

Tekstistä saa tolkkua – toisinaan

Kielten konekäännös tekee tuhruista jälkeä

Nykyiset konekäännökset murjovat yhä pahasti suomen kieltä. Tavutusmuodot löytyvät jo aiempaa paremmin, mutta jäsentämisen puutteellisuudet heikentävät tulosta. Silti konekäännöksestä on toisinaan hyötyä.

TEKSTI: JUHO USKI
KUVA: ERIC LERAILLES

Tietokoneet osaavat kääntää kieliä aiempaa paremmin, mutta laatu ei ole parantunut samassa tahdissa kuin suorittimien teho. Konekäännöksen suuri ongelma ei ole enää tunnistumattomat sanat vaan kielellisen logiikan puute.

Koska konekäännöksen tekoäly on hiukan hakusalla, lähdetekstin kirjoittajan kannattaa auttaa konekääntäjää parhaansa mukaan.

Kone ei ymmärrä asiayhteyksiä assosiativisesti, joten kirjoittajan on pitädyttävä mahdollisimman yksinkertaisissa muodoissa. Käännöksessä on kuitenkin useita ongelmia, joihin ei löydy yksiselitteistä ratkaisua.

Esimerkiksi englannissa käytetään määrättyä ja epämääräistä artikkelia erottamaan subjektiivin ”yleisyyden aste”. Lisäksi tietyt lauserakenteet määrittävät artikkelin automaattisesti.

Suomen kieli ei käytä tällaisia määritteilyjä, joten jos lauseyhteydestä ei voida suoraan päätellä kummasta on kyse, on mää-

rittelyt ”vedettävä hihasta”.

SUOMESSA EI OLE SEKSIÄ

Suomea on väitetty vähiten seksistiseksi kieleksi maailmassa. Suomen kieli ei tee eroa taivutuksessa tai persoonapronomeineissa tekijän (subjektin) sukupuolen mukaan.

Suomessa genetiivi voidaan ilmaista vain sanan perään liitettävällä possessiivisuffiksilla, joka on monesta kirjoituksesta nykyään kadoksissa ”meidän koira”, ”minun kengät”. Koska tällainen muoto ei ilmaise omistajan sukupuolta, on se suomesta englantiin käännettäessä kehitettävä asiayhteydestä.

Nimiä käytettäessä tämä edellyttäisi melkoista tietokantaa nimistä, joiden sukupuoli voidaan tunnistaa. Ongelmia tuottavat myös mm. koirat ja kissat, joista brittienglannissa ensimmäisiä kutsutaan miespuolisella pronominilla, jos sukupuoli ei ole tiedossa, jälkimmäisiä taas kohdellaan oletuksena naaraina.

JÄSENTÄMISESSÄ ON JÄRKEÄ

Jotta lauseesta tulisi järkevää, se pitäisi myös osata jäsentää. Tätä varten ohjelman tulisi kyetä erottamaan verbit – ja juuri se predikaatti, eli persoonamuodossa taipunut verbi – substantiivit, adjektiivit ja muut sanaluokat toisistaan.

Varsinkin suomenkielisessä tekstissä sanojen taipumisen takia muodostuvat homonyymit tuottavat tässäkin tunnistamisessa ongelmia.

Siinä, missä indo-eurooppalaisissa kielissä ilmaistaan asiat prepositioilla, suomessa lisätään sanaan taivutuspäätteitä. Taivutuksen tunnistamisessa on edetty jo varsin hyvälle tasolle.

Suomi on kuitenkin eksoottinen kieli, ja maailmalta löytyy käännöskoneita, joiden tekijöillä ei ilmeisesti ole mitään ymmärrystä sanojen taivutuksesta.

Suomen kielen suhteen kannattaa luottaa kotimaisiin ohjelmiin. Tökkivinäkin ne ymmärtävät taivutusmuotoja käsittämättömän paljon paremmin kuin ulkomaiset viritelmät.

AMMATTILAISET EPÄILEVÄT

Hype on kuollut konekäännöksenkin osalta. Netin maksullisten konekäännöspalveluiden määrä tuntuu jopa vähentyneen, eikä niistä enää pidetä yhtä kovaa meteliä kuin kolmisen vuotta sitten.

Monet alan suuret nimet kuten IBM ovat suosiolla jättäneet omat konekäännösprojektinsa ja niiden tulokset julkistamatta. Taustalla on huoli siitä, millaisen kuvan kömpelö konekäännös antaisi yrityksestä.

