

Retorisia siirtoja ja metodologisia ongelmia crowdsourcing-lokalisoinnin tutkimuksessa

Tommi Nieminen
Itä-Suomen yliopisto

Abstract

Translating is often thought of as transfer of meanings from one linguistic-cultural sphere to another. Often, however, in order for the transfer to succeed the translator needs to revise the target system itself. This is particularly true in software localization. When the concepts in the text are new even for the source culture, translators become the very tools with which the target culture expands and changes. The route to this are the ever-changing norms of language. This article presents some preliminary findings from a study of the Finnish localization team of the KDE SC desktop environment. I will demonstrate the kinds of rhetorical 'moves' localizers use when tackling with the linguistic norms and also briefly consider the methodological problems one encounters in investigating this 'normative work'.

Keywords: localization, free and open source software, linguistic norms, emergence of norms, languaging

Avainsanat: lokalisointi, vapaat ohjelmistot, kielen normit, normien emergenssi, kielentäminen

1 Kielen normit crowdsourcing-kääntämisessä

Kääntäminen on joskus enemmän kuin kääntämistä. Annetun semioottisen muodon siirtäminen uutta yleisöä varten uudeksi semioottiseksi muodoksi ei aina suoraan onnistu, koska muuntamisen resurssit eli tarvittavat kielenaineokset puuttuvat kohdekielestä. Tässä artikkelissa tarkastelen, miten crowdsourcing-tyyppisesti toteutettuun lokalisointiin osallistuvat kääntäjät suhtautuvat kielen eri tasoisiin normeihin ja niiden uudentamiseen. Pohdin sitä, missä määrin lokalisoijat hyväksyvät valmiit, yleisesti hyväksytyihin sanakirjoihin ja kielenhuollon oppaisiin kirjatut kielen deskriptiiviset ja preskriptiiviset normit, missä määrin taas tekevät *normatiivista työtä* arvottaessaan, uudentaessaan tai luodessaan uusia tai päivittyviä normeja. Lokalisointiaineiston avulla on mahdollista osoittaa, että kääntäjä toimii pikemmin aktiivisena kielen ylläpitäjänä ja luojaana – ”kielentäjänä” (Cowley 2005) – kuin kahden valmiin systeemin välisenä välittäjänä, kuten kääntäminen sanana antaisi ymmärtää.

Crowdsourcing (Howe 2006) on uudehko käsite, joka on syntynyt internetin sosiaalisemman ja vuorovaikutteisemmän, vuonna 2004 Web 2.0:ksi käsitteistetyn (Collis & Moonen 2008: 94) toisen vaiheen myötä. Tavallisesti sillä tarkoitetaan tietokoneohjelmistojen lähdekoodin tai sen määräosien kuten lokalisointitiedostojen avaamista laajan yleisön vapaaehtoistyölle, mutta se voi myös viitata laajemmin ”virtuaalisesti yhteiseen” sisällöntuotantoon (engl. *virtual co-creation*, ks. esim. Harwood & Garry 2010), joka tyypillisesti muttei yksinomaan on vapaaehtoisluontoista. Kun esimerkiksi

Nogueira ja Semolini (2010) esittelevät crowdsourcingin käsitettä *Translation Journal*-verkkolehdessä, he lähtevät luontevasti liikkeelle Wikipedian sana-artikkelista, sillä wikit ovat osa samaa sosiaalisten medioiden todellisuutta kuin crowdsourcingin. Ikävimmillään crowdsourcing on osa suuryritysten ulkoistamisstrategiaa, jossa suuritöiset ja potentiaalisesti kalliit työvaiheet kuten ohjelmistojen lokalisoiminen pienille kielille tai pieniin kulttuureihin tuotetaan ilmaiseksi vapaaehtoisilla, mutta parhaimmillaan se muodostaa kiinteän osan Web 2.0:n myötä syntyneestä uudesta, avoimesta sisällön tuottamisen kulttuurista, jossa raja tuottajan ja kuluttajan tai käyttäjän välillä hämärtyy (Collis & Moonen 2008: 94–96; Bell & Loane 2010: 218–221; Karasavvidis 2010: 220; Lievrouw 2010: 220–221) ja jopa sisällön omistussuhteet saattavat sumeta (Harwood & Garry 2010: 296–297; Karasavvidis 2010: 222–223). Parhaimmillaan crowdsourcingin voi jopa väittää mahdollistaneen pienten ja keskisuurten sisällöntuottajien pääsyn kansainvälisille markkinoille, missä mukautuminen kulloisenkin yleisön tarpeisiin on elinehto (Bell & Loane 2010: 223–224).

Tarkasteluni kohteena on kokonaisuudessaan vapaaehtoisvoimin toteutettava KDE Software Compilation eli KDE SC (<http://kde.org/>). Se on vapaa (engl. *free/libre*), ilmainen, usealle laitealustalle saatava graafinen työpöytäympäristö ja ohjelmistokokonaisuus. KDE oli alkuaan lyhenne sanoista ”K Desktop Environment”, ja se luotiin avoimen lähdekoodin vaihtoehdoksi kaupalliselle, mm. Hewlett Packardin, IBM:n, Novellin ja Sunin yhdessä kehittämälle CDE:lle (Common Desktop Environment, <http://www.opengroup.org/cde/>). Nykyisen tuotteistuksen mukaan KDE on ohjelmistokokonaisuutta kehittävän sisällöntuottajaryhmän nimi eikä lyhenne mistään. Itse tuote, KDE SC, tarjoaa paitsi työpöytäympäristön myös laajan hyöty- ja viihdeohjelma-kokonaisuuden, jota ajetaan jossakin ikkunointijärjestelmässä, tyypillisimmin Linuxin X11-ikkunoinnissa, vaikka valtaosa ohjelmista toimii myös esim. Microsoft Windowsissa. Aineistoni koostuu sen suomalaisen lokalisointiryhmän työstä, jossa olen itse ollut mukana muutaman vuoden. Aineistona ovat erilaiset kääntäjien keskinäisen viestinnän muodot, joista tässä artikkelissa käytän ainoastaan .po- eli käännöstiedostojen kommenttijonoja (ks. lukua 2).

