A Comparative Evaluation of Localisation Tools: Reverso Localize and SYSTANLinks

GRAY, Charlotte

Abstract
This thesis compares and analyses two free, online tools, called Reverso Localize and SYSTRANLinks, which have both machine translation and website localisation capabilities. The aim of this work was to subjectively evaluate the performance of the tools, as well as execute an end-user satisfaction evaluation. In order to do so, an evaluation centering around three EAGLES characteristics (functionality, usability, and efficiency) was designed. From the results it was possible to create recommendations for both tools in order to improve their systems as well as their end-user satisfaction scores, and to state which one more effectively localised and translated the text website provided.

Reference

Available at:
http://archive-ouverte.unige.ch/unige:55939

Disclaimer: layout of this document may differ from the published version.
Charlotte Gray

MA Thesis

A Comparative Evaluation of Localisation Tools: Reverso Localize and SYSTRANLinks

Supervisor: Prof. Pierrette Bouillon

Jury: Dr Lucía Morado Vázquez

Multilingual Information Processing Department
Faculty of Translation and Interpreting
University of Geneva
August 2014
Acknowledgements

I would like to thank Prof. Pierrette Bouillon who helped and supported me in my 11th hour decision to change my specialisation to translation technologies, as well as throughout writing this thesis. Thanks also to Dr Lucía Morado Vázquez for being on my jury and for her notes throughout, as well as giving me the localisation bug in her fantastic classes.

Last, but not least, a huge thank you goes to my parents for their unconditional love and support, and to Ewan, for all those Skype conversations which kept me sane! Without them the last two years would not have been possible.
## Contents

### Introduction

<table>
<thead>
<tr>
<th>1 Localisation</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Concept</td>
<td>6</td>
</tr>
<tr>
<td>1.2 History and Development</td>
<td>10</td>
</tr>
<tr>
<td>1.3 Localisation Process</td>
<td>11</td>
</tr>
<tr>
<td>1.4 Conclusion</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 Website Localisation</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>18</td>
</tr>
<tr>
<td>2.2 Websites and HTML</td>
<td>19</td>
</tr>
<tr>
<td>2.3 Localisation Methods</td>
<td>21</td>
</tr>
<tr>
<td>2.3.1 Types of Website and their Localisation</td>
<td>21</td>
</tr>
<tr>
<td>2.3.2 Difficulties for Localisers</td>
<td>25</td>
</tr>
<tr>
<td>2.4 Cultural Issues within Website Localisation</td>
<td>25</td>
</tr>
<tr>
<td>2.5 Conclusion</td>
<td>26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 Machine Translation</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>28</td>
</tr>
<tr>
<td>3.2 Machine Translation System Types</td>
<td>29</td>
</tr>
<tr>
<td>3.2.1 Direct Linguistic Systems</td>
<td>29</td>
</tr>
<tr>
<td>3.2.2 Indirect Linguistic Systems</td>
<td>31</td>
</tr>
<tr>
<td>3.2.3 Statistical Systems</td>
<td>33</td>
</tr>
<tr>
<td>3.3 Machine Translation Evaluation</td>
<td>34</td>
</tr>
<tr>
<td>3.3.1 Manual Evaluation</td>
<td>35</td>
</tr>
<tr>
<td>3.3.2 Automatic Evaluation</td>
<td>36</td>
</tr>
<tr>
<td>3.4 Conclusion</td>
<td>37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4 Localisation Tools Used</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Reverso Localize</td>
<td>38</td>
</tr>
<tr>
<td>4.2 SYSTRANLinks</td>
<td>41</td>
</tr>
<tr>
<td>4.3 Conclusion</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5 Methodology</th>
<th>46</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Purpose and Objectives</td>
<td>47</td>
</tr>
<tr>
<td>5.2 Task Model and Parameters</td>
<td>47</td>
</tr>
<tr>
<td>5.2.1 Language Parameters</td>
<td>47</td>
</tr>
<tr>
<td>5.2.2 Platform Parameters</td>
<td>48</td>
</tr>
<tr>
<td>5.2.3 Corpus Used</td>
<td>48</td>
</tr>
<tr>
<td>5.2.4 Participants</td>
<td>50</td>
</tr>
<tr>
<td>5.3 Characteristics Evaluated</td>
<td>51</td>
</tr>
<tr>
<td>5.4 Evaluation Method</td>
<td>53</td>
</tr>
<tr>
<td>5.5 Conclusion</td>
<td>55</td>
</tr>
</tbody>
</table>
Introduction

This master’s thesis will be a comparative evaluation of two online localisation tools, Reverso Localize and SYSTRANLinks, in order to ascertain to what extent they meet users’ needs. The tools focus on two main aspects: localisation and machine translation, two aspects which are particularly interesting to me, and which therefore make up the evaluation of the two tools.

The reason for carrying out this work is not just because of my interest in localisation and machine translation. The growing localisation domain means that companies such as Reverso-Sofiásmó and Systran are starting to respond to a real need in the market: that of small companies, as well as individuals, who wish to expand their customer base and therefore need to localise their website. These localisation tools are, to a certain extent, cutting-edge technology, offering limited services for free that many companies would charge for. It is necessary to evaluate these tools, to really see if they help or hinder the people that have no linguistic or technological skills. This type of evaluation has been carried out on Reverso Localize in the last year in Peron (2013), which has been incredibly useful and has been used as a base for this master’s thesis.

The evaluations of Reverso Localize and SYSTRANLinks were carried out through scenario testing, where a simulation was used in order to obtain results on how well the tools work. From these results, my aim is to point out the strengths and weaknesses of both the tools, with an aim to state which tool would be better for end users and why.

This thesis is separated into seven chapters. Chapters 1, 2, and 3 are theoretical in nature, with chapter 1 outlining the general localisation domain, chapter 2 going into more detail on website localisation, and chapter 3 explaining machine translation.

Chapter 4 will present the functionalities of Reverso Localize and SYSTRANLinks, and point out their similarities and differences. Chapter 5 will set out the methodology used for the evaluation, with chapter 6 dedicated to the design and execution of the evaluation. Chapter 7 will display the results and provide an analysis of all of them. Finally, the conclusion will summarise the results, state which tool was more effective in the evaluation, give recommendations for improvement, and describe the limits to this thesis and ideas for further research. The annexes include the material used as part of the evaluation, as well as extra information, results, and calculations.
1 Localisation

This thesis will be evaluating two tools for their localisation abilities. Localisation is a complex subject which covers a broad range of translation and technological issues, and therefore this chapter will be giving an introduction to the concept. It will start by defining and discussing the general concept of localisation (section 1.1) before moving onto its history and development (section 1.2). Section 1.3 will detail the localisation process and the tools that a localiser might use during this process, before the conclusion in section 1.4.

1.1 Concept

It is first necessary to make a distinction between localisation, globalisation and internationalisation. The Localization Industry Standards Association (LISA), which operated until 2011 and was one of the industry’s oldest associations, defined localisation as “the adaptation of a product to make it appropriate to the target locale where it will be used and sold” (LISA 2007c). ‘Locale’ in localisation terms is used in a technical context, meaning a certain combination of language, region, and character encoding (Esselink 2000).

Globalisation takes this concept further. It is defined as “the process of making all the necessary technical, financial, managerial, personnel, marketing and other enterprise decisions necessary to facilitate international business” (LISA 2007a).

Internationalisation, on the other hand, is “the process of designing a product which is as culturally and technically ‘neutral’ as possible, and which can therefore easily be localized for a specific culture or cultures” (LISA 2007b).

It is possible to see how these three concepts link together. The Globalisation, Internationalisation, Localisation, and Translation (GILT) cycle is a model which shows how localisation and internationalisation can be perceived as being part of a bigger globalisation scale, while also demonstrating the interdependence of all stages (Jiménez-Crespo 2013). Figure 1 shows how localisation is placed within the wider concept of globalisation, with it being the second step in the globalisation process, after internationalisation, which normally occurs during the developmental stages of a digital product (Jiménez-Crespo 2013). Note that translation is seen as a separate step to localisation.
It is important to remember that localisation is “the translation and adaptation of a software or web product” (Esselink 2000, p. 1), and that “the objects that are processed in localization are ‘products’ and not ‘texts’” (Jiménez-Crespo 2013, p. 13), meaning that it differs greatly from traditional document translation as it includes a digital product. Attention must be drawn to Esselink’s separation of localisation into the concepts of translation and adaptation. Jiménez-Crespo (2013, p. 15) points out that the translation of text “generally constitutes the bulk of a localization project,” but adaptation is the “additional component that localization provides,” with the adaption being of the product itself to the locale.

There are many problems that need to be addressed when localising products. LISA (2007a) split localisation into four issues, called ‘linguistic,’ ‘business and cultural,’ ‘technical,’ and ‘physical,’ as shown in the list below:

1 Linguistic issues must be overcome through the process of linguistic adaptation. This includes translating documentation related to the product, but also marketing materials, webpages, and subtitles. Linguistic adaptation can also have an impact on product design, where the interface of a product may need to be adapted to suit significantly longer or shorter character chains. Figure 2 below shows that the developers of Microsoft Word 2010 would have had to make sure that the French language version of the software had specified enough space to fit the chain “Enregistrer sous” in the menu.
2 Business and cultural issues are so commonplace that people tend to overlook them. Problems from local payment options to local address and telephone formats could plague international companies on a daily basis if their website or documentation was not suited to the locale. Cultural mistakes are often made simply because there has been no adequate research into the culture for which a product has been localised. Consider Figure 3 below, as taken from Nauert (2007). It shows Swarovski’s wedding jewellery page, as localised for China. However, the localiser or the company is clearly not aware of the local market, that is, that white is not associated with weddings in China, and therefore the marketing message is lost on the Chinese culture.

Figure 3: Swarovski wedding jewellery website for China (Nauert 2007, p. 3)
3 Technical issues are mostly related to supporting local languages. For example, Greek, Russian, and Arabic have different alphabets to English, so design must be adapted accordingly. Furthermore, Arabic is read from right to left instead of left to right, so user interfaces must be changed. However, these bigger technical issues do not overshadow the importance of smaller ones: thought must be put into language conventions, such as number separators and date formats. Figure 4 below shows this with the Arabic version of Microsoft Word 2010. Note how the interface has been changed to accommodate users reading from right to left.

![Figure 4: Arabic version of Microsoft Word 2010](image)

4 Physical issues do not affect localisation as much, but LISA (2007a) states that it is still necessary to mention them. Physical modification is not required for most digital products, but the modifications made to the items they depict can play a role. For example, plugs and sockets change from country to country, and documentation related to a localised digital product must reflect these changes.

These localisation issues can be applicable to any type of localisation. Issues surrounding website and software localisation have already been briefly mentioned, but videogame localisation, small device localisation, and multimedia localisation can all equally be affected by cultural, business, and technical issues during the localisation process, such as by local laws defining how much violence is permitted in videogames (Peron 2013).
1.2 History and Development

The localisation industry is relatively young, and can be traced back to the late 1970s and early 1980s, when personal technology became more mainstream (Jiménez-Crespo 2013). The industry emerged properly in the 1980s, when software publishers based in the United States of America grew to understand the need for localised and international products in order to grow their market and make a larger profit. Historically, software publishers have typically localised their products into FIGS (French, Italian, German, and Spanish) first, as well as into Japanese.