Konekäännöksen näytteitä löytyy vielä jokunen internetistä, jopa kielenkäännöstoimistojen etusivuilta. Ainakin Translations.com on kuitenkin ottanut konekäännökseen hiukan nurinkurisen asenteen.

Yhtiöstä kerrotaan, ettei konekäännöstä käytetä edes apuna heidän käännösprosesseissaan, mutta etusivulla oleva surkea konekäännös on omiaan myymään todellisten, koulutettujen ihmiskääntäjien palveluita. Mielenkiintoista markkinointia.

DIREKTIIVIT KÄÄNTYVÄT KELVOLLISESTI

Tietokoneavusteiseen kieltenopiskeluun erikoistunut professori **John Bennett** (ks. MikroPC 11 / 1999, s. 17) ei ole tänä päivänä aivan yhtä pessimistinen kieltenopetuksen suhteen. ”Konekäännös kehittyi jatkuvasti”, hän toteaa, mutta vielä tällä hetkellä on pitkä matka sujuvan arkipäivän keskustelun yleiskäännökseen.

Konekääntäjät eivät selviä kulttuurisidonnaisesta tai erityisalojen tekstistä kovinkaan hyvin. ”EU-direktiivien kääntämiseen ne kyllä tuntuvat soveltuvan varsin hyvin”, Bennett toteaa.



Tämä joutune siinä, että direktiivien kirjoittamisesta on olemassa tarkat ennalta sovitut säännöt. Tällaiset säännöt helpottavat käännöstyötä, sillä konekääntäjä ”tietää mitä odottaa”.

Bennettin uumoilema kännykkään turisteja varten sijoitettu käännöspalvelu tuntuu vielä olevan aivan yhtä pitkän matkan päässä kuin kolme vuotta sitten.

MIHIN TARPEESEEN?

Konekäännöksellä on kuitenkin käyttökohteensa. Jos ei ole mitään käsitystä artikkelin aihepiiristäkään, on konekäännöksellä helppo selvittää, kannattaako artikkelin kanssa kääntyä ammattilaisten puoleen.

Konekäännöstä voi myös käyttää lähtökohtana selkeän kirjoittamisen opiskelulle. Jos kirjoittamasi manuaali kääntyy ymmärrettäväksi englanniksi konekääntäjän avulla, se epäilemättä on tarpeeksi helposi kirjoitettu.

Omaa järkeään ei kannata kuitenkaan aliarvioida. Vaikka konekäännökset tuntuvat usein hupaisilta ja kömpelöiltä, voi omaa alaansa käsittelevästä artikkelista helposti löytää aihealueen, jota artikkeli käsittelee.

Vieraastakin aiheesta voi saada jotain selkoa, ja on parempi ymmärtää hiukan kuin ei ollenkaan. Tosin samaan hengen vetoon on muistettava, että voi olla pa-

rempi olla ymmärtämättä kuin ymmärtää täysin väärin.

MYÖS KOTISIVUT KÄÄNTYVÄT

Netistä löytyvien ilmaisten käännöspalveluiden tavamerkinä tuntuu olevan kotisivuista tehtävät käännökset.

Pari vuotta takaperin näitä palveluita löytyi useampikin, nyt ne ovat käytännössä kutistuneet Systranin konekäännöksen avulla toteutettuun käännökseen (babel.altavista.com).

Kotisivujen käännöspalvelu on konekäännöksen hyötykäyttöä parhaimmillaan. Vieraskielisistä sivuista saa konekäännöksen avulla helposti hahmotettua sen, onko sisällössä jotain, joka houkuttelee tekemään sivujen kanssa lähempääkin tuttavuutta.

Tämä tietenkin edellyttää, että linkit ovat tekstilinkejä. Kvalinkeihin kirjoitettuun tekstiin eivät käännöskoneet pysy. Tosin, jos tekijät käyttävät viisaasti ALT-tageja kuvalinkeissä, apua löytyy tähänkin.

ENTÄ SE OIKEA KÄÄNNÖS?

Konekäännöksen koettua pientä inflaatiota on oikea kielenkäännös nostanut päätänsä. Harva yritys luottaa yksinomaan konekäännökseen dokumenttiansa käännöstyössä.

Varsinaisiakin käännöspalveluita löytyy maailmalta useita. Hinnat ja kielivalikoima vaihtelevat suuresti, eivätkä kummatkaan käy välttämättä ilmi yhtiön etusivulta.

