

Frekvenssejä ja reseptiota: Kaksi näkökulmaa kotouttamiseen ja vieraannuttamiseen

Jukka Mäkisalo
Itä-Suomen yliopisto, Joensuu
Käännöstiede

Hannu Kemppanen
Itä-Suomen yliopisto, Joensuu
Venäjän kieli ja kääntäminen

Abstract

This study has a three-fold objective. First, it aims at examining whether it is possible to identify domesticating and foreignising features in translated texts by using corpus-based methods. Secondly, the study seeks to discover which features of these texts are considered as domesticating and which as foreignising by translation students and translation scholars. Thirdly, the study discusses whether the results of the two analyses correlate with each other. The material consists of translated and non-translated texts on the political history of Finland. The results show that the corpus analysis does not allow us to recognise translational features in a text to the same extent as the readers of translations do in a reception test. The analysis reveals that the concept *foreign* can be regarded as a marked feature. This finding questions the dichotomic nature of the concept pair domestication vs. foreignisation.

Avainsanat: kotouttaminen, vieraannuttaminen, kääntäminen, korpuslingvistiikka, reseptio

1 Taustaa

Käännöstekstien kielellisiä ja tekstuaalisia ominaisuuksia on tutkittu empiirisesti korpusmenetelmin Suomessa paljonkin. Käännössuomen osalta on tutkittu käännöstekstien sanastoa puhtaasti tilastollisten tunnuslukujen valossa (Nevalainen 2005), mutta myös lingvistisemmin kollokaatioiden muodostumista (Jantunen 2004b, 2005) tai sanaston avulla käännösten ideologisia kytkentöjä (Kemppanen 2005, 2008). Käännöstekstien kieliopillisia ominaisuuksia on selvitetty kieliparivertailuin, mm. Eskola (2002, 2005) venäjä–suomi- ja englantia–suomi-kielipareilla, samoin saksa–suomi- ja englantia–suomi-vertailulla (Mauranen & Tiittula 2005). Vähemmän sen sijaan on tutkittu tekstien kotouttavia ja vieraannuttavia elementtejä, eikä ollenkaan korpusmenetelmin. Kun näiden käännöstekstien ominaisuuksien korpuslingvistiseen tarkasteluun yhdistetään vielä reseptio, on kyseessä pilottiluonteinen tutkimushanke. Tämä artikkeli on pilottitutkimuksen väliraportti.

Tavoitteet ovat sekä empiirisiä että käsitteellis-teoreettisia. Empiirisenä tavoitteena on ensinnäkin tarkastella sitä, kuinka hyvin kotouttaviksi ja vieraannuttaviksi oletetut käännöstekstien piirteet ovat korpusmenetelmin löydettävissä. Onko käännettyillä teksteillä tilastollisia ominaisuuksia, jotka voisivat olla yhteydessä kotouttavaan tai vieraan-

nuttavaan käännösstrategiaan? Toisena empiirisenä tavoitteena on reseptiotutkimuksen menetelmin kartoittaa sitä, millaisia tekstiominaisuuksia kääntämisen opiskelijat ja opettajat pitävät kotouttavina ja vieraannuttavina. Kolmanneksi katsomme, kuinka nämä kaksi näkökulmaa kotouttaviin ja vieraannuttaviin käännöksiin suhteutuvat toisiinsa, ts. korreloivatko ne mitenkään keskenään. Lopuksi mietimme tulosten pohjalta, millaisia käsitteitä *kotouttaminen* ja *vieraannuttaminen* voisivat olla ja mitä tulokset teoreettisesti tarkoittavat.

2 Kotouttamisen ja vieraannuttamisen käsitteet

Käännöstieteellisessä kirjallisuudessa nykyisin usein vastaan tulevat kotouttamisen ja vieraannuttamisen käsitteet toi tieteelliseen keskusteluun yhdysvaltalainen käännös-tutkija Lawrence Venuti (1995). Hän kritisoi angloamerikkalaista kaunokirjallista käännös-traditiota sen kotouttavasta luonteesta. Kääntäjät pyrkivät tuottamaan käännöksiä, joissa lähtökielestä ja -kulttuurista peräisin olevat kielelliset, etniset ja ideologiset elementit häivytetään niin, että vastaanottaja-lukija tuntee lukevansa alun perin kohdekulttuurissa kirjoitettua tekstiä. Kotouttavaa strategiaa käyttävä kääntäjä pyrkii tekemään tekstistään mahdollisimman sujuvaa ja helppolukuista. Hän välttää epäidio-maattisia ja arkaaisia ilmauksia, samoin jargonia ja toistoa.

Venuti (1995) puolustaa innokkaasti vieraannuttavan strategian käyttöä vaihtoehtona kotouttamiselle. Vieraannuttavat käännösratkaisut antavat lukijalle mahdollisuuden kohdata lähtökieli ja -kulttuuri käännetyssä tekstissä. Vieraannuttavassa kääntämisessä niin ikään kääntäjän aktiivinen rooli korostuu. Venuti puhuu tässä yhteydessä kääntäjän näkyvyydestä (*visibility*).

Venuti on kyseenalaistanut erityisesti sujuvan kohdetekstin tuottamisen ainoana tavoiteltavana kääntäjän taitona. Kuitenkin hänen lähestymistapansa on herättänyt myös kritiikkiä mm. kotouttamisen ja vieraannuttamisen käsitteiden epämääräisyyden ja niiden dikotomisen luonteen vuoksi (Tymoczko 2000, Boyden 2006).

