



Tässä kirjoituksessa käsitellään tapoja, joilla Avoimen koodin käännösprojekti voi pyrkiä tavoittelemaan suurta käyttäjämäärää ja minimoida käyttäjätuen tarpeen. - Jos käännät pienelle, rajatulle kohdeyleisölle ja haluat maksimoida tulot käyttäjätuesta, niin toimi päinvastoin kuin alla esitetään.

Yhteenveto

Lokalisoinnissa (kotoistamisessa) on kaksi perusvaihtetta:

1. Käännöstyö
2. Käännöksen laadun kohottaminen
 - Tavoite: kohdeyleisö ymmärtää nopeasti viestin ja ymmärtää sen oikein.

1. Alkutoimet:

- Silmäilläään alkuteksti
- Tehdään sanakirjakäännös

2. Suodatetaan seuraavien 'suotimien läpi':

- Sovitetaan asiayhteyteen
- Lisätään asiantuntemus aiheesta tai yleissivistys
- Lisätään alkuperäisen kulttuurin tuntemus
- Sovitetaan kohdeyleisön puhe- tai ilmaisutapaan
- Vältetään erikoisanal sanastoa
- Vältetään lainasanoja
- Tarvittaessa: luodaan uusi sana tai annetaan sanalle uusi merkitys
- Sovitetaan suomen kielen oikeinkirjoitukseen
- Sovitetaan ilmaus tilaan joka on tarjolla (esim. 20 merkkiä korkeintaan)
- Sovitetaan muuhun rajoitteeseen (esim. ei 'äökkösiä')

- Yhtenäistetään ilmaukset (sama sana, sanonta kun samasta asiasta kyse)
- Yhtenäistetään kirjoitusasu/esitystapa ilmaukselle

Ohjelmistotuotteen kotoistaminen (lokalisointi)

'Käännös'projektin alussa on oltava vastaus ainakin näihin kysymyksiin:

Kohdeyleisö ja muut tavoitteet

- Kenelle lopputuote on tarkoitettu? = Kuka tuotetta tulee käyttämään?
- Kohdistetaanko tuote rajatulle ryhmälle, esim. tietyn koulutusalan/-tason henkilöille vai mahdollisimman laajalle käyttäjäkunnalle? = Onko tarkoitus palvella rajattua joukkoa vai saavuttaa mahdollisimman suuri käyttäjäkunta?
- Pyritäänkö tekemään tuote helppotajuiseksi vai maksimoidaanko käyttäjäkoulutuksesta saatavat tulot tekemällä siitä vaikeatajuisempi, tavalla tai toisella?

Avoim koodi -projektit

Avoim koodi (Open Source) on erinomainen ilmiö. Sillä on lukuisia etuja sekä kehittäjille että lopputuotteen käyttäjille. Kaikki voivat osallistua, mutta tämä etu on samalla sen Akilleen kantapää. Aktiiviset ja hyvin johdetut projektit kehittyvät pienin askelin ja suurin harppauksin paremmiksi, jopa kaupallisia tuotteita paremmiksi. Mutta hyvin suurta osaa kannattaa innokkuus osallistua ja tehdä.

Avoim koodi projektit ovat tunnettuja puuttuvasta käyttäjätuesta ja -koulutuksesta. Sen pohjalta voisi olettaa, että erityisesti Avoimen koodin tuotteet pyrittäisiin lokalisoimaan seuraavasti:

- Käännetään mahdollisimman [helpoiksi käyttää](#), jotta tuen määrä olisi pienin mahdollinen.
- Tavoitellaan mahdollisimman [laajaa käyttäjäkuntaa](#) - suuri käyttäjämäärä tuo projektille elinvoimaa pitkällä aikavälillä.



Laatu

Laatua on helppo tavoitella ja siitä puhua, mutta laadukkaaseen lopputulokseen on yllättävän vaikea yltää. Lokalisoinnin laadulle voidaan esittää esim. seuraavat erotteluperusteet (kriteerit):

- Sovitettu kohdeyleisölle.
- Ymmärrettävyys: ymmärtämisen maksimointi, väärinkäsitysten minimointi.
- Yhtenäinen kieliasu ja yhtenäinen kirjoitusasu (= muotoseikat).