At first, software publishers used their in-house translation departments, or outsourced their translation work to freelance translators. As localisation projects increased in size, importance, and complexity, companies had to find an outsourcing model, as they simply did not have the time or knowledge to manage localisation projects (Jiménez-Crespo 2013).

This led to the emergence of multi-language vendors (MLVs) in the 1990s, who started to sell not just translation services, but also engineering, testing, and support services, and the localisation industry as it is now known was formed. In 1990, LISA was formed in Switzerland, with the goal to promote “the localization and internationalization industry and provide a mechanism and services to enable companies to exchange and share information on the development of processes, tools, technologies and business models connected with localization, internationalization and related topics” (Esselink 2000, p. 8).

Two more organisations – the Localisation Research Centre (LRC) and the Software Localisation Interest Group (SLIG) – were founded in Ireland during the 1990s in order to establish contacts within the industry. Ireland has traditionally been seen as the leader in the localisation industry, as during the 1990s the Irish government decided to focus on attracting foreign investment from the research and development sector, and targeted hi-tech, blue chip companies for government grants.

By this point, the localisation industry had been well defined as an industry in its own right. During the 1990s and 2000s there were changes in the industry where companies started to merge in order to provide more services, and in the early 2000s web localisation overtook software localisation as the largest localisation type (Jiménez-Crespo 2013).

Nowadays, the localisation sector is worth around $1 billion annually in Ireland, with estimates of 14,000-16,000 jobs in the Irish localisation and translation sector. More than 150 researchers are working on advancing the industry, as well as founding some of the most well-known localisation specialisation qualifications in the world (CNGL 2012).
The Common Sense Advisory, a market research company for the localisation, globalisation, and internationalisation industry, estimates that the worldwide translation and localisation market was worth $12 billion in 2010, with the top ten language service providers earning $2.7 billion between them in 2012 (DePalma and Kelly 2012).

1.3 Localisation Process

The localisation process is much more complex than the simple translation of text linked to a product; it encompasses and is encompassed by a broad range of competences. As mentioned in section 1.1, localisation must overcome linguistic, business, cultural, and technical issues, and therefore a successful localisation project must include localisers who have linguistic, business, cultural, and technical competence.

There are many people involved in the localisation process, and each person is linked to what is normally called an account manager. Figure 5 below shows the hierarchy of people in a localisation team.

![Figure 5: People involved in the localisation process (Esselink 2000, p. 14)](image)

Within each phase of the workflow, each person involved can make use of certain tools at their disposition, as shown in Figure 6. It is interesting to note that there are just as many tools needed to manage the localisation itself as there is to help with the translation of the documentation.
Quah (2006) points out that the operational and management tools indicated can be used for all translation management, not just within localisation, and can be used to manage many languages at any time for the same source material.

The translation tools are used for the translation and management of the source files, while the management tools help monitor the localisation process as a whole, and enable team members to keep updated with each stage of the localisation process.

**Figure 6: Tools used in the localisation process (Quah 2006, p. 114)**

Although the localisation process varies greatly depending on the size of the project, the languages involved, and the type of product being localised, Esselink (2000) has demonstrated a typical localisation workflow, as shown in Figure 7.
For Esselink, the process starts long before the source material is delivered – a client must first approach a localisation team to ask for their product to be localised. The client would usually provide some source material so that it can be evaluated in order for a quote to be created. There is then a kick-off meeting for members of the localisation team, in order to provide an overview of the project and to provide training of the particular product to be localised, if needed.

Schäler (2009) breaks down the next steps of the localisation process into five key components: analysis, preparation, translation, engineering and testing, and project review. If this were to be linked with Esselink’s workflow, it may look like Figure 8 below.

In the following descriptions of the localisation workflow (analysis, preparation,
translation, engineering and testing, and project review), the tools that a localisation
project uses (see Figure 6) and the people involved (see Figure 5) will be pointed out at
the appropriate point.

Analysis

The main analysis is of the source material, which can take many forms. For
example, a project manager may receive a live version of the original product, a build
environment, a documentation folder, or, if necessary, previously localised versions of the
product, or existing glossaries or translation memories (Esselink 2000). All aspects of
the source material are evaluated by the lead linguist and the engineering manager, with
problem areas being identified. It is also necessary to analyse the source material to find
out if it is suitable for localisation, or if a new version needs to be created that better fits
the locale (Schäler 2009).

Schäler (2009, p. 158) mentions that a process called ‘pseudo-translation’ can
occur in the analysis stage, where “original strings are automatically replaced with
strings expanded by a certain percentage and containing characters from the target
language according to pre-defined algorithms in order to mimic a translation.” This
process determines how a translation would affect a product and is akin to a small
localisation project on the product, to determine what issues may arise and how to
approach its localisation.

Preparation

On the basis of the analysis, the project is given a go-ahead, and members of the
team such as the project manager, the engineering manager, the lead linguist, the desktop
publishing (DTP) manager, and any localisation engineers, assemble a project plan and
a localisation kit (Schäler 2009).

The project plan includes scheduling and budgeting plans, which are constantly
reviewed as the project develops. The localisation kit contains all the material needed
for the project, such as the source material, tools, guidelines, and instructions specific to
the client and project, which is made available to all members of the localisation team.

At this stage it is also highly likely that hardware and software may need to be
reconfigured, and the product may have to undergo an internationalisation process in
order to make it localisable (Quah 2006).

During this part of the process, project managers will start to make use of project
management and workflow management tools, and will continue to update them as the
process advances. These tools help the team keep updated with the project; the team
can see at what stage they are and improve their time management.
Translation

According to Schäler (2009), there are many tasks in addition to translation in this step. It is necessary for computer-assisted translation (CAT) tool engineers to prepare terminology databases, maintain translation memories, analyse and pre-translate text with automated translation systems, use and maintain machine translation applications and resources, as well as manage the source and target files.

As briefly mentioned, many localisation projects will involve tools such as machine translation in the form of either automatic or computer-assisted translation, as glossaries and term banks can be used and memories can be updated. With the somewhat repetitive nature of some localisation projects, in particular with software localisation, such as similar menus and commands being repeated throughout a piece of software, machine translation can save time.

Figure 9 below shows the translation process using a machine translation system, a translation database, and a terminology database.

![Diagram of translation process](Quah 2006, p. 117)

**Figure 9: Translating using a machine translation system (Quah 2006, p. 117)**

In this process, a project starts with the source language texts or files to be translated (in the case of strings, these are extracted prior to the translation process). The source language text is compared to the translation database to find similar resources, before being pre-edited using the terminology database to clarify or simplify the text as needed. In the pre-editing stage, as well as using the terminology database to verify that the correct terms are being used, there tends to be a set of linguistic rules and patterns that must be followed to make a text as simple and as unambiguous as possible for the system that will be used for the translation (Bouillon 2014a).
This pre-edited source language text is then machine translated, the process of which will be explained in Chapter 3, to produce a draft target-language text, before being post-edited. Post-editing is the process where the raw machine translation output is corrected by linguists, editors, or translators to make the final text publishable (Bouillon 2014c). During post-editing, any new terms are put into the terminology database. Once the final target language text has been produced, it is put into the translation database, where it is aligned with the source language text to create translation segments, which the translation system can easily manage for its memory.

After the translation is complete, linguistic quality testing is carried out by proofreaders and the lead linguist. This happens before the engineering and testing phase, as it needs to be as perfect as it can be, considering the time constraints, before being used as part of the localised product.

Engineering and Testing

The engineering and testing process is mostly unrelated to the linguistic side of the localisation, as quality testing of the translations would have happened before this stage. Changes may be made to the product using document and graphic design tools, and the localised files will be compiled by localisation engineers into a running application (Esselink 2000).

In the testing stage, quality control engineers test the functions of the localised product using testing and validation tools as well as test plans, test scripts, and error reporting and tracking procedures to ensure the quality of the product. These test plans, test scripts, and procedures are supplied in the preparation stage (Schäler 2009).

Once the quality assurance testing has met the required standards, the final material is delivered to the client.

Project Review

At the end of the project, the process undergoes a thorough review by both the localisation team and the client to assess the team’s work, the technical and linguistic quality of the delivered product, and to identify and suggest areas for improvement. Any localised files and documents will be archived by the localisation team for any potential future use for the client, and the project will be officially closed.

1.4 Conclusion

This section has presented the concept and history of localisation as a discipline that encompasses translation, and has also presented it as part of the bigger globalisation and internationalisation cycle. Issues that localisers must constantly be aware of (which
LISA (2007c) called linguistic, business and cultural, technical, and physical) have been discussed, and examples have been used to show where localisers can go wrong or how products must be changed to accommodate to other locales. The general localisation process has been described, with further reference made to the people and tools involved in the process. The next chapter will go into further detail on the localisation of websites.
2 Website Localisation

This chapter will be going into more detail on the concept of website localisation, as this is what part of the evaluation will be based on. Section 2.1 will start by giving an introduction on the reasons for localising a website, before defining the concept of a website and detailing its main component – that of HTML (section 2.2). Section 2.3 will go on to describe the different types of websites that exist and localisation methods for each, as well as outlining problems that web localisers may face within the industry. Section 2.4 will briefly describe cultural issues within localisation, before presenting a conclusion in section 2.5.

2.1 Introduction

Since the 1990s, localisation has become a big part of companies’ marketing plans. There are many different types of localisation - the main ones are software localisation, videogame localisation, and web localisation. Software has traditionally held the biggest part of the localisation market, but since the introduction of the World Wide Web and globalisation strategies, web localisation has become increasingly important. In 2013, web localisation received the biggest volume of business, followed by videogames and software (Jiménez-Crespo 2013).

A website can be seen as a company’s main marketing tool, and localising a company website proves to prospective clients that an organisation is willing to accommodate their needs (de Bortoli and Maroto 2001): according to the Common Sense Advisory (2006), 52.4% of people only buy from websites where information is presented in their native language.

In 2001, Theo Schewe, a Swedish researcher, conducted a study which established a close link between a company’s marketing policy and its language choice for its web presence, meaning that website localisation is a function of international marketing strategies (Sandrini 2005). In the study, Schewe presents what he calls ‘web site language design strategies,’ with three different types of website: monolingual, bilingual, and multilingual. Within each type, the language choice shows a certain marketing strategy that stretches from domestic, with a monolingual website in a company’s native language, to a global marketing strategy, where a company has a central website in its native language, with local websites in local languages (Sandrini 2005). This study clearly bears some weight: 65% of multinationals believe that localisation is important for achieving higher revenue (GALA 2014).

Website localisation is clearly a necessity for companies wishing to have a global
presence, but, as this chapter will demonstrate, each website is different and many strategies must be taken into account during the localisation process.

2.2 Websites and HTML

Website localisation is “the process of modifying a website for a specific locale according to the goals outlined by the client” (Sandrini 2005, p. 2). In order to start this process, however, the concept of a website must be defined. A website “encompasses all web pages which are accessible under a common Web address (domain name),” and is made up of “a number of documents, graphics, programs, […] each of which is identified by a uniform resource identifier (URI),” or in simpler terms, a website is “a container with an address and a domain name” (Sandrini 2005, p. 1).

It is impossible not to talk about websites without mentioning Hypertext Markup Language (HTML), because it is “the standard file format for displaying pages on the World Wide Web” (Esselink 2000, p. 205). Thanks to standards in the localisation industry such as XML Localization Interchange File Format (XLIFF), there is less need nowadays for localisers or translators to become experts in various formats when localising a website (Esselink 2006), but people working in the field are encouraged to have a general understanding of it.