Tutkimuskiinnostukseni kohteena on yleisimmällä tasolla se, miten kieli koetaan normatiivisesti oikeaksi ja miten tämä sosiaalisten, konventionaalisten normien emergenssi (Määttä 1999, 2000, MacWhinney 2001 ja yleisemmällä tasolla Bendor & Swistak 2001) tapahtuu. KDE SC edustaa uutta, nopeasti kehittyvää käyttöliittymätekniikkaa, jossa syntyy jatkuvasti uusia käsitteitä, joille ei ole vakiintunutta kielellistä ilmaisua välttämättä edes käännösten lähdekielessä englannissa, saati kohdekielessä suomessa. Lokalisoinnin tavoitteena on kuitenkin toimiva, asiaan vihkiytymättömänkin käyttäjän käyttöön sopiva järjestelmä, jossa käyttöliittymän käsitteet kielennöksineen toimivat johonkin muuhun tähtäävän toiminnan ongelmattomina apuvälineinä. Lokalisoijien ongelmaksi muodostuu ristiriita toisaalta tarpeesta uudistaa kieltä uutta käsitteistöä myötäillen, toisaalta pyrkimyksestä olla nostamatta kieliasua käyttäjän huomiokynnyksen yli ohjelmiston käytön hidasteeksi. Kieli toimii graafisissa käyttöliittymissä parhaiten silloin, kun se itsessään on mahdollisimman näkymätöntä.

Aktiivisia suomalaisia KDE SC -lokalisoijia on vähän, alle kymmenen; tarkkaa määrää on mahdotonta sanoa, koska lokalisointiin voi osallistua satunnaisesti, esim. yksittäisen korjauksen verran, ja monet aktiivisimmistakin lokalisoijista saattavat pitää osallistumisessaan pitkiä taukoja. Crowdsourcing-lokalisointi on harrastus, joka ei sido. Asian voi ilmaista myös niin, että vaikka varsin moni käyttää suomenkielistä KDE SC –ympäristöä, harva antaa käännöksestä palautetta ja vielä harvempi tulee itse osallistuneeksi lokalisointiin. Ääripään harvenelevassa jatkumossa muodostavat ne pari ihmistä, joilla on oikeus julkaista käännöksiä eli lähettää ne yleisesti saataville. Joukkovoimaa hyödyntävä crowdsourcing on käytännön tasolla pitkälti yksityistä työtä ja siinä käytetty valta meritokraattista, työn mukana ansaittua ja muiden osallistujien arvioimaa. Lokalisoijien käytännön päätökset ovat heidän omiaan, eikä ratkaisuista yleensä neuvotella muiden lokalisoijien kanssa etukäteen – laajemmasta yleisöstä puhumattakaan.

Tutkimuskysymyksiksi nousevat: Miten kielen(käytön) normit lokalisoitityössä näkyvät? Miten ja millä ehdoin niihin viitataan? Miten normit syntyvät? Kysymyksiä voi edelleen tämentää. Miten metalingvistinen neuvottelu – joka käytännössä ei edes ole ”neuvottelua” sanan varsinaisessa mielessä vaan jonkinlainen julkisen diskurssin piiloprosessi – toteutuu käytännössä? Miten kieli vakiintuu ja miten vakiintuneita normeja voidaan muuttaa? Edelleen, kun käännetään, minkä kaikkien tekijöiden välille syntyy suhteita ja miten ne aktualisoidaan? Millaiset tekijät vaikuttavat diskurssiyhteisön valtasuhteisiin ja miten valtaa signaloidaan?

2 Kääntäjän karu käytäntö

Crowdsourcing-käännöstyöhön ryhdytään tyypillisesti sen kautta, että itse käytetään jotakin vapaata ohjelmistoa, jonka lokalisoinnissa havaitaan puutteita. Ohjelmiston dokumentaatiosta (esim. ohjelman Ohje-valikosta) tai web-haulla löydetään tieto oman kielen lokalisoitiryhmästä, johon otetaan yhteyttä, ja käännöstiedostot noudetaan itse versionhallintaohjelmistolla verkosta; jälkimmäiset vaiheet voivat tapahtua yhtä hyvin käänteisessä järjestyksessä. Lokalisoitiryhmään liittyminen on tyypillisesti ilmoitusasia, jolla lähinnä tiedotetaan siitä, että on halukas osallistumaan työhön. Ryhmään kuulumisesta ei sinänsä seuraa erikoisoikeuksia tai erityisvelvoitteita.

Ns. **alavirtaan**, ohjelman tekijöiltä käyttävään yhteisöön päin, vapaan ohjelmiston käännöstiedostot ovat yleensä avoimesti jaossa eikä niihin pääsyyn tarvita erikoisoikeuksia. **Ylävirtaan** eli yksittäisiltä tekijöiltä ohjelman jakelukanaviin päin välitysoikeudet sen sijaan rajataan ”portinvartioihin” eli esimerkiksi lokalisoitiryhmien ylläpitäjiin ja erityisen luotettuihin vanhempiin jäseniin. Sikäli kuin käännösten laatua valvotaan, se tapahtuu tässä vaiheessa. Yksittäinen lokalisoija joutuu näin tekemisiin käännösryhmän ja mahdollisesti laajemman yhteisön kanssa lähinnä silloin, kun lähettää käännöksiään ylävirtaan.