Kritiikki puuttuu siihen, että Venuti tarkastelee kotouttamisen ja vieraannuttamisen käsitteitä yhteiskuntien ja kulttuurien tasolla, mutta hän ei kuvaa sitä, kuinka kotouttaminen ja vieraannuttaminen realisoituvat käännöksissä. Jotkut tutkijat ovatkin tehneet ehdotuksia näiden käsitteiden operationaalistamiseksi. Esimerkiksi Pedersen (2007) on jaotellut kahden päästrategian sisällä erilaisia konkreettisia käännöskeinoja, joita kääntäjä käyttää valitsemansa strategian toteuttamiseen. Kotouttaminen ja vieraannuttaminen eivät siten ole pelkästään koko yhteiskunnan tasolla esiintyviä yleisiä strategioita, vaan myös konkreettiseen viestintätilanteeseen tai tiettyyn tekstilajiin liittyviä tapoja lähestyä käännettävää tekstiä ja kohdekulttuurin lukijoita.

Konkreettiset käännöstilanteet kyseenalaistavat myös jyrkän dikotomian. Esimerkiksi lastenkirjallisuuden kääntämistä tutkinut Haidee Kruger (2009) on todennut, että kotouttaminen voi olla juuri sopiva ratkaisu lapsille käännettäessä. Kuitenkin sellaisissa monikulttuurisissa yhteisöissä kuten esim. Etelä-Afrikassa lapsille tarjoutuu vieraan-

nuttavien käännösratkaisujen kautta mahdollisuus lisätä tuntemustaan muista samassa yhteiskunnassa asuvista kansoista, niiden kielistä ja kulttuureista.

Tämän tutkimuksen kannalta ovat kiinnostavia lisäksi tutkimukset, joissa tekstin kotoisuutta ja vierautta on selvitetty reseptiotestien avulla (Vehmas-Lehto 1989, Tirkkonen-Condit 2002, Jantunen 2004b). Vaikka kyseisiä testejä ei olekaan tehty varsinaisesti vieraannuttamisen ja kotouttamisen ilmiöiden analysoimiseksi, niissä on kuitenkin pohdittu kysymystä siitä, erottavatko lukijat käännöksen ja ei-käännetyn kohdekielisen tekstin toisistaan. Näissä testeissä kotouttamista ja vieraannuttamista on lähestytty empirian kautta. Testejä voidaan siten tietyin osin pitää yrityksinä operationaalistaa tämä käsitepari.

3 Avainsanat ja korpusanalyysi

Korpusanalyysin lähtökohtana oli idea siitä, että tekstin ns. avainsanat jollakin tavalla ilmentäisivät sellaisia ominaisuuksia, jotka koetaan vieraaksi. Kemppanen (2008: 210) on havainnut, että käännettyissä teksteissä on supisuomalaisiin¹ teksteihin verrattuna yhtenä ryhmänä sellaisia avainsanoja, jotka viittaavat interferenssiin: ne ”implikoivat käännösratkaisuja, joita tehdessään kääntäjä on käyttänyt kohdekielille epätyypillisiä leksikaalisia aineksia.” Havainto johti siihen hypoteesiin, että *WordSmith Tools* -korpusohjelman avainsana-toiminnon (*keywords*) avulla voisi edes jollakin tavalla päätellä, onko teksti käännetty vai ei, tai kenties jopa päätellä, kuinka vieraan oloinen käännetty teksti on.

Avainsana-toiminto vertaa jotakin haluttua tekstiä (tai tekstimassaa) ja sen sanastollisia ominaispiirteitä suhteessa laajempaan tekstimassaan, esim. yhtä kaunokirjallista teosta laajempaan kaunokirjallisten teosten joukkoon. Toiminto tekee kummastakin tekstimassasta taajuusanalistan ja sen jälkeen poimii tarkasteltavasta tekstistä ne sanat, jotka esiintyvät siinä tilastollisesti merkitsevästi useammin kuin vertailukorpuksessa. Kuva 1 on esimerkki yhden aineistoomme kuuluvan tekstin avainsanoista suhteessa vertailukorpuksena käyttämäämme sanomalehtikorpukseseen.

N	Key word	Freq.	%	Freq.	RC. %	Keyness	P
1	SKP	7	0,69	2		87,09	0,0000000000
2	SUOMEN	23	2,27	1 614	0,16	79,14	0,0000000000
3	KANSANHALLITUKSEN	3	0,30	0		41,40	0,0000000000
4	TERIJOEN	3	0,30	0		41,40	0,0000000000
5	NEUVOSTOLIITON	6	0,59	71		40,81	0,0000000000
6	VALLANKUMOUKSEN	4	0,39	8		39,94	0,0000000000
7	NEUVOSTOLIITOSSA	4	0,39	9		39,17	0,0000000000
8	VALLANKUMOUKSELLISEN	3	0,30	1		36,90	0,0000000000
9	KUUSINEN	3	0,30	3		33,08	0,0000000059
10	PUOLUEEN	6	0,59	152	0,02	32,09	0,0000000118
11	KOMMUNISTISEN	3	0,30	4		31,84	0,0000000138
12	ALKAMISTA	3	0,30	7		29,19	0,0000000626
13	JOHDON	4	0,39	38		28,86	0,0000000750
14	ASEELLISIA	2	0,20	0		27,59	0,0000001466
15	MANNERIN	2	0,20	0		27,59	0,0000001466
16	PROLETARIAATIN	2	0,20	0		27,59	0,0000001466
17	OLI	22	2,17	5 737	0,57	26,47	0,0000002644
18	KOMMUNISTIEN	3	0,30	16		24,85	0,0000006157
19	TALVISODAN	3	0,30	16		24,85	0,0000006157
20	VASTAAN	7	0,69	491	0,05	24,00	0,0000009625

Kuva 1. Esimerkki *WordSmith Toolsin* avainsanalista, jossa testikorpuk-
 sena on yksi aineiston teksti (Baryshnikov; ks. alla s. 7) ja vertailukorpuk-
 sena sanomalehtikorpus. Taulukossa vasen *Freq.* on sanan frekvenssi testi-
 korpuksessa ja oikea *Freq.* vertailukorpuksessa.