There are two basic types of markup language, as defined by Esselink (2000):

1. **Procedural markup** mainly defines the layout of a text, for example, the font, font size, and colour of each heading style in a document.

2. **Descriptive markup** defines the purpose and content structure of a document, that is, it defines what each element in a text is, rather than how it is displayed.

HTML is broadly a descriptive markup. It is used to create hypertext documents, which are used by Web browsers, and they define everything in a document other than content (Esselink 2000; Raggett 2005; W3C 2003).

Two of the most important things related to HTML is the hypertext and tags (Schiller 2006). Schiller (2006) defines hypertext as the most important feature of a webpage, which allows a user a high level of flexibility on a website. It is made up of items of information which are what is known as hyperlinked to each other electronically; the information items can be on the same webpage, on different webpages, or even on different websites. Schiller (2006) also says that hypertext includes hypermedia, which describes the linking of different types of media to each other – she uses the example of linking the medium of text to the medium of video or audio.
Because of hyperlinks and hypermedia, Sandrini (2005) points out that a hypertext's main feature is that the text is not put in a sequential order, and that instead short sections of text are offered to the reader, who can then decide in what order to read them.

HTML tags provide instructions to webpage documents relating to font size, colour, and layout, among other things, and therefore give the page structure (Schiller 2006). HTML tags have their own structure: it is necessary to correctly use them so that the page is displayed correctly. There are many types of tag, such as those used to define the body of the document, the title, and the separation of paragraphs. Below is an example, as adapted from W3Schools (2014b).

```
<!DOCTYPE html>
<html>
<body>
  <h1>Heading</h1>
  <p>Paragraph one</p>
  <img src="uniqelogo.jpg" alt="uniqelogo" width="104" height="142"/>
  <a href="http://www.unige.ch">Link to UNIGE</a>
</body>
</html>
```

**Figure 10:** Example HTML tags and their display on a web browser

This example uses the tags `<html>` (called the root element), `<body>`, `<h1>`, `<p>`, `<img>`, and `<a>`. As can be seen here, most of the tags have an opening and closing tag, that is, the opening tag `<html>` has the closing tag `</html>`, `<body>` has `</body>`, `<h1>` has `</h1>`, `<p>` has `</p>`, and `<a>` has `</a>`. The closing tag tells the document that the particular element defined by the opening tag has finished. These elements will still work without the closing tag, as they can be considered optional, although W3Schools (2014a) advises against this.

There are some elements, such as for images, which do not need a closing tag, and are called standalone tags. In this case, the image element defines the source for the image (`<img src=" "`), the alt text (`alt=" "`), that is, the alternate text that will appear if the image cannot be displayed, and the image's height and width in pixels (`width=" " height=" "`).

The tags `<h1>` and `</h1>` denote that the text in between the two tags should be labelled as "the most important" (Raggett 2005) heading, as headings in HTML are
labelled from 1 to 6, with 1 being the most important (and with the biggest text) and 6 being the least. `<p>` and `</p>` denote a block of text as a paragraph, and without these tags the html document will not recognise line breaks as the beginning of a new paragraph. Finally, the tags `<a>` and `</a>` denote a hyperlink, with the `href=" "` attribute being the link’s destination. The text after the href attribute is the text which will be made into a hyperlink.

Going into more detail on HTML is not under the scope of this thesis. However, every localiser must have a basic knowledge of HTML, as there are many parts which should not be translated. In the example above, only the text “Heading,” “Paragraph one,” the alt text “unigelogo,” and “Link to UNIGE” should be translated. If any other parts where translated, deleted, or changed, the structure of the webpage may be damaged.

2.3 Localisation Methods

This section will start by describing the main types of website, and outlining methods for their localisation. It will then go on to talk about difficulties that localisers face, with particular reference to the skills expected of them from the website localisation industry.

2.3.1 Types of Website and their Localisation

There are three main types of webpage: static, semi-static, and dynamic, and the localisation method changes depending on what type of page is to be localised. A static website is a group of files, the documents of which are written in HTML, and which are organised into folders and linked to each other (Rodríguez Vázquez 2012). As the name suggests, its architecture is very simple, which can be seen in Figure 11. A web browser sends a request to a web server, which looks for the file in the relevant folder. Once retrieved, the information is sent to the web server, which is then sent to the web browser, which displays the relevant information.
Static webpages use relative and absolute paths within the HTML to make sure that the correct information is being displayed. A relative path is a path relative to the file system of a certain operating system, whereas the more common absolute path points to the same location on one file system regardless of its location (Morado Vázquez 2014).

A semi-static webpage is very similar to a static webpage, with the difference being that they have interactive and dynamic content (Rodríguez Vázquez 2012) which are written by HTML complements such as JavaScript, ShockWave, and Flash (Morado Vázquez 2014). It is important to remember, however, that this new dynamism comes from within the HTML content (Rodríguez Vázquez 2012), unlike a dynamic webpage.

The same localisation method is used for static and semi-static webpages. Rodríguez Vázquez (2012) suggests the following method:

1. Download the source code.
2. Replicate the source structure.
3. Link parallel pages by language (the original language page must have an extra link inserted in order to link the pages together at the end).
4. Localise the HTML files and any other files which include text, using CAT tools.
5. Localise graphic files such as images (in most cases this involves modification of the source files).

Dynamic webpages, on the other hand, are much more complex, as they have the ability to constantly change. Instead of containing a fixed set of data, dynamic webpages are generated in real-time with the help of a database. This means that they are easier to update, as in order to change content on a static webpage it is necessary to open the page, edit it, and publish it as a new version, but all that a webmaster needs to do to
update a dynamic webpage is to change the relevant database record without touching the webpage itself (Torres del Rey and Rodríguez V. de Aldana 2014). This method makes a dynamic webpage’s architecture much different to that of a static webpage’s (see Figure 12).

Figure 12: Dynamic webpage architecture (Torres del Rey and Rodríguez V. de Aldana 2014)

Nowadays, Content Management Systems (CMS) are often used to create dynamic webpages. A CMS is a “software tool that allows you to create, edit, and publish content” with an aim to “provide an intuitive user interface for building and modifying webpage content” (TechTerms 2013). Figure 13 below shows how dynamic webpages work with a CMS.

Figure 13: Dynamic pages and CMS (Torres del Rey and Rodríguez V. de Aldana 2014)
The CMS shown in Figure 13 does everything that is needed to generate a dynamic webpage, as seen in Figure 12. As shown in steps 2 and 3 in Figure 13, the CMS database does not store full webpages, but instead stores segments of information which it then uses to create the requested webpage. With CMS, it is possible to maintain the website from the front end (the part of the webpage the user sees) if needed: these tools allow website administrators to log in to the webpage and make changes as they wish.

Torres del Rey and Rodríguez V. de Aldana (2014) suggest one of three methods to localise webpages using CMS:

**Method 1**

1. Download HTML files, as if they were part of a static website.
2. Use a CAT tool to translate the translatable content.
3. Localise any other files as necessary.
4. Place localised content into appropriate database tables and records.

**Method 2**

1. Translate and localise files directly on the CMS using the CMS editor.
2. Localised content will be automatically updated on the database.

**Method 3**

1. Export the HTML files into a CAT-friendly format (for example, XLIFF).
2. Translate content and import content to the CMS.
3. Localise any other files as necessary and import content to the CMS.

Each of these methods has its own advantages and disadvantages. It is up to the localiser and the webmaster in charge of the website to decide which method is best suited to the localisation task. However, it must be noted that the localisation methods presented here are generic: each website and CMS has its own structure and unique elements, and these methods must be adapted accordingly.
2.3.2 Difficulties for Localisers

The problem for localisers is that they are rarely involved in the overall webpage design process, especially in the case of outsourced localisers, who are only given absolutely necessary material. This does not stop clients demanding a variety of skills from localisers (Gouadec 2003), skills which include, but are not limited to:

- **Specialised translation:** the localiser must be able to translate information in any domain and be ready at any moment to do so, and must be able to have complete knowledge over all aspects of translation, such as documentation and terminography.

- **Revision (and, sometimes, rewriting):** websites are constantly changing, and the localised website must change with the source website, too. Localisers must always be able to rewrite advertising campaigns, user guides, and, if necessary, controlled languages.

- **Knowledge of ergonomics:** a localiser must be aware of how the localised version of the site will look to users, even if she will never see the finished product before publication.

- **Quality management:** a localiser must be able to define, maximise, and apply rigorous quality assurance procedures.

- **Technology knowledge:** a localiser must have up-to-date and in-depth knowledge on all technology that is being used in the localisation industry, and must be able to know how to use them when necessary. A localiser must be completely at ease in a technological environment.

2.4 Cultural Issues within Website Localisation

Cultural issues within localisation have been traditionally linked to issues such as colours, icons, pictures, and date and number formats. Nowadays, while these issues must be respected and adapted to a locale, there are four main cultural elements that must be respected within cultural adaption (Jiménez-Crespo 2013):

1. Linguistic-technical aspects, such as intertextuality, register or macrotext.

2. Visual-iconic aspects.

3. Technical aspects.
4 Cognitive aspects, such as navigation, metaphors, mental models or interaction.

It has been pointed out by de Bortoli and Maroto (2001) that for each individual case it is worth considering whether there is a need to provide country specific content. It has, however, been pointed out that the aim of website cultural adaptation is, rather than making a user believe that a website is a local production, to make a user think that the company in question is conscious and respectful of the receiving culture (Jiménez-Crespo 2013). It is important that multinational companies wishing to expand their business consider these aspects, at least as part of their greater marketing strategy.

Unfortunately, cultural adaptation goes against the cost-efficiency favoured by companies. There are therefore five website localisation categories as specified by Nitish Singh and Arun Pereira (Jiménez-Crespo 2013), from which a company could choose to satisfy their needs:

1 **Standardised websites**: where a multinational company offers a site in one language for all.

2 **Semi-localised websites**: specific local content is a contact page in the target language with information about local branches and contacts.

3 **Localised websites**: where most content is localised, but the original functionalities and back end (where the website information is stored) are not.

4 **Extensively localised websites**: there is global localisation, and all content, as well as site structure and functionalities, are adapted to the target locale.

5 **Culturally adapted websites**: at the highest level of localisation, there is complete immersion in the target locale, and the website is adapted to main cultural aspects such as perception, symbolism, and behaviour.

These categories and approaches to culturally adapting websites must not overshadow the fact that the “linguistic quality of a translation is the most specialised and integral part of any localisation effort” (de Bortoli and Maroto 2001), and that localisers and translators must not be out of touch with the language, and particularly the culture, that they are localising and translating into.

### 2.5 Conclusion

This section has aimed to give a general introduction to web localisation. It has started by giving a general introduction to why it is necessary to localise websites, before
going on to describe websites and HTML. More detail has been given on the main types of website that exist, along with ideas and methods on how to localise them and the difficulties that localisers face. Finally, there has been a short discussion on cultural issues to be faced when localising a website. The next section will now discuss an area that is becoming more and more important within website localisation: machine translation.
3 Machine Translation

This chapter will be describing the concept of machine translation, as this is how the localisation tools evaluated in this thesis will translate the website material used. It will start with a general introduction, before moving on to explain the main differences between machine translation systems. Section 3.3 will be dedicated to discussing different methods of evaluating machine translations.