Suomessa ammattikäännöstä edustavasti tekevä BGS Finland Oy kertoo, että dokumentti, jossa on 50 A4-sivua, kääntyy viikossa, parissa. Moisen hinta on 2 000–5 000 euroa. Kiireellisemmät työt, kielten eksoottisuus ja asiakkaan omat toivomukset muuttavat hintoja, joten tämä hinta on vain ohjeellinen.

Tätä artikkelia varten pyydettiin käännösnäytteet samoista pitkistä teksteistä, joita käytettiin yleiskuvan saamiseksi testattavien konekääntäjien tasosta. Ainakin toistaiseksi ihmiskääntäjien ei tarvitse olla huolissaan. Eroa on yhä kuin yöllä ja päivällä.

Suomi on maailman kielten joukossa eksoottisimmasta päästä. Suomalaisen kannattaakin luottaa toiseen suomalaiseen käännöstyössä, käyttipä konekäännöstä tai oikeita käännöspalveluita. Viestin perilleminen mahdollisuus kasvaa, kun asialla ovat vaikeaa kieltämme äidinkielenään käyttävät asiantuntijat. ■

SANASTO

Konekääntäjä: Machine translation (MT). Kielenkäännösaunomaatti, jonka toiminta perustuu laajaan sanastoon ja taivutussääntöihin. Konekäännös ei välttämättä toimi kumpaankin suuntaan kahden kielen välillä.

Käännösmuisti: Translation Memory (TM). Käännös, joka perustuu koneelle valmiiksi opettuihin kielen rakenteisiin. Toimii kaksisuuntaisesti.

Morfologia: Kielen sanojen taivutussäännöt.

Synonyymi: Samaa tarkoittava sana, jolla voi olla eri merkitysvaihtoehto.

Homonyymi: Samalla tavalla kirjoitettava sana, jonka merkitys on erilainen.

Syntax 9.0 Classic Plus vs. TranSmart

Omaan koneeseen asennettavista konekääntäjistä testattiin kaksi, Blue Ball -yhtiön Syntax 9.0 Classic Plus ja Kielikoneen TranSmart. Kolmatta markkinoilla olevaa tuotetta ei testiin saatu, mutta Eurotranslator onkin enemmän sanakirja kuin käännösohjelma, ja sanakirjatestin yhteydessä mainitun ohjelman käännöstyö oli pelkkää sanakirjakäännöstä.

Paljoa silkasta sanakirjakäännöksestä eivät testiin osallistuneetkaan ohjelmat poikenneet. Lähestymistavat ja käyttöliittymä vain poikkesivat jonkin verran toisistaan.

Molempien ohjelmien painopisteenä tunnetaan käännöstyön nopeus paremminkin kuin täsmällisyys.

Loppukäännöstyötä tarkistavia algoritmejä parantamalla voisi käännösaika toki moninkertaistua, mutta kukaan ei tunnu huomanneen, että hyvä käännös viidessä minuutissa voi olla parempi vaihtoehto kuin surkea muutamassa sekunnissa.

Syntax 9.0 Classic Plus

Syntax 9.0 on tarkoitettu työasemaympäristöön. Toisin kuin TranSmart, Syntax 9.0:ssa käännöstä voi tehdä myös englannista suomeen päin. Tämä vastaa huomattavasti suurempaan tilaukseen kuin pelkkä yksisuuntainen käännös.

Ohjelman käyttöliittymän kieli on valittavissa, toisin kuin TranSmartissa.

Syntax on tehty perustekstin muotoon ja työkaluriveiltä löytyvätkin tavallisimmat tekstin ulkoasun muokkaamiseen tarvittavat työkalut. Silmiinpistävä erona tekstinkäsittelyohjelmiin on, että ohjelman päänäyttö on jaettu pystysuunnassa kahtia.

Vasemmalle puolelle tulee suomenkielinen teksti, oikealle englanninkielinen. Kumpaa tahansa ruutua voi käyttää lähdeaineiston syöttämiseen.

Käännöstyö on varsin nopeaa, ja sen edisty-

mistä lause lauseelta on helppo seurata.

Järjestelmä on sinänsä sovellusriippumaton, lähdeaineistoksi kelpaavat tavalliset tekstitiedostot ja rtf-tyyppiset tiedostot, jolloin yhteensopivuus tavallisten perustekstureiden kanssa on taattu. Lisäksi tekstin voi suoraan kirjoittaa ohjelman päänäyttöön.