Hyvin monen asian pitää osua kohdalleen jotta lopputulos on laadukas. Eräs tapa esittää asia:

1 Mistä on kyse:

Ohjelmistotuote on työväline. Sen käyttöönnotolla tavoitellaan usein työn **tuottavuuden kohottamista**. Tärkeäksi nousee ohjelmiston ja sen käyttäjän välinen viestintä. Sen tulisi tietysti olla nopeaa, helposti ymmärrettävää ja väärinkäsitykset minimoivaa. - Asiaa voidaan tarkastella esimerkiksi viestinnän käytännöläheisestä (pragmaattisesta) näkökulmasta, seuraavasti:

Sovitetaanko sanoma kuulijakunnalle - ihmisille joille sanoma kohdistetaan?

On kaksi vastakkaista tapaa esittää asia:

1. Esittäjä esittää sanoman käyttämällä **oman taustansa sanastoa**.
 - Helppo esittäjälle. Tarjoaa usein mahdollisuuden 'briljeerata sivistys'sanoilla.
 - Lisää etäisyyttä kuulijoihin ja pidentää aikaa joka tarvitaan sanoman 'perille menoon'.
 - Esitys voi mennä hukkaan - se ei tavoita yleisöä.
 - Yrittää pakottaa yleisön oman sanaston muottiin - vaikka heillä olisi jo sanasto?
 2. Esittäjä esittää sanomansa käyttäen **kuulijakunnan käyttämää sanastoa**.
 - Vaikeampi esittäjälle koska asia on todella hallittava niin hyvin, että osaa pukea asian 'toiselle kielelle'.
 - Vaikeampi, koska esittäjän on tunnettava yleisön 'kieltä'.
 - Helpompi yleisölle - viesti 'uppoaa' eli menee 'paremmin perille'.
 - Kuulijakuntaa kunnioitetaan ja se usein myös huomataan.
- Esimerkki:
 - Kun kuulijoina on lingvistiikan tai tietojenkäsittelyn osaajia/harrastajia, voidaan käyttää sanoja: kategoria, subkategoria, taksonomia, attribuutit, variaatiot.
 - Kun sama 'asia' esitetään kaupan alalla toimiville kauppiaille, kannattaa se esittää toisilla sanoilla: esim. osasto, tuoteryhmä, luokittelu, [tuotteen]ominaisuudet, [tuote]muunnelmat.
 - Miksi 1? Koska näitä konkreettisia sanoja käytetään suomen kielessä päivittäin eikä samojen rakenteiden siirtäminen ohjelmistotuotteeseen edellytä välttämättä niiden esittämistä uuden sanaston avulla - vaikka ohjelmiston tekijä olisi työssään käyttänyt toisia sanoja.
 - Miksi 2? Koska vaikeuttaa ymmärtämistä. Esimerkki:
 - "Before adding variations, add and save some attributes on the Attributes tab."
 - Ennen variaatioiden lisäämistä, lisää ja tallenna attribuutteja, Attribuutit -välilehdellä.
 - Ensinnäkin lisää vähintään yksi ominaisuus ja tallenna se 'Ominaisuudet' välilehdellä. Vasta sen jälkeen voit lisätä muunnelman."

Siis, kumartaako esittäjä yleisölle vai onko yleisön kumarrettava esittäjälle?

Jos esitys on hyvä, kuulijakunta esittää aplodit esityksen lopussa. Esittäjän tehtävä on saada 'yleisö mukaan' ennen esityksen loppua. Paljon on kiinni käytetystä sanastosta.



2 Kuinka laatu tehdään:

Tutustutaan alkutekstiin

- Yritetään ymmärtää minkä tiedon teksti yrittää siirtää katsojalle.

Tehdään ns. sanakirjakäännös

- Tehdään ensimmäinen käännös jota aletaan työstää. Käytetään apuna jotain sanakirjaa, vakiosanastoa tai jopa konekäännöstä.
- Käännös voi olla kohtalainen jos se on yleiskieltä. Jos käytetty sanasto on erikoisalan sanakirja ja tuote suunnataan laajalle yleisölle niin käännös voi olla käyttökelvoton.
- Ensimmäinen käännös voi olla harhaanjohtava, vaikeaa kapulakieltä ja sisältää muita puutteita.
 - Avoin koodi -käännöksissä tämä on liian usein ainoa tarjolla oleva käännös. Jos tilanne jää pysyväksi, ohjelman leviäminen hidastuu.