3.1 Introduction

Machine translation is the name used for the method of automating all or part of the translation of one human language to another (Arnold et al. 1994). Despite common misconceptions of the science of machine translation, there is no one method used to machine translate a document - there is a scale based on human involvement and mechanisation (Hutchins and Somers 1992). This scale starts at one end with what could be be called ‘traditional’ human translation, and works its way up to machine-aided human translation (MAHT), and human-aided machine translation (HAMT) - both of which come under the heading of computer-assisted translation (CAT).

There is one final stage to this scale, which is called fully automatic high quality machine translation (FAHQMT). FAHQMT may have been on the machine translation scale since the 1950s, but it is an unattainable dream for most people, and translations are rarely fully automated except in some specific domains, such as in the case of METEO, used by the Canadian Meteorological Center (Hutchins and Somers 1992). Normally, machine translation systems need one or all of the following human intervention stages in order to achieve a translation of acceptable quality, and thus putting it at the HAMT point on the scale mentioned above: pre-editing, interaction during the translation process, and post-editing.

Pre-editing normally includes checking source texts to be translated for problems that the translation system may come across, in order to create good input for the translation system. These checks can range from checking on basic writing and grammatical rules, to adding symbols that certain systems will recognise as denoting certain grammatical categories (such as with the SUSY system, where N stands for noun, and V stands for verb), to using controlled languages to edit out ambiguities (Hutchins and Somers 1992).

The second human intervention stage is interaction during the translation process, where the aim is to resolve any ambiguities, particularly lexical ones, found by the machine
within the text during translation (Hutchins and Somers 1992; Arnold et al. 1994). The machine will stop during the translation process to ask the translator involved for help with resolving these ambiguities.

Finally, post-editing is where output from a machine translation system is corrected manually to an agreed standard (Hutchins and Somers 1992). These standards are normally set to the requirements of the client, with particular reference to the aim of the translation and the time available to post-edit, but a lot of the post-editing is dictated by the quality of the machine translation output. This output is usually quite easily edited, as post-editors become familiar with a machine translation system’s error patterns (Arnold et al. 1994), thus making machine translation and post-editing less time-consuming than human translation.

3.2 Machine Translation System Types

Machine translation systems can be split into two main types: the first being linguistic, the second statistic. Linguistic machine translation tools have changed since their creation, from the older, direct systems that did not necessarily separate rules for analysis and transfer, to the indirect systems that have clearly separated components which work with different linguistic levels (Hutchins and Somers 1992). The difference between them and statistic systems is that linguistic systems will aim to understand the source text with the help of grammatical and lexical rules, amongst others, whereas a statistical system does not try to understand a text but instead bases its translation decisions on a corpus of previous translations, where the translation chosen has the highest probability of being correct.

This chapter separates the different types of system into three sections: section 3.2.1 will describe direct linguistic systems, section 3.2.2 will describe the two types of indirect linguistic systems, and section 3.2.3 will describe statistical systems.

3.2.1 Direct Linguistic Systems

Linguistic systems can be separated into direct systems and indirect systems, with indirect systems being further split into transfer and interlingual systems. They will be described in the order given, that is, from the least complex system to the most.

Both localisation tools used in this study (Systran and Reverso) are direct linguistic systems, although some of Systran’s translation stages “resemble a transfer-based MT [machine translation] system because translation also involves the use of rules to direct
syntax generation” (Clough and Sanderson 2003, p. 3), which a direct system would not necessarily do.

The direct system is very simple; it has no grammatical knowledge. It is called a direct system, as there are no intermediate steps within the translation process: the source language text is processed, which then leads directly to the target language output. It is important to point out that due to its simplicity, a direct system is intended for the translation of only one pair of languages, and in only one direction (Arnold et al. 1994) - which is interesting to note considering that both Reverso and Systran boast many languages in both directions as part of their services.

A direct system has two main characteristics. The first is, as mentioned above, its minimal comprehension of the source text. The majority of the source text analysis carried out is at a lexical level in order to resolve any ambiguities before translation. Source words are then replaced with their target language equivalents as set out in a bilingual dictionary, then rearranged in order to conform to target language rules.

Figure 14: Direct MT system (Hutchins and Somers 1992, p. 72)

Figure 14 above shows the process used by direct systems. It can be clearly seen that this is a word-for-word translation process, with some arrangement according to the target language as necessary, for example, in the case of rearranging the word order of nouns and adjectives. The analysis briefly outlined above tends to be morphological and grammatical, in that word endings would be identified and inflected word forms would be changed into their original, infinitive forms in order to be input into a bilingual dictionary - there is no complete analysis of structure or semantic relationships. It is simply a surface analysis to gather essential information (Hutchins and Somers 1992).

The word-for-word translation given by this type of system means that heavy human intervention is needed in the pre and post-editing stages, but because of the minimal analysis done by the system, it is almost impossible that a system would ask for human intervention during the translation process to help with problems, as it is not designed to ‘think’ in a way that other systems are.

However, direct linguistic systems do have their advantages. The main advantage is that they are robust, in that a direct linguistic system will always give a translation, and not give many ‘error’ warnings like other linguistic systems may do (Arnold et al. 1994). This is a particularly important advantage for general-purpose machine translation.
3.2.2 Indirect Linguistic Systems

Indirect linguistic systems differ from direct linguistic systems in that they go further to solve syntactic, structural, and lexical ambiguity, by doing a complete syntactic analysis of a sentence using its own grammar, or representation (Bouillon 2013b). These systems do not link words and their equivalents in other languages together as in direct systems, but instead link a representation of meanings and concepts together. Transfer and interlingual systems are both types of indirect linguistic systems, with a transfer linguistic system having one representation for each language, and an interlingual linguistic system having a representation which is separate from the languages used by the system.

As mentioned above, transfer systems have independent and separate grammars for each language it uses. These grammars are specialised to the language it represents, to the extent that they may be developed by different people with different language specialities (Arnold et al. 1994). As shown in Figure 15, a transfer system is made up of three main steps: analysis, transfer, and synthesis, also called generation.

![Transfer System Diagram](Arnold et al. 1994, p. 72)

Figure 15: Transfer system (Arnold et al. 1994, p. 72)

In the first step, analysis, the 'interface structure' is essentially the product of a syntactic analysis using the grammar for the source language. The transfer step finds equivalences between the source and target language using transfer rules. There are three main rules used by the transfer system: lexical (which translate content words), structural (which translate sentence structure), and semi-lexical (which translate while
taking context and structure into account) (Bouillon 2013b). The final step, synthesis, changes the target language representation created in the transfer step to a true target language translation, by using the target language grammar.

A disadvantage to a transfer system is that for every new language a company decides to add to its repertoire, new grammar modules must be added for that language, as well as new transfer modules for each language pair to be translated. This means that more monolingual and bilingual specialists may need to be recruited to create the new transfer models, costing a company time and money. Another disadvantage is the ambiguities that can arise during the translation process: analysis can lead to ambiguities (where, for example, a system assumes a word is a noun when in fact it is a verb), which can be difficult to fix.

With the interlingual system, there are only two steps: the source text is analysed using a neutral grammar, which then generates, or synthesises, the target text from an interlingua, which is entirely language independent. This process is shown in Figure 16.

![Figure 16: Interlingual system (Arnold et al. 1994, p. 83)](image)

The analysis stage is the traditional syntactic analysis that takes place in machine translation systems. However, unlike with other systems where a transfer is made, each word is changed into an abstract concept, and syntactical relationships are changed into semantic roles. The theory behind abstracting the source text in this way is that the interlingua is a much more universal grammar and language, which links many concepts together for many languages.
This leads to the second step of the interlingual system: the synthesis, or generation, of the target text. The system does not look back at the syntax of the source text in order to generate the target text; instead it uses the interlingua's abstract representation to produce the target text. It must be pointed out, however, that the very nature of the interlingua means that the end product of the machine translation is in fact a paraphrase of the source text, as its goal is to attribute a meaning rather than syntactic information to words (Bouillon 2013c).

It is easily understood that this system is quite complex, as there are overlapping and fragmented concepts according to each culture and language, all of which should be expressed within one interlingua. This can lead to much more time being used to create an interlingua which might not even satisfy all concepts in all languages (Arnold et al. 1994). However, it should be noted that, when comparing interlingual to transfer systems, interlingual systems are more attractive for multilingual systems, as it involves only adding two new modules (the analysis and generation grammars) to the system, as the interlingua is independent.

The three types of linguistic system described above each have their advantages and disadvantages. Direct systems offer much lower quality translations than indirect systems, as it is much more word-to-word based. Indirect systems have been taught to 'think' to a certain extent, with the use of grammars and representations put into the systems themselves, and are therefore less dependent on the source language to produce a translation. However, the deeper analysis carried out by indirect systems means that they are less robust than direct systems.

3.2.3 Statistical Systems

Statistical systems are far from being the new machine translation system many think them to be: it was one of the first techniques used for machine translation in the 1950s, but was abandoned in the 1960s in favour of linguistic systems (Bouillon 2013a). The first general machine translation in recent years is Google Translate, which has helped statistical systems become more popular and more accessible. Statistical systems aim to translate a source sentence by searching a corpus for the target sentence which has the highest probability of being the translation of a source sentence. This is based on probabilities taken from the corpus, which are based on a language model and a translation model (Bouillon 2013a).

A language model determines the likelihood of a sequence of words being used by a native speaker of a given language, in order to create a target sentence that is fluent (Koehn 2010). This is done by using n-gram language modelling and the Markov chain, and using these n-grams to search for the number of times strings of words appear in
a monolingual target-language corpus. This determines the fluency of a target-language sentence: the more times a certain string of words appear together, the more likely it is to be a fluent target-language sentence.

A translation model differs from a language model in that it measures how faithful a translation is, instead of how fluent a translation is. It does this by determining the likelihood that a word in a target language is the translation of the word given in the source language (Bouillon 2013a), by searching through an aligned bilingual corpus. To further determine the likelihood of a word being translated in a certain way, a translation model also uses the criteria of fertility and distortion, with fertility being how many words could be used as a translation, and distortion being the probability of the position of a target word within a sentence.

The main disadvantage of statistical machine translation systems is that there is no analysis involved: where a linguistic system has been programmed to analyse certain grammars and words in order to try to find the correct equivalent, statistical machine translation systems have not. Instead, as already mentioned, they base their translations on probabilities, which are more likely to go wrong, particularly if the corpus the probabilities are based on is limited.

On the other hand, statistical machine translation systems require much less human involvement and effort than linguistic machine translation systems, as all that is needed is corpus maintenance, even if another language pair were to be added. It could also be argued that, if the corpus has been well chosen, the translations provided by a statistical system are more true to life, as the corpus comes from genuine human productions.

3.3 Machine Translation Evaluation

There are many types of evaluation designed for the performance of a machine translation system, including prototype evaluations (testing during the development and design of the system), development evaluations (testing linguistic and operational capabilities before release), operational evaluations (evaluation by potential buyers), translator evaluations (testing by the users of the system), and recipient evaluations (normally comparative evaluations carried out by the recipient of the translation) (Hutchins and Somers 1992).

As this study includes comparing the quality of two machine translated versions of the same text, this section will be describing the quality assessment of a machine translation system’s raw output - that is, its translation of a text before pre or post-editing, which comes under the heading of recipient evaluation.
There are two ways to evaluate the output from a machine translation system: manual evaluation (section 3.3.1) and automatic evaluation (section 3.3.2). This section will be describing both, as this thesis will be using both methods and then attempting to correlate and compare the results.

