

**TARTU UNIVERSITET**  
**Institutionen för germansk, romansk och slavisk filologi**  
**Avdelningen för skandinavistik**

**Översättning från svenska till estniska**  
**med hjälp av Google Translate**

**Kertu Hiire**

**Bacheloruppsats**  
**VT 2013**  
**Handledare: Maiu Elken, MA**

# Innehåll

Förklaringar för förkortningar .....	3
1. Inledning .....	4
1.1 Syfte och frågeställningar .....	4
1.2 Avgränsning .....	4
2. Bakgrund .....	5
2.1 Maskinöversättningens historia och typer .....	5
2.2 Google Translate och statistisk maskinöversättning .....	6
2.4 Andra undersökningar .....	8
3. Material och metod .....	9
3.1 Texter .....	9
3.2 Metod .....	11
4. Resultat .....	14
4.1 Text 1 – lärobokstexten .....	14
4.2 Text 2 – nyhetstexten .....	17
4.3 Text 3 – den skönlitterära texten .....	21
4.4 Sammanfattning av felen .....	23
5. Analys .....	26
6. Sammanfattande diskussion .....	28
Litteraturförteckning .....	30
Bilagor .....	33
Bilaga 1 .....	33
Bilaga 2 .....	34
Bilaga 3 .....	35
Resümee .....	37

## Förklaringar för förkortningar

GT – Google Translate

MT – Maskinöversättning (från engelskans *Machine Translation*)

GT-översättning – översättningen som Google Translate skapar

KT – källtexten

Ö – översättningen som Google Translate skapar från källtextens mening (används vid exemplen)

Korrekt – min översättning som mänsklig översättare. Det kan finnas flera korrekta översättningar av en mening. Vid exemplen har jag valt att visa den som mest liknar GT-översättningen (Ö).

# 1. Inledning

Denna uppsats kommer att handla om Google Translate (GT) som hjälpmedel för att översätta från svenska till estniska. Google Translate är som ett magiskt program som översätter hela meningen eller även hela texten med bara ett klick. Och just därför är det så populärt bland studenter. Men som hos alla goda saker finns det brister även i Google Translate. Mitt intresse för att skriva om dessa brister väcktes när vår svenska lektor vid Tartu Universitet förbjöd oss att använda det när vi skulle skriva texter för en kurs. Men det fanns alltid någon som hade gjort detta och lektorn visste det i alla gånger. Och det gav mig idén att undersöka om det finns några vanliga fel som alltid förekommer när man använder Google Translate.

## 1.1 Syfte och frågeställningar

Syftet med undersökningen är att visa hur många fel det förekommer när man använder Google Translate för att översätta hela texten från svenska till estniska. Dessutom ska undersökningen presentera vilka språkliga fel som är vanliga och mest frekventa i texterna och erbjuda några förklaringar för dem baserade på olikheter mellan estnisk och svensk grammatik. Eftersom det finns tre olika typer av text som används i undersökningen vill undersökningen även antyda på vilken grad man kan använda Google Translate för att förstå vad texten skriver om. Följande frågor vill jag besvara i undersökningen:

- Vilka språkliga fel förekommer i översättningen från svenska till estniska?
- Hur många fel förekommer i olika typer av text?
- Vilka feltyper är vanliga?
- Varför finns det fel i GT-översättningar?

## 1.2 Avgränsning

Även om det skulle vara intressant att undersöka både översättningar från svenska till estniska och tvärtom är mitt intresse att bara visa fel i det första fallet. De som studerar svenska som främmande språk använder Google Translate för båda: för att skriva en text på svenska och också för att förstå vad texten talar om. Men det finns ju också många andra som använder Google Translate för att förstå vad texten som är skriven på

svenska innebär på estniska. Och deras antal är definitivt större än studenter, som redan läser svenska. I min undersökning läser jag först alla GT-översatta meningar som en vanlig människa, som vill begripa vad texten handlar om, som om den är översatt av en mänsklig översättare. Sedan hittar jag alla språkliga fel som finns i GT-översättningen.

## **2. Bakgrund**

### **2.1 Maskinöversättningens historia och typer**

Maskinöversättning (MT) innefattar översättning från ett språk till annat med hjälp av en maskin som vanligen är datorn. Enligt John Hutchins kan någon idé av maskinöversättning anteckna två patent från 1933: en av Georges Artsrouni och den andra av Petr Trojanskij. Artsrounis patent var för en maskin som också skulle fungera som ett mekaniskt flerspråkigt lexikon. Trojanskij's patent var också liksom Artsrounis, ett mekaniskt lexikon, men han hade också några idéer om kodering och användning av allmänna symboler för att tolka grammatiska funktioner. (Hutchins 2006:375)

John Hutchins skriver att en mer realistisk idé av maskinöversättare kan vara Warren Weavers memorandum i 1949 när han stimulerade början av maskinöversättningsforskning i USA. Den första maskinöversättare demonstrerades 31:e maj 1949 när The New York Times hade “‘Electric Brain’ able to translate foreign languages is built” som huvudrubrik. (Hutchins 1997: 195-196)

1960-talet inställde maskinöversättningsforskning eftersom Automatic Language Processing Advisory Committee (ALPAC) gjorde analyser för maskinöversättningens kvalitet och påstod att det skulle bli bättre om man gjorde ordböcker och andra hjälpmedel för mänskliga översättare (Hutchins 2006: 377). Men maskinöversättningens undersökning fortsatte ändå och det gjordes många nya program och uppdateringar. Dess betydelse i vår tid är även större än det var när man först gjorde maskiner för att hjälpa människor med främmande språk.

Enligt Ping kan man generellt uppdelat maskinöversättning i två: korpusbaserad och regelbaserad. De första maskinerna var regelbaserade – basen för översättningen var språkets regler. Det användes två olika steg för att göra översättning: direkt och indirekt. System som utvecklades före 1980-talet använde direkt synsätt. (Ping 2009: 162)

Anna Sågvall Hein klargör att man använde översättningslexikon för att översätta ord eller fraser och därtill användes olika regler för att översätta korrekt. Tyvärr gjorde man ingen meningsanalys och det kom några fel eftersom maskinen hade svårigheter att skilja mellan till exempel mellanobjekt och subjekt. (Hein 2005: 25)

På 1980-talet dominerade ett mer förfinade sätt vid MT – indirekt maskinöversättning. Enligt Routledge Encyclopedia of Translation Studies kan indirekt maskinöversättning delas in i två olika kategorier: transfer och interlingua MT. Den första kategorin består av tre olika steg: först analyseras källtextens syntaktiska struktur till en formell betydelseavbild, som fortfarande behåller källtextens djupstruktur, sedan kommer överföringsfas (transferfas), där denna formella betydelseavbild överförs så att den behåller målspråkets djupstruktur, och sist kommer generering, där en målmening skapas från denna betydelseavbild. (Ping 2009: 162-163)

Den andra kategorin av indirekt MT är interlingua MT som betyder att källspråkets abstrakta betydelseavbild är konstruerad med hjälp av ett slags mellanstruktur (interlingua), d.v.s. ett avbild oberoende av käll-/målspråket, från vilket måltexter på olika språk kan skapas. Den innebär att översättningsprocessen består av två steg: analysator omvandlar källtexten till interlingua och generator omvandlar den till måltext. (Ping 2009: 163)

Den andra lösningsmodellen för maskinöversättning – korpusbaserad MT – blev populär sedan 90-talet. Den kan också delas in i två: exempelbaserad MT och statistisk MT. Korpusbaserad MT använder stora mängder av översatta texter d.v.s. parallella korpusar. Exempelbaserad MT innebär matchning av input fraser med likvärdiga eller likartade källspråks fraser i en korpus av parallella tvåspråkiga texter, utdragning av målspråkets likvärdiga eller likartade fraser, och – den mest komplicerade processen – sättning de utvalda målspråkets fraser samman för att skapa en flytande och grammatiskt korrekt text (Hutchins 2006: 380).

## **2.2 Google Translate och statistisk maskinöversättning**

Statistisk maskinöversättning är den andra typen av MT där stora korpusar är använt. Hein beskriver den statistiska varianten av maskinöversättning såhär:

Det specifika med metoden är att översättningslexikonet skapas automatiskt genom datamaskinell bearbetning av stora mängder översatt text. Käll- och måltext länkas ihop

meningsvis, frasvis och ordvis, huvudsakligen med hjälp av statistiska metoder. Utmärkande för detta så kallade länklexikon är att det innehåller många olika alternativ men ingen lingvistisk information. Förutom länklexikonet innehåller de statistiska översättningssystemen en enkel ordföljdsmodell för målspråket, som också den är automatgenererad på statistisk grund. Därtill kommer metoder för att välja ut den mest sannolika översättningen bland de många alternativ som skapas. (Hein 2005:26)

Ping nämner också att eftersom översättningssystemet använder korpusar, måste man bygga tvåspråkiga korpusar av hög kvalitet för att lyckas med statistisk maskinöversättning (Ping 2009:163).

Arnold förtydligar statistisk översättning med ett exempel: ”When presented with a French sentence  $f$ , we imagine that the original writer actually had in mind an English sentence  $e$ , but that  $e$  was somehow garbled in translation so that it came out as  $f$ . The job of the MT system is just to produce  $e$  when presented with  $f$ .” (Arnold 2003:138)

Han jämför statistisk översättning med en utsändning genom en bråkig kanal (liksom en telefonledning), och det finns en standard teknik som kan användas för att återfinna den ursprungliga inputen (se Arnold 2003:138).

Det är svårt att förklara precis hur Google Translate fungerar eftersom materialen ingående systemets struktur är ganska allmänna på GT:s egen webbsida:

När Google Översätt skapar en översättning görs en sökning efter mönster i miljontals dokument för att hitta den bästa översättningen. Genom att identifiera mönster i dokument som redan har översatts av mänskliga översättare kan Google Översätt göra kvalificerade gissningar om vad som är en lämplig översättning. (Google Translates webbsida)

Moberg påstår att svårfunna förklaringar ”kan ha att göra med att Google ser tjänstens utformning som en företagshemlighet som de inte gärna sprider till konkurrenter” (2012: 4). Moberg använder därför populärvetenskapliga reportage för att ge en skymt av detta, och av samma skäl använder också jag i mina följande förklaringar om GT samma reportage och nyhetsartiklar.