AIHEALUEEN VOI MÄÄRITELLÄ

Ohjelma ei ilmoita käännösvirheitä selkeästi. Osa jää suoraan kääntämättöminä sanoina tekstiin, osa sentään erotetaan suluilla. Käyttäjälle olisi mukavampaa huomata virheet helpommin. Varsinkin pitkistä tekstistä voi virheitä olla hankala havaita.

Käännöksen tuloksista löytyy ohjelmassa pieniä mutta kiusallisia oikeinkirjoitusvirheitä. Varsinaista ongelmaa näistä ei tule, mutta tällaiset mokat nakertavat käyttäjän luottamusta ohjelman toimivuuteen.

Ohjelma sisältää mahdollisuuden määrittää

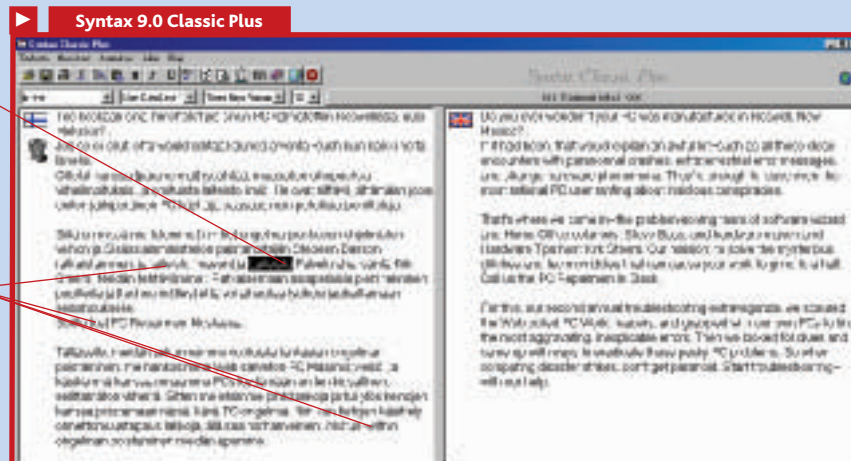
Syntaxin käännöksen tarkkuus ei ole parhaasta päästä: laiteisto pitäisi olla laitteisto.

Ohjelma ei korosta tarpeeksi selvästi erilaisia käännöksessä tapahtuneita virheitä, eikä erottele niitä kyllin selvästi.

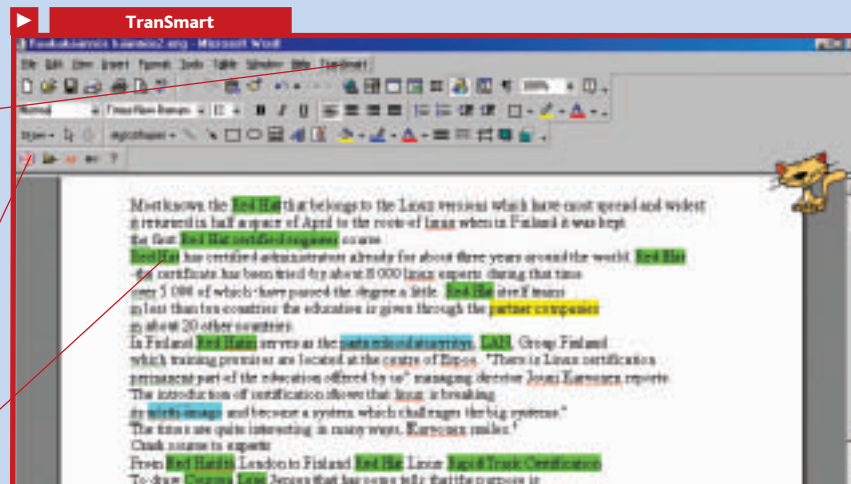
TranSmart luo oman valikon Wordiin, jota kautta kielenkäännöksen ominaisuuksia pääsee muokkaamaan.

TranSmart tekee Wordiin myös oman työkalurivin, jolla käännöksen voi helposti käynnistää.

Käännöksessä ilmenneet erilaiset ongelmat ohjelma ilmoittaa väreillä.