Käännöstiedostotkaan eivät ole kovin avulias lähde aloittavan lokalisoijan työlle. Vaikka lokalisoititiedostojen muotoja on monia, vapaissa ohjelmistoissa tavallisin ja myös KDE SC -projektin käyttämä on GNU Gettext, jossa käännetään .pot-tiedostojen seuraavan kaltaisia merkkijonoja:

```
msgid "String to be translated"
```

.pot-tiedoston – missä ”t” viittaa englannin sanaan *template*, aihio tai mallipohja – jokainen jono on samalla sekä ohjelmakoodin tietyn sijainnin tunniste että lähdekielinen käännettävä jono. KDE SC:ssä kuten useimmissa muissakin vapaan ohjelmiston projekteissa lähdekieli on aina englanti.

Lähtökohtana olevista .pot-tiedostoista lokalisoijan tehtävänä on tuottaa .po-tiedostoja, joissa jokaista lähdekielen jonoa vastaa kohdekielen jono avain–arvo-pareittain:

```
msgid "String to be translated"  
msgstr "Käännettävä jono"
```

Koska lähdekielen jonot toimivat myös ohjelmakoodin tunnisteinä (*msgid*, *message ID*), niiden tulee olla ainutkertaisia. Jos jotakin merkijonoa tarvitaan useampia kertoja eri konteksteissa, se on ainutkertaistettava, koska *msgid* ei kerro vain kääntäjälle, millainen ilmaus täytyy kääntää, se myös kertoo *ohjelmakoodille*, mikä nimenomainen jono kaikista mahdollisista on kyseessä.

Ainutkertaistaminen on mahdollista kahdella tapaa. Vanhempi tapa oli käyttää &-merkkiä, jonka varsinainen tehtävä *msgid*:ssä on osoittaa, mitä pikanäppäinyhdistelmää (Alt + &-merkkiä seuraava kirjain) käyttäen kohdistus siirtyy kyseiseen käyttöliittymä-elementtiin. Tämä on kömpelöä ja epäselvää, joten ratkaisuksi kehitettiin valinnainen kolmas tunniste *msgtxt*, jonka tehtävä on ainutkertaistaa tunniste aiheuttamatta sivuvaihtuksia:

```
msgtxt "Something that together with msgid is unique in the file"  
msgid "Some string"
```

Ainutkertaisena tunnisteenä toimivat nyt *msgtxt*- ja *msgid*-jonot yhdessä. Tämä on *msgtxt*:n päätehtävä, jonka ohessa se voi esittää lukijalle selittävän kommentin (mistä nimi *msgtxt*, *message comment text*). Lokalisoijan kannalta on toivottavaa – ja onneksi vakiintuneehko käytäntö – että tunnistetta käytetään myös välittämään tiedostoa lukevalle ihmiselle mielekästä tietoa *msgid*:n käyttöyhteydestä. Varsinaista todellista tekstikontekstia ei .po-tiedostossa silti milloinkaan näy.

Lokalisoijan kannalta *msgtxt* voi siis olla kontekstia mahdollisesti avaava selitys, mutta tämä ei ole sen päätarkoitus eikä sitä voi lisätä .po-tiedostoon kuin ohjelman tekijä. Omia huomioitaan ja työmerkintöjään lokalisoija voi lisätä tiedostoon #-merkin avulla:

```
# Tämä on kääntäjän kommentti
```

Gettext käyttää myös tämän merkintätavan muunnelmia omiin tarkoituksiinsa, mutta kun #-merkkiä välittömästi seuraa välilyönti, se aina aloittaa lokalisoijan itsensä tai muiden luettavaksi tarkoittaman kommentin, joka tällaisena myös esitetään käännös-ohjelmissa kuten KDE SC:n virallisessa käännöstyökalussa Lokalizessa. Nämä jonot muodostavat käsillä olevan aineiston.

Lokalisointityön kannalta .po-tiedosto on karu lähde ja crowdsourcing välinpitämätön työnantaja. Käännösjonot eivät esimerkiksi ole määräjärjestyksessä: peräkkäiset jonot eivät välttämättä esiinny käyttökonteksteissaan lähekkäin. Käyttöä valaisevat msgtxt-jonot ovat suhteellisen vähälukuisia. Ellei ohjelman käännettävää, usein vasta kehityksessä olevaa ja epävakaa versiota käytä itse, käännettävien jonojen toimivuuden arviointi jää myöhemmän palautteen ja muun lokalisointiyhteisön niukan avun varaan. Tavallista on, ettei varsinaista palautetta saa lainkaan, vaan huonoiksi tai virheellisiksi käytössä koetut käännösjonot yksinkertaisesti myöhemmässä vaiheessa korvataan hiljaisesti toisilla.

Mielenkiintoista onkin nyt tarkastella sitä, miksi ja miten lokalisointiyhteisö puuttuu eksplisiittisesti kielen normeihin. Millaiset ongelmat koetaan riittäviksi tarttua .po-tiedoston kommenttien muodossa yhteisön lähes automatisoituun kielelliseen toimintaan? Miten puuttumista motivoidaan ja miten se ilmenee?

3 Lokalisointityön retorisia siirtoja

Lähdin alustavasti luokittelemaan lokalisoiden keskinäisiä kommentteja sen mukaan, millaista toimintaa kommentti tuntuu asettuvan odottamaan. Semioottisesti ajatellen etsin siis kunkin kommentin **energeettistä tulkitsinta** (Peirce, *CP*: 5.475, *EP*: 2.409, 2.418): kommentin yhteisöllisesti riittävä tulkinta on jokin toiminta, jonka odotettu muoto kommentista toiseen vaihtelee. Vuorovaikutus on prosessi, jossa kaksi semioosin osapuolta, merkki ja tulkitsin, etualaistetaan yhdessä (Kockelman 2007: 377). Vuorovaikutuksen kannalta kommentin voi sanoa olevan vierusparin aloittava eli etujäsen. Koska tarkastelemani kommentit odottavat laajasti samanmuotoista vastausta, kutsun niitä tässä **retorisiksi siirroiksi** (vrt. Leiwo ym. 1992: 67–74). ”Siirron” voi ymmärtää pelin siirron tapaan: se tapahtuu määräsuntaan ja siihen on vastattava määrätavalla.