Sovitetaan asiayhteyteen (kontekstiin, context)

- Katsotaan ohjelmasta (käännettävästä ohjelmistosta) paikka, jossa ilmausta käytetään.
- Onko käännös sopiva kyseiseen kohtaan.
- Esimerkki: englannin sanat To: ja From: on alla käännetty kuin kyseessä olisi postilähetys. Tässä kohdassa kyse on kuitenkin alku- ja loppupäivämäärästä (alempi kuva).

Lähetäjä:	2013-01-01	Vastaanottaja:	2013-01-01	Näytä
Alku pvm.:	2013-01-01	Loppu pvm.:	2013-01-02	Näytä

Lisätään asiantuntemus aiheesta tai yleissivistys

- On otettava huomioon, että käännettävä teksti voi olla (alun pitäen alkukielellä) huonosti muotoiltu ja harhaanjohtava.
- Ilmaistaan asia tarvittaessa tarkemmin, ymmärtämistä kohottaen ja väärin ymmärtämisen mahdollisuutta vähentäen.
- Tarkastetaan oma asiantuntemus - onko asia todella ymmärretty, vai kätketäänkö oma osaamattomuus ympäröivään sanahelinän, koukeroisen kielen tai 'sivistys'sanojen käytön taakse?

Lisätään alkuperäisen kulttuurin tuntemus tarvittaessa ja jos sitä on

- Joskus käännettävä ilmaus olettaa alkuperäisen [kulttuuri]ympäristön tuntemuksen, jotta se on täysin ymmärrettävissä.
- Yksinkertainen esimerkki: USA:sta peräisin olevissa ohjelmissa kohdataan osavaltiot (State). Jos ohjelmaa käytetään esim. maailmanlaajuiseen postitukseen, on parempi kääntää esim. sanalla 'alue', koska useassa muussa maassa on vastaavia hallinnollisia alueita, mutta niitä kutsutaan eri nimellä, esim. provinssi, territorio, maakunta, jne.

Sovitetaan kohdeyleisön puhe- tai ilmaisutapaan

- Kun käytetään kohdeyleisölle tuttuja sanoja ja lauserakenteita, nopeutetaan ymmärtämistä.
- Esimerkki:
 - "In order to demonstrate your lack of understanding of the Finnish language, do not fail to use English word order and sentence structures."
 - Osoittaaksesi ymmärtämättömyytesi suomen kieltä kohtaan, älä ole käyttämättä englannin kielen sanajärjestystä ja lauserakenteita. - Kapulakieltä?
 - Käytä aina englannin kielen sanajärjestystä ja lauserakenteita kun haluat osoittaa, että et ymmärrä suomen kieltä [kielenä].
 - Kun haluat osoittaa, että et ymmärrä suomen kieltä [kielenä], niin käytä aina englannin kielen sanajärjestystä ja lauserakenteita.



Vältetään erikoisan sanastoa

- Vältetään erikoisan sanastoa (jargon tai slang, murre, jne.) ellei pääasiallinen kohdeyleisö hallitse kyseistä erikoisan sanastoa, toimi päivittäin sen parissa ja sen käyttö helpottaisi ohjelmiston käyttöä.
- Yhden erikoisan sanastoa ei 'pakkosyötetä' toiselle erikoisalalle.
- Erikoisan sanastoa ei 'pakoteta' yleiskieleen jos nopeampi ymmärrettävyys saavutetaan muulla tavalla.