Syntax 9.0 Classic Plus käyttää omaa käyttöliittymää, jossa käännettävä teksti ja käännöksen tulos näkyvät kumpikin eri puolella ruutua. Tavaliseenkin tekstin muokkaamiseen löytyy ohjelmasta itsestään paljon ominaisuuksia.



TranSmart-ohjelman toiminta integroituu osaksi Microsoftin Word-textruria. Ohjelma tukee Wordin versioita 6.0, 7.0, 97 ja 2000. Käännösohjelman sitomisen toiseen sovellukseen pakottaa ihmisen tämän sovelluksen käyttöön.

tekstin aihealue, jolloin voi yrittää saada tarkempia tuloksia mm. tietotekniikan, lääke- tai lakitieteellisestä tekstistä. Useimmiten käännös ei kuitenkaan tunnu paljoa eroavan yleisasetuksella tehdystä käännöksestä.

Lisäominaisuutena ohjelmassa on internet interface -käännösohjelma, jonka avulla on mahdollista saada ainakin aihe selville englanninkielisten sivujen sisällöstä.

Ongelmana tässä palikassa on huono käytölliittymä. Olisi mukavaa, jos www-sivuston teksti saisi suoraan ja automaattisesti lähdekieleltä toiselle, näin ei kuitenkaan ole, vaan käyttäjän on nähtävä itse vaivaa siirtääkseen sivuista osiakin ohjelmalle käännöstyötä varten.

Syntax 9.0 Classic Plus maksaa 260 euroa, lisätietoja www.blueball.fi ■

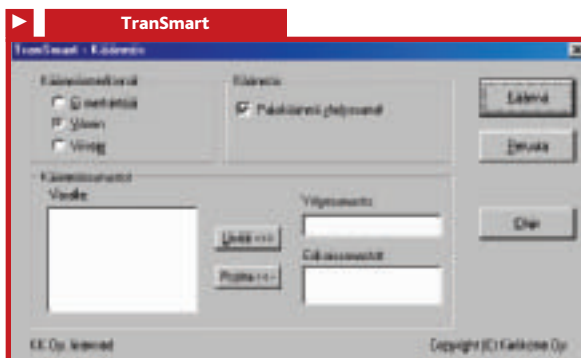
▶ TranSmart

Kielikone Oy:n TranSmart löytyy niin verkkopalveluna kuin omalle koneelle asennettavana pakettinakin. Ohjelma on selvästi suunnattu pieniin verkkoympäristöihin eikä pelkkään kotikäyttöön.

Ohjelmasta on tehtävä kaksi asennusta, ja jos käytössä on vain yksi kone, on nämä molemmat tehtävä samaan koneeseen. Palvelinasennus asentaa TranSmart-käännöspalvelimen, joka jää työasemassa taustalle odottamaan käännöstyötä.

Työasemapalikka lisää MS Wordin käyttöliittymään uuden työkalurivin, jolta käännös suomesta englantiin voidaan käynnistää.

Palvelinosa asennetaan oletusarvoisesti käynnistymään automaattisesti käyttöjärjestelmän käynnistyessä. Palvelimen käynnissäolon huomaat paitsi pienestä ikonista windowsin tehtävälaskissa, myös siitä, että se tuntuu hidastavan järjestelmän toimintaa.



TranSmartin käännöksen alkaessa käyttäjä voi valita, miten käännösvirheet ilmoitetaan, miten yhdyssanat käännetään ja valita omia sanastoja dokumenttikohtaisesti.

Kotikoneeseen ohjelmaa asennettaessa kannattaakin valita asennusvaiheessa jokin muu kuin Käynnistä-ryhmä (startup) ohjelman pikakuvakkeen asennuspaikaksi, tai siirtää se sieltä pois jälkikäteen ja käynnistää palvelin vain käännöstä tarvittaessa.

PARI SIVUA KÄÄNTYY PARISSA SEKUNNISSA

Sitouttaminen Microsoftin Wordiin on kaksi-piippuinen juttu. Toisaalta Word on pc-maailmassa epäilemättä käytetyin tekstinkäsittelyohjelma, mutta toisaalta tällainen valinta vähentää käyttäjän valintoja.

Paras vaihtoehto olisi ollut, että käännöstyöhön olisi tarjottu sinänsä näppärän työkalurivin lisäksi ohjelman oma käyttöliittymä, johon tekstin olisi myös voinut siirtää.

Käännökset sujuvat sukkelaasti, parisivuinen dokumentti kääntyy muutamassa sekunnissa ja useita kymmeniä sivujakin pitkä teksti vie vain jokusen minuutin.