Aineistoni rajoittuu tässä lokalisoiden .po-tiedostoihin lisäämiin kommenttiriveihin, jotka ovat julkista materiaalia ja vapaasti luettavissa. Tuonnempana pyrin laajentamaan käsittelyä muihin, sulkeisempiin merkitystyön muotoihin kuten lokalisoiden sähköpostikirjeenvaihtoon, jonka retorinen toiminta mahtuu yleensä samoihin luokkiin kuin kommenttirivienkin mutta joka ei ole samalla tavoin julkista ja jää tavallisesti yksin lokalisoiden tietoon, ainakin sikäli kuin he eivät itse käytä kirjeenvaihtoaan perusteluna käännöspäätöksilleen jollakin julkisella kanavalla.

Lokalisointikommenttien retoriset siirrot muistuttavat mahdollisilta seuraamuksiltaan läheisesti toisiaan ja siten luokkina lomittuvat, mikä toisaalta pätee mihin tahansa viestinnän muotoon. Vaikka esimerkiksi (tyypillinen) kysymys asettuu odottamaan vastausta, liian vaikea, tilanteeseen sopimaton tai muuten ongelmallinen kysymys voidaan sivuuttaa tai siihen vastaaminen siirtää tuonnemmas. Etenkin kun kyse on kirjallisesta viestinnästä, jossa viestinnän osapuolet eivät ole välittömästi tekemisissä keskenään, viestien välinen odotusjakso sisältyy implisiittisesti jokaiseen viestiin. Odotusjakso ei toisaalta aseta ehtoja sille, kuinka pitkä se saa olla: kirjoitettu, vastausta odottavaksi tulkittu kysymys voi odottaa vastaustaan periaatteessa loputtoman kauan.

Alustavasti päädyin luokittamaan siirrot neljiä. Nimitän luokkia *kutsuksi, valinnaksi, perusteluksi ja selitykseksi*. Lyhyesti ne voi kuvata seuraavasti:

- kutsu:** Kutsutaan yhteisöä ratkaisemaan jokin ongelma (vaikka malliratkaisukin voidaan esittää päätöksenteon pohjaksi).
- valinta:** Tehdään valinta eri vaihtoehdoista mutta jätetään auki eriävän mielipiteen mahdollisuus.
- perustelu:** Tehdään valinta ja perustellaan se.
- selitys:** Annetaan selitys muuten vaikeasti avautuvalle valinnalle.

Annan seuraavissa alaluvuissa esimerkin tai pari kustakin retorisesta siirrosta ja pohdin samalla sitä, millainen on kunkin suhde kielen normeihin: mitä normeja nostetaan esiin, millainen suhde normeihin etualaistetaan ja mitkä mahdolliset julkiset motiivit johtavat juuri tiettyyn siirtoon.

3.1 Kutsu

- lähdejono:** Favicons
käännös: Webbisivukuvat
kommentti: Webbisivukuvake, verkko-osoite-kuvake, pikavalintakuvake, kirjanmerkkikuvake

Kutsu on lokalisointikommenttien retorisista siirroista yksinkertaisin ja perustavin. Se luettelee käännösvaihtoehtoja ottamatta niihin kantaa ja samalla kutsuu yhteisöä tekemään esille nostetusta asiasta päätöksen. Toisaalta koska koko crowdsourcing-konsepti perustuu sille, että kääntäminen mahdollistetaan lähdekoodin mukana julkisesti jaettavien .pot-tiedostojen avulla, kutsua voi pitää lähinnä emfaattisena eleenä: sillä vain nostetaan ongelma paremmin esille.

Kuten annetussa esimerkissä, kutsu liittyy yleisesti tilanteeseen, jossa lähdekielessä tietty käsite on termiytynyt mutta kohdekielestä vakiintunut vastine puuttuu. Kutsu suuntautuu siis tilanteesta normiin eli on normihakuinen. Haettu normi on tasoltaan mahdollisimman yleinen. Kutsussa esitetyt vaihtoehdot saattavat ottaa kantaa erilaisiin ala- tai osanormeihin kuten lokalisoinnin konventioon ("näin on tehty tähän saakka") tai jossakin toisaalla vallitsevaan mutta toistaiseksi yleisempää hyväksyntää vailla olevaa normiin ("GNOMEssa tämä käännetään näin"), mutta kutsun päämääränä on nimenomaisesti valita ja nostaa jokin muodoste mahdollisimman yleiseen käyttöön. Kutsun perimmäinen tavoite on yleiskielen tai vähintäänkin ATK-jargonin normi.

Ääritapauksessa kutsu ei esitä vaihtoehtoja lainkaan vaan ainoastaan korostaa ilmauksen kääntämättömyyttä. Nollavaihtoehtona on aina lähdekielestä otetun sitaattilainan käyttö sellaisenaan.

3.2 Valinta

- lähdejono:** Suffixes
käännös: Loppuliitteet
kommentti: Tässä kohdassa ”Päätteet” olisi oikea käännös, mutta koska päätte viittaa tietoliikenteessä vähän muuallekin ja saattaa olla tässä asiayhteydessä harhaanjohtava, niin muutin tämän ”Loppuliitteet” -sanaksi.

Valinta muistuttaa kutsua siinä, että se luettelee eri käännösvaihtoehtoja, mutta toisaalta eroaa siitä siinä, että lokalisoiija tekee valintansa ja perustelee sen. Yhteisölle osoitetun kutsun voi kuitenkin yhä väittää jäävän auki, mitä tulkintaa puoltaa se, että kommentti jätetään paikalleen merkiksi tehdystä valinnasta. Lokalisointityön kannalta valinta seuraa (implisiittistä tai eksplisiittistä) kutsua mutta edeltää lopullisesti vakiintunutta normia.