Yleensä tarkasteltava teksti on lyhyt, kun vertailukorpus taas on laaja ja muodostuu vähintään sadoista samanlaisista teksteistä. Toimintoa on käytetty mm. lingvistik-
 tyylintutkimuksessa (ks. Opas 1995). Luonnollisesti se, mitä avainsanoiksi valikoituu,
 vaihtelee sen mukaan, mistä vertailukorpus koostuu. Tässä tutkimuksessa verrataan alun
 perin suomeksi kirjoitettuja ja suomennettuja historia-aiheisia tietokirjallisia tekstejä
 yhteen ja samaan sanomalehtikorpukseen (Karjalaisen korpus).

Teoreettisesti avainsanojen ajatellaan näin edustavan leksikaalista poikkeamaa ns.
 normista. Normi on määritelty tässä yhteydessä Mike Scottin (1998) esittämällä tavalla
 frekvenssinormina. Kiinnostavaa on erityisesti se, kuinka paljon käännosten
 avainsanalistaat eroavat suomalaisten tekstien avainsanalistoista, kun niitä verrataan
 yhteiseen suomalaiseen korpuspohjaan. Oletus tietysti on, että ainakin suomalaiset
 historiatekstit ovat lähempänä suomalaista sanomalehtitekstiä kuin suomennetut
 historiatekstit, ja sen pitäisi näkyä niiden pienempänä avainsanojen määränä. Hypoteesi
 siis on, että mitä enemmän avainsanoja, sitä vieraampi teksti. Näin yhdeksi mittariksi
 valittiin avainsanojen määrä, joka esim. kuvassa 1 on 20.

Avainsanojen määrän lisäksi mittariksi valittiin toiseksi ns. keyness-arvon tai -indeksin
 maksimi ja kolmanneksi keskiarvo. *WordSmith Tools* tulostaa suoraan keyness-arvon

jokaiselle avainsanalle erikseen (ks. kuva 1). Se ilmaisee siis tilastollista todennäköisyyttä², jolla sana päättyy avainsanalistalle. Keynes-arvon maksimi taas kuvaa avainsanaisuuden huippuarvoa: mitä korkeampi, sitä merkitsevämpi korkeimman arvon saaneen avainsanan tilastollinen ero vertailukorpukseen on. Esim. kuvassa 1 korkein arvo on 87,09, jonka saa avainsana *SKP*. Keynes-keskiarvo taas on kaikkien avainsanojen keyness-arvojen keskiarvo, ja se kuvaa kaikkien mukaan tulleiden avainsanojen keskimääräistä eroavuutta vertailukorpukseen nähden. Välttämättä kaikki kolme avainsana-arvoa eivät aseta tekstejä samaan järjestykseen, sillä etenkin keyness-arvon maksimiarvo on tilastollisena tunnuslukuna kapea-alainen, kun se sanan kaikista keyness-arvoista poimii vain yhden eli maksimiarvon.

Keyness-arvojen lisäksi kustakin tekstistä poimittiin vielä sekä niiden type/token-suhde (TTR)³ että virkepituuden keskiarvo. Näin kukin teksti saa viisi tilastollista tunnuslukua.

Avainsanoihin liittyvät arvot kuvaavat tekstin leksikaalista poikkeavuutta vertailukorpukseen nähden. Jos sen ajatellaan edustavan jonkinlaista normikielenkäyttöä, niin silloin korkeat indeksit merkitsevät poikkeamista normista. TTR taas liittyy sanastolliseen yksinkertaisuuteen tai luovuuteen: mitä korkeampi TTR-arvo, sitä leksikaalisesti monipuolisempi teksti. Virkepituus taas on yleensä liitetty tekstin luettavuuteen: mitä pitemmät virkkeet, sitä kuormittavampi työmuistin kannalta eli vaikealukuisempi teksti⁴.

Korpusanalyysiin valittiin kuusi tietokirjallista tekstiä, jotka käsittelevät Suomen poliittista historiaa. Mukaan otettiin sekä alkuperältään suomalaisia että suomennettuja venäläisiä tekstejä. Kaksi teksteistä on supisuomalaisia (Apunen ja Tarkka) ja neljä suomennettua (Bartenjev, Baryshnikov, Holodkovskij ja Komissarov). (Ks tarkemmin lähteet, tutkimusaineisto.)

Reseption arviointia varten kunkin tekstin alkupuolelta otettiin tasan 1000 sanan otos, jota käytettiin myös korpusanalyysissä. Näin sekä korpusanalyysi että reseption arviointitesti kohdistuvat tarkalleen samoihin tekstipätkiin.

4 Kotoisuuden ja vierauden arviointitesti

Käännösten kotoutuneisuuden⁵ ja vieraantuneisuuden (eli kotoisuuden ja vierauden) operationaalistaminen reseptiotutkimuksen avulla on toinen tavoitteemme. Suomalaisessa tutkimuksessa on useampaankin otteeseen koehenkilöiltä kysytty, tunnistavatko he suomalaisen tekstin suomennoksesta (Vehmas-Lehto 1989; Tirkkonen-Condit 2002; Jantunen 2004b), tosin aina se ei ole ollut päätavoitteena. Myös tähän arviointitestiin sisältyy tekstin luokittelu alun perin suomeksi kirjoitettuun ja suomennokseen, vaikka se ei päätavoite olekaan.