### 3.3.1 Manual Evaluation

There are three types of manual evaluation: intuitive judgement, comparative evaluation, and number of errors found (Bouillon 2014b; Koehn 2010).

With intuitive judgement, the most important decisions to make are how to choose the evaluators, and how to choose the criteria against which the translation will be judged. Depending on the type of source text and the translation given, many types of evaluator can be chosen to carry out an intuitive judgement evaluation. It is normally deemed best to ask people who understand both languages involved to evaluate the target text, but people who only understand the target language could evaluate a translation, if given a reference translation to work from (Koehn 2010).

As for the criteria, it is popular opinion that there is no set way to judge machine translation system output but, despite discussions in the area, there does seem to be some agreement that the best criteria to measure machine translations against can be taken from the categories of accuracy (also called adequacy), and fluency (also called intelligibility) (Arnold et al. 1994; Hartley and Popescu-Belis 2004; Hutchins and Somers 1992; Koehn 2010). Accuracy is taken to mean that the output “conveys the same meaning as the input sentence,” and checking for accuracy ensures that no part of the message is “lost, added, or distorted,” (Koehn 2010, p. 218). Fluency is whether the output is good, clear, and fluent in the target language. Either one, or both, of these criteria can be used; it depends upon the needs of and the reasons behind the evaluation.

Koehn (2010) advises that these criteria are put on a scale of 1 to 5, with 5 being the best (for example, ‘flawless English’), and 1 being the worst. This fits in with survey guidelines, which suggests using an odd scale with either 5 or 7 options to choose from (Fink 1995).

Instead of rating a translation on a scale and set criteria, the second manual evaluation method is to use a comparative evaluation, where evaluators are asked which translation is better (when based on certain criteria) out of a choice of two translations. Koehn (2010) claims that this may be a better evaluation method than scoring two translations individually, as comparative evaluation questions are generally answered more consistently then separate evaluation questions.

The third and final type of manual evaluation is the weighting of the number of errors found in a translation, where the more serious errors, such as corrections of groups
of words, have a higher weighting than minor errors, such as choice of synonyms. The advantage to this method is that less time and resources are needed to find suitable evaluators, as it can be done by one person.

An example of this third type of manual evaluation is the SAE J2450 Quality Assurance Model, which was created for the automobile industry but has been generalised to cover all types of translation. In this case, there are seven error categories, split into minor and serious errors, with different weighting scores (SAE 2001):

1. **Wrong term**: a target language term that is in violation of a client glossary, or denotes a concept that is clearly different to the source language (minor error weighting is 2, serious is 5).

2. **Syntactic error**: a target language counterpart is assigned the wrong part of speech, or the text has poor structure (minor error weighting is 2, serious is 4).

3. **Omission**: a block of source text has no corresponding target text (minor error weighting is 2, serious is 4).

4. **Word structure or agreement error**: a target word is in the incorrect tense, case, gender, or any other inflection (morphological error), or there is an agreement error between target words (minor error weighting is 2, serious is 4).

5. **Misspelling**: a target word violates spelling as set out in a client glossary, or goes against generic spelling conventions (minor error weighting is 1, serious is 3).

6. **Punctuation error**: a target text contains an error according to the punctuation rules of the target language (minor error weighting is 1, serious is 2).

7. **Miscellaneous error**: any other linguistic error which does not clearly belong to any other category (minor error weighting is 1, serious is 3).

The document’s overall score is calculated by adding all of the weighted scores together and dividing the total by the number of words in the source text. The lower the score, the better the quality of the translation.

### 3.3.2 Automatic Evaluation

BLEU is the most popular method for the automatic evaluation of machine translation output. There are other methods apart from BLEU, such as word error rate (WER), where the main requirement is that words must match in an order given by a reference translation. However, BLEU was the first automatic machine translation evaluator that included a brevity penalty, thus making it superior to other automatic evaluation methods (Koehn 2010).
The main idea behind the BLEU system is that “the closer a machine translation is to a professional human translation, the better it is” (Papineni et al. 2001, p. 311); that is, it judges the accuracy of a machine translation. It works by automatically comparing translations to one or more reference translations (normally done by a human) using the n-grams that the machine translation and the reference translation have in common (Bouillon 2014b). As mentioned above, the BLEU method also includes a brevity penalty, so that in order to receive a good BLEU score, a translation must “match the reference translations in length, in word choice, and in word order” (Papineni et al. 2001, p. 315) when compared to the reference translation. The highest BLEU score a translation can receive is one, with the lowest possible score being zero.

3.4 Conclusion

In this chapter, the concept of machine translation has been described in detail. A general overview of the machine translation process has been given, along with an explanation of the different types of machine translation systems – linguistic and statistic – and their subcategories. A final section described the evaluation of machine translation output, along with different methods and how they work. This is the final theoretical chapter of this thesis - it will now move on to the tools that will be evaluated.
4 Localisation Tools Used

This chapter will introduce the two tools to be evaluated, which will be used to localise the test website. Both of the tools are online platforms dedicated to website localisation; the first is called Reverso Localize, and the second is SYSTRANLinks. This chapter will be presenting the versions of these two platforms as they were at the time that the evaluation was carried out, that is, June 2014.

4.1 Reverso Localize

Reverso Localize is an online platform developed in 2012 by Reverso-Softissimo within the EU-funded Flavins programme. It has been constantly updated since its first release, particularly to improve the website and the translation engine for users.

Figure 17: Reverso Localize homepage

Figure 17 shows the site’s homepage, with the four main uses of the website, along with an instructional video. From this page, it is possible to create a free account, or connect with a Facebook account, and from there more information can be read about how the tool works. In April 2013, the interface was available in five languages (Peron 2013), but now it is only available in three: French, English, and Spanish. I believe this to be a reflection of the type of client that uses the platform: out of the 1851 websites localised at the time of the evaluation, 98% of them had been translated from either French, English, or Spanish.

As I have mentioned, Reverso Localize has four main functions: the translation of documents, websites, applications, and data. It is suggested that the function for
translating websites is only to be used for ‘classic’, or static, websites, and the other functions are essentially a support for the basic website translation tool, which allow translations of linked documents, web and mobile applications, and data.

Users wishing to translate a website for free receive credit for 8000 words, which is approximately 20 pages of text, and are limited to two languages and a website size of 3500kb, which Reverso Localize states is enough to translate the main pages of a website. There are 13 languages available as source and target language: English, French, Spanish, German, Italian, Romanian, Polish, Swedish, Portuguese, Chinese, Arabic, Russian, and Hebrew. There are also two different translation domains to choose from (General and Travel) and two ways to translate the website, called File translation and Mirror site. The mirror site is a copy of the source website which is created, and then prepared for localisation and localised by Reverso Localize. In this way Reverso Localize guarantees that no changes are made to the source website. The mirror site method is what this thesis will be focusing on; Figure 18 below shows its features in more detail.

As seen in Figure 18, the process of localising a website is made much easier for non-experts by simply entering the URL of the website to be localised. Reverso Localize will extract the website’s translatable content automatically, and will also detect any spelling and grammar errors in the original website. All the user has to do is select the source and target languages, as well as the whether they want it to be translated for a general domain or a travel domain. The user is then informed by a pop up window that the translation is being carried out in the background, as shown in Figure 19, and that they will be notified by email and on the platform’s dashboard when it is finished.
Figure 19: Reverso Localize ‘job created’ alert

Once finished, the translation can be revised by the user, a reviser that has been invited by the user, or a professional reviser. The revision, or post-edition, is done from within the mirror site, using a pop up revision box when text is clicked on, as shown in Figure 20. Text can be changed and formatted in this environment, and any modifications made can be applied as the reviser wishes. It is also possible for a reviser to add a comment on the translation for the creator of the project to read. This environment is clearly a What You See Is What You Get (WYSIWYG) environment, with no extra knowledge needed than what the user can see and understand on the screen.

Figure 20: Reverso Localize post-editing environment

It is also interesting to note that the machine translation of the website can be made into a personal user translation memory for future use in other website localisation projects.

After the translation has been corrected, the user can publish and index the translated website, on the condition that the user states that they have permission to
do so from the owner of the source website. Instructions are given on how to link the source and target websites together by inputting HTML code given by Reverso Localize into the source code of both the source and target websites, as seen in Figures 21 and 22, and there are more simple and detailed instructions specifically for WordPress or Overblog sites.

![Figure 21: HTML link for source language website](image)

![Figure 22: HTML link for target language website](image)

It has been noted that although there is an option to contact Reverso Localize, there is no feature that allows a user to talk to an employee or get any outside, real-time help during the localisation process.

4.2 SYSTRANLinks

SYSTRANLinks was released as an online tool in 2012 by Systran in order to make “website translation faster, easier and more cost-effective than current solutions” (SYSTRAN 2012) in a user-friendly way for as many clients as possible: small to large businesses, eCommerce sites, language service providers, web agencies, and finally, individuals. It is with this last user in mind that the functions of SYSTRANLinks will be described. The platform has had several updates since its release, some within formal re-releases, and others added as fixes to improve the user experience.

SYSTRANLinks’ interface is available in 11 languages: English, German, Spanish, French, Italian, Dutch, Portuguese, Russian, Japanese, Chinese, and Korean, which I believe to be a reflection of the number of clients Systran expects to want website localisation services. Using Systran machine translation, it is possible to translate to and from 15 languages: Arabic, Chinese, Dutch, English, French, German, Greek, Italian, Japanese, Korean, Polish, Portuguese, Russian, Spanish, and Swedish.
As with Reverso Localize, SYSTRANLinks has a free account for clients to use, but also has three levels of paid service. Figure 23 shows the difference between the editions. This section will be describing the free version of SYSTRANLinks, which is described as being tailored for small companies with low-traffic websites or blogs, or anyone who wants to test the power of what the platform can do.

As Figure 23 shows, users wishing to translate websites for free can do so an unlimited number of times, as long as the localised website is hosted by Systran under a SYSTRANLinks.net domain, and the total page views per year only go up to 6,000. However, SYSTRANLinks will not replace images, files, and external links: this must be done manually. Users with a free subscription have their websites translated by the Systran machine translation system, and they may create a user dictionary with up to 500 entries, and use a translation memory created by SYSTRANLinks of up to 750 segments, but it must stay within the SYSTRANLinks system. Users can also invite people to help manage the project, as one of SYSTRANLinks’ philosophies is collaboration.

After a user has created their account, they are directed to the SYSTRANLinks user area, as shown in Figure 24, which shows their subscription usage. The interface is designed to be as user-friendly as possible, and if more than one localised website is being managed by the platform, they are shown in individual squares.
Users wishing to localise a website do so by clicking on ‘add a website,’ and entering the source website URL and the source language of the website. Before being able to continue, SYSTRANLinks will check to make sure that the website exists. The target language for the website to be localised is then selected, and the final step is choosing the ‘language selector’ type, which is a banner that SYSTRANLinks inserts so that people can switch between the localised website’s languages. It is possible to change this feature after the website has been translated.

The automatic extraction and translation of translatable content then takes place, during which time the user must wait. A notification then appears on the screen telling the user that the website has been localised, and gives the URL of the new website, as seen in Figure 25.