Annetussa esimerkissä kommentin retorinen toiminta lähtee liikkeelle yleiskielen normista (”olisi oikea käännös”), siirtyy siitä jargonnormiin huomauttaen samalla yleiskielen normin olevan riittämätön tilanteen ratkaisemiseksi (”viittaa tietoliikenteessä vähän muuallekin”) ja kohdentaa huomion lopuksi nimenomaiseen tekstikontekstiin (”tässä asiayhteydessä”). Tekstikonteksti, jossa viestinnällinen merkityksellisyys ratkeaa, osoittautuu normien suhteen ongelmalliseksi. Koska jargonnormi implikoidaan yleiskielen normia relevantimmaksi, kääntämisen keskeiseksi kontekstiksi nousee ”rel-text”, ”related documents and other resources” (Melby & Foster 2010: 6).

Valinnassa voi huomata lähde- ja kohdeilmausten olevan molemmin puolin vähintään jossakin määrin termiäilyneitä ilmauksia. Normit ovat siis olemassa ja vakiintuneita eikä niitä suoranaisesti kiistetä. Lokalisoiija ei toisin sanoen väitä normia ongelmalliseksi tai että sitä täytyisi muuttaa. Toisin kuin kutsussa, kyse ei ole termiäilymisestä eikä normin emergenssistä vaan tilanteisesta viestinnällisyydestä, joka asettaa normit kilpasille. Lopullinen normi olisi ratkaisu normien väliseen kilpailuun, mutta valinnan siirrossa tähän ei päästä.

Tutkijan kannalta valinta on erityisen mielenkiintoinen siirto, koska siinä normien monitasoisuus parhaiten paljastuu.

3.3 Perustelu

- lähdejono:** KDE logout was canceled
käännös: KDE:n uloskirjautuminen kumottiin
kommentti: Tästä ei saa oikein oikeinkielisyysoppaissaakaan neuvoa. Omassa kielenkäytössäni autoa peruutetaan ja tietokoneohjelmistojen toimintoja perutaan. Parempi käyttää sanaa ”kumota”

Perustelu muistuttaa valintaa muttei liiku normien ja tilanteen välillä vaan kohdistuu normiin ja asettaa sen kyseenalaiseksi. Esimerkissä käsitellyn englannin *cancel*-verbin

käyttöliittymäkäytön jargonnormi ei ole täysin vakiintunut: verbi voidaan suomentaa joko *perua* tai *peruuttaa*, kuten kommentoija huomauttaa. Jargonnormia tarkempi on käyttöliittymäperinteen normi. Microsoft Windows -käyttöjärjestelmään sekä tässä kohteenani olevaan vapaaseen KDE SC -työpöytäympäristöön on vakiintunut vastineeksi *peruuttaa* (tyypillisesti yksikön 2. persoonan imperatiivissa *Peruuta*), toiseen vapaaseen työpöytäympäristöön GNOMEen puolestaan *perua* (*Peru*). Jargonnormeista poikkeavaa käännöstä *kumota* perustellaan tässä yleiskielen normilla kielioppaiden avulla, jolloin käytännössä kyseenalaistetaan termiytyymisen normi.

Perustelu on väline, jolla lokalisoija kykenee avaamaan vakiintuneenkin normin uudennettavaksi. Kielen vakiintuneisuutta voi sinänsä pitää lokalisoinnin normina: loppukäyttäjän kannalta paras ja tehokkain on ilmaus, joka ei vaadi erillistä selitystä vaan toistuu kontekstista toiseen automaattisen oloisesti. Lokalisoinnin päämääränä on tuottaa kieltä, joka ei kiinnitä huomiota itseensä vaan on mahdollisimman läpinäkyvä väline johonkin muuhun tähtäävässä toiminnassa. Jo vakiintuneen termin muuttaminen edellyttää siis erityisen vahvaa perustelua, minkä takia perustelut ovat tyypillisesti pitkiä ja yksityiskohtaisia. Erityisen hyvin tämä ilmenee seuraavasta (kommentista on ilmeisimmät kirjoitusvirheet stilisoitu pois):

lähdejono: About Bug Reporting – Help
käännös: Vian ilmoittamisesta – Opaste
kommentti: Suomentamisessa sanavalinnalle voi olla monia perusteita. Esi-merkiksi Help-sanan suomennos ”Ohje” koetaan säännöksen tai määräyksen kaltaiseksi ehkä negatiivisia mielleyhtymiä herättäväksi sanaksi ja siksi ”Opaste” on vakiintumassa suomennokseksi. [...]

Tämä oli erityisen pitkä kommentti, joka jatkui lukuisilla lisäesimerkeillä. Se perustelee vetoamalla vakiintuvaksi esitettyyn mutta voimassaolevasta poikkeavaan normiin. *Ohje* on hyvin vahvasti vakiintunut KDE SC -järjestelmien valikoihin, ja melkeinpä tyypillisin aloittelevalta lokalisoijalta korjattava virhe on muuttaa englannin *Help*-sanan muut mahdolliset käännösvastineet muotoon *Ohje*. Tällaisen jo lokalisoinnin käytännön toiminnan tasolle juurtuneen normin muuttaminen selittänee riittävästi tarpeen liki esseemäiseen perusteluun.

Perustelu on normin muuttamisen väline, joka voi sijoittua mille vain normitasolle. Yleisimmin lokalisoija ei pyrkine muuttamaan yleiskielen eikä ehkä jargoninkaan normia vaan ainoastaan osakulttuurin tai ”ekosfäärin” (tässä KDE SC:n) alanormia, mutta koska normit eivät ole toisistaan irrallisia, alanormin muuttaminen luo aina mahdollisuuden myös yleisemmän normin muuttumiselle (vrt. Nieminen 1998: 137).