Pilotinomaisesti olemme järjestäneet tähän asti kaksi käännösten reseptioon liittyvää testiä, joissa lukija-koehenkilöitä on pyydetty lyhykäisestään asteikolla 1–5 arvioimaan sitä, kuinka kotoisaksi tai vieraaksi he tuhannen sanan tekstipätkät kokevat. Tähän väli-

raporttiin on koottu tulokset toisesta testistä, jossa koehenkilöinä toimi viisi kääntämisen syventävien opintojen vaiheen (4.–5. vsk.) opiskelijaa⁶. Lisäksi koehenkilöitä pyydettiin poimimaan juuri esimerkkitekstistä ainakin yksi sellainen piirre, joka heidän mielestään tekee sen kotoiseksi tai vieraaksi. Näin arviointitestiin sisältyi sekä kvantitatiivinen mittaus että kvalitatiivinen arviointi.

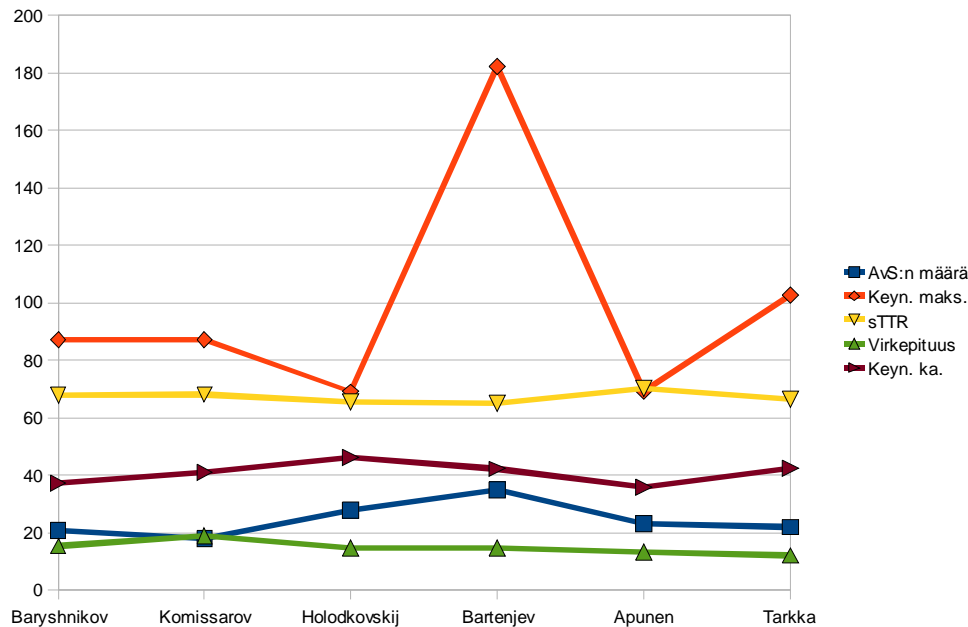
Arviointitestien lomakkeiden ohjeistus ja kysymykset/väittämät olivat seuraavat:

Lue oheiset historia-aiheiset tekstit, ja vastaa seuraaviin kysymyksiin ja väittämiin.

1. Arvioi, ovatko tekstit mielestäsi käännöksiä vai supisuomalaisia tekstejä?
2. Arvioi tekstejä niiden kotoisuuden ja vierauden mukaan asteikolla 1–5 (1 = kotoinen ja 5 = vieras).
3. Mainitse jokaisesta tekstistä yksi mielestäsi keskeinen piirre, joka tekee siitä kotoisen/vieraan oloisen?

5 Korpusanalyysin tulokset

Kahta alun perin suomeksi kirjoitettua historiateosta ja neljää historia-aiheista suomenosta verrattiin siis viiden tilastollisen ominaisuuden (virkepituus, TTR, avainsanojen määrä sekä avainsana-indeksin maksimi ja keskiarvo) suhteen niin, että niiden vertailukorpuksena käytettiin sanomalehtiaineistoa. Kaikkiaan keskeisin tulos lienee se, että eri piirteet korreloivat vain heikosti keskenään. Kaavio 1 kokoaa piirteet graafisesti.



Kaavio 1. Korpusanalyysiin sisältyvien ominaisuuksien (virkepituus, standardoitu type/token-suhde sekä avainsanojen määrä, maksimiarvo ja keskiarvo) keskinäinen vertailu teksteittäin.

Ensimmäiseksi voidaan todeta, että neljän suomennoksen ja kahden suomalaisen tekstin keskinäiseen kvantitatiiviseen vertailuun tekstien määrä on liian pieni. Ainoa selvästi havaittava ero on se, että kaikkien suomennosten virkepituuudet ovat suurempia kuin suomalaisten tekstien. Näin suomennosten voitaisiin olettaa arviointitestissä erottuvan vaikealukuisempina kuin suomalaisten tekstien. Sen sijaan esim. leksikaalisesti varioivin on Apunen (sTTR = 70,4), mutta sen jälkeen järjestys onkin Komissarov, Baryshnikov, Tarkka, Holodkovskij ja Bartenjev. Suomalaiset historiatekstit eivät mitenkään erityisesti erotu suomennoksista. Eri avainsanaindeksien suhteen voidaan todeta sama.