Fastän Google Translate skapar en språkmodell som beskriver det elementära utseendet av en välformad mening, sker ingen lingvistisk tolkning av språket (Kingsley 2011).

Det som GT skapar är tack vare korpusens storlek. Enligt Adams använder Google Translate en stor del av transkriptioner från Förenta Nationernas möten, som är översatta av mänskliga översättare till sex språk samt transkriptioner från Europaparlamentet, som är översatta till 23 språk. Google har också fått sitt material i samband med dess bokscanningsprojekt som ger den försprång mot konkurrenterna Microsofts Bing och Yahoos Babel Fish, som är baserade på i stort sett samma principer. (Adams 2010)

Det som Google Translates egen sida inte säger, är att den använder ”engelska som övergångsspråk när det inte finns tillräckligt med text som underlag för att göra tillförlitliga statistiska beräkningar för andra språk” (Domeij 2011). På grund av detta kan GT-översättningar mellan svenska och norska vara svårt att förstå. Stefan Johansson tar upp följande exempel när han intervjuas av Domeij:

[...] "Hur upplever läsaren texten?" översätts till "Hvad gjør leseren teksten?". I det här fallet kan man gissa att det är det engelska do som ställer till det. [...] Meningen "Hur upplever läsaren texten?" översätts alltså först till engelska och blir, kryptiskt nog, "What do the reader the text?". Sedan översätts meningen från engelska till norska, och blir "Hvad gjør leseren teksten?". (Domeij 2011)

Google Translate förbättras kontinuerligt, antalet språk som stöds idag är 71 och översättningskvalitet förbättras också under tiden. Man måste ändå förstå att Google Translates syfte inte är att bli en ersättning för mänsklig översättare utan hjälpa människor att förstå vad texten handlar om, även om det finns några fel i resultatet.

## **2.4 Andra undersökningar**

Det finns inga bekanta undersökningar angående översättningar från svenskan till estniskan eller tvärtom med hjälp av Google Translate. Men det finns undersökningar angående översättningar till eller från engelskan. En av dem är Ida Stefanssons C-uppsats "Översätt den här sidan" *The advancement of Google Translate and how it performs in the online translation of compound and proper nouns from Swedish into English*. Undersökningens resultat var inte förvånande – det finns brister när det gäller översättning av namn och sammansättningar.

En annan undersökning beträffande Google Translate och dess kvalitet är *Taking on new challenges in multi-word unit processing for machine translation* skriven av Johanna Monti, Anabela Barreiro, Annibale Elia, Federica Marano och Antonella Napoli. Deras undersökning utforskade Google Translate i jämförelse med Open Logos, en regelbaserad maskinöversättnings system och resultatet var att Open Logos fungerar bättre när det gäller uttryck bildade av flera ord, till exempel idiom.

Från förra året (2012) kommer en undersökning av Björn Moberg kallad *Googleöversättning - Om Google Translate:s möjligheter och begränsningar vid översättning till svenska*. Hans slutsats var att Google Translate verkar översätta konventionella fraser bättre än statistiskt mindre förekommande fraser och ju mer kreativ text desto svårare är den att översätta för GT (Moberg 2012: 28-29). Mobergs källtexter kom inte från olika genrer men en av hans texter var en blandning av ekonomisk text, som skulle översättas bra med GT, och journalistisk kreativ text, som är svår för GT. I min undersökning försöker jag bedöma om någon genre är lättare att översätta för GT, eftersom mitt material består av tre olika typer av text som jag klargör i nästa kapitel.

En annan undersökning från år 2012 var *Maskinöversättning – bra eller dåligt? – En översättningsvetenskaplig studie om typiska fel och begriplighet hos maskinöversatta facktexter* av Amanda Lindqvist, som undersökte två översättningstjänsten – Google Translate och Systran – vid översättningar från engelska till svenska och resonerade att ”Google Translate producerar begripliga översättningar i stor utsträckning medan Systran för det mesta genererar obegripliga översättningar enligt både lekmän och experter” (Lindqvist 2012: 2). I min undersökning kommer jag liksom Lindqvist göra en analys om typiska felen, men jag gör ingen vetenskaplig analys om begripligheten.

### **3. Material och metod**

#### **3.1 Texter**

Eftersom syftet med undersökningen är att visa vilka och hur många fel förekommer när man använder Google Translate har jag valt tre texter av olika typer. På det sättet kan man se om det finns fel som förekommer i olika slags texter och om det finns fel som

man bara kan hitta i till exempel skönlitterära texter. Eftersom jag vill visa vilka fel som förekommer i alla texternas estniska översättningar finns det 6 texter: tre texter på svenska och deras tre översättningar till estniska med hjälp av Google Translate. Emedan Google Translate förändras och förbättras kontinuerligt måste man lägga märke till att alla översättningarna är gjorda den sextonde april 2013.

Texterna är valda ganska slumpmässigt eftersom användning av texter som man tänker skulle ha många problem ger ensidiga resultat. Genrerna är dock valda med tanke på att de skulle representera olika grader av svårighet: en lättare text (lärobokstexten), en mellanliggande text (nyhetstexten) och en svårare text (den skönlitterära texten).

Den första genren av text kommer från en lärobok i svenska för ester. Den heter På Svenska! – Rootsi keele õpik och är skriven av Ulla Göransson, Annika Helander och Mai Parada, boken är utgiven år 2007. Texten som jag har valt är från avsnitt 14 av 16, vilket betyder att studenten har läst svenska cirka 4 månader dvs. den är inte krånglig men inte så lätt heller. Texten innehåller även idiom, till exempel *kärlek vid första ögonkastet* och *ha tur*, samt ord som kan vara svåra att översätta, till exempel *farfar* och *morfar* och *en tvåa*, men hela texten är texten lätt att läsa. Första texten består av 265 ord. Textens Google-översättning på estniska innehåller 224 ord.

Den andra genren av text är nyhetstext. Denna text karakteriseras av olika bisatser, person- och ortnamn. Artikeln kommer från Dagens Nyheters webbsida och publicerades den sextonde april 2013 klockan 13:21, uppdaterad 14:26. Texten, benämnd ”Hundratals befaras döda i skalv” (17:51 uppdaterad till ”Jätteskalv skakade Iran och Pakistan” med tillsatta författarna och förändringar i text), handlar om en jordbävning i Iran och artikeln är en liten rapport av händelserna med några kommentarer. Man kan lägga märke till att det finns några förkortningar, till exempel *UD* (Utrikesdepartementet), *USGS* (USA:s geologiska institut) och även *USA* i texten, som kan vara svåra för Google Translate ifall det inte är säkert om man måste översätta förkortningar eller inte. Text 2 innehåller 273 ord och dess översättning genom Google Translate till estniskan innefattar 217 ord.

Den tredje och sista genren av text är skönlitteratur. Texten kommer från Sven Delblancs bok *Åminne – en berättelse från Sörmland* från sidan 62 och består av några stycken som tillsammans utgör 277 ord. Avsnittet talar om hur Märta möter Urse och Tok-Harry och beskriver tankar och känslor som de har. Därför innehåller texten fraser

som är typiska för poetiska texter till exempel liknelse – *hennes hjärta flämtar till som en ljuslåga*, ovanligt ordförråd- *oförbindlig*, metaforer – *hon vindar som en nykalvad kviga*, ovanlig ordföljd och ellips – *Så hård hans hand* och synonymer – *tomma, blanka blick*. De alla kan bli krångliga att tolka, i synnerhet för Google Translate. Textens originella tryck innehåller ett felstavat ord - *nyklavad*, som jag i min undersökning ersätter med det korrekta ordet. Då blir *nyklavad* till *nykalvad* i min källtext. Textens översättning till estniska innehåller 197 ord, som är 80 ord mindre än originalen.

### 3.2 Metod

För att utvärdera om maskinöversättning är riktig används olika metoder. Flera av dem är program- och korrelationbaserade och hör mer till matematiker och programmerare. Därför är min metod att se vilka och hur många fel det finns i översättningarna ganska enkel. Först tar jag en text på svenska och kopierar den till Google Translates webbsida <http://translate.google.com/>, då låter jag Google Translate att översätta texten till estniska. Google Translate har också möjlighet att ersätta ord i den översatta texten genom att välja ett mer passande ord bland dem som Google Translate erbjuder eller skriva det rätta ordet själv, men detta använder jag inte.

Sedan läser jag den översatta texten som om den var en riktig estnisk text och bestämmer om jag kan förstå vad var och en sats betyder. Det finns ingen vetenskaplig metod bakom denna process eftersom undersökningens fokus är på språkliga fel och inte begripligheten. Bedömningarna om begripligheten är för att ge läsaren en bättre bild om hur många av dessa GT-översatta meningar och i hur stor utsträckning en est kan förstå.

Efter detta sätter jag båda texterna intill varandra och markerar fel som finns i de GT-översatta texterna. Språkliga fel är räknade i jämförelse med en korrekt översättning, som i exempelmeningarna är betecknade som ”Korrekt”. Det kan finnas många översättningar av en mening på svenska som kan vara korrekta, men i den undersökningen visar jag den översättning som liknar mest den Google-översatta meningen.

Felen hittar jag med hjälp av ”Rootsi-eesti sõnaraamat/Svensk-estnisk ordbok: 100 000 märksõna/ uppslagsord” av Aidla, et al. (2004) som jag använder för att kontrollera om orden har översatts rätt. Dessa fel grupperar jag enligt ”Error Analysis of Statistical

Machine Translation Output” av Vilar et al. (2006), men jag gör några förändringar för att tillämpa dessa grupper till mina språkpar.