3.4 Selitys

- lähdejono:** Possessive
käännös: Kuukaudennimen obliikvimuoto
kommentti: ”Possessive” on ilmeisesti englanninkielisen kuvitelma siitä muodosta, joka kuukaudennimestä tulee päivämäärään, jollei se ole nominatiivi (se on genetiivi esim. venäjässä)

Selitys edustaa aivan toisen tyyppistä retorista siirtoa kuin edellä mainitut kutsu, valinta ja perustelu. Kutsun tavoin se on silti yksinkertainen ja perustava siirto: selityksessä lokalisoija nimittäin perustelee toimintansa näennäistä mielivaltaisuutta eli sitä, miksi msgid- ja msgstr-jonot eivät pikaisella silmäyksellä näy kuuluvan yhteen.

Esimerkin kommentti – joka valitettavasti on tämän artikkelin kirjoittajan oma – on tarkoitettu opastamaan muita lokalisoijia ymmärtämään poikkeuksellista ilmaisutapaa. Sen voi sinänsä tulkita myös kutsuksi parempaan käännösvastineeseen. Selityksen tapauksessa termityminen ei liene näköpiirissä kummassakaan kielessä eikä kyse ole kohdekielen normeista vaan pikemmin tilanteesta, jossa lähdekielisessä jonossa pikemmin kuvaillaan jotakin käsitettä kuin varsinaisesti nimitään se.

Normien kannalta selitys ei siis ole kovin kiinnostava retorinen ele. Sen sijaan kiinnostavaa on, että kaikki toistaiseksi tarkastellut kommentit on se mukaan lukien mahdollista pakottamattomasti sijoittaa näihin neljään siirron luokkaan. Koska normit nousevat tavalla tai toisella esiin kolmessa tapauksessa neljästä, .po-tiedostojen kommenttijonoilla on selvästikin kiinteähkö yhteys lokalisoinnin normatiiviseen työhön.

4 Normit nousevat näkyviin

Lokalisointityössä voi elävästi havaita kielen normien moninaisuuden. Koska crowdsourcing-lokalisoijat eivät ole alan ammattilaisia eivätkä välttämättä aloittaessaan edes harrastajia vaan pelkästään aktiivisia (tai aktiivistuvia) käyttäjiä, monen lähtökohtana käännöstyötä aloittaessaan on yksinkertaisesti arkinen kirjoitusviestinnän normi: ”näin minä sen sanoisin”. Kosketuspinta on ensi sijassa omaan normistoon, jonka avulla arvioi julkista kielenkäyttöä. Tilanteista viestinnällisyyttä painottavat argumentit painavat tässä vaiheessa perinteistä kirjoitetun kielen preskriptiota enemmän.

Innon ja huolellisuuden työn myötä kasvaessa tulee seuraavaksi yleensä vastaan jo koulusta tuttu yleiskielen preskriptiivinen normi, joka helposti nousee semmin liiaksikin esiin lokalisoijan kommentteissa. Suomalaiset on kouluopetuksessa varsin pitkälti ehdollistettu ajattelemaan, että kielessä olisi olemassa jokin ehdoton, absoluuttisesti ”tosi” korrektiuden normi; esimerkiksi aloittelevia suomen kielen yliopisto-opiskelijoita opettaessa opettajan ensimmäinen tehtävä on yleensä murtaa tämä oletus ja osoittaa, että koulun (ja yliopiston valintakokeiden) aikana vaivalla opitut preskriptiiviset säännöt ovat aina jonkun päättämiä ja laajemmasta kielen ”luontaisten” normien joukosta valittuja.

Kielessä ja kielenkäytössä ei olekaan yhtä normia vaan normien moninaisuus. Aineistoissani siitä retorisisina siirtoina aktivoituvat ainakin seuraavat:

- yleiskielen (= kielenhuollon) normi
- jargonnormit (ATK, Linux, KDE SC jne.)
- ryhmädynamiikkaan kuuluvan konvention normi ("Näin meillä on ollut tapana")
- käyttökelpoisuus- tai viestinnällisyysnormi ("Näin on helpompaa")
- perinteen säilyttämisen normi: maltillisen konservatiivisuuden arvo kielen automaattistajana
- etsittävyys- eli googlausnormi: läpinäkyvyys suhteessa lähdekieleen internet-hakujen helpottamiseksi

Kaksi ensimmäistä, yleiskielen ja jargonin normit, ovat selkeimmin preskriptiivisiä ja painottavat yleistä erityisen eli tilanteisen sijaan. Näitä tasoja hallitsee vaihtoehdottomuus: yksi muoto tai ilmaus on oikein, jokin toinen väärin. Konvention, käyttökelpoisuuden, perinteen ja etsittävyyden kriteerit painottavat enemmän vaihtoehtoisuutta. Konventioita ja perinteitä on monia (esim. tässä: kilpailevat työpöytäympäristöt KDE SC ja GNOME), käyttökelpoisuus mitataan suhteessa tilanteeseen eikä yleiseen normiin (vaikkakin jos tarkkoja ollaan, tilanteen normiin eli sen yleistykseseen, ei tiettyyn tilanteeseen) ja etsittävyys liittyy kielen kykyyn sanoa sama asia käänös- tai supisuomeksi.

Lokalisoijan on tärkeä ymmärtää, ettei ainutkaan normi, tasostaan riippumatta, missään ehdottomassa mielessä *ole* jotakin vaan jokainen niistä *luodaan* siinä nimenomaisessa toiminnassa, johon lokalisoija paraikaa osallistuu. Kieli on sosiaalista, dialogista toimintaa, jota kuljetetaan samanaikaisesti eteenpäin usealla aikatasolla (Kravchenko 2009: 528): tietystä tekstikontekstista (yksittäinen käänösjono), tekstin kokonaisuudessa (tietyn ohjelman käänös), tekstien muodostamassa kokonaisuudessa (ohjelmistokokonaisuuden käänös), tekstien historiallisesti koostamassa osakulttuurissa (KDE SC historiallisesti elävänä kielen kontekstina) ja näiden koostamisissa suuremmissa kulttuureissa (tietotekniikka, länsimainen arkielämä, josta ne ovat osa, jne.). Lokalisoijien kommentteja tarkastelemalla tätä dynamiikkaa on mahdollista havainnollistaa.