Toiseksi voidaan todeta, että suurimmalla osalla piirteitä ei ole korrelaatiota toisiinsa nähden. Voimakkaimmin (Pearson $p = 0,62$) korreloivat keskenään TTR ja avainsanaindeksin keskiarvo, mutta nekin käänteisesti: 62 % aineistosta toteuttaa sitä suhdetta, että mitä korkeampi TTR, sitä alhaisempi avainsanaindeksin keskiarvo on, eli mitä vaihtelevampaa teksti on leksikaalisesti ja siis mitä harvemmin kutakin sanaa käytetään, sitä pienempi on todennäköisyys, että tekstin sanat päätyvät avainsanalistalle. Tämä suhde on luonnollinen teoreettisesti asiaa tarkasteltuna, sillä sanat päätyvät avainsanoiksi korkean frekvenssinsä ansiosta, eikä tämä korrelaatio näin ollen kerro teksteistä mitään normaalista poikkeavaa. Hieman heikommin (Pearson $p = 0,51$) korreloivat keskenään avainsanojen määrä ja avainsanaindeksin maksimi-arvo: noin puolessa aineistossa avainsanojen korkeaa määrää seurasi myös avainsanaindeksin korkea maksimi-arvo. Myös tämä suhde on luonnollinen jo teoreettisestikin: mitä enemmän avainsanoja on, sitä korkeammaksi niiden keyness-arvot nousevat.

6 Arviointitestin ja korpusanalyysin vertailu

Lukija-koehenkilöiden arvioinneilla kuusi tekstiä yritettiin saada järjestykseen kotoisuuden ja vierauden suhteen. Erilaisten tilastollisten tunnuslukujen suhteen tekstit asettuivat aina samaan järjestykseen (ns. ranking; ks. taulukkoa 1).

Taulukko 1. Arviointitestin analyysin tunnusluvut (mediaani, vaihteluväli sekä suomalaisiksi tai suomennokseksi luokittelu), kun koehenkilöt arvioivat tekstien ominaisuuksia asteikolla 1–5.

Teksti	Arvioinnin med.	Vaihteluväli	Käännös/supisuom.
Tarkka	2	1–3	2/3
Apunen	3	1–3	1/4
Bartenjev	3	2–5	3/2
Holodkovskij	4	3–4	4/1
Komissarov	4	4–5	5/0
Baryshnikov	4	4–5	5/0

Testi toimi myös sikäli hyvin, että tekstien välille syntyi kategorisia eroja juuri suomalaisten ja suomennosten väliin. Ensinnäkin lukijat erottivat tilastollisesti erittäin merkittävästi toisistaan suomalaiset tekstit suomennoksista (Khiin neliö Yatesin korjauksella $p < 0,001$): oikeiden vastausten osuus oli 80 %. Tulos on parempi kuin Tirkkonen-Conditin (2002: 211) testissä (63,1 %), joka on hyvin samansuuntainen kuin Jantusen (2004b: 561) tulos (62,6 %); tosin testimme tulokseen liittyy monta varausta, mm. pieni koehenkilömäärä, tiukka tekstilajikategoria jne. Mutta myös mediaanien ja vaihteluvälien suhteen suomalaisten ja suomennosten ryhmien välille muodostuu selvä ero⁷. Toisin sanoen tämän otoksen lukijat pystyvät tilastollisesti luotettavasti erottamaan arvioinnissa alun perin suomeksi kirjoitetut tekstit suomennoksista.

Arviointitestin suhde korpusanalyysiin onkin huomattavasti ongelmallisempi tulkittava. Vahvin korrelaatio ($p = 0,59$) tekstien järjestyksen ja korpusanalyysin välille syntyi jälkimmäisen virkepitäisyyden suhteen: mitä pitempi virke, sitä varmemmin käännös. Tuloksen voi tulkita merkitsevän myös sitä, että kun virkepituus korreloi vaikealukuisuuden kanssa, niin käännökset koettiin nimenomaan vaikealukuisiksi.

Sen sijaan on todettava, että sTTR:n tai avainsanaisuuden kanssa arviointitestin tulos ei korreloi. Hypoteesi, jonka mukaan avainsanaisuus jollakin tavalla paljastaisi käännökset ei-käännöksistä, on tämän tuloksen valossa hylättävä.

Arviointitestin kvalitatiivinen osa, jossa koehenkilöiltä pyydettiin mainitsemaan vähintään yksi tekstiä kotouttava tai vieraannuttava piirre, tuotti mielenkiintoisempia tuloksia. Suurin osa maininnoista liittyi kieleen, mutta myös tyyliin ja sisältöön liittyviä piirteitä mainittiin. Taulukko 2 kokooa tulokset.

Taulukko 2. Kotoisaksi tai vieraaksi mainitut tekstien ominaisuudet ja mainintojen määrät.

Piirre	Kotouttava	Vieraannuttava
Tekstin näkökulma	2	10
Lauserakenne		10
Yksittäinen sana		4
Fraasi/kollokaatio	3	4
Sanajärjestys		2
Selittävä osa		1
Ortografia		1
Tyyli (sujuva/ei sujuva)	7	1
Mainintoja yhteensä	12	33

Mainintoja vieraannuttamisesta kertyi 33: kutakin käännettyä tekstiä kohti 8,25 ja kutakin koehenkilöä kohti 1,65. Mainintoja kotouttamisesta kertyi 12: kutakin tekstiä kohti 6,0 ja kutakin koehenkilöä kohti 1,2. Mainintoja vieraannuttamisesta kertyi siis merkitsevästi useammin (yksisuuntaisessa Khiin neliö -testissä $p < 0,01$), ts. tekstin vieraantuminen on näkyvämpi piirre kuin kotoutuminen. Pohdimme tämän eron merkitystä kotouttamisen ja vieraannuttamisen käsitteille päätännässä.