Vältetään lainasanoja

- Lainasanat ovat piirre useissa kielissä, kun kehitys ja innovaatiot tulevat 'rajan takaa'.
- Kun kieli täydentyy, alussa on nopeinta ja ymmärrettävää käyttöä lainasanaa.
- Sen jälkeen voi tulla uudissana yritelmiä, joiden sisältöä on yhtä vaikea ymmärtää. Tässä vaiheessa voi käyttää kumpaakin, alkuperäinen ilmaus esim. suluissa. Maksimoi ymmärrettävyyden.
- Lopulta joku muunnelma tai kotoperäinen ilmaus vakiintuu. Esimerkkejä:
 - plastic -> plastiikki -> muovi
 - magnetophone -> magnetofoni -> (äänikkö) -> nauhuri
 - widget -> vidgetti -> vimpain (-> ?)
 - domain -> domaini -> verkkotunnus
- Esimerkki: "Shows a custom attribute in a widget which lets you narrow down ..."
 - Näyttää kustomi attribuutin vimpaimessa, joka antaa sinun rajata näytettävien ...
 - Mukautetut ominaisuudet näytetään vimpaimessa jonka tehtävä on rajata hakua ...

Tarvittaessa: luodaan uusi sana tai annetaan sanalle uusi merkitys

- Uusi tuote, ilmiö tai käsite voi tarvita uuden sanan jolla se saadaan erottumaan vanhasta tai liittymään johonkin vanhaan ja tuttuun. Uusi yhdyssana on usein helppo ratkaisu.
- Esimerkkejä:
 - Kuinka otsikoida ohjelmistoon liitettävät [verkko]palvelut kun alkup. teksti on "Integration"? Sopiva lokalisoitu sana on esim. 'Palvelusidokset' tai lyhyemmin 'Sidokset'. 'Integraatio' on lainasana ja sen merkitystä on kohdassa vaikeampi hahmottaa kuin palvelusidoksen. Lisäksi sanoilla liitokset ja liittymät on jo yleistajuiset merkitykset.
 - Face-to-face kommunikaatio -> kasvokkain viestintä.
 - Sana 'mikro' on palvellut ja palvelee yhä hyvin laajasti: ollut esim. mikro(aaltouuni - Microwave oven); mikro(shortsit - Hot pants); mikro(tietokone eli PC); ...

Sovitetaan suomen kielen oikeinkirjoitukseen

- Ensisijainen tavoite on käyttää suomen kielen oikeinkirjoitusta. Mutta:
- Ohjelmistotuotteen lokalisoinnissa kohdataan tilanteita, joissa se on lähes mahdotonta. Esimerkiksi alla esitetyt tekstikentän max pituus ja ei ääkkösiä.
- Suuri osa tilanteista on kuitenkin sellaisia, joista selvittää 'joustamalla tai venyttämällä' oikeinkirjoitusta. Tällöin kannattaa nähdä päätös vaikka seuraavasti:
 - Kieli kehittyy väistämättä.
 - Ilmauksen 'graafinen esitystapa' voi ja saa poiketa oikeinkirjoituksesta**.
 - Tarve ilmaisuuden selvyydelle ja napakkuudelle pakottaa joskus poikkeamaan perinteisestä oikeinkirjoituksesta ja kieliopista, ymmärrettävyyden nimissä.
- Hakuteos, jos joskus tarvitsee: Jukka K. Korpela: [Nykyajan Kielenopas](#).

Sovitetaan ilmaus tilaan joka on tarjolla

- Esimerkiksi: tekstikentässä on tilaa vain 20 kirjoitusmerkillä.
- Tämä voi olla lopulta määräävin suodin. Suomen kielessä on pitkiä sanoja. Kun tavoitellaan napakkaa ilmaisua, se voi tarkoittaa seuraavaa:
 - Sanan perusmuotoa käytetään taivutetun muodon paikalla.
 - Sana katkaistaan.
 - Käytetään lyhennettä.
 - Käytetään muuta ilmausta asian esittämiseen. Esimerkki:
 - "Page 5 out of 12 pages" -> Sivun 5 / (12) --- tai Sivun 5 (12)