Metalingvisticissä ilmauksissa, jollaisina tässä tarkasteltuja retorisia siirtoja voi myös pitää, normi asettuu tavallisesti selittävän asemaan. Sen avulla selitetään, miksi tietty ilmaus on tietyn muotoinen. Samalla niissä kuitenkin myös vähintään vahvistetaan normia tai jopa luodaan se takautuvaksi selitykseksi jo muodostetulle ilmaukselle. Kaiken tämän metalingvistisen toiminnan tarkastelu on vaikeaa arkikielenkäytössä, missä harvoin julkisesti palataan jo sanottuun normien kannalta. Lokalisoinnin tutkimus tarjoaa tähän väylän – muttei ongelmatonta.

5 Metodologisia ongelmia (ja paljonko niillä on väliä)

Kuten jo edellä totesin, lokalisoijat tekevät käännösratkaisunsa tavallisesti itse ja yksin, nostamatta normeja julkisesti näkyviin. Normit toimivat tällöin samoin kuin ne toimivat arkikielenkäytössä, piilossa ja ”tiedostamattomasti”. (Pidän itse asiassa ”tiedostamatonta” tässä harhaanjohtavana ja jopa varsin perustellusti vääränä sanana, mutta käytöllä on vankka kielitieteellinen historia, johon tässä ei ole tilaisuutta mennä.) Vain harvoin normit asettuvat huomion kohteeksi niin, että voitaisiin sanan varsinaisessa mielessä puhua merkitysneuvottelusta.

Kuitenkin lokalisoinnissa on merkittävässä määrin kyse paitsi valmiiden normien käytöstä myös uusien normien jatkuvasta luomisesta. Käännösten ”ekosfääri”, tietyn työpöytäympäristön uusi tekniikka, on kohdekielen – ja usein lähdekielenkin – kannalta tuore eikä valmista normia ole olemassa. Tämän normien luomisen nostaminen aineistosta esiin ei ole helppoa. Kommenteissa toteutuvat retoriset siirrot ovat hyvä mutta suppea aineisto. Valtaosa normatiivisesta työstä tapahtuu niiden ulkopuolella.

Voikin kysyä, mikä tämän tutkimuksen kontekstissa lopulta muodostaa relevantin aineiston. Lokalisointitiedostot ovat ilmeinen ja helppo aineiston muoto, mutta sen ohessa kulkee paljon mielenkiintoista dataa. Yksi tällainen muoto ovat .po-tiedostojen muutoslokit. Käännettävä materiaali on sijoitettu versionhallintajärjestelmään, joka kirjaa kunkin tehdyn muutoksen ja kykenee tarvittaessa palauttamaan muutoksia ajassa taaksepäin. Yksi kiintoisa mahdollisuus tutkimukseen tarjoutuu tämän tehtyjen muutosten historiaan liittyvän oheismateriaalin tarkastelusta. Muutoksesta ei tällöin tarvita .po-tiedoston kommenttiin merkittyä motiivia: pelkästään se, että jokin on muutettu muodosta toiseen antaa aiheen kysyä, mitä normeja sovellettiin ja miksi. Kommenttidataan verrattuna ongelmana tietysti on, että käytetyt normit on muodoista pääteltävä, koska niihin ei eksplisiittisesti vedota.

Edelleen aineistona voi pitää muilla kanavilla käytyä lokalisoijien välistä diskurssia, mutta sen kerääminen on monistakin syistä ongelmallista ja työlästä. Suurena metodologisena ongelmana on sen transiensi – lyhytikäisyys ja häilyväisyys. Valtaosa lokalisoijien siitä keskinäisestä vuorovaikutuksesta, jota ei sijoiteta .po-tiedostoihin, käydään erilaisissa ei-säilyvissä viestimissä kuten IRCissä, Skypen chat-toiminnossa, Facebookissa jne. Toki pysyvämpiäkin muotoja on kuten sähköposti, sähköpostilistat, foorumit ja wikit (esim. sanastot), mutta usein pysyvämmät välineet tulevat käyttöön normien luomisen kannalta myöhäisessä vaiheessa.

Toisenlainen metodologinen ongelma on se, mikä muodostaa käännettävän *tekstin*. Käännöstieteessä on tavattu korostaa tekstien kääntämistä merkityskokonaisuuksina (Bassnett 2002: 23–30; Munday 2001: 19–22). Tämä on ymmärrettävää, koska teksti muodostaa merkityksen sulkeuman (Nieminen 2010: 79), jossa sille voidaan esittää mielekäs kokonaistulkinta. Tekstin merkityssuhteet ulottuvat sen itsensä ulkopuolellekin mutta nimenomaan semanttis-pragmaattisina viittauksina, ei välttämättöminä tekstuaalisen muodon kytköksinä. Tekstin rajat muodostavat sen kokonaisuuden rajat, joka kääntäjän on tulkittava semioottiseksi yksiköksi eli viestiksi. Jo yksittäisen tieto-

koneohjelman – kokonaisesta käyttöympäristöstä puhumattakaan – samastaminen perinteiseen tekstiin on haastavaa (vrt. esim. Lehtovaara 2010: 26–39) ja johtaa pahimmillaan löysiin analogioihin. Tätä puolta lokalisoinnin ja kääntämisen suhteesta ei kuitenkaan ole vielä juuri tutkittu. Esimerkiksi arvioidessaan käännöstieteen teorialokalisoinnissaan lokalisoinnin ja muun kääntämisen ongelmallista suhdetta Anthony Pym (2010) ei huomioi lokalisoinnin tekstikäsitteen eriävyyttä, vaikka hän toisaalla (2011: 6) mainitsee sen ongelman tai rasisitteen luonteisena teknisenä kysymyksenä. Kyse ei kuitenkaan ole vain tekniikasta vaan käsitteellisestä erosta: yksittäiset lokalisoitavat jonot eivät ole tekstien tapaan merkityksen sulkeumia vaan rinnastuvat pikemmin käännöstekstin irrallisiin, tekstistä irrotettuihin lauseisiin. Tekstiksi taas hahmottuu parhaiten lokalisoitujen ympäristön lopullinen dynaaminen käyttö, joka ei missään olosuhteissa ole lokalisoijan välittömästi tavoitettavissa. Koska tutkimukseni kohteena ei kuitenkaan ole kääntäjän työ tekstissä (= mitä käännetään) vaan tekstin kautta (= mitä käännöksen avulla kielessä saavutetaan), metodologinen vaikeus voidaan tässä ohittaa.