Vieraannuttaviksi ja kotouttaviksi nostettiin myös eri piirteitä. Näkyvimpinä vieraannuttavina piirteinä pidettiin tekstin asiasisältöön liittyvää näkökulmaa ja vierasta lauserakennetta. Ne saivat 61 % kaikista vieraannuttamista koskevista maininnoista. Sujuva tyyli taas mainittiin yleisimpänä kotouttavana piirteenä, ja se sai 58 % maininnoista.⁸

7 Päätäntä

Tähän väliraporttiin on koottu neljän suomennoksen ja kahden supisuomalaisen historiatekstin korpusanalyysi sekä sen tulosten vertailu kotouttamista ja vieraannuttamista mittaavan arviointitestin kanssa. Tutkimusprojektissa on tähän mennessä keskitytty menetelmien kehittämiseen; tätä pilottivaihetta täydennetään jatkossa suuremmalla koehenkilömäärällä. Haluamme kuitenkin nostaa esiin muutamia kvalitatiivisia tuloksia. Ensimmäinen ja tärkein on vastaus siihen kysymykseen, voiko käännökseen tunnistaa yhtäältä tekstin valittujen tilastollisten piirteiden ja toisaalta lukijoiden subjektiivista arviointia käyttävän testin pohjalta. Edelliseen kysymykseen on näiden testien pohjalta vastattava ”ei” ja jälkimmäiseen ”kyllä”. Tekstien valituilla tilastollisilla piirteillä, jotka mittaavat vaikealukuisuutta (virkepituus), sanaston monipuolisuutta (sTTR) ja sanastollista eroa ns. normista (avainsanaisuuden indeksit), ei suomennoksia voi erottaa mitenkään selvästi ja systemaattisesti suomalaisista teksteistä. Sen sijaan kysymällä lukijoilta – jotka tässä toki edustavat kääntämisen opiskelijoita eivätkä maallikoita –

arviointitestillä saadaan suomennosten ja suomalaisten tekstien välille selvä ero, ts. lukijat tunnistavat käännökset luotettavasti. Suomennosten ja suomalaisten tekstien keskeisiksi piirteiksi myös nousevat eri ominaisuudet.

On vaikea sanoa kokeilematta, erottaisivatko maallikot yhtä hyvin kuin kääntämisen opiskelijat (tai opettajat) suomennokset suomalaisista teksteistä; sitä eivät Tirkkonen-Condit (2002) tai Jantunenkaan (2004) ole testanneet. Sen sijaan on melkoisen varmaa, että maallikot eivät pystyisi likimainkaan yhtä hyvin luokittelemaan tekstejä kotoisiin ja vieraisiin, saati erittelemään niihin liittyviä piirteitä. Onkin tärkeää ymmärtää, että kysymyksellä maallikkolukijoista ei juuri ole merkitystä tämän tutkimusasetelman kannalta, ainakaan vielä; tässä on ensisijaisesti kyse kotoisen ja vieraan operationaalistamisesta ja sen problematiikasta.

Keskeinen tulos viittaa siihen, että tilastollinen korpusanalyysi ei tavoita tekstin käännösluonnetta yhtä hyvin kuin lukijoiden subjektiivinen arviointi. Luonnollisesti on niin, että valitut mittarit ratkaisevat koko mittauksen luotettavuuden, ja kannattaa muistaa, että tässä on kyse menetelmiä testaavasta pilottitutkimuksesta. Sen toteaminen ei kuitenkaan ole sama, kuin että olisimme valmiita hylkäämään ensimmäiset tulokset merkityksettöminä. Päinvastoin, haluamme pohtia niiden valossa keskeisimpiä käsitteitä, tekstien kotoisuutta ja vierautta.

Näyttäisi ensinnäkin siltä, että kotoisen ja vieraan käsitteet eivät ole tavanomainen rinnakkainen käsitepari, vaan niistä toinen, nimittäin *vieras*, on tunnusmerkkinen – lainataksemme strukturaalisen kielentutkimuksen vanhaa käsitettä. Se huomataan, se tunnistetaan ja sille osataan luetella erottavia piirteitä. Sen rinnalla *kotoinen* on tunnusmerkitön, neutraali lähtökohta, ja tunnusmerkkistä käsitettä kuvataan suhteessa siihen.

Se myös oikeastaan tarkoittaa sitä, että käsitepari *kotoinen* ja *vieras* ei olisi tiukasti dikotomia. Arviointitestissä (ks. taulukkoa 1) jokin teksti luokiteltiin yhtäältä suomalaiseksi, toisaalta käännökseksi, mutta siinä on kyse eri koehenkilöiden antamasta luokituksesta. Vaikka käsitteellisesti toinen sulkee pois toisen, niin haluamme jatkossa testata, voiko teksti olla yhtä aikaa sekä *kotoinen* että *vieras*. Tosin vieläkin luontevampi testauskohde on se, muodostavatko *kotoinen* ja *vieras* lukijoiden mielessä jonkinlaisen jatkumon, jolloin kyse ei ole dikotomiasta.