Felen är grupperade hierarkiskt (se lista 1) och delade in i 5 kategorier, som har underkategorier. Första kategori är **förkommet ord** som delas in i två: först dessa ord som inte finns i de översatta texterna men har innehållsmässig betydelse i meningen, till exempel *unistas lauljaks (drömde om att sångerska)* i stället för *unistas lauljaks saamisest (drömde om att bli sångerska)*, och den andra underkategorin är för fyllande ord som måste vara i satsen för att den skulle bli grammatiskt korrekt men man kan förstå även om dessa ord saknas, till exempel *sajad surnud (hundratals döda)* istället av *sajad on surnud (hundratals är döda)*. Den andra kategorin är **ordföljd**, till exempel *seismoloog Ülikooli Uppsala* istället för *seismoloog Uppsala Ülikoolis*. Den tredje kategorin är **felaktiga ord** som delas in i sex underkategorier. Först enligt betydelse, som också har underkategorier: när systemet väljer ett felaktigt ord, till exempel *king (sko)* istället för *suss (toffel)*, och när systemet disambiguerar inte korrekt ord i källtexten – det vill säga om det finns dubbelbottnade ord i källtexten och Google Translate använder det som inte passar i kontexten, till exempel *ülesanne (uppgift, uppdrag)* istället för *info (uppgift, information)*.

Sedan finns det en underkategori i **felaktiga ord** för ord som har fel form. Till exempel hos substantiv ordet *buss (bus)* istället för *bussis (i bussen)*, och hos verb *laulis (han/hon sjöng)* istället för *laulavad (de sjunger)*. I denna underkategori sätter jag också ord som har rätt stam men fel form, till exempel när ett adverb bildas av ett adjektiv. Den tredje kategorin är extra ord som inte bör finnas i de översatta texterna, den fjärde är fel i stilen, som betyder att meningen i satsen är bevarad, men ord är inte perfekta för den satsen, till exempel *ma arvan, et sa oled kasvanud? (jag tänker du har vuxit)* i stället för *mulle tundub, et oled kasvanud (jag tror du har vuxit)*. Den femte kategorin inom **felaktiga ord** är idiom som Google Translate översätter fel, till exempel *som minne av = mälestuseks*, inte *nagu mälu*, som det översätts när man tar ord för ord. Och den sista kategorin är för namn som är översatta fel, till exempel personnamn är översatta som vanliga ord. I den här kategorin sätter jag också namn som inte har rätt böjningsändelse, till exempel när *till Ulla* är översatt bara som *Ulla* och inte *Ullale*, som är korrekt.

Den fjärde huvudkategorin är *kommatering* – till exempel när det måste finnas ett kommatecken i satsen men det finns inte eller när det inte måste vara ett kommatecken men det ändå finns i den GT-översatta meningen. Och den sista kategorin är för *obekanta ord* – ord som Google Translate inte vet och som finns i den översatta texten i samma form som de är i källtexten.

#### Lista 1: Kategorier av felen

1. Förkommet ord
  - 1.1 Kontenta ord
  - 1.2 Fyllande ord
2. Ordföljd
3. Felaktiga ord
  - 3.1 Betydelse
    - 3.1.1 Systemet väljer felaktigt ord
    - 3.1.2 Systemet disambiguerar inte
  - 3.2 Fel form
    - 3.2.1 Fel form av substantiv
    - 3.2.2 Fel form av verb
    - 3.2.3 Fel form av adjektiv
  - 3.3 Extra ord
  - 3.4 Stil
  - 3.5 Idiom
  - 3.6 Namn
4. Kommatering
5. Obekanta ord

Dessutom analyserar jag kort varför felen även finns. Det kan finnas språkliga skäl för felens förekomst. För att hitta vilka skillnader mellan svenska och estniska som kan vara anledningen för fel använder jag två böcker: ”Lär känna estniskan: en introduktion till estniska språket” av Raimo Raag (2007) och ”Rootsi-eesti abiraamat = Svensk-estnisk hjälpbok” av Björn Collinder, Joh. Aavik och Oskar Loorits (1944).

Därtill använder jag också Doug Arnolds artikel ” Why translation is difficult for computers”, avsnitt ”Machine Translation” av Ke Ping från Routledge Encyclopedia of

Translation Studies och Anna Sågvall- Heins artikel ”Datorn behöver statistik och grammatik” som beskriver närmare fel som medförs med statistisk maskinöversättning.

I analysen räknar jag alla felaktiga ord som ett fel, även om samma fel förekommer genom hela texten. När orsaken till ett fel är felaktig ordföljd räknar jag ett fel per sats, inte ett fel per varje ord som står på fel plats. Till exempel när en mening har en bisats räknar jag fel i både huvudsatsen och bisatsen. Ytterligare fäster jag också uppmärksamhet till ord och namn som är översatta olika i olika meningar.

Dessutom ska jag också göra en kvantitativ analys av alla texter. Först ska jag räkna alla fel i en text och då beräkna hur många fel det finns per ord i var och en typ av text. Det ger en översikt över vilken text som är mest komplicerad för Google Translate att översätta. För övrigt ska jag också se vilka typer av fel är mest frekventa i var och en text samt i alla texter tillsammans.

## 4. Resultat

### 4.1 Text 1 – lärobokstexten

I den första texten finns det 22 meningar. Av dessa 22 meningar finns det 3 som är översatta korrekt. Dessa meningpar är:

- |  |   |
|--|---|
| 1. Drömmar och verklighet                | Unenäod ja reaalsus                     |
| 2. Det började så bra                    | See algas nii hästi:                    |
| 3. Det blev kärlek vid första ögonkastet | See oli armastus esimesest silmapilgust |

Det finns 11 meningar vars huvudbetydelse jag kan förstå, men det finns några ord som är översatta fel och därför kan man bara begripa halva meningen. Tillsammans med de 3 helt korrekta meningarna förstår man då ungefär 64 % av alla meningar. En av dem kan man förstå halvt, men dess innebörd är faktiskt olik från källtexten.

Källtext (KT): *De hade tur och fick en plats på dagis när Harald var ett år.*

Översättning: *Nad olid õnnelikud ja sai koha lasteaias kui Harald oli üks aasta.*

Korrekt översättning: *Neil oli õnne ja nad said koha lasteaias, kui Harald oli ühe aastane.*

I denna mening kan man förstå att de fick en plats på dagis även om det finns några grammatiska fel, men första halvan av meningen som är grammatiskt korrekt är

egentligen fel eftersom den översatta meningen säger att ”de var lyckliga”, inte att ”de hade tur”.

Betydelsen i 8 meningar är obegriplig eftersom det finns för många fel. Totalt antal fel i text 1 var 98, vilket betyder att det finns 37 % fel i texten – ett fel efter varje 2,7 ord i genomsnitt. Det finns 22 meningar i texten och statistiskt innebär det att det finns ungefär 4,5 fel i en mening i genomsnitt.

Antal fel enligt kategorierna i lista 1:

Kategori	1.1	1.2	2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3	3.4	3.5	3.6	4	5
Antal fel	5	2	8	7	11	31	9	2	3	5	1	9	2	3

Man kan se att mest förekommer fel i kategori 3 - *felaktiga ord*. Sedan följer *förkommet ord*, som följs av kategorin *ordföljd*.

Det mest typiska felet i text 1 var fel form av substantiv som illustreras av följande meningspar i källtexten (KT) och översättningen (Ö):

KT: *På kvällarna sjöng hon i en kör och drömde om att bli sångerska.*

Ö: *Õhtuti laulis **koor** ja unistas **laulja**. (På svenska: På kvällarna sjöng **kören** och drömde om en **sångerska**)*

Korrekt översättning för denna mening är:

*Õhtuti laulis ta **kooris** ja unistas **lauljaks** saamisest.*

Fel form är även frekvent hos verb:

KT: *Daniels kusin, Sofia, slutade skolan efter årskurs nio och började arbeta i en affär.*

Ö: *Daniel nõbu, Sofia, koolist **lahkunud** pärast üheksanda klassi ja alustas tööd kauplus. (Daniels kusin, Sofia, **slutad** skolan efter årskurs nio och började arbete affär)*

Korrekt: *Danieli nõbu, Sofia, **lahkus** koolist pärast üheksandat klassi ja alustas tööd kaupluses.*

Det andra mest typiska felet är när Google Translate inte disambiguerar och väljer fel ord. Till exempel:

KT: *De fick gamla möbler från vänner och släktingar, och Tobias, som är händig, målade och reparerade dem.*

Ö: *Neil oli vana mööbli sõbrad ja sugulased ja Tobias, mis on mugav, värvitud ja parandas neid. (De hade gamla möbler vänner och släktingar, och Tobias, som är bekväm, målad och reparerade dem)*

Korrekt: *Nad said sõpradelt ja sugulastelt vana mööblit ja Tobias, kes on kätega osav, värvis ja parandas neid.*

I den meningen finns det tre fel där svenska ord har dubbelmeningar och Google Translate inte vet vad är korrekt. För det första finns det pronomen *som*, som hänvisar till Tobias som är en person men samma ord kan också hänvisa till ett livlöst ting. I den meningen har GT valt det sista. Det andra felet är *händig*, som betyder båda att något är praktiskt eller trivsamt att använda och att någon är skicklig. I den här meningen måste det bli det sista men GT översätter att Tobias är bekväm som betyder det samma att han är lat som är motsatsen till meningens egentliga betydelse. Sista felet som gäller denna kategori är ordet *målade*, som kan vara både preteritumform av verbet måla (någon målade något) och också perfekt particip av verbet ”måla” som till exempel i substantivfrasen *den målade dörren* där ordet används som adjektiv i frasen. I den här meningen måste det vara ett verb i preteritum men GT har valt den andra betydelsen.