Jos käännös tökkii sanoista, joita ohjelma ei tunnista, on asiakirjakohtaisesti helppo lisätä uusia sanoja sinänsä jo laajaan sanastoon.

Ohjelmalla on helppo tehdä käännöksiä, joista tekstin sisällön runkoajatus käy selville. Näitä käännöksiä voi käyttää pohjana lopulliselle käännökselle.

VAIN SUOMESTA ENGLANTIIN

Vaikka TranSmart oli työasemaan asennettavista kääntäjistä testissä paremmin menestynyt, ei moisesta saavutuksesta vielä voi ylpeillä.

Lopulliseen sujuvasti luettavaan ja kauttaaltaan ymmärrettävään käännökseen on vielä pitkä matka. Lisäksi ohjelma kääntää vain suomesta englanniksi.

Ohjelman ohjeet ovat kattavat, ja sisältävät myös ohjeet siitä, kuinka voi kirjoittaa helpommin käännettävää tekstiä. Näitä ohjeita noudattamalla on mahdollista vähentää käännösvirheitä.

TranSmart löytyy myös web-palveluna osoitteesta <https://websmart.kielikone.fi/>. Palvelu on maksullinen, mutta sitä pääsee kokeilemaan helposti.

Tämä web-palvelu antoi toisinaan selkeästi parempia käännöstuloksia kuin omalle koneelle asennettava versio. Ilmeisesti webissä olevaa versiota on helppo pitää jatkuvasti ajan tasalla.

Www-palvelun kone käänsi suomea englantiin ehdottomasti parhaiten testin ohjelmista. Ymmärrettävää käännöstä aikaansaadakseen pitää tekstin tuottajan kuitenkin tässäkin tapauksessa skarpata paljon. ■

▶ KIELENKÄÄNNÖSPALVELUITA

KOTISIVU	TYYPPI	HINTA	KIELET	MUUTA
www.dagostini.it/translationservice	Tarkistettu konekäännös	9,33\$/100 sanaa	eng/esp/ger/fra/ita	
www.tranexp.com	Konekäännös/websivut	Ilmainen	n. 30	Jälki huonoa
www.foreignword.com/technology/mt/mt.htm	Linkkilista eri käännöspalveluihin/ohjelmiin	Ilmainen		Osin vanhentunut
websmart.kielikone.fi	Konekäännös	Laskutus sopimus	suomi>englanti	
http://translator.dictionary.com/	Konekäännös/websivut	Ilmainen	ger/eng/esp/por/fra	Systranin kone
http://babel.altavista.com/	konekäännös/websivut	Ilmainen	ger/eng/esp/por/fra/kor/jap/chi/rus	Systran
www.translations.com	Konekäännös/websivut/ihmis	Kone ilmainen, ihmis sop.mukaan	Kone: ger/eng/esp/por/fra	Systranin kone
www.worldlingo.com/	Ihmiskäännös	Sopimuksen mukaan		
www.mendez.fi/	Ihmiskäännös	Sopimuksen mukaan	Useita!	*
* Bowne Global Solutions = ent. Mendez				

Happotesti ja tulokset

1. Haavasta vuosi verta.
2. Ensi vuosi on erilainen.
3. Soita ja tilaa!
4. Suomesta löytyy soita ja tilaa.
5. Suihkukone jätti jäljen taivaalle.
6. Jätti laski taakkansa maahan.

SYNTAX 9.0 CLASSIC PLUS

Year blood from the cut.
Next year is different.
Marshes and orders!
From Finland can be found ring and orders.
Jet left to the sky of the mark.
Left to calculate his burden to country.

TRANSMART JA WEBTRANSMART

From wound year of blood.
The next year is different.
Ring and order!
In Finland swamps are found and order.
The jet left the trace on the sky.
Its burden left, it calculated to the country.

Emästesti ja tulokset

1. Verta vuoti haavasta.
2. Ensi vuosi tulee olemaan erilainen.
3. Soita ja tilaa!
4. Suomesta löytyy suomaata ja avaruutta.
5. Jättiläinen pudotti taakkansa alas.
6. Tyttö kampasi hiuksiaan.

SYNTAX 9.0 CLASSIC PLUS

Blood leaked from the cut.
Next year will come to be different.
Marshes and orders!
From Finland can be found marshland and space.
Giant dropped his burden down.
Girl your comb his hairs.