Toiseksi näyttäisi siltä – tai ainakin ajatus luontevasti syntyy tulosten pohjalta –, että tekstien tilastollinen tutkimus ei välttämättä suuresti tavoita tekstien sitä luonnetta, joka on lukijoille olemassa. Arviointitestissä (myös aivan ensimmäisessä, kääntämisen opettajille tehdyssä) kiinnitti ensimmäiseksi huomiota se, että kvalitatiivisesti esiin nostetut ominaisuudet ovat sellaisia, että niitä ei tilastollinen kuvaus ainakaan suoraan tavoita, nimittäin yksittäisiä sanoja, fraaseja, kollokaatioita, lauserakenteita ja tyylipiirteitä. Jälkikäteen ajateltuna se on luontevaakin: lukija havaitsee kerrallaan yhden lauseen, ja sen ominaisuudet hän tietoisesti tavoittaa; mieleen jäävät yksittäiset sanat, outo ortografia, ehkä tyylipiirteinä pitkät virkkeet, koska lukeminen alkoi tuntua vaikealta.

Nämä ajatukset haluamme tässä esittää jatkohypoteesien muodossa, ja käsityksiä on korjattava, jos tulokset eivät pidä paikkaansa laajemmilla otoksilla.

Lähteet

Tutkimusaineisto

Tekstiotteiden lähdemateriaali:

Apunen = Apunen, Osmo 1977. *Paasikiven–Kekkonen linja*. Ulkopoliittisen instituutin julkaisuja. Helsinki: Tammi.

Bartenjev = Bartenjev, Nikolai. & Komissarov, Vladimir 1977. *Kolmekymmentä vuotta hyvää naapuruutta*. Suom. Ulla-Liisa Heino. Helsinki: Otava.

Baryshnikov = Baryshnikov, Nikolai & Baryshnikov, Vladimir 2001. *Terijoen hallitus. Uusien asiakirjojen kertomaa*. Suom. Esa Adrian. Helsinki: Johan Bäckman Instituutti.

Holodkovskij = Holodkovskij, V.M. 1978. *Suomi ja Neuvosto-Venäjä 1918–1920*. Suom. Marja-Leena Jaakkola. Helsinki: Tammi.

Komissarov = Komissarov, Juri 1985. *Linja*. Helsinki: Otava.

Tarkka = Tarkka, Jukka 1987. *Ei Stalin eikä Hitler. Suomen turvallisuuspolitiikka toisen maailmansodan aikana*. Helsinki: Otava.

Referenssikorpus:

Karjalaisen korpus. 1990-luvun suomen sähköinen tutkimusaineisto. [‘Karjalainen newspaper corpus. The 1990’s Finnish computerized database.’] Koostanut Joensuun yliopiston yleisen kielitieteen oppiaine prof. Jussi Niemen johdolla. [SGML-muotoon saattanut Helsingin yliopiston yleisen kielitieteen laitos] Saanti: Kielipankki-palvelin, CSC – Tieteellinen laskenta Oy. <http://www.csc.fi/kielipankki>.

Tutkimuskirjallisuus

Boyden, Michael 2006. Language politics, translation, and American literary history. *Target. International Journal of Translation Studies*, 18 (1), 121–137.

Eskola, Sari 2002. *Syntetisoivat rakenteet käännessuomessa. Suomennetun kaunokirjallisuuden ominaispiirteiden tarkastelua korpusmenetelmillä*. Joensuun yliopiston humanistisia julkaisuja, 30. Joensuun yliopisto: Joensuu.

Eskola, Sari 2005. Lauserakenteiden käytön piirteitä suomennetussa kaunokirjallisuus-

nessa. – Ks. Mauranen & Jantunen (toim.) 2005.

Jantunen, Jarmo 2004a. *Synonymia ja käännessuomi. Korpusnäkökulma samamerkityksisyyden kontekstuaalisuuteen ja käännesskielen leksikaalisiin erityispiirteisiin*. Joensuun yliopiston humanistisia julkaisuja, 35. Joensuun yliopisto: Joensuu.

Jantunen, Jarmo 2004b. Suomennosten kieliasu puntarissa. *Virittäjä*, 4, 531–558.

Jantunen, Jarmo 2005: *Hyvin kummallista vai kovin hyvää?* – Esimerkkejä käännesssuomen epäkonventionaalista kontekstuaalisista valintasuhteista. – Ks. Mauranen & Jantunen (toim.) 2005.

Kemppanen, Hannu 2005. Avainsanoja, ideologiaa, ja lojaaliutta: venäjistä suomennettujen historiatekstien korpusanalyysi. – Ks. Mauranen & Jantunen (toim.) 2005.

Kemppanen, Hannu 2008: *Avainsanoja ja ideologia: käännessettyjen ja ei-käännessettyjen historiatekstien korpuslingvistinen analyysi*. Joensuun yliopiston humanistisia julkaisuja, 51. Joensuun yliopisto: Joensuu.

Kruger, Haidee 2009. The concepts of domestication and foreignization in the translation of children's literature in the South African educational context. Judith Inggis & Libby Meintjes (eds) *Translation Studies in Africa*. London/New York: Continuum, 161–179.

Marton, Klara & Richard G. Schwartz & Lajos Varkas & Valeriya Kastnelson 2006. Effect of sentence length and complexity on working memory performance in Hungarian children with specific language impairment (SLI): a cross-linguistic comparison. *International Journal of Language Communication Disorders*, 41 (6): 653–673.

Mauranen, Anna & Jarmo Jantunen 2005. *Käännessuomeksi. Tutkimuksia suomennosten kielestä*. Tampere Studies in Language, Translation and Culture, A1. Tampere University Press: Tampere.