The user can then choose to look at the website as it will be published and edit it as they browse in a limited manner (that is, just change the wording of the translation) as shown in Figure 26, or can review, manage, and verify the translations of each sentence from a dashboard, which looks more like a CAT tool environment. From here changes can also be made to images and other localisable content.
The possible advantage to the localised site post-editing environment is that, as with Reverso Localize, it has a WYSISYG interface, in which case no knowledge of post-editing or any technical domain is needed. The CAT tool-like environment for post-editing the localised website shows each translated sentence next to its source sentence. The translated segment can then be compared to the source and changed if necessary, as well as marked as an acceptable translation.

In order to publish the localised website, the user must be verified as the owner of the website, as shown in Figures 27 and 28. SYSTRANLinks does this in one of two ways: by uploading HTML verification file to the source website (‘Default method’), or by inserting a meta tag into the source code of the website’s homepage (‘Alternative method’). In order to do either of these and therefore have the source website verified as belonging to the user, SYSTRANLinks is assuming that the user has administrative rights to the website and can therefore change the source code as they wish.

Figure 27: Default method for website verification
4.3 Conclusion

This chapter has presented the two localisation platforms that will be evaluated in this thesis, Reverso Localize and SYSTRANLinks. It can be seen that in some ways the tools are quite similar; for example, they both make use of a localised site as a post-editing environment. In other ways, however, they are quite different: Reverso Localize gives more post-editing options in its localised site, whereas SYSTRANLinks includes a dashboard for further post-editing options. There is also a difference in publishing the localised and source websites together, with SYSTRANLinks offering a slightly more secure method.

The chapter has briefly described why and in what context the platforms were developed, in order to further describe the features that free subscriptions provide. These features will be discussed in more detail in later chapters, with some being assessed.
5 Methodology

This chapter will be presenting the methodology used to evaluate Reverso Localize and SYSTRANLinks. Developing a methodology for evaluating translation tools can be difficult, as there is no standard evaluation method which is reliable and acceptable, despite much research being done in the field (Quah 2006). Quah (2006) points out that there have been, however, several notable attempts, such as the 1966 ALPAC report, and Georges Van Slype’s modern methodology for the European Commission, whose aims were to present an outline of evaluation methods, provide a critical appraisal of the evaluation methods, and to make recommendations as to which evaluation method to adopt.

Despite a lack of uniform evaluation methodologies for evaluating translation tools, this study will be using the EAGLES (Expert Advisory Group on Language Engineering Standards) evaluation method, as created in 1996 and documented in their Evaluation of Natural Language Processing Systems Final Report. The EAGLES method originally took ISO (International Organisation for Standardisation) standard 9126, which is concerned with defining quality characteristics used in software evaluation (EAGLES 1996a), and adapted it to a translation environment in order to produce an evaluation in which different features and attributes could be combined to reflect the needs of an end user (Quah 2006). The EAGLES framework is a general one, but has been purposefully made in that way in order to guide individual evaluations and make it easier to understand and compare results (Hovy et al. 2002).

The EAGLES methodology is split into what is called the EAGLES 7-step recipe (EAGLES 1999), which describes the major stages necessary to carry out a successful evaluation of language technology systems or components. The evaluation will be using these seven steps in order to create a suitable methodology. As well as leaning heavily on EAGLES, the methodology used in this work will also be following the methodology structure used in Peron (2013), which has been greatly useful as her work also centred around evaluating localisation tools.

This chapter will be structured in a similar way, in order to demonstrate how the methodology for the evaluation has evolved. Therefore, this chapter will continue with the first step, which explains the purpose of the evaluation (section 5.1), and will then elaborate a task model (section 5.2) defining the system parameters and the corpus used, amongst other things. This chapter will then go on to set out the third and fourth steps in section 5.3, by defining the features of each localisation platform which will be evaluated. The fifth step will devise the methods that will be used to evaluate the tools (section 5.4), before moving onto step six and the next chapter, where the evaluation will be designed.
The seventh and final step, where the evaluation is carried out and results are analysed, will also be in the following chapters.

5.1 Purpose and Objectives

The aim of the evaluation is to compare and analyse the free localisation platforms Reverso Localize and SYSTRANLinks from the point of view of end-user satisfaction, in order to create recommendations as to which platform is the best to use, and how both platforms can improve. In this thesis, the end users are non-professionals with some localisation knowledge. The evaluation will be a black-box evaluation, focussing on the overall performance and process of both systems, as it is the best type of evaluation for end users (Quah 2006). The evaluations will take place in the context of translating and localising a hotel’s website from French into English, which will be further explained in sections 5.2.1 and 5.2.3.

5.2 Task Model and Parameters

Both Reverso Localize and the version of SYSTRANLinks evaluated in this thesis are online and free platforms dedicated to localising websites, with Reverso Localize also translating documents and mobile and web applications. As they are both free, several types of user are envisioned, from companies to private individuals. Both platforms claim to be easy-to-use, with SYSTRANLinks (2013) boasting “no need for HTML knowledge” and an “intuitive interface,” and Reverso Localize claiming that a user doesn’t “need any technical skills, Reverso Localize does everything” (Reverso-Softissimo 2013).

5.2.1 Language Parameters

The source language for the evaluation is French. There are several reasons for this choice, one of the main ones being that a Swiss (preferably Genevan) website was desirable, as this study is being written from within the University of Geneva. It was imperative that the French-language website was originally written in French, and not translated from another language into French, as a translation of a translation could have repercussions on the target language.

The target language is English. As a native English speaker, it would be possible for me to take a more active role in the evaluation, particularly in the analysis of the raw machine translation. The source and target languages also have similar roots, in that they are both Indo-European languages, which gives both localisation platforms an opportunity to perform well.
Another reason for the choice of source and target languages is the platforms to be evaluated. To date, Reverso Localize has localized 1851 websites, of which 1441 were translated from French into English (Reverso-Softissimo 2014), therefore meaning the scenario created for this thesis is quite close to the real-life situations for which other have used Reverso Localize. While data of this type is not available for SYSTRANLinks, it is relevant to point out that Systran developed, and is continuing to develop, a French-English machine translation system just after the English-French version in 1976 (Hutchins and Somers 1992), and it could be logical to say that its online localisation platform would use all of that amassed linguistic knowledge and developments.

5.2.2 Platform Parameters

The evaluation will be carried out on Reverso Localize and SYSTRANLinks, in the frame of localising a website. The free, English-language, online version of both platforms will be evaluated.

For the evaluation of Reverso Localize, the Mirror site method of localisation will be used, as it is specifically advertised as the solution for complete website localisation, with a What You See Is What You Get (WYSIWYG) interface for revising the translations, as described in section 4.1. The General domain has been chosen for the translation of the text, as the Travel domain is only compatible with translations from English into French and German.

When evaluating SYSTRANLinks, the only option given for localising the website is similar to that of Reverso Localize's mirror site. As for viewing the localised website, there are two ways to do so. The user can use either the WYSIWYG interface for viewing and editing the raw translation, or they may view the translations as if they were using a CAT (Computer Assisted Translation) tool, as described in section 4.2. Users in the evaluation will be invited to test both.

5.2.3 Corpus Used

The corpus to be used is the website for Lake Geneva Hotel (http://www.lakegenevahotel.ch), a hotel that opened in early 2014 and has been written in French, translated into German, and at the time of contacting the company, was due to be translated into English (and at the time of the evaluation, it had been partially translated into English). It was created using the open-source content management system TYPO3, in order to better manage its content. It has been chosen because:
• it is a Swiss website, written in French, and with a need for it to be translated in English,

• it has dynamic content (such as the type of website explained in section 2.3.1), for example, interactive links and forms, which make it suitable to be used in a localisation evaluation,

• it is written for the tourist domain, which means that the language used is basic enough for a translation machine to translate without encountering too many difficulties.

I contacted the owners of the website in February 2014 to ask for permission to use the website, which I received both verbally and by email and letter in April 2014. I was unable to get any more information due to confidentiality issues, although I have made the assumption that English speakers regularly visit the website, as the owner stated that there is a need for the website to be translated into English.

The entirety of the website will be used in the localisation process, in order to make the study as broad and fair as possible. Table 1 below shows the main content to be translated and localised:

<table>
<thead>
<tr>
<th>Text</th>
<th>Localisable Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>4072 words</td>
<td>Hyperlinks &amp; buttons</td>
</tr>
<tr>
<td>Bullet points &amp; menu headings</td>
<td>Images</td>
</tr>
<tr>
<td>Long sentences (more than 30 words)</td>
<td>Text formatting</td>
</tr>
<tr>
<td></td>
<td>Interactive calendar</td>
</tr>
<tr>
<td></td>
<td>Menus</td>
</tr>
</tbody>
</table>

Table 1: Corpus content

As can be seen, the website is made up of a variety of translatable and localisable content. There are 4072 words in the website, used in a variety of ways, such as bullet points, menus, and longer text. The localisable content consists of images, text formatting, and interactive content such as forms. The images do not need to be changed as they are generic images of a hotel; however, the interactive content, hyperlinks, text formatting, and menus do need to be localised correctly.

As for evaluating the quality of the translation, the translation of the Accueil page (see Figure 29) will be used, as it has plenty of text in sentences, as well as some in lists. The Accueil page is made up of 365 words, split into 61 segments.
5.2.4 Participants

The profile of the participants used in the study is quite fluid, as both Reverso Localize and SYSTRANLinks are aimed at a range of people. Although both platforms advertise the possibility of post-editing translations, this study is concerned primarily with the localisation process itself, so it must be reiterated that the participants in the study will not be testing this feature.

Six participants were chosen for the evaluation, all of whom being students in the Faculty of Translation and Interpreting at the University of Geneva. Students were chosen as they have not yet become acquainted with translation or localisation at a professional level, and can therefore evaluate the systems with no bias. However, as they are students in the Masters in Translation programme at the university, they have the basic knowledge needed to properly evaluate such systems for the processes used.

The six participants were split into two groups of three, with one group evaluating the localisation properties of the two localisation tools, and the other group evaluating the machine translation output of the two systems.

Information about the localisation participants was acquired through the use of a questionnaire, as seen in Annexe III: it was imperative that the participants evaluating the localisation tools had English as one of their translation languages, as the platforms were to be used in their English language version. It was also preferable for them to
be acquainted to a certain extent with the theories and practicalities of localisation, and the responses from the questionnaire were used to ascertain if this was the case. In this evaluation, all participants had a general overview of localisation concepts.

The participants evaluating the machine translation output had to be English native language speakers, with French in their language combination, in order to understand the source French text and compare it to the target English text.

5.3 Characteristics Evaluated

The third stage of the EAGLES methodology stipulates that top level quality characteristics must be defined in order to evaluate a system. There are six top level quality characteristics defined by EAGLES, as shown in the list below:

1 **Functionality**: a set of attributes that bear on the existence of a set of functions and their specified properties. The functions are those that satisfy specified stated or implied needs.

2 **Reliability**: a set of attributes that bear on the capability of software to maintain its level of performance under stated conditions for a stated period of time.

3 **Usability**: a set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users.

4 **Efficiency**: a set of attributes that bear on the relationship between the level of performance of the software under stated conditions for a stated period of time.

5 **Maintainability**: a set of attributes that bear on the effort needed to make specified modifications.

6 **Portability**: a set of attributes that bear on the ability of software to be transferred from one environment to another.