TRANSMART JA WEBTRANSMART

Verta leaked from the wound.
The next year will be different.
Ring and order!
In Finland a marshland and space are found.
The giant dropped its burden down.
The girl combed her hair.

▶ KÄÄNNÖSTEN ARVIOINTI

HAPPOTESTI	Syntax 9.0	TranSmart
Lause 1	vuosi tunnistui substantiiviksi "the"puuttui	vuosi tunnistui substantiiviksi
Lause 2	futuraa ei havaittu	futuraa ei havaittu
Lause 3	ok	ok
Lause 4	"tilaa" tunnistui verbin imperatiiviksi	"tilaa" tunnistui verbin imperatiiviksi
Lause 5	täysin sekava	the trace p.o. "a trace"
Lause 6	täysin sekava	täysin sekava
EMÄSTESTI	Syntax 9.0	TranSmart
Lause 1	Hyvä!	verta ei tunnistanut enää
Lause 2	Hyvä, mutta hiukan sekava.	ok
Lause 3	ok	ok
Lause 4	Lähes hyvä. "In Finland" ei from "	"a marshland" – ei artikkelia!
Lause 5	ok (the vai ei - his vai its mielipidekysymys)	ok
Lause 6	Sekava	ok

Mikä meni vikaan?

Käännökoneille suoritettiin ensin happotesti, johon ujutettiin tarkoituksellisesti homonymyjä. Tarkoituksena oli tarkistaa, pystyvätkö kääntäjät huomaamaan asiayhteydestä, mitä sanaa haetaan tai tunnistamaan sanaluokan. Myös aikamuotojen löytämistä testattiin.

Emästestissä oli tarkoitus osoittaa, että oikean käännöksen löytäminen koneellisesti on mahdollista, jos luovasti osataan käyttää synonyymejä. Lisäksi aikamuotoja kierrettiin kiertoilmauksilla.

Sekä happo- että emästestissä TranSmart ja sen www-versio WebTranSmart antoivat yhtäläiset tulokset. Kaikkien testattavien ohjelmien käännöstarkkuus pa-

rani, kun kirjoitettiin mahdollisimman yksiselitteistä tekstiä.

Virheitä tuli kuitenkin kummassakin testissä, ja ohjelmat pärjäivät eri lauseissa vuorotellen. TranSmartin yleisarvosana on kuitenkin parempi kuin Syntax 9.0:n.

Käännökoneiden testin aikana vakavimmaksi puutteeksi osoittautui koneiden kyvyttömyys jäsentää lauseita.

TAIVUTUKSESTA TOLKKUA

Äidinkielenopettajani kertoi aikoinaan yläasteella, että suomen kielen lauseen tunnistaa siitä, että siinä on persoonamuodossa taipunut verbi.

Niinpä maallikon silmin tuntuisi järkevältä, että ennen kuin koneen annetaan etsiä sanojen eri taivutuksia, etsittäisiin lauseesta ne verbit.

Jos moisia ei löydy, voisi sen jälkeen turvautua yhä sana sanalta käännökseen. Samoin lopputuloksesta voisi tarkistaa, löytyykö lauseesta verbiä.

Tällöin välttyttäisiin edes osin happo- ja emästestissä esiintyneiltä virheiltilä, joissa verbin preesensmuoto tunnistuu substantiiviksi (kampasi: yksikön ensimmäinen persoona, preesens verbistä kammata, mutta myös substantiivi kampa, jossa kammam omistajana on yksikön toinen persoona). ■

Babel Fish ei osaa suomea, mutta käännös vaikkapa kiinasta englantiin voi olla iso apu. Näyte on osoitteesta www.wushu.com.cn.



Näin helpotat käännöstä

1. Vältä tuotemerkejä. Jos teksti on täynnä erilaisia tuotemerkejä, vaikeutuu käännösohjelman työ huomattavasti. Korvaa tuotemerkit jollain selkeällä merkijonolla, joka ei ole mikään sana ja lisää oikeat tuotemerkit jälkeensä tekstiin oikeassa muodossaan.

2. Käytä yleiskieltä. Murteiden ja slangin kääntäminen ymmärrettävästi on usein vaikeaa ihmisillekin. Murteellinen kieli vaikeuttaa kääntämistä tunnistamattomien sanojen ja outojen kielen rakenteiden vuoksi.