Mauranen, Anna & Liisa Tiittula 2005. MINÄ käännesssuomessa ja supisuomessa. – Ks. Mauranen & Jantunen (toim.) 2005.

Mikka, Jaan 2008: Sentence length for revealing the cognitive load reversal effect in text comprehension. *Educational Studies*, 34 (2): 119–127.

Nevalainen, Sampo 2005. Köyhtyykö kieli käännessettäessä? – Mitä taajuuslistat kertovat suomennosten sanastosta. – Ks. Mauranen & Jantunen (toim.) 2005.

Opas, Lisa Lena 1995. New approaches to computer applications in literary studies. *Literary and Linguistic Computing*, 10 (4), 261–291.

Pedersen, Jan 2007. How is culture rendered in subtitles? Teoksessa: S. Nauer (ed.) *Challenges of Multidimensional Translation*. Proceedings of the Marie Curie Euro-

conferences *MuTra: Challenges of Multidimensional Translation* – Saarbrücken 2–6 May 2005.

http://www.euroconferences.info/proceedings/2005_Proceedings/2005_proceedings.html

Scott, Mike. *WordSmith Tools Manual Version 3.0*. Oxford: Oxford University Press. [Saatavissa osoitteessa; <http://lexically.net/wordsmith/version3/manual.pdf>, luettu 10.9.2010.]

Tirkkonen-Condit, Sonja 2002. Translationese – a myth or an empirical fact? *Target*, 14, 2, 207–220.

Tymoczko, Maria 2000. Translation and political engagement, activism, social change and the role of translation in geopolitical shifts. *The Translator*. Volume 6, Number 1, 23–47.

Vehmas-Lehto, Inkeri 1989. *Quasi-correctness. A Critical Study of Finnish Translations of Russian Journalistic Texts*. Neuvostoliittoinstituutti: Helsinki.

Venuti, Lawrence 1995. *The Translator's Invisibility. A History of Translation*. London & New York: Routledge.

¹ *Supisuomalaisella* tarkoitetaan käännöstutkimuksessa 'alkuperäisesti suomeksi kirjoitettua' tekstiä.

² Todennäköisyys on tässä kumuloituvaa arvoa, jonka ohjelma laskee käyttäen khiin neliö -testiä ja loglineaarista mallia.

³ TTR on näin samalla ns. standardoitu TTR, joka normaalisti *WordSmith Tools*issa lasketaan kullekin tekstilelle tuhannen sanan otosten perusteella, kun tässä käännösten tekstiotokset ovat juuri 1000 sanan pituisia.

⁴ Virkepituutta käytetään paljon yksinkertaisena tekstin vaikeutta ilmaisevana mittarina. Lukemisen kannalta tekstin virkkeillä on optimipituus, joka esim. englannissa on toistakymmentä sanaa (Mikkan (2008) mukaan 140 merkkiä), jota pitemmillä virkkeillä lukemisen ymmärtäminen vaikeutuu selvästi, kun virkerakenteen muut tekijät pysyvät samana. Käytännössä virkerakenteen kompleksisuus kuitenkin hankaloittaa lukemista monissa koejärjestelyissä virkepituutta enemmän, ja myös kielten välillä on eroja (Marton ym. 2006).

⁵ Englannin termin *domestication* suomentamisen taival on ollut vähintäänkin polveileva. Ensimmäinen suomennos oli *kotouttaminen*, ja hetken aikaa sen rinnalla eli *kotiuttaminen*, ennen kuin kielitoimisto ehdotti *kotoistamista*. Tutkijoiden käyttöön on kuitenkin ensimmäinen *kotouttaminen* jo ehtinyt vakiintua siinä määrin, että jatkamme tässä vastarantaista perinnettä kielitoimistoon nähden. Mutta kuten lukija artikkelistamme voi huomata, käyttämämme takaperoisjohdos *kotoinen* ja *kotoisa* luontevammin kyllä liittyisivät verbiin *kotoistaa*. *Vieras*-sanana johtosuhte verbiin *vieraannuttaa* on kokonaan toinen, joten siitä ei ole termiparilleen malliksi.

⁶ Ensimmäisessä arviointitestissä, joka toimi koko tutkimusasetelman pilottina, olivat arvioitavana vain neljä ao. suomennosta, ja sen koehenkilöinä toimi kuusi kääntämisen opettajaa.

⁷ Varsinaisen tilastollisen merkitsevyyden laskemiseen (esim. Khiin neliöllä) aineistossa on liikaa kategorioita, jotka saavat arvokseen 0 eli ei yhtään tapausta.

⁸ Ensimmäisessä arviointitestissä koehenkilöinä toimivat kääntämisen opettajat, ja sen tulokset ovat vieraannuttavien piirteiden suhteen hyvin samansuuntaiset, vaikka eivät olekaan täysin vertailukelpoiset tämän kanssa sen vuoksi, että mukana ei ollut suomeksi kirjoitettuja tekstejä ollenkaan. Niinpä ensimmäisessä testissä ei kotouttavia piirteitä mainittu juurikaan (2 mainintaa: tyyli ja lyhyet virkkeet); useimmin vieraannuttaviksi piirteiksi nostettiin yksittäiset sanat (11 mainintaa) ja lauserakenne (6), jonkin

verran myös sanajärjestys (3), tekstin näkökulma (2) ja ortografia (2). Ensimmäisen ja toisen arviointitestin eroista voisi muodostaa sen jatkohypoteesin, että sisällölliset seikat kiinnittävät huomiota vasta sitten, kun arvioitavana on sekä käännöksiä että supisuomalaisia tekstejä.