Based on what has been described about the two localisation platforms, it becomes quite a simple decision to evaluate the functionality, usability, and efficiency characteristics, as these are the most important characteristics from the point of view of the context of this study (the same conclusion was drawn in Peron (2013)). It is unnecessary to evaluate reliability, maintainability, and portability, for the following reasons. In a black box evaluation such as this one, evaluating the reliability of a platform can be integrated into the accuracy and suitability subcharacteristics of functionality (see Table 2 below). It is also unnecessary to evaluate maintainability, as the evaluation will be carried out at a certain time and for a certain task, where it is
unnecessary to make modifications to the software. The characteristic of portability is a particularly moot point for this evaluation, as both platforms are available online only, and therefore are more suited to be transferred from environment to another.

In order to properly test these characteristics, it is necessary to decide how each one will be tested. Each top level quality characteristic chosen also has its own subcharacteristics, which can help with the evaluation process. These are shown in the table below.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Subcharacteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functionality</td>
<td>Suitability</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
</tr>
<tr>
<td></td>
<td>Interoperability</td>
</tr>
<tr>
<td></td>
<td>Compliance</td>
</tr>
<tr>
<td></td>
<td>Security</td>
</tr>
<tr>
<td>Usability</td>
<td>Understandability</td>
</tr>
<tr>
<td></td>
<td>Learnability</td>
</tr>
<tr>
<td></td>
<td>Operability</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Time behaviour</td>
</tr>
<tr>
<td></td>
<td>Resource behaviour</td>
</tr>
</tbody>
</table>

Table 2: Subcharacteristics of chosen top level characteristics

Within the functionality characteristic, the two subcharacteristics suitability and accuracy will be evaluated, as these can be measured in a valid and reliable way, as demanded by EAGLES. The subcharacteristics to be tested within the usability category are understandability, learnability, and operability. Within the efficiency characteristic, the time behaviour subcharacteristic will be evaluated, as it is necessary to see if the two platforms are useful, and if they will save time, and therefore money, for users.

A tentative final evaluation table is proposed below.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Subcharacteristic</th>
<th>Evaluation question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functionality</td>
<td>Suitability</td>
<td>Are the platforms’ attributes suitable for the task to be carried out? Do they work correctly?</td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td>Do the platforms’ attributes produce accurate localisation and translation results, as expected?</td>
</tr>
<tr>
<td>Usability</td>
<td>Understandability</td>
<td>Do users find it easy or difficult to understand the platforms’ concepts and how they work?</td>
</tr>
<tr>
<td></td>
<td>Learnability</td>
<td>Can users easily learn how to use the practicalities of the platforms?</td>
</tr>
<tr>
<td></td>
<td>Operability</td>
<td>Are the platforms’ functions easy or difficult to operate?</td>
</tr>
<tr>
<td>Efficency</td>
<td>Time behaviour</td>
<td>Do the platforms take a relatively long or short time to carry out tasks?</td>
</tr>
</tbody>
</table>

Table 3: Evaluation questions for EAGLES characteristics chosen

Now that the characteristics and sub-characteristics for the evaluation have been decided upon, I will now go on to talk about how the evaluation will be carried out.

5.4 Evaluation Method

For user-oriented software tests such as the one carried out here, EAGLES (1996b) suggests using one of three tests. These are:

1. **Scenario test**: a test case which aims at a realistic user background for the evaluation of software. It involves putting the system into its intended use by its envisaged user performing a standardised task. The two types of scenario test are field tests and laboratory tests.
2 **Systematic test**: a test which examines the behaviour of software under specific conditions with particular results expected. It is possible to carry out this type of test with or without user participation. User-oriented systematic tests can be split into task-oriented testing, menu-oriented testing, and benchmark testing.

3 **Feature inspection**: a test whose aim is to describe the technical features of a piece of software in as detailed a way as possible, so that it allows the comparison between systems of the same type.

The evaluation in this case will be carried out using both scenario and systematic testing. Scenario testing has been chosen as a way of evaluating the system in a subjective way from the users’ point of view, and will be video recorded on-screen, whereas systematic testing will be used to objectively evaluate some functions of the platforms which do not necessarily need a user opinion. Table 4 sets out the evaluation methods used with respect to each characteristic.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>To be evaluated</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functionality</strong></td>
<td>Interaction</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Localisation system</td>
<td>Number of localisation errors (SA)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of omissions and silences (SA)</td>
</tr>
<tr>
<td></td>
<td>Machine translation</td>
<td>BLEU, human evaluation &amp; quality assurance</td>
</tr>
<tr>
<td></td>
<td>Localised site</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Publishing and linking</td>
<td>Questionnaire</td>
</tr>
<tr>
<td><strong>Usability</strong></td>
<td>Understandability</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Learnability</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Operability</td>
<td>Number of clicks per task (VA)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number and which type of keys pressed (VA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of tabs open (VA)</td>
</tr>
<tr>
<td><strong>Efficency</strong></td>
<td>Time behaviour</td>
<td>Time taken to extract data (VA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time taken to achieve tasks (VA)</td>
</tr>
</tbody>
</table>

*(SA)=localised site analysis
**(VA)=video analysis

Table 4: Evaluation methods for EAGLES characteristics

5.5 Conclusion

This chapter has presented the first five steps of the EAGLES 7-step evaluation method. The purpose and objectives of the evaluation have been explained, with task model parameters being set. The features of the localisation platforms which will be evaluated have been set out and described, and an evaluation method has been devised for each feature. The next chapters will explain how the evaluation was designed, and how it was carried out, before moving onto the results and analysis.
6 Evaluation Design and Execution

This chapter will be setting out how the evaluation was designed and carried out, as specified in EAGLES steps six and seven. It will start with an explanation of the evaluation’s structure, before describing the materials and tools used for the evaluation, as well as the aims of the pilot tests. The chapter will then be split into two further sections: section 6.5 will describe how the pilot test and evaluation for the performance of the localisation tools proceeded, with section 6.6 doing the same for the evaluation of the machine translations done by each localisation tool.

6.1 Evaluation Structure

The evaluation is made up of two sections which, from the point of view of the end user, are very important: using the software, and evaluating the quality of the machine translation output. Therefore, two different evaluations were designed with two different sets of evaluators: one set to evaluate the localisation software performance, with the help of a scenario test, with the other set to evaluate the machine translation output.

The test scenario is therefore a user who will be creating a localisation project with both Reverso Localize and SYSTRANLinks, before publishing the final website. It was decided that each participant would create localisation projects with both the tools, in order to be able to compare the questionnaire results with greater ease.

After the project creation, the translations from Reverso Localize and SYSTRANLinks for the Accueil page described in section 5.2.3 will then be passed on to another person who will evaluate the quality of the translation done by the two platforms. The way in which this process will be carried out means that all available data and opinions will be collected, which will be shown and analysed in the next chapter.

Each stage of the evaluation has been supported by documentation in order to try to create a situation that is as uniform as possible. A uniform situation is necessary, because if the evaluation variables are not the same for each user, it will be extremely difficult to correlate and analyse responses and data.

The second part of the evaluation, that of evaluating the machine translations of a certain page of the website to be localised, will be evaluated by using a Microsoft Excel 2010 spreadsheet, where participants can rate, on a scale of one to five, the accuracy of the machine translations provided. This part of the evaluation is by its very nature uniform, as the participants have all similar profiles, and the same translations are given to each participant.
6.2 Materials

Various materials were prepared in English for the evaluations:

- participant information (Annex I)
- consent form (Annex II)
- instructions - for Reverso Localize (Annex IV), SYSTRANLinks (Annex V), and the machine translation evaluation (Annex VI)
- questionnaires to get information on the localisation participants (Annex III), plus questionnaires given after the scenario tests of Reverso Localize (Annex VII) and SYSTRANLinks (Annex VIII)

The participant information and consent form were based on guidelines provided by the Commission d'éthique at the University of Geneva. The participant information stated the objectives of the research, along with what the participant should expect from the evaluation process and what they are expected to do, including how long the evaluation was expected to last, benefits and risks to the participant, data protection issues, and contact details of the researcher. The consent form asked that the participants agreed to what was in the participant information, and gave their permission for data collected to be used in this thesis.

Afterwards, the separate user instructions were prepared for Reverso Localize and SYSTRANLinks. Taking instructions from classes given by the Multilingual Information Processing Department (TIM) to help students with their practical work, a basic structure was formed where each step was set out clearly, in simple sentences and with the help of screenshots where necessary. In cases where it was felt that the participants should think about the consequences of different steps, extra explanations or suggestions were given in italics. This structure was decided upon as the participants had all taken TIM classes, and were therefore used to instructions like this. As Reverso Localize had been evaluated in Peron (2013), some of the structure of her instructions was also used as an aide.

Much more of a challenge was the creation of the questionnaires for Reverso Localize and SYSTRANLinks. Creating questionnaires meant leaning heavily on academia, such as Briony J Oates’ Researching Information Systems and Computing and Arlene Fink’s How To Ask Survey Questions. A questionnaire is a “pre-defined set of questions, assembled in a pre-determined order” which “must be carefully designed and constructed so that valid, reliable data can be generated,” with “each question relating to research questions” (Oates 2006, p. 219). Oates (2006) also suggests that
each question be brief (20 words or less), relevant to the overall questionnaire, unambiguous, specific and objective, with Fink (1995) cautioning against what she calls biasing words, as well as two-edged and negative questions - the researcher must not lead the person taking the questionnaire into an answer through the wording of a question.

As learning the opinions of the participants was the main objective of the questionnaires, both types of question (open and closed) were used when designing them. This is because using open questions all the time can put people off who are not sure of their opinion, whereas closed questions can always elicit some kind of response (Fink 1995). Scale questions were very often used, along with Lickert scale questions (where a participant rates their degree of agreement or disagreement), yes/no questions, and opinion questions; it is normally considered a good idea to vary the structure of a questionnaire as it can keep a participant’s attention (Oates 2006). With the Lickert scale questions, the ‘neither agree nor disagree’ category was erased, in an attempt to take people off the fence as pointed out in Oates (2006). The questionnaire was split into the sections which would be assessed in the analysis section, in an attempt to keep each question as relevant as possible to the overall questionnaire. The sections led the participants to be asked their opinions on: interface and usability, interaction, the localisation tools, the localised site, publishing, and time taken, with a section for any final opinions.

The questionnaire to get information on participants was designed to assess the localisation and technology profile of each participant. They were straightforward questions, such as sex, age, maternal language, knowledge of certain localisation aspects, and computer experience. The responses given to this questionnaire meant that the profile of each participant could be put in relation to the answers given on their questionnaires for Reverso Localize and SYSTRANLinks. It was decided to not create a participant questionnaire for the machine translation evaluation, as the desired profile was much more fluid: all that was needed from the participants was that they were translation students and native English speakers. The opinion and participant questionnaires in Peron (2013) inspired the creation of these questionnaires used, as the information required from the participants was very similar.

The participant information questionnaire was printed out and then filled in by hand by each participant before the evaluation. This was done so that the participant’s profile was known when reading through the opinion questionnaires. The opinion questionnaires were completed by each participant at the end of the Reverso Localize and SYSTRANLinks evaluations, when the experience was still fresh in their minds.
6.3 Tools

For the evaluation of the localisation platform, a personal laptop was used. It was decided that the same laptop should be used, in an attempt to create consistency. The laptop was a Toshiba Satellite L755-15R with an Intel(R) Core(TM) i5-2410M CPU @ 2.30GHz processor, running a Microsoft Windows 7 64-bit 2009 version, service pack 1 system. The laptop’s interface and programmes were set in English, with a QWERTY English-language keyboard.