Namn har också blivit svåra för GT att översätta. Många fel kommer när GT inte har lagt böjningsändelse till namn, till exempel:

KT: *De åkte på bröllopsresa i Sofias gamla Volvo till Tobias mormor i Gränna.*

Ö: *Nad läksid edasi mesinädalad Sofia vana **Volvo Tobias** vanaema **Gränna**. (De åkte vidare bröllopsresa Sofias gammal **Volvo Tobias** mormor **Gränna**)*

Korrekt: *Nad läksid mesinädalatele Sofia vana **Volvoga Tobiase** vanaema juurde **Grännasse**.*

I föregående exempel har GT inte alls valt någon böjningsändelse, men det finns också en mening där den rätta ändelsen är på plats:

KT: *Efter två månader hyrde de en liten tvåa i en förort till **Stockholm**.*

Ö: *Kahe kuu pärast, nad üüritud väike korter **Stockholmi** eeslinnas.*

Korrekt: *Kahe kuu pärast üürisid nad väikese kahetoalise korteri **Stockholmi** eeslinnas.*

Men böjningsändelse är inte det enda problemet som förekommer med namn. Det finns en mening där GT har översatt ett namn till till dess motsvarighet på engelska:

KT: När **Sofia** var 18 år träffade hon Tobias, en 19-årig TV-reparatör från Småland.

Ö: Kui **Sophia** oli 18, ta kohtus Tobias, 19-aastane TV-remondimees alates Småland.

Korrekt: Kui **Sofia** oli 18, kohtus ta Tobiasega, 19-aastase teleriremondimehega Smålandist.

Den andra huvudkategorin som har många fel är *ordföljd* – 8 fel i 22 meningar, vilket betyder att det finns felaktigt ordval i genomsnitt efter varje 2,75 meningar. I en mening där det finns bisatser kan man se även två felordningar. Ifrågavarande mening är följande:

KT: Först tyckte Sofia att han var löjlig, men sedan blev hon trött på hans misstänksamhet, och efter ett tag bestämde de sig för att skiljas

Ö: Esimene arvas, Sofia, et ta oli naeruväärne, kuid siis sai **ta väsinud** tema kahtlused, ja mõne aja **pärast nad otsustasid** lahutada. (Första tyckte, Sofia, att han var löjlig, men sedan fick hon trött hans misstänksamheter, och efter ett tag bestämde de sig för att skiljas).

Korrekt: Esmalt arvas Sofia, et ta oli naeruväärne, kuid siis **väsis ta** tema kahtlustest, ja mõne aja **pärast otsustasid nad** lahutada.

I text 1 finns det fel i alla kategorierna, men exempel på de mest intressanta felen kommer jag att redogöra i kapitel 4.4 där jag kommenterar alla texternas fel som inte var mest frekventa tillsammans.

## 4.2 Text 2 – nyhetstexten

I den andra texten finns det 273 ord i 19 meningar. Det finns inga meningar som är korrekta. Det finns 12 meningar vars huvudbetydelse jag kan förstå (63 % av det totala antalet meningar) och två meningar som är någorlunda begripliga för mig, men betydelsen är fel i jämförelse med källtexten. Dessa två meningspar är:

KT: Även på den pakistanska sidan av gränsen har det krävts dödsoffer.

Ö: Ka Pakistani pool piiri on **vaja** inimohvleid. (Även på den pakistanska sidan av gränsen **behövs** dödsoffer)

Korrekt: *Ka Pakistani pool piiri on see **nõudnud** inimohvleid.*

KT: *Minst fem personer har omkommit, uppger en pakistansk tjänsteman till nyhetsbyrån Reuters.*

Ö: *Vähemalt viis inimest on **tapetud**, vastavalt Pakistani ametnik rääkis Reutersile. (Minst fem personer har **mördats**, uppger en pakistansk tjänsteman talade till Reuters.)*

Korrekt: *Vähemalt viis inimest on **hukkunud**, teatab Pakistani ametnik uudisteagentuurile Reuters.*

I båda exemplen är verben det som bär meningen om situationen, men i översättningen har verbets betydelse förvandlats så att det betyder något annat.

Det totala antalet fel i text 2 är 94, som betyder att det finns 34 % fel i texten – i genomsnitt ett fel efter varje 2,9 ord. Det finns 19 meningar i texten och statistiskt innebär det att det finns i genomsnitt 4,9 fel i en mening.

Antal fel enligt kategorierna i lista 1:

Kategori	1.1	1.2	2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3	3.4	3.5	3.6	4	5
Antal fel	5	12	10	20	2	15	10	1	4	1	0	9	2	3

I text 2 finns de mest frekventa felen också i kategori 3, men i motsats till text 1, finns det mest fel i underkategori 3.1.1. där systemet väljer ett felaktigt ord. Detta illustrerar även de två exemplen som jag omnämnde ovan. Därtill finns det också fel när det gäller det viktigaste med artikeln – vad var det som egentligen hände. Det finns sex ord i källtexten som används för att beteckna jordbävning: *skalv, skalvet, skalven, jordskalv, jordskalvet* och *jordbävning*. GT klarar av med *jordbävning, jordskalv, jordskalvet* och *skalven*, men när det finns *skalv* eller *skalvet* är det översatt fel. Till exempel:

KT: *Det finns inget samband mellan **skalven**, det är bara tillfällighet, säger Reynir Bödvarsson, seismolog vid Uppsala universitet.*

Ö: *Puudub seos **maavärinad**, see on lihtsalt kokkusattumus, ütles Reynir Bödvarsson, seismoloog Ülikooli Uppsala. (Det finns inget samband **skalven**,*

*det är bara tillfällighet, sa Reynir Bödvarsson, seismolog Universitetets Uppsala.)*

Korrekt: *Puudub seos **maaväri**nate vahel, see on lihtsalt kokkusattumus, ütles Reynir Bödvarsson, sesmoloog Uppsala Ülikoolis*

KT: *Hundratals befaras döda i **skalv**.*

Ö: *Sajad kartsid surnud **värisema** (Hundratals befarade de döda att **skälva**)*

Korrekt: *Kardetakse, et sajad on **maaväri**nas surnud.*

Liksom i text 1 finns det också många fel när det gäller form av substantiv. Till exempel i den här meningen:

KT: *Hundratals människor befaras ha omkommit i en kraftig jordbävning i sydöstra Iran, uppger iranska myndigheter enligt Reuters.*

Ö: *Sajad inimesed kartsid surnud võimas **maaväri**n kagu Iraan vastavalt Iraani **ametivõimude** sõnul Reuters. (Hundratals människor befarade omkommit en kraftig jordbävning sydöstra Iran enligt iranska myndigheter efter Reuters)*

Korrekt: *Kardetakse, et sajad inimesed on surnud võimsas **maaväri**nas Kagu-Iraanis, teatavad Iraani **ametivõimud** Reutersi andmeil.*

Man kan också se att den meningen har brister även i kategori 3.2.2 där verbets form är fel. Estniskt passiv är skapad med frasen *kardetakse, et ... on surnud* som Google Translate inte klarar av. Dessutom är också supinum översatt fel på estniska eftersom ordet *on* är förkommet.

Detta bringar oss till nästa felgrupp som är ganska frekvent i text 2. Det finns 10 fel i kategori 1.2 där ett fyllande ord är förkommet.

KT: *Strax före två på eftermiddagen fanns ännu inga uppgifter om att någon svensk skulle vara berörd av skalvet.*

Ö: *Just enne kahte päraslõunal, ei olnud teateid kõigist Rootsi mõjutaks värisema. (Strax före två på eftermiddagen, fanns inga uppgifter om alla Sverige skulle påverka att skälva.)*

Korrekt: *Just enne kahte päraslõunal ei olnud **veel** teateid, et mõni rootslane **oleks** maaväri*nast puudutatud.

Även om det inte är lätt att bedöma om ett ord är fyllande eller har en innehållsmässig betydelse utgår jag från att när man lämnar ut dessa tre ord, kan man ändå förstå vad är meningens betydelse.

I text 2 finns det förkortningar som GT klarar inte av. Det finns två exempel på detta:

KT: *Skalvet, som mätte 7,8 på Richterskalan, inträffade i Iran i på tolv mils djup, enligt USA:s geologiska institut USGS.*

Ö: *Värisema, mis mõõdeti 7,8 Richteri skaalal, toimus Iraanis kaksteist miil sügav, vastavalt US Geological Institute USGS. (Att skälva, som mätte 7,8 på Richterskalan, inträffade i Iran tolv mils djup, enligt US Geological Institute USGS)*

Korrekt: *Maavärin, mis mõõdeti 7,8 palli Richteri skaalal, toimus Iraanis 120 km sügavusel, vastavalt USA geoloogilisele instituudile USGS.*

KT: *Vår ambassad i Teheran undersöker detta nu och vi kommer säkert få mer information snart men så här långt finns inga uppgifter om skadade svenskar, säger Catarina Axelsson på UD:s presstjänst.*

Ö: *Meie saatkond Teheranis uurib seda nüüd ja me kindlasti saada lisateavet varsti, kuid seni ei ole infot vigastatud rootslased, ütleb Catarina Axelsson MFA pressiteenistus.*

Korrekt: *Meie saatkond Teheranis uurib seda nüüd ja me kindlasti saame lisateavet varsti, kuid seni ei ole infot vigastatud rootslastest, ütleb Catarina Axelsson välisministeeriumi pressiteenistusest.*

I det första exemplet kan man se att förutom förkortningen *US* finns det också några andra ord som förekommer på engelska och inte på estniska. Dessutom innehåller meningen en svensk term *mil*, som kan översättas som *miil* men då betyder ordet vanligtvis en annan måtenhet. I denna mening måste det översättas med *120 km* eftersom en mil här är lika med 10 km.

I det andra exemplet finns det också spår av engelskan eftersom UD:s (Utrikesdepartementet) motsvarighet på engelska är MFA (Ministry of Foreign Affairs). På estniska kan man använda förkortningen VM, men den är inte välkänd och oftast används det vanliga ordet *välisministeerium*.