Moninaisista metodologisista ongelmista huolimatta katsoisin, että lokalisoinnissa tapahtuvien retoristen siirtojen (sekä julkisten että piilosiirtojen) tutkimus avaa uudenlaisen mahdollisuuden tarkastella kääntämistä kielen normatiivisena työnä, kielentämisenä. Tässä kontekstissa kääntäjä ei ainoastaan välitä tekstejä lähettäjältä vastaanottajalle (Bassnett 2002: 43–44) vaan tekee paljon enemmän: muodostaa sitä koodia, jonka avulla tuo välittäminen on mahdollista.

Kirjallisuuskäsitteet

Bassnett, Susan 2002. *Translation studies*. 3. laitos. London & New York: Routledge.

Bell, Jim & Loane, Sharon 2010. 'New-wave' global firms: Web 2.0 and SME internationalisation. *Journal of Marketing Management* 26:3–4, 213–229.

Bendor, Jonathan & Swistak, Piotr 2001. The evolution of norms. *American Journal of Sociology* 106:6, 1493–1545.

Collis, Betty & Moonen, Jef 2008. Web 2.0 tools and processes in higher education: Quality perspectives. *Educational Media International* 45:2, 93–106.

Cowley, Stephen J. 2005. Linguaging: How humans and bonobos lock on to human modes of life. *International Journal of Computational Cognition* 3:1, 44–55.

Harwood, Tracy & Garry, Tony 2010. 'It's mine!': Participation and ownership within virtual co-creation environments. *Journal of Marketing Management* 26:3–4, 290–301.

Howe, Jeff 2006. The rise of crowdsourcing. *Wired* 14.06/2006. Artikkelin löytyy myös web-arkistosta: <http://www.wired.com/wired/archive/14.06/crowds.html>.

Karasavvidis, Ilias 2010. Wiki uses in higher education: Exploring barriers to successful implementation. *Interactive Learning Environments* 18:3, 219–231.

Kockelman, Paul 2007. Agency: The relation between meaning, power, and knowledge. *Current Anthropology* 48:3, 375–401.

Kravchenko, Alexander V. 2009. The experiential basis of speech and writing as different cognitive domains. *Pragmatics & Cognition* 17:3, 527–548.

Lehtovaara, Aatu 2010. *Norsunluutornista basaariin: Nordin tekstianalyysimalli lokalisoinnin tarkastelun pohjana*. Pro gradu -tutkielma. Tampere: Käännöstiede (englanti), kieli- ja käännöstieteiden laitos, Tampereen yliopisto.

Leiwo, Matti, Luukka, Minna-Riitta & Nikula, Tarja 1992. *Pragmatiikan ja retoriikan perusteita*. Jyväskylä: Jyväskylän yliopiston viestintätieteiden laitoksen julkaisuja 8.

Lievrouw, Leah A. 2010. Social media and the production of knowledge: A return to little science? *Social Epistemology* 24:3, 219–237.

MacWhinney, Brian 2001. Emergentist approaches to language. Teoksessa: Joan Bybee & Paul Hopper (toim.) *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: John Benjamins, 449–470.

Melby, Alan K. & Foster, Christopher 2010. Context in translation: Definition, access and teamwork. *Translation & Interpreting* 2:2.

Munday, Jeremy 2001. *Introducing translation studies: Theories and applications*. London & New York: Routledge.

Määttä, Urho 1999. Kielen sosiaalinen perusta ja sääntöjen emergoituminen. Teoksessa: Urho Määttä, Pekka Pälli & Matti K. Suojanen (toim.) *Kirjoituksia sosiolingvistiikasta*. Folia fennistica & linguistica 22. Tampere: Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos. 23–52.

Määttä, Urho 2000. Mistä on pienet säännöt tehty? *Virittäjä* 104, 203–221.

Nieminen, Tommi 1998. Merkityksen evoluutiomallista. Teoksessa: Urho Määttä & Klaus Laalo (toim.) *Kirjoituksia muoto- ja merkitysopista*. Folia fennistica & linguistica 21. Tampere: Tampereen yliopiston suomen kielen ja yleisen kielitieteen laitos, 117–141.

Nieminen, Tommi 2010. *Lajien synty: Tekstilaji kielitieteen semioottisessa metateoriassa*. Jyväskylä Studies in Humanities 136. Jyväskylä: Jyväskylän yliopisto.

Nogueira, Danilo & Semolini, Kelli 2010. Crowdsourcing. *Translation Journal* 14:2. Verkkoasiakirja: <http://translationjournal.net/journal//52crowd.htm>.

Peirce, Charles S. *CP = Collected papers of C. S. Peirce*. Toim. Charles Hartshorne & Paul Weiss (osat 1–6, 1931–35), Arthur W. Burks (osat 7–8, 1958). Cambridge: Harvard University Press.

Peirce, Charles S. *EP = The essential Peirce: Selected philosophical writings*. Toim. Nathan Houser & Christian Kloesel (osa 1, 1992), The Peirce Edition Project (osa 2, 1998). Bloomington & Indianapolis: Indiana University Press.

Pym, Anthony 2010. *Exploring translation theories*. London: Taylor & Francis.

Pym, Anthony 2011. What technology does to translating. *Translation & Interpreting* 3:1, 1–9.