Sovitetaan muuhun rajoitteeseen

- Jotkut tekniset rajoitteet ovat mahdollisuuksia.
 - Esimerkiksi UTF-8 merkkikoodauksen käyttö verkkosivustolla ja "utf8_general_ci" koodaus [MySQL] tietokannassa laajentaa huomattavasti käytettävissä olevien **kirjoitusmerkkien ja kirjoituskielten valikoimaa** samassa asiakirjassa. - ISO-8859-1 tai ISO-8859-15 koodaus on rajoite.
- Toiset tekniset rajoitteet toimivat nimensä mukaan ja aiheuttavat sekaannusta ja vähentävät ymmärtämistä. Esimerkiksi:
 - Ei ääkkösiä eli skandinaaviset merkit tai 'skandit'. UTF-8 koodausta käytettäessä eivät ääkköset enää ole ongelma mutta sähköpostiosoitteissa ja vanhoissa ohjelmistoissa ne voivat olla ylitseppäsemätön este. Esimerkki: väärä = vaara.
 - Mutta:** sähköpostiosoitteiden globaalisti yhtenäinen kirjoitustapa (ASCII merkit) on yksi globaalien talouden tukijaloista. - Jos kiinalainen yritys ilmoittaisi sähköpostiosoitteensa kiinalaisilla merkeillä, moniko osaisi kirjoittaa sen omalla näppäimistöllään? Eikä valtaosa maailman väestöstä osaa kirjoittaa ä, ö ja å kirjaimia. Kansainvälinen viestintä on niin pienestä asiasta kiinni.
- Käännös ja lokalisointi on sovitettava rajoitteiden muottiin, mutta ainakin osa rajoitteista voidaan nähdä myös positiivisessa valossa.

Yhtenäistetään ilmaukset

- Ohjelmiston käyttö helpottuu, kun:
 - käytetään samankaltaista ilmaisua jokaisessa tietynlaisessa tilanteessa,
 - käytetään napakkaa ja ytimekästä ilmaisua,
 - pyritään minimoimaan käytettyjen ilmausten määrä.
- Tähän on teknisiä apuvälineitä, esim. käännösmuisti (TM, Translation Memory)

Yhtenäistetään kirjoitusasu/esitystapa ilmaukselle

- Sanan tai usean sanan voi kirjoittaa usealla tavalla. Suurella tai pienellä alkukirjaimella, yhteen tai erikseen, väliviivalla yhdistäen. Näiden lisäksi on mainittava tavat joita suomen oikeinkirjoitus ei tunnusta: alleviivausviivalla yhdistäminen ja ison kirjaimen sijoittaminen yhdyssanan keskelle.
- Ilmauksen esittäminen aina samalla tavalla, eri paikoissa nopeuttaa ymmärtämistä. Yhtenäistämisen tarve ja edut koskevat sekä hyväksytyjä että uusia kirjoitustapoja.

Tässä on käsitelty

Käännöksen ja kotoistamisen (lokalisoinnin) laatu riippuu useasta tekijästä. Työssä joudutaan toimimaan sellaisten rajoitteiden parissa, joita ei ollut kun suomen kielen oikeinkirjoitussäännöt on luotu.

Lokalisointia tekevä henkilö joutuu ajoittain luomaan uutta - viemään suomen kieltä ja sen käyttämistä eteenpäin globaalissa maailmassa. Hän ei pyri tuhoamaan suomen kieltä, vaan pyrkii saamaan sen toimimaan teknisen ja tieteellisen kehityksen kelkassa, hyvin käytännönläheisesti. Joskus uusi ratkaisu tarttuu, joskus 'menee metsään'.

Laadun mittarina joudutaan käyttämään ymmärrettävyyttä ja ymmärrettävyyden nopeutta ja helppoutta.

** Tämä tarkoittaa: joskus muinoin pankki (Kansallis-Osake-Pankki, Merita Pankki) alkoi kirjoittaa nimensä tavalla, joka ei ollut Suomen 'Kotimaisten kielten keskuksen' mieleen. Suomen kielen perinteistä oikeinkirjoitusta ei käytetty joten Kielitoimisto älähti. Silloin elettiin vielä siinä uskossa, että Kielitoimiston suositusta tai ohjetta oli noudatettava kuuliaisesti.

Pankki piti kiinni asiasta väittäen (argumentoiden): yritys saa ja sen pitää saada päättää kuinka se esittää yrityksensä tunnukset, nimi mukaan lukien. Toisin sanoen, yritys saa määrätä yrityksen nimen "graafisen esitystavan" - siis miltä nimi näyttää kirjoitettaessa.

Pankki voitti, graafinen ala sai vapaammat kädet, kielen holhous höllentyi hiukan ja Kielitoimistosta voi sanoa "Haukkuva koiri ei pure" vaikka se kuinka **metelöi** (kun huomasi, että siltä puuttuu hampaat?).