Several components were added to the system in order to carry out the evaluation of the localisation platforms: Google Chrome as the internet browser (with Gmail accounts set up for each evaluator), Notepad++ as the text editor, BB FlashBack Express to record the evaluation, and a folder with necessary information.

All information and instructions necessary were installed on the desktop. The instructions were in .pdf format, so that they couldn’t accidentally be changed, and were put there in case the evaluators preferred reading the information on the screen. Having the .pdf files also meant that any links the evaluators were asked to go to could easily be reached from the file, as the links were hyperlinks in the .pdf files. Paper versions of the instructions and information were also given to the participants.

Google Chrome was used as the internet browser, as it was already installed on the laptop in English, and was reported to be the most liked and used by the evaluators. Notepad++ was chosen as the necessary text editor because it is commonly used at FTI during localisation classes, and therefore less time would be taken by the evaluators in learning how it works. It is also very user-friendly, with colours that display different types and parts of text.

Finally, a piece of software called BB FlashBack Express (see Figure 30) was installed, as it video and audio-records what happens on-screen, allowing data to be extracted for analysis. This is particularly useful as it records information such as mouseclicks and key touches, and also allows the person viewing the recorded video to jump back and forth between mouseclicks and key touches, which is particularly useful for analysis.
The recorder was started and stopped by the researcher, at the beginning and end of each evaluation after the evaluator said that they were ready to begin and that they had finished. Each participant was told before consenting to the evaluation that the screen and noises would be recorded, and were reminded of it by the instructions.

Email addresses were also set up for the evaluators, as both localisation platforms required the use of an email address to sign in, and, in the case of Reverso Localize, notifications would be sent to the user’s email address. In order to keep the evaluations as anonymous as possible, Gmail accounts and passwords were created for each participant, in the following format (where the number changed depending on the participant number):

Email: localisation.user1@gmail.com
Password: localisationuser1

The example above shows that the email address would start with ‘localisation,’ followed by the user number. The password follows the same format, without the full stop between ‘localisation’ and the user number.

After the creation of the email addresses, accounts were created by the researcher on both Reverso Localize and SYSTRANLinks, with the same password as for the email address, in order to make it as simple as possible for the evaluation participants. These accounts were also activated from the email addresses, so that the evaluators would only have to log onto the localisation platform websites in order to start the evaluations.

The tools used for the machine translation evaluation were fewer and more flexible in their specifications. Participants could use their own laptop, or the same laptop as provided for the localisation performance evaluation, as they wished, with the only requirement being that they had Microsoft Excel 2010 or newer installed, as the evaluation would be recorded on a spreadsheet.
6.4 Pilot Test Objectives

It was decided to create pilot tests to test the material created and the tools chosen. The type of pilot test evaluator chosen had to be as close as possible to the type of evaluator chosen for the actual evaluation, in order to keep the tests consistent.

The questionnaire material was tested to find out whether there were problems such as those defined in Oates (2006, p. 226). It is important to ask whether the participant is:

- having difficulties answering certain questions
- finding some questions ambiguous or vague
- following the instructions for how to answer the different types of questions

as well as finding out whether the pre-defined responses cover all desired answers, and how long it takes for the participant to complete the questionnaire.

Although these points to consider are only particularly relevant to questionnaires, they were a good place to start from for evaluating the instructions given as part of the material. Therefore, pilot test participants were asked if they had difficulties in fully understanding the instructions, and whether the instructions were too ambiguous or vague to follow easily.

The next part of this chapter will be split into two sections: one to describe the pilot and the execution of the localisation tool evaluations (section 6.5), and the other to describe the pilot test and the execution of the machine translation evaluations (section 6.6).

6.5 Localisation Tool Performance Evaluation

This section will describe the pilot test and execution of the performance evaluation of Reverso Localize and SYSTRANLinks. It will start with the results from the pilot test, before moving onto the evaluation itself.

6.5.1 Pilot Test

Once the tools were installed and the material prepared, the pilot test was carried out to identify any possible problems.

Firstly, the material was proofread by a person who had localisation knowledge and who was also a native language English-speaker, in order to make sure that the material
provided was correct from a linguistic point of view. This meant that the pilot test would for the most part be concentrated on the wording of the instructions and questionnaires and how well they related to the execution of the real evaluations.

The person doing the pilot test was asked to be critical and say anything they were not happy with or that they thought could be improved, particularly in terms of clarity. To do this, they were particularly asked to follow the criteria as set out by Oates (2006) in section 6.4, and point out anywhere where further instructions or visual aids could be useful.

The pilot test evaluator for the localisation tool evaluation suggested some improvements: mostly that more screenshots were added, but there was one point where there were problems with linking SYSTRANLinks’ localised website to the original website. This was because the evaluator had closed the tab containing the localised website and its address, which caused some loading problems when the evaluator tried to re-open it, so a warning note was added to the instructions, telling the evaluators to not close SYSTRANLinks’ localised website.

The pilot test showed some fundamental issues with the original questionnaires used for the evaluations. The participant felt that the questions skipped too much from area to area within the evaluation. Therefore, a structure was decided upon where each question related to an EAGLES characteristic, as well as the function, that was being evaluated, as described in section 6.2.

It can be seen from the comments above that the pilot test was very useful: any problems were addressed and resolved, and it was possible to estimate how long the evaluations should take and thus plan accordingly.

6.5.2 Execution of the Evaluation

As it is a place that all three participants know, are used to, and can access easily, the evaluation took place at the University of Geneva’s Uni-Mail building, on the 2nd floor, which is where FTI’s library is situated. To keep the evaluation consistent, this place never changed, and it was essential that there was a consistent internet connection, as both platforms to be evaluated could only be accessed online.

The evaluations took place over a space of a week, between 1st and 8th June 2014, in order to fit in with the evaluators’ availabilities. In this time, the pilot test as well as the three evaluations took place. Each evaluation took place in a reserved room in the FTI library, in order to give the participants time and space to relax and work without pressure, and to ask the researcher questions as necessary. There was no time limit for
the localisation evaluation, although from the pilot test it was estimated that it would take no more than an hour and a half.

All information required (such as the participant information and the instructions for each localisation platform) was already installed on the laptop before the evaluations were started. When entering the room, each participant received the participant information (Annex I), which contained all necessary information on how the evaluation would proceed, and a consent form (Annex II), which they were asked to read and sign. They then were given and asked to complete a participant questionnaire (Annex III), which asked information on any localisation knowledge they had prior to the evaluation. Afterwards, each participant received the following:

- username and password that would work for Gmail, Reverso Localize, and SYSTRANLinks
- Reverso Localize instructions (Annex IV)
- Reverso Localize questionnaire (Annex VII)
- SYSTRANLinks instructions (Annex V)
- SYSTRANLinks questionnaire (Annex VIII)

It was originally planned that each set of instructions and questionnaires would be given to the participant separately, but the pilot test participant preferred having the freedom to choose which localisation platform to evaluate first, and thus all material was given at the same time, in three sets: the first set had the username and password details, the second had the Reverso Localize instructions and questionnaire, and the third had the SYSTRANLinks instructions and questionnaire.

At this point, the evaluation itself started. The researcher started BB Flashback Express, then left the laptop with the participant to start the evaluation, following the instructions as seen in Annexes IV and V.

The first part of the evaluation was one of the most important, as it allows participants to get used to the site to be localised, as well as the interface of whichever platform is being used as the localisation tool. To do this, they were asked to use Google Chrome, and open separate tabs for the localisation tool and the website to be localised so they could go back and forth between the websites as they wished. The participants were to spend some time familiarising themselves with the interface of the localisation tool, and to find out how it works by either reading the material provided on the website, or, in the case of Reverso Localize, by having the option to watch an information video.
Once this initiation phase was over, the participants were asked to login to the localisation tool and, in the case of Reverso Localize, their Gmail account, with the information provided at the beginning of the evaluation. They were then asked to create a new project, and put in the information as explained in the instructions, such as the target language of the website to be localised. In this way, the website to be localised is localised by the tool as requested, and the participant can view the localised website. There is also the option to invite a reviser, which was not used, and to publish the localised website.

The final stage for the participants was to link the localised website to the original, by downloading the original website’s source code and changing the source code to Notepad++ so that the original website will link to the English-language version of the site as translated by either Reverso Localize or SYSTRANLinks. If done correctly, it was possible to open the source code of the original website in Google Chrome, but this time the English-language version of the website would be the one as localised by either Reverso Localize or SYSTRANLinks. As this was the end of the evaluation, the participant would tell the researcher that they had finished, the researcher would then stop BB Flashback Express’ recording, and thus the recorded part of the evaluation would be finished. The participants were then asked to fill in a questionnaire relating to their experience.

The same process would repeat itself with the second localisation tool, and the participants would complete a second questionnaire relating to the evaluation of the second tool.

6.6 Machine Translation Evaluation

This section will be describing the pilot test and evaluation carried out for the evaluation of the machine translations done by Reverso Localize and SYSTRANLinks.

6.6.1 Pilot Test

The pilot test for the machine translation evaluation was mostly to check that the instructions provided (Annex VI) were written in such a way to provide a clear idea of how the translations were to be evaluated by the participants, as well as to check the clarity of the spreadsheet, as shown in Figure 31.
The pilot test participant was asked to be as critical as possible in their reading of the instructions, as it was imperative that the instructions made it clear that accuracy was the factor for which the translations were being evaluated, and to define what was meant by the term ‘accuracy.’ In this way, some changes to the definition of ‘accuracy’ were made in order to make the definition clearer.

During the pilot test, it was decided to lock the source and target text in the spreadsheet, as the pilot test participant accidentally deleted some of the source and target text given for one of the localisation tools, by putting a ‘1’ into the target text area instead of the scoring area. Due to the nature of Microsoft Excel, this meant that the entire cell’s content was deleted.

This pilot test was helpful, as it allowed for a clearer explanation of ‘accuracy,’ as well as ensuring against any errors made on the spreadsheet itself.

6.6.2 Execution of the Evaluation

As with the evaluation of the localisation tools’ performance, the evaluation of the machine translations took place at the University of Geneva’s Uni-Mail building, on the 2nd floor, in the FTI library, as the area is known to and accessible by all three participants. The evaluations took place between 6th and 11th June 2014, again to fit in with the availabilities of all. To keep the work space relaxed, a reserved room was used, so that the participants could take their time and ask any questions as necessary. Participants were asked to keep the evaluation time to less than an hour, in order to spare overthinking in the process.

The laptop as specified in section 6.3 was at the participants’ disposition, but they were allowed to use their own laptop if wished, on the condition that they had Microsoft
Excel 2010 or newer installed. Despite this, all participants chose to use the laptop provided. The laptop already had the relevant information and instructions installed on the desktop. When entering the room, each participant received a participant information sheet, which detailed all necessary information on the objectives of the evaluation, and a consent form (Annex II), which they were asked to sign. Afterwards, they were given a user number, and the instructions on how to carry out the evaluation, where they were advised to have a look at the website for which the translations had been done. They were then invited to open the spreadsheet to view the translations.

As specified in the instructions, the participants were to score each translation on a scale of one to five on the translation’s accuracy. All participants were invited to look