3. Käytä lyhyitä lauseita. Mitä pitempi lause, sen monimutkaisempi se on kääntää. Käytä mahdollisuuksien mukaan vain päälauseita, vältä varsinkin peräkkäisten, toisilleen alisteisten sivulauseiden käyttöä. Pronominien viitetaussuhteet ovat usein vaikeita ihmisillekin.

4. Varo kulttuurisidonnaisia ilmauksia. Erilaiset sanonnat eivät useinkaan käänny ymmärrettävästi. Hyvätkin vitset ovat usein kulttuurisidonnaisia, ja niiden ymmärtämiseen ei pelkkä hyvä kielentuntemuskaan riitä. Lisäksi monilla erityisaloilla on ilmaisuja, joiden kääntäminen jopa saman kielen sisällä alaa tuntemattomalle on hankalaa. Tällaisten ilmaisujen oikea kääntäminen vaatii usein erikoissanaston ja rakenteiden opettamista käännösohjelmalle.

5. Älä tee peräkkäiskäännöksiä. Käännöstä ei kannata suoraan kääntää toiselle kielelle. Tällaisella operaatiolla moninkertaistaa tekstissä esiintyvien virheiden määrää. Käännetyt tekstin oikeellisuus pitää tarkistaa mahdollisuuksien mukaan jokaisen käännöksen jälkeen. Jos haluat kääntää useille eri kielille, käytä kaikkien lähteenä mahdollisuuksien mukaan samaa alkukielistä tekstiä.

6. Käytä synonyymejä, vältä homonyymejä. Käännöskoneen on mahdollista sotkea kaksi samalla tavoin kirjoitettavaa mutta eri asiaa tarkoittavaa sanaa keskenään. Tällöin on järkevää korvata sana jollain synonyymillä, eli ulkoasultaan erilaisella, mutta merkitykseltään samanlaisella sanalla.

7. Kirjoita idiootille. Muista aina, että käännösohjelmat ovat varmasti tyhempinä kuin yksikään ihmiskääntäjä. Pyri kirjoittamaan niin, ettei väärinkäsityksille ole mahdollisuuksia.

8. Muista tarkistaa käännös. Älä jätä käännöstä koneen varaan. Jos käytössäsi on käännöksen kohdekieltä huonosti osaava henkilö, lueta teksti hänellä. Ihminen antaa usein anteeksi moniakkin virheitä. Tarkista myös epäilyttävät kohdat sanakirjan kanssa. ■

Ota tolkkouudosta kielestä

Jos on tarve selvittää jokin teksti aivan ummikkona, automaattikäännöskin auttaa. Ainakin sen avulla voi arvioida, kannattaako teksti käännättää oikeasti.

Systranin käännösaunomaatti osoitteessa <http://babelfish.altavista.com> tuntee eksoottisikielisiä. Tällä hetkellä valikoimassa on molemmat käännösuunnat englannin ja näiden kielten välillä: espanja, italia, japani, kiina, korea, portugali, ranska, saksa. Lisäksi on käännökset saksan ja ranskan välillä sekä venäjämä englantiin.

Vaikka suomi ei ole tuettuna, käännös englantiin voi riittää monessa tapauksessa.

Sivustossa voi käännättää enintään 150 sanan mittaisia tekstinpätkiä tai kehyskellisiä nettisivuja. Tekstinpaljoja voi siirtää kääntäjään leikepöydän kautta.

Koekäännökset kiinasta tuottivat jokseenkin ymmärrettävän tuloksen, mutta japanin rönnyilevät rakenteet saavat automaatin pyörälle päästään. ■

MIRJA PAATERO

ASIAN YDIN

Kone kääntää luonnollista kieltä kömpelösti

Moni hyötyisi valtavasti siitä, jos saisi tietokoneensa hetkessä kääntämään tekstin kuin tekstin laadukkaasti jollekin toiselle kielelle. Tietotekniikan ja kielitieteen kehityksestä huolimatta tämä haave ei ole vielä toteutunut.

Pientä apua on sentään saatavilla jo nykyisistäkin käännösohjelmista ja www-palveluista. Ammattitaitoisen ihmisen tekemän käännöksen ja koneen tekemän käännöksen välinen ero on kuitenkin yhä valtava: ihminen päihittää koneen 6–0.

Jos haluaa kääntää koneellisesti tekstiä suomen kielestä tai suomen kieleen, kannattaa valita kotimaisia tuotteita. ■