Exempel på den svenska måtenheten *mil* finns mer i texten, men dess översättning är olik från exemplet ovan. I följande exempel har GT inte förstått *mil* är en måtenhet och har valt översättningen *miljon* istället:

KT: *Det inträffade i sydöstra delen av landet, nio **mil** sydöster om staden Khash klockan 10.44 lokal tid, enligt USGS.*

Ö: *See toimus kaguosas, üheksa **miljonit** kagu linna Khash kell 10:44 kohaliku aja järgi USGS. (Det inträffade i sydöstra delen, nio miljoner sydost staden Khash klockan 10:44 lokal tid USGS)*

Korrekt: *See toimus USGSi sõnul riigi kaguosas, **90 km** Khashi linnast kagus kell 10:44 kohaliku aja järgi.*

Dessutom finns det många fel i kategori 3.6 – fel i namn, eftersom namnen i de estniska meningarna behöver böjningsändelser. Till exempel:

KT: *Kärnkraftverket vid iranska Bushehr ska inte ha skadats av skalvet, enligt det ryska företaget som byggt det.*

Ö: *Tuumaelektrijaama **Bushehr Iraan** ei tohi olla kahjustatud värisema, vastavalt Vene ettevõtte, mis ehitasid. (Kärnkraftverket Bushehr Iran ska inte vara skadad att skälva, enligt det ryska företaget, som byggde.)*

Korrekt: ***Busheheri** tuumaelektrijaam **Iraanis** ei tohiks olla maavärinast kahjustatud, nagu väidab Vene ettevõtte, mis selle ehitas.*

Andra intressanta språkliga fel som jag hittade i text 2 som faller i kategorier som inte var mest frekventa kommer jag att kommentera i kapitel 4.4.

### 4.3 Text 3 – den skönlitterära texten

I den tredje texten finns det 277 ord i 22 meningar. Det finns en mening som är översatt korrekt, men meningen är inte en fullständig mening, utan en fras:

KT: *Nästan vuxen.*

Ö: *Peaaegu täiskasvanud.*

Jag kan förstå huvudbetydelse i 5 GT-översatta meningar som tillsammans med den helt korrekta frasen gör 27 % av alla meningar, d.v.s. man förstår ungefär bara 27 % av texten.

Det totala antalet fel i text 3 är 126, vilket betyder att det finns 45 % fel i en text – i genomsnitt ett fel efter varje 2,2 ord. Det finns 22 meningar i texten och statistiskt innebär det att det i genomsnitt finns ungefär 5,7 fel i en mening.

Antalet fel enligt kategorierna i lista 1:

Kategori	1.1	1.2	2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3	3.4	3.5	3.6	4	5
Antal fel	8	11	7	25	2	26	9	8	9	8	1	4	4	4

Liksom i text 1 finns det mest fel i kategori 3.2 där ordets form är felaktig. I de första två texterna finns det få fel i kategori 3.2.3, men eftersom text 3 är skönlitterär, finns det 8 fel även när det gäller adjektiv och adverb. Till exempel:

KT: *Tankspridd torkar han sig om händerna, i en dunkel förnimmelse att ha vidrört något naket och vått.*

Ö: *Hajameelne* kuivab ise oma kätega, vaid *ebamäärane* tunne ta puudutanud midagi alasti ja *niiske*. (*Tankspridd* torkar själv med egna händer, bara en *dunkel* förnimmelse han vidrört något oklätt och *vått*. Båda adjektiven i slutet av den estniska meningen saknar böjningsändelser som är obligatoriska men som inte kan visas i den här svenska översättningen.)

Korrekt: *Hajameelselt* kuivatab ta oma käed, *ebamäärase* tundega, et ta on puudutanud midagi paljast ja *niisket*.

I exemplet innebär det första felet att Google Translate inte har valt ett adverb, utan ett adjektiv i översättningen. Anledningen till två andra fel är att adjektiven har ingen böjningsändelse.

En annan kategori med många fel är 3.1.1 vilket innebär att GT väljer ett felaktigt ord. En av meningarna har till och med 4 dylika fel:

KT: *Nynnande sin eviga sång börjar han lunkande förfölja henne som en kelen tjurkalv.*

Ö: *Koolibrid* oma igavest laulu alustab *oma kõmpima* kiusavad teda *kaisu Mullikka*. (*Kolibrier* sin eviga sång börjar *sin att pallra* mobbar henne en *kram-Mullikka*.)

Korrekt: *Ümisedes* oma igavest laulu hakkab *ta sörkides* teda jälitama nagu *hellust igatsev pullvasikas*.

I text 3 finns det ett stort antal fel i kategorierna 1 – förkommet ord och 3.3 – extra ord. Ett exempel på detta ger följande mening:

KT: *Med ens är hon maktlös och tung av sin kvinnlighet, benen ger vika under henne, och hela hennes viljekraft är hjälplöst samlad och innesluten i hans hårda hand.*

Ö: *Järsku ta on jõuetu ja raske koos oma naiselikkust, tema jalad andma teed all tema ja kõik tema tahte jõud on abitu kogutakse ja tema kõva käsi. (Med ens är hon maktlös och tung med sin kvinnlighet, hennes ben går ur vägen under henne och hela hennes viljekraft är hjälplös samlas och hans hårda hand.)*

Korrekt: *Järsku on ta jõuetu ja raske enda naiselikkusest, jalad annavad järele ja kogu tema tahte jõud on abitult kogutud ja kinni mehe kõvade käte embuses.*

I den översatta meningen har jag framhävt orden som är extra och i den korrekta meningen har jag framhävt orden som egentligen måste finnas där.

I skönlitterära texter finns det ofta synonymer i samma sats för att göra texten mer poetisk. Om GT översätter båda ord som är synonymer till varandra med samma ord kan det vara grammatiskt korrekt, men skapa fel i stilen. I min undersökning finns det ett exempel på synonymer som GT har översatt som samma ord, därtill finns det fel i böjningen:

KT: *Tok-Harry reser sig och betraktar henne med sin tomma, blanka blick.*

Ö: *Tok-Harry püsti ja vaatab teda oma tühi, tühi vahtima. (Tok-Harry upp och betraktar henne sin tomma, tomma att stirra.)*

Korrekt: *Rumal-Harry tõuseb püsti ja vaatab teda oma tühja, sisutu pilguga.*

#### 4.4 Sammanfattning av felen

De tre texterna tillsammans har 815 ord i 63 meningar. Det totala antalet fel i texterna är 318, som betyder att det finns ungefär 5 fel i en mening, ett efter varje 2,6 ord.

Distributionen av fel enligt kategorierna i lista 1:

Kategori	1.1	1.2	2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3	3.4	3.5	3.6	4	5
Antal fel	18	25	25	52	15	72	28	11	16	14	2	22	8	10

Mest typiska fel är form av substantiv, som följs av fel som förekommer eftersom GT väljer ett felaktigt ord beträffande dess betydelse. Många fel finns också i kategorier

3.2.2 och 2 som är fel i form av verb respektive ordföljd. Exempel på de mest frekventa felen har jag redan beskrivit i avsnitt 4.1, 4.2 och 4.3, men i viss mån är fel som förekommer inte så ofta även mer intressanta. Till exempel i kategori 5 – obekanta ord. Domeij (2011) skriver att Google Translate använder engelska som övergångsspråk och jag hittade spår av detta vid 8 gånger: när GT inte översätter ett ord och sätter ett engelskt ord i översättningen och också när den översätter svenska namn till deras engelska motsvarigheter. Det mest komplicerade i detta område för GT var den skönlitterära texten. I följande exempel finns båda – ett engelskt namn och ett ord som är på engelska:

KT: *På trappan störtar Märta ofrivilligt i famnen på Urse, och han griper henne om armen för att hon inte ska fall.*

Ö: *Trepil **plunges Martha** tahtevastaselt sülle Urse ja ta haarab tema poolt käe, nii et ta ei ole asjas. På trappan **plunges Martha** ofrivilligt i famnen Urse och han griper från hennes sida armen så att hon inte är i saken)*

Korrekt: *Trepil **kiirustab Marta** tahtmatult Urse sülle ja mees haarab ta käest kinni, et neiu ei kukuks.*

Dessutom uppmärksammade jag att det finns spår av finskan i en av de översatta meningarna när ordet är översatt på finska istället för estniska:

KT: *Nynnande sin eviga sång börjar han lunkande förfölja henne som en kelen tjurkalv.*

Ö: *Koolibrid oma igavest laulu alustab oma kõmpima kiusavad teda kaisu **Mullikka**. (Kolibrer sin eviga sång börjar sin att pallra mobbar henne en kram-**Mullikka**.)*

Korrekt: *Ümisedes oma igavest laulu hakkab ta sörkides teda jälitama nagu hellust igatsev **pullvasikas**.*

Obekanta ord som förekommer i dess originella form finns i bara en mening. Ord som GT inte vet är sammansättningen ”sångarvänner” som även i källtexten har apostrofer. Ordet är inte listad i Svenska Akademiens Ordlista heller, vilket betyder att det inte är ett välkänt ord.

Därtill finns det intressanta fel i kategori 3.4 – stil. Till exempel i följande mening från text 1 finns orden ”farfar” och ”morfar”, som bara har en motsvarighet på estniska, men man behöver förklara för att undvika upprepning av ordet:

KT: *De fick en son som de döpte till Harald, Emil, Gustav (Emil efter **morfar** och Gustav efter **farfar**).*

Ö: *Neil oli poeg, keda nad nimega Harald, Emil, Gustav (pära**st vanaisa** Emil ja Gustav pära**st vanaisa**). (De hade en son, som de med namn Harald, Emil, Gustav (efter **morfar** Emil och Gustav efter **morfar**)).*

Korrekt: *Nad said poja, kelle nad ristisid Haraldiks, Emiliks, Gustaviks (Emil **ema isa järgi** ning Gustav **isa isa järgi**).*

Kända idiom klarar GT ganska bra, till exempel *kärlek vid första ögonkastet*, men det finns ändå meningar där Google Translate har försökt att översätta ord för ord. Till exempel i den här meningen från den skönlitterära texten:

KT: *Med ens är hon maktlös och tung av sin kvinnlighet, benen **ger vika** under henne, och hela hennes viljekraft är hjälplöst samlad och innesluten i hans hårda hand.*

Ö: *Järsku ta on jõuetu ja raske koos oma naiselikkust, tema jalad **andma teed** all tema ja kõik tema tahtejõud on abitult kogutakse ja tema kõva käsi. (Med ens är hon maktlös och tung med sin kvinnlighet, hennes benen **går ur vägen** under henne och hela hennes viljekraft är hjälplös samlas och hans hård händ.)*

Korrekt: *Järsku on ta jõuetu ja raske enda naiselikkusest, jalad **annavad järele** ja kõik tema tahtejõud on abitult kogutud ja kinni tema kõvade käte embuses.*

Fel i kommatering var inte frekventa i de GT-översatta texterna, och större delen av felen förekommer när ett ord eller konstruktion behöver ett kommatecken på estniska men inte på svenska. Men det finns också en mening i nyhetstexten där inget kommatecken finns i den svenska källtexten och det behövs i den estniska översättningen inte heller:

KT: *Strax före två på eftermiddagen fanns ännu inga uppgifter om att någon svensk skulle vara berörd av skallet.*

Ö: *Just enne kahte pära**stlõunal**, ei olnud teateid kõigist Rootsi mõjutaks värisema.*

Korrekt: *Just enne kahte pärastlõunal ei olnud veel teateid, et keegi rootslane oleks maavärinast puudutatud.*

När det gäller begriplighet kan jag resonera att den skönlitterära texten (text 3) var mest komplicerad (27 % begriplig) för Google Translate medan nyhetstexten (text 2) och lärobokstexten (text 1) var ganska lika: 63 % respektive 64 % begripliga.

Det finns också olikheter mellan texterna när det gäller distribution av fel. I den skönlitterära texten och lärobokstexten är det mest frekventa felet *fel form i substantiv*, men i nyhetstexten är det *systemet väljer felaktigt ord*.

## 5. Analys

I min undersökning kan man se att det finns olika fel i texter, men deras skäl är okända. I detta kapitel söker jag ett svar för detta, men man måste påminna sig att många av dessa svar är bara gissningar eftersom det är omöjligt att se varifrån Google Translate tar ord som den sätter ihop till lösningen.

När Doug Arnold skriver om varför översättningen är svår för datorn säger han såhär: ”Part of the reason why translation is difficult for computers is that translation is just difficult: difficult even for humans (Arnold 2003:119)”

Översättningens korrekthet beror på språkpar, både av språkliga skäl samt Google Translates egenskaper. Översättningar från engelskan till svenskan är mer begripliga och har färre fel eftersom dessa språk är mer likadana än svenskan och estniskan, samt eftersom det finns antagligen mer material mellan engelskan och svenskan för GT att analysera. Det betyder att en orsak för felen är källmaterialets storlek. Även om Google har tillgänglighet till ett enormt antal av text är det inte tillräckligt.

Ping delar upp utmaningarna för maskinöversättning i språkliga och icke-språkliga. Icke-språkliga problem har att göra med intuitiv slutledningsförmåga som datorer saknar. Till exempel för att sluta sig till ett svar för en fråga som ”är 35 000 X 58 000 större eller mindre än 1?” använder människan sina kunskaper av världen, medan en dator måste besvara genom en uträkning. Icke-språkliga problem ses som svårare än språkliga eftersom extra-lingvistiska vetenskap är svår att få en dator att upptäcka (Ping 2009:165).

När Ping talar om språkliga utmaningar delar han de upp till två huvudtyper: tvetydigheter (ambiguity) och mismatchningar mellan käll- och målspråk. Ord som Google Translate inte disambiguerade satte jag till en egen underkategori och det förekom några fel i GT-lösningarna. Ping förklarar tvetydighet såhär:

*Lexical ambiguity is typically caused by polysemy and homonymy. Structural or grammatical ambiguity arises where different constituent structures (underlying structures) may be assigned to one construction (surface structure). [...] Uncertain anaphoric reference occurs when an expression can refer back to more than one antecedent, as in the following example: "There's a pile of inflammable trash next to your car. You are going to have to get rid of it." Here, it is not possible to determine, without reference to the context, whether "it" refers anaphorically to "trash" or "car". To disambiguate the second sentence, MT system developers must encode a great deal of real-world knowledge and develop procedures to use such knowledge. (Ping 2009: 166)*

Ping påstår att mismatchningar mellan källspråket och målspråket beror på lexikala och grammatiska skillnader mellan språken. Lexikala mismatchningar förekommer på grund av olikheter mellan hur språken klassificerar världen. (Ping 2009:166) Detta exemplifierar orden "morfar" och "farfar", som har samma motsvarighet på estniska. När det bara finns ett av de orden i meningen, som man vill översätta från svenskan till estniskan, finns det inget problem, men när meningen innehåller båda, som i min källtext (se avsnitt 4.4), måste man i detalj redogöra för ord. När man vill byta språkpar och översätta från estniskan till svenskan skulle GT:s lösningarna bli även sämre eftersom ordet "vanaisa" kan översättas både som "farfar" och "morfar". Ord som GT väljer i denna situation beror på ordens statistiska förekomst i relation till föregående ord (se Arnold 2003: 139-140).

Från det språkliga perspektivet kan man nämna några olikheter mellan svenskan och estniskan. I min undersökning förekommer de flesta av felen i kategorin om substantivens form. Estniskan är ett agglutinerande språk (d.v.s. affix fogas till en ordstam) och har 14 kasus medan svenskan har bara 2 eller 3 (nominativ, genitiv och "objektskasus" (*me, dig*)) och använder prepositioner. Dessutom har estniska adjektiv samma kasusböjning som substantiv: "när ett adjektiv står som attribut före ett substantiv, ersättes terminativen, essiven, abessiven och komitativen av genitiven, men i

övrigt står adjektivet i samma kasus som sitt huvudord” enligt Collinder, Aavik och Loorits (1944:139). Kasusen kan därför förklara varför det finns så många fel i substantiv, adjektiv och även namn, eftersom böjningsändelser fogas till alla.

Dessutom saknar estniska substantiv kön, som först syns som ett förenklande medel, men i några fall behöver översättaren ersätta pronomen med ett ord som antecknar kön för att visa meningens betydelse. Till exempel i meningen ”På trappan störtar Märta ofrivilligt i famnen på Urse, och han griper henne om armen för att hon inte ska falla” måste man till exempel ersätta ”han” med ett ord som betecknar en man eller ersätta det första ”hon” med namnet, annars vet läsaren inte vem som griper vem och vem som faller. Då kan en av översättningarna vara *”Trepil kiirustab Märta tahtmatult Urse sülle ja mees haarab tema käest kinni, et neiu ei kukuks.”* Google Translate översätter dock alla pronomen till estniska könlösa pronomen: *”Trepil plunges Martha tahtevastaselt sülle Urse ja ta haarab tema poolt käe, nii et ta ei ole asjas.”*

Språkliga olikheter kan även i någon utsträckning förklara felen i ordföljden eftersom estniskan har en annorlunda ordföljd i jämförelse med svenskan. Ett exempel är frasen ”nära den pakistanska gränsen” där i den estniska versionen placeras ordet ”nära” (”lähedal” på estniska) i slutet av frasen – ”pakistani piiri lähedal”.

Men fel finns även i meningarna där estniskan och svenskan har likartad ordföljd. För att inse varför, måste man veta vilken övergångsspråk GT har använt, men min undersökning visar att för det mesta används engelskan, som i dessa meningar har en annorlunda ordföljd.

## 6. Sammanfattande diskussion

Min undersökning har visat att nyhetstexten och lärobokstexten är lättare för Google Translate att översätta och man kan förstå textens innehåll på generell nivå. När man vill GT-översätta en skönlitterär text måste man bereda sig för olika fel och obegripliga meningar.

Dessutom distribueras felen olika i olika genrer. I den skönlitterära texten finns det många fel i adjektivets form, samt finns det mer fel när det gäller extra ord. I

nyhetstexten kan GT-lösningarna ha felaktiga ord samt fel i förkortningar. I lärobokstexten finns större delen av felen i formen av substantiv.

Anledningarna av felen är olika, de har att göra med språkliga och icke-språkliga problem. Bland icke- språkliga problem finns intuitiv slutledningsförmåga som datorer saknar, samt ett alltför litet antal översatta texter, som GT kan använda och analysera. Språkliga problem innebär olikheter mellan hur språken klassificerar världen, samt olikheter mellan ordföljd, kasus och även kön.

Man måste vara försiktig med användning av Google Translate eftersom det kan skapa obegripliga meningar. Jag rekommenderar använda GT för enkla texter som text 1 (lärobokstexten) eller för att uppskattningsvis begripa vad nyhetstexter (som text 2) talar om. Min undersökning har visat att skönlitterära texter är för invecklade för GT.

Hutchins ger följande anvisningar för användning av maskinöversättning i relation till mänskliga översättare: mänskliga översättare för litteratur och andra kulturellt-känsliga texter samt för verbala språket, och maskinöversättning för tekniska texter som är kulturellt neutrala samt för realtidsöversättningar på Internet (Hutchins 2011:52).

Av mina resultat kan man bland annat utröna att förutom engelskan använder GT finskan som övergångsspråk när den översätter från svenskan till estniskan.

Det finns många frågor kvar efter denna studie. Framför allt skulle framtida forskningar ta begripligheten av meningar i fokus eftersom min undersökning fokuserar på språkliga fel och bara har några meningar om begripligheten. Ett annat förslag till en vidare undersökning kan vara att klargöra vilken effekt användning av Google Translate i hemuppgifter har på studenternas kunskaper.

## Litteraturförteckning

Adams, T., 2010. Can Google break the computer language barrier?. *The Guardian* 19/12 2010. <<http://www.guardian.co.uk/technology/2010/dec/19/google-translate-computers-languages>>[Hämtad 2013-04-28]

Aidla, M., Cousins, S., Elken, M., Kanarbik, M., Mullamaa, K., Mullamaa, T., Pertšjonok, M., Praizner, M., Raag, R., Raag, V., Tuldava, J., 2004. *Rootsi-eesti sõnaraamat = Svensk-estnisk ordbok: 100 000 märksõna / uppslagsord*. Tallinn: Valgus.

Arnold, D. J., 2003. Why translation is difficult for computers. I: H. Somers, red. 2003. *Computers and Translation: A translator's guide*. Amsterdam: John Benjamins. s. 119-142.

Collinder, B., Aavik J. & Loorits, O., 1944. *Rootsi-eesti abiraamat = Svensk-estnisk hjälpbok / Gunnar Björkman'i "Svensk Språklära't" kasutades koostanud Björn Collinder, Joh. Aavik ja Oskar Loorits*. Uppsala: Almqvist & Wiksell.

Delblanc, S., 1973. *Åminne: en berättelse från Sörmland*. 2:a uppl. Stockholm: Bonnier.

Domeij, R., 2011. Datorn god översättare i snäv mening. *Språktidningen* 3, s. 44-49. <<http://www.spraktidningen.se/artiklar/2011/03/datorn-god-oversattare-i-snav-mening>> [Hämtad 2013-04-28]

Google Translates webbsida: <[http://translate.google.com/about/intl/sv\\_ALL/](http://translate.google.com/about/intl/sv_ALL/)> [Hämtad 2013-04-28]

Göransson, U., Helander, A., Parada, M., 2007. På svenska! Rootsi keele õpik. Tallinn: Kirjastus Ilo.

Hein, A. S., 2005. Datorn behöver statistik och grammatik. *Språkvård* 2005:1, s. 23–30. <<http://sprakteknologi.se/vad-aer-sprakteknologi/sagvallhein.pdf>> [Hämtad 2012-04-08]

Hutchins, J., 1997. From first conception to first demonstration: the nascent years of machine translation, 1947–1954. A chronology. *Machine Translation* 12(3), s. 195–252. <<http://www.hutchinsweb.me.uk/MTJ-1997.pdf>> [Hämtad 2012-04-28]

Hutchins, J., Brown, K. ed., 2006. Machine Translation: History. *Encyclopedia of Language & Linguistics, Second Edition*, vol 7, s. 375-383. <<http://www.hutchinsweb.me.uk/EncLangLing-2006.pdf>>[Hämtad 2012-04-08]

Hutchins, J., 2011. *Machine Translation systems and applications: past and present*. [pdf] Presentation på 11 maj 2011 i Imperial College London. <<http://www.hutchinsweb.me.uk/ImpColl-2011-ppt.pdf>>[Hämtad 2013-05-03]

Jordan, T., Kleberg., C. F., 2013. Jätteskalv skakade Iran och Pakistan. Dagens Nyheter. (Senast uppdaterad 2013-04-16 17:51) <<http://www.dn.se/nyheter/varlden/kraftigt-skalv-skakar-asien>> [Hämtad 2013-04-16 15:32]

Kingsley, J., 2011. It already speaks 57 languages as well as a 10-year-old. How good can it get?. *Slate* 31/10 2011. <[http://www.slate.com/articles/technology/technology/2011/10/google\\_translate\\_will\\_google\\_s\\_computers\\_understand\\_languages\\_be.single.html?utm\\_medium=referral&utm\\_source=pulsenews](http://www.slate.com/articles/technology/technology/2011/10/google_translate_will_google_s_computers_understand_languages_be.single.html?utm_medium=referral&utm_source=pulsenews)>[Hämtad 2013-04-28]

Lindqvist, A., 2012. *Maskinöversättning – bra eller dåligt? – En översättningsvetenskaplig studie om typiska fel och begriplighet hos maskinöversatta facktexter*. [Magisteruppsat] Göteborgs Universitet. <[https://gupea.ub.gu.se/bitstream/2077/32504/1/gupea\\_2077\\_32504\\_1.pdf](https://gupea.ub.gu.se/bitstream/2077/32504/1/gupea_2077_32504_1.pdf)> [Hämtad 2013-04-28]

Moberg, B., 2012. *Googleöversättning - Om Google Translate:s möjligheter och begränsningar vid översättning till svenska*. [examensarbete]. Lunds Universitet. <<http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=2776647&fileOId=2776649>>[Hämtad 2013-04-16]

Monti, J., Barreiro, A., Elia, A., Marano, F., & Napoli, A., 2011. Taking on new challenges in multi-word unit processing for machine translation. I F. Sánchez-Martínez, J.A. Pérez-Ortiz, red. *Proceedings of the Second International Workshop on*

*Free/Open-Source Rule-Based Machine Translation*. Barcelona, Spanien januari 2011. s. 11-19.

<<http://hdl.handle.net/10609/5646>> [Hämtad 2012-04-26]

Ping, K., 2009. Machine Translation. I: M.Baker, G. Saldanha, red. 2009. *Routledge Encyclopedia of Translation Studies*. 2:a uppl. New York: Routledge. s. 162-169.

Raag, R., 2007. *Lär känna estniskan: en introduktion till estniska språket*. 2:a uppl. Uppsala Universitet, Institutionen för moderna språk.

Stefansson, I., 2011. "Översätt den här sidan" *The advancement of Google Translate and how it performs in the online translation of compound and proper nouns from Swedish into English*. [C-uppsats]. Karlstads Universitet.

<<http://kau.diva-portal.org/smash/get/diva2:440909/FULLTEXT01>> [Hämtad 2012-04-25]

Vilar, D., Xu, J., d'Haro, L. F., & Ney, H., 2006. Error analysis of statistical machine translation output. I: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*. Genoa, Italien. s. 697-702.

<[http://hmk.ffzg.hr/bibl/lrec2006/pdf/413\\_pdf.pdf](http://hmk.ffzg.hr/bibl/lrec2006/pdf/413_pdf.pdf)> [Hämtad 2012-04-25]

# Bilagor

## Bilaga 1

På svenska! s. 107:

Drömmar och verklighet

Det började så bra:

Daniels kusin, Sofia, slutade skolan efter årskurs nio och började arbeta i en affär. På kvällarna sjöng hon i en kör och drömde om att bli sångerska.

När Sofia var 18 år träffade hon Tobias, en 19-årig TV-reparatör från Småland. Det blev kärlek vid första ögonkastet.

Efter två månader hyrde de en liten tvåa i en förort till Stockholm. De fick gamla möbler från vänner och släktingar, och Tobias, som är händig, målade och reparerade dem. Till slut blev lägenheten riktigt fin. Trots att de inte hade mycket pengar köpte de en CD-spelare och en dator. Efter sex månader som sambor gifte de sig, för Sofis väntade barn. De åkte på bröllopsresa i Sofias gamla Volvo till Tobias mormor i Gränna. Som minne av bröllopsresan har Sofia sparat en jättestor polkagris.

De fick en son som de döpte till Harald, Emil, Gustav (Emil efter morfar och Gustav efter farfar). Först var Sofia mammaledig, och sedan var Tobias pappaledig. De hade tur och fick en plats på dagis när Harald var ett år.

Sofia åkte med sin kör på turné till de baltiska länderna. Där fick hon många nya "sångarvänner", och efter en tid började det komma vykort och brev, mest från unga män (tyckte Tobias). Han blev svartsjuk, och de började gräla. Först tyckte Sofia att han var löjlig, men sedan blev hon trött på hans misstänksamhet, och efter ett tag bestämde de sig för att skiljas.

Hur kunde det sluta så här?

Nu bor Sofia och Harald kvar ensamma i lägenheten, och Sofia tänker ofta: "Gjorde jag rätt eller fel?"

GT-översättning 16:e april 2013:

Unenäod ja reaalsus

See algas nii hästi:

Daniel nõbu, Sofia, koolist lahkunud pärast üheksanda klassi ja alustas tööd kauplus. Õhtuti ta laulis koor ja unistas laulja.

Kui Sophia oli 18, ta kohtus Tobias, 19-aastane TV remondimees alates Småland. See oli armastus esimesest silmapilgust.

Kahe kuu pärast, nad üüritud väike korter Stockholmi eeslinnas. Neil oli vana mööbli sõbrad ja sugulased ja Tobias, mis on mugav, värvitud ja parandas neid. In the end, korter tõesti kena. Kuigi nad ei ole palju raha, et nad ostsid CD-mängija ja arvuti.

Pärast kuus kuud kooselude abiellus, Sofia oli rase. Nad läksid edasi mesinädalad Sofia vana Volvo Tobias vanaema Gränna. Kuna mälu on mesinädalad, Sofia salvestatud ilmatu pulgakommi.

Neil oli poeg, keda nad nimega Harald, Emil, Gustav (pärast vanaisa Emil ja Gustav pärast vanaisa). Kõigepealt oli Sofia sünnituspuhkuse ning seejärel oli Tobias isapuhkus. Nad olid õnnelikud ja sai koha lasteaias kui Harald oli üks aasta.

Sofia läks koos oma koori visiiti Balti riikidesse. Ta sai palju uusi "sångarvänner," ja mõne aja pärast hakkas postkaardid ja kirjad, peamiselt noored mehed (arvasin Tobias). Ta sai armukade, ja nad hakkasid tüli. Esimene arvas, Sofia, et ta oli naeruväärne, kuid siis sai ta väsinud tema kahtlused, ja mõne aja pärast nad otsustasid lahutada.

Kuidas seda lõpetada sel viisil?

Nüüd elab Sofia ja Harald üksi korteris, ja Sofia arvavad sageli: "Kas ma õige või vale?"

## Bilaga 2

”Hundratals befaras döda i skalv”

Hundratals befaras döda i skalv

Hundratals människor befaras ha omkommit i en kraftig jordbävning i sydöstra Iran, uppger iranska myndigheter enligt Reuters.

Skalvet, som mätte 7,8 på Richterskalan, inträffade i Iran i på tolv mils djup, enligt USA:s geologiska institut USGS. Det inträffade i sydöstra delen av landet, nio mil sydöster om staden Khash klockan 10.44 lokal tid, enligt USGS. Området ligger i provinsen Sistan Baluchistan, nära den pakistanska gränsen.

Iransk statlig tv rapporterar att minst 40 personer har dödats i skalvet. Även på den pakistanska sidan av gränsen har det krävts dödsoffer. Minst fem personer har omkommit, uppger en pakistansk tjänsteman till nyhetsbyrån Reuters.

– Det var det största jordskalvet i Iran på 40 år, säger en iransk tjänsteman till Reuters.

Senast förra veckan drabbades Iran av ett jordskalv som mätte 6,3 på Richterskalan. 30 personer omkom och över 800 skadades.

– Iran är ett land som ofta drabbas av jordskalv. Det finns inget samband mellan skalven, det är bara tillfällighet, säger Reynir Bödvarsson, seismolog vid Uppsala universitet.

Han menar att det nu bara går att hoppas på att skadorna för invånarna blir så små som möjligt, men det kommer att komma efterskalv.

– Det kommer att bli efterskalv, det är bara att hoppas på att de blir mindre, säger Reynir Bödvarsson.

Strax före två på eftermiddagen fanns ännu inga uppgifter om att någon svensk skulle vara berörd av skalvet.

– Vår ambassad i Teheran undersöker detta nu och vi kommer säkert få mer information snart men så här långt finns inga uppgifter om skadade svenskar, säger Catarina Axelsson på UD:s presstjänst.

Kärnkraftverket vid iranska Bushehr ska inte ha skadats av skalvet, enligt det ryska företag som byggt det.

Uppdateringar följer

GT-översättning 16:e april 2013:

Sajad kartsid surnud värisema

Sajad inimesed kartsid surnud võimas maavärin kagu Iraan vastavalt Iraani ametivõimude sõnul Reuters.

Värisema, mis mõõdeti 7,8 Richteri skaalal, toimus Iraanis kaksteist miil sügav, vastavalt US Geological Institute USGS. See toimus kaguosas, üheksa miljonit kagu linna Khash kell 10:44 kohaliku aja järgi USGS. Piirkond asub provintsis Sistan Baluchistan lähedal Pakistani piiri.

Iraani riigiteleviseiooni teatas, et vähemalt 40 inimest on hukkunud värisema. Ka Pakistani pool piiri on vaja inimohvreid. Vähemalt viis inimest on tapetud, vastavalt Pakistani ametnik rääkis Reutersile.

- See oli suurim maavärin Iraanis 40 aastat, ütles Iraani ametnik Reutersile.

Just eelmisel nädalal Iraani tabas maavärin, milles mõõdeti 6,3 Richteri skaalal. 30 inimest hukkus ja üle 800 vigastada.

- Iraan on riik alati maavärinate. Puudub seos maavärinate, see on lihtsalt kokkusattumus, ütles Reynir Bödvarsson, seismoloog Ülikooli Uppsala.

Ta usub, et see on nüüd võimalik vaid loota, et kahju elanikele on nii väike kui võimalik, kuid see saab toimuda pärast värisema.

- Tekib järeltõuked, siis tuleb loota, et nad on väiksemad, ütleb Reynir Bödvarsson.

Just enne kahte pärastlõunal, ei olnud teateid kõigist Rootsi mõjutaks värisema.

- Meie saatkond Teheranis uurib seda nüüd ja me kindlasti saada lisateavet varsti, kuid seni ei ole infot vigastatud rootslased, ütleb Catarina Axelsson MFA pressiteenistus.

Tuumaelektrijaama Bushehr Iraan ei tohi olla kahjustatud värisema, vastavalt Vene ettevõtte, mis ehitasid.

Uuendused järgmiselt

### **Bilaga 3**

Sven Delblanc Åminne – en berättelse från Sörmland, s. 62

Gubbarna suckar och blickar tungsint ner i sina händer. På trappan störtar Märta ofrivilligt i famnen på Urse, och han griper henne om armen för att hon inte ska fall. Hon står för ett ögonblick innesluten i hans famn, och hennes hjärta flämtar till som en ljuslåga. Med ens är hon maktlös och tung av sin kvinnlighet, benen ger vika under henne, och hela hennes viljekraft är hjälplöst samlad och innesluten i hans hårda hand.

Han luktar cigarr och eau de cologne. Hans blå ögon i det stränga, oförbindliga ansiktet. Sår hård hans hand.

- Är det inte... Är det inte Märta? Jag tror du har vuxit?

Hon vindar som en nykalvad kviga för att undvika hans hårda blick, gör sig fri, kommer sig inte för med att niga och skyndar i väg på skälvande ben. Och förklät som är lortigt, och inte ser ja väl klok ut i håret, tänker hon oredigt.

Tok-Harry reser sig och betraktar henne med sin tomma, blanka blick. Nynnande sin eviga sång börjar han lunkande förfölja henne som en kelen tjurkalv.

Så har ingen sett på mig förut: i en mans blick blev jag till kvinna. Släktets och mödrarnas röster viskar i mitt blod: nu är du inte längre något barn, var redo... Och jag är maktlös och tung av min längtan att lyda dessa röster, följa den väg de visar, underkasta mig och bära...

Urse blev stående ett ögonblick förströdd. Han drog näsduken ur bröstfickan och slog ut den i ett doftande moln. Tankspridd torkar han sig om händerna, i en dunkel förnimmelse att ha vidrört något naket och vått. Märta, Svenssons flicka. Kände inte igen henne först. Nästan vuxen.

GT-översättning 16:e april 2013:

Poisid ohkab ja otsib melanhoolia nende kätte. Trepil plunges Martha tahtevastaselt sülle Urse ja ta haarab tema poolt käe, nii et ta ei ole asjas. Ta seisab hetkel sisaldas tema süles ja tema süda ohe kui küünal leek. Järsku ta on jõuetu ja raske koos oma naiselikkust, tema jalad andma teed all tema ja kõik tema tahtejõud on abitult kogutakse ja tema kõva käsi.

Ta lõhnab sigari ja odekolonnid. Tema sinised silmad karm, sidumata nägu. Haavad kõvasti käes.

- Kas see pole mitte ... Kas pole Martha? Ma arvan, et sa oled kasvanud?

Ta kerib nagu äsja poeginud mullikas vältida tema raske pilku, vabastab ennast, ei hooli võttes kniks ja kiirustab kaugusel värisevad jalad. Ja põll, mis on räpane, ja ei näe Nojah naeruväärne tema juuksed, ta arvab, korrastamata.

Tok-Harry püsti ja vaatab teda oma tühi, tühi vahtima. Koolibrid oma igavest laulu alustab oma kõmpima kiusavad teda kaisu Mullikka.

Nii et keegi on näinud mind enne: a mehe silmad, olin naine. Perekond ja emade hääled sosistavad minu veres: nüüd sa ei ole enam laps, oli valmis ... Ja ma olen jõuetu ja raske minu soov kuuletuma neid hääli, järgige tee show, esitama ise ja kandma ...  
Urse jäi seisma hetkel segane. Ta tõmbas oma rätiku taskust ning palli läbi lõhnava pilv.  
Hajameelne kuivab ise oma kätega, vaid ebamäärane tunne ta puudutanud midagi alasti ja niiske. Martha, Svensson tüdruk. Ei tundnud teda alguses. Peaaegu täiskasvanud.

## Resüme

Käesoleva bakalaureusetöö pealkirjaks on „Tõlkimine rootsi keelest eesti keelde Google Translate'i abiga“. Huvi teema vastu tärkas sellest, et mõningates olukordades keelatakse Google Translate'i kasutus, sest arvatakse, et sellest tekib liiga palju vigu. Kuigi keeleõppes on vigade esinemine tavaline, võib näiteks keelt emakeelena rääkiv inimene märgata, et sellegipoolest on kasutatud just Google Translate'i tõlget. Nii tekkiski mõte, kas vead, mida Google Translate teeb, on kuidagi erilised ja korduvalt pidevalt.

Just seepärast uuringi enda töös, millised vead tekivad kui lasta Google Translate'il tõlkida paarisajasõnalise teksti rootsi keelest eesti keelde. Et tuua välja ka erinevusi, missuguseid vigu võib eri tüüpi tekstides tekkida, olen valinud algtekstideks 3 teksti erinevatest žanritest: rootsi keele õpikutekst, artikkel ajalehest ning ilukirjanduslik tekst. Töös vaatlen, millist tüüpi vead on kõige sagedasemad, kuidas erinevad vead žanrite vahel ning samuti pakun mõningaid seletusi sellele, miks need vead võivad olla tekkinud.

Bakalaureusetöös paljastub, et ilukirjanduslik tekst on Google Translate'i jaoks kõige keerulisem, ajalehetekst aga sisaldas kõige vähem vigu. Samas ei saa öelda, et õpikutekst oleks tõlgitud palju rohkemate vigadega kui ajalehetekst, nende veaprotsent on üsna sarnane.

Võib öelda, et erinevused on ka selles, millist tüüpi vigu kõige enam ühes või teises tekstis tekib. Õpikutekstis ning ilukirjanduslikus tekstis tekkis kõige rohkem vigu käänetega, ajalehetekstis aga oli kõige sagedasem see, et Google Translate'i tõlkes oli sõna vale tähendusega.

Veel saab tööst järeldada, et lisaks inglise keelele, mida Google Translate kasutab kui üleminekukeelena, on kasutusel ka soome keel, vähemalt sel juhul, kui tõlgitakse rootsi keelest eesti keelde.

Bakalaureusetöö analüüsisiosas järeldub, et vigade põhjuseid on mitmeid, ning neid võib jaotada keelelisteks ja mittekeelelisteks. Mittekeelelised vead on põhjustatud sellest, et arvutitel puudub inimestele omane argimõistus, samuti on probleemiks see, et kasutatavat algmaterjali, mida Google Translate saab enda tõlgete jaoks kasutada, on liiga vähe. Keeleliste põhjustena võib nimetada seda, et keeled liigitavad sõnu erinevalt, samuti on rootsi ja eesti keele vahel grammatilisi erinevusi nagu käänded, sugu ja sõnade järjekord lauses.

**Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks**

Mina \_\_\_\_\_ Kertu Hiire \_\_\_\_\_

(*autori nimi*)

(sünnikuupäev: \_\_\_\_\_ 18 aprill 1990 \_\_\_\_\_)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose  
\_\_\_\_\_ Översättning från svenska till estniska med hjälp av Google  
Translate \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

(*lõputöö pealkiri*)

mille juhendaja on \_\_\_\_\_ Maiu Elken \_\_\_\_\_,

(*juhendaja nimi*)

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, \_\_\_\_\_ Kertu Hiire \_\_\_\_\_ „22“ mai 2013