologie zu treffen.

Bei institutionellen Bezeichnungen versagten beide Systeme. EU-Richtlinien haben offizielle finnische Namen, die in der EUR-Lex-Datenbank veröffentlicht sind. ChatGPT ließ die Richtliniennamen vollständig unübersetzt und kopierte den deutschen Text, während DeepL versuchte, direkt zu übersetzen, ohne die offiziellen Bezeichnungen zu verwenden. Ähnliche Probleme zeigten sich bei der Benennung von Gerichten: DeepL übersetzte “Bundesgerichtshof” inkonsistent als “liittovaltion tuomioistuin” und “liittovaltion korkein oikeus”, beides unpräzise Verallgemeinerungen. Die korrekte Lösung, wie sie der menschliche Übersetzer wählte, lautete: “Bundesgerichtshof (liittovaltion ylin yleinen tuomioistuin, Saksa)”.

Auch standardisierte institutionelle Phrasen wurden mechanisch übersetzt. Beide Systeme übersetzten “Pressekontakt” als “Lehdistökontakti”, obwohl die finnischen Pressemitteilungen des EuGH standardmäßig die Überschrift “Lisätietoja” verwenden.

Morphosyntaktische Interferenz

Strukturelle Interferenz aus dem Deutschen erwies sich als das beherrschende morphosyntaktische Problem und war in nahezu allen analysierten Segmenten (37 von 40) bei mindestens einem der beiden Systeme feststellbar. Syntaktische Probleme waren dabei deutlich häufiger als rein morphologische Fehler.

Die charakteristische deutsche Satzklammer wurde systematisch und ohne Anpassung an die syntaktischen Konventionen des Finnischen übertragen. In Segment L12 produzierte ChatGPT die Struktur “Bundesgerichtshof [...] haluaa [...] muun muassa tietää”, wobei das Hilfsverb “haluaa” vom Hauptverb “tietää” durch zahlreiche Satzglieder getrennt wird. Im Finnischen sollten Verben unmittelbar aufeinander folgen.

Auch die Informationsstruktur wurde durch mechanische Übernahme deutscher Wortstellung verzerrt. In Segment L14 platzierte ChatGPT das Subjekt fehlerhaft am Satzende (“ovatko unionin oikeuden kanssa yhteensopivia Saksan oikeussäännökset”), während DeepL der deutschen ob-Satz-Struktur folgte. Im Finnischen sollte das Subjekt als bekannte Information (Thema) am Satzanfang nach dem Fragewort stehen.

Lange Relativsatzketten und eingebettete Nebensätze, die für das Deutsche typisch sind, wurden nahezu unverändert übernommen und führten zu schwer lesbaren Strukturen. In Segment L25 erhielt DeepL die deutsche Relativsatzkette: ”äänitallenteesta otetun äänifragmentin käyttö, josta voidaan tunnistaa teos, josta se on otettu”. Solche Verkettungen sollten im Finnischen durch Partizipialkonstruktionen oder separate Sätze aufgelöst werden.

Morphologische Fehler waren seltener, hatten aber teilweise schwerwiegende juristische Konsequenzen. Am kritischsten war der Modusfehler in Segment L10: Beide Systeme produzierten den Indikativ "on kopioinut" (hat kopiert), obwohl der Kontext den Konditional "olisi kopioinut" (hätte kopiert) erforderte. Das deutsche "habe kopiert" (Konjunktiv I) drückt indirekte Rede aus, die im Finnischen typischerweise durch den Konditional markiert wird. Die Verwendung des Indikativs suggeriert, dass das Kopieren eine feststehende Tatsache ist, während es sich lediglich um eine noch zu prüfende Behauptung einer Partei handelt – ein juristisch relevanter Unterschied.

Weitere morphologische Probleme betrafen Kasusfehler, insbesondere bei der Übertragung deutscher Präpositionalstrukturen in finnische Kasusformen. In Segment L33 verwendeten beide Systeme den Adessiv ("toisella äänitallenteella"), wo der Inessiv ("toisessa äänitallenteessa") idiomatisch korrekt gewesen wäre.

Funktionale Ebene: Register und Rechtswirkung

Auf funktionaler Ebene zeigten sich Defizite in der Einhaltung juristischer Registerkonventionen und institutioneller Kommunikationsnormen. Die maschinellen Übersetzungssysteme litten unter Registerinstabilität mit einer Tendenz zur Umgangssprache.

In den Segmenten L16 und L17 verwendete ChatGPT das Verb "muistuttaa" (erinnern) als Äquivalent für das deutsche "hinweisen", was in einem Kontext, in dem der EuGH auf gefestigte Rechtsprechung verweist, zu informell ist. In den Segmenten L12 und L14 verwendete ChatGPT die Wendung "haluaa tietää" (will wissen) für ein Vorabentscheidungsersuchen, was stilistisch zu umgangssprachlich ist; etablierte Formulierungen wie "tiedustelee" oder "pyytää selvitystä" wahren die formale Amtlichkeit.

Kulturgebundene Höflichkeitskonventionen bereiteten ChatGPT Probleme. In den Segmenten L9-L11 wiederholte das System unnötig "herra" (Herr) vor Namen, obwohl im finnischen Amtsstil die bloße Nennung der Namen üblich ist. Die Titelwiederholung erzeugt einen Eindruck übertriebener Förmlichkeit.

Quantitative Ergebnisse und Post-Editing-Bedarf

Die Studie belegt, dass nur ein geringer Anteil der maschinell übersetzten Segmente ohne Nachbearbeitung publikationsreif ist (ChatGPT 12,5 %, DeepL 7,5 %). In der quantitativen Bewertung erzielte ChatGPT-5.2

eine Gesamtpunktzahl von 231 MQM-Punkten, während DeepL Next-Gen mit 285 Punkten bewertet wurde (wobei eine höhere Punktzahl eine geringere Qualität widerspiegelt). Obwohl DeepL mit insgesamt 113 Fehlern geringfügig weniger Fehler aufwies als ChatGPT (115 Fehler), lag der Anteil schwerer (*major*) und kritischer (*critical*) Fehler bei DeepL höher, was das schlechtere Gesamtergebnis erklärt. Die überwiegende Mehrheit der Übersetzungen erforderte mindestens Korrekturen der Stufe 3 oder 4 der entwickelten Skala, was substantielle fachliche Eingriffe bedeutet.

Der zentrale Unterschied zwischen den beiden Systemen liegt nicht im Gesamtumfang des Post-Editing-Bedarfs, sondern in dessen Art: ChatGPT erfordert vom Post-Editor vor allem Fachkompetenz zur Identifikation und Korrektur terminologischer Ungenauigkeiten, Modalfehler und Abweichungen vom juristischen Register. Die Texte sind sprachlich flüssiger, weisen jedoch 24 terminologische Fehler auf und erfordern substantielle Korrekturen der fachlichen Inhalte. DeepL hingegen ist terminologisch präziser (trotz 30 Fehlern in mehr Segmenten handelt es sich oft um weniger schwerwiegende Abweichungen), verlangt aber primär sprachliche Kompetenz zur Auflösung strukturell ungelinker Sätze und zur Beseitigung deutscher Interferenzerscheinungen.

Schlussfolgerungen

Die Untersuchung unterstreicht, dass die juristische und sprachliche Fachkompetenz professioneller Übersetzer nach wie vor unverzichtbar für die Qualitätssicherung maschinell erstellter juristischer Übersetzungen ist. Post-Editing verbessert signifikant die Lesbarkeit, Zielgruppenadäquatheit und juristische Präzision der Texte. Die Ergebnisse deuten darauf hin, dass sich die Rolle des Übersetzers zunehmend von der primären Textproduktion hin zur strategischen Qualitätssicherung und zum sprachlichen Risikomanagement verlagert.

Die Studie zeigt, dass oberflächlich flüssig wirkende maschinelle Übersetzungen subtile Fehler enthalten können, die die Rechtswirkung des Textes beeinträchtigen. Der professionelle Übersetzer muss über die grammatikalisch korrekte Oberfläche hinausblicken und jene Stellen identifizieren können, an denen die juristische Logik fehlerhaft ist. Dies erfordert den Übergang vom reinen Post-Editing zur strategischen Qualitätssicherung, bei der der Übersetzer die letztendliche Verantwortung für die juristische Integrität des Textes trägt.

Die Ergebnisse haben auch wichtige Implikationen für die Übersetzerausbildung. Das für Fachübersetzungen erforderliche Spezialwissen wird durch die technologische Entwicklung nicht obsolet, sondern zur unverzichtbaren Voraussetzung für die Qualitätssicherung KI-gestützter Übersetzungen. Übersetzer müssen die Unterschiede zwischen Rechtssystemen und Rechtssprachen sowie die juristische Terminologie beherrschen, um als bewertende Fachexperten für maschinelle Ausgaben fungieren zu können.

Für das Finnische als kleine Sprachgemeinschaft mit strukturell großer Distanz zu den Hauptsprachfamilien besteht zudem das Risiko, dass strukturelle Interferenz aus dem Deutschen oder Englischen den finnischen juristischen Sprachgebrauch prägen könnte, wenn maschinelle Übersetzungen ohne ausreichende Kritik verwendet werden. Die Arbeit empfiehlt daher nachdrücklich, dass maschinelle Übersetzungssysteme im juristischen Bereich ausschließlich in Verbindung mit professionellem Post-Editing durch qualifizierte Fachübersetzer eingesetzt werden sollten. Für die Zukunft ist es entscheidend, ein Gleichgewicht zwischen technologischer Effizienz und menschlicher Expertise zu finden, damit die finnische Rechtssprache auch im Zeitalter der künstlichen Intelligenz präzise und eigenständig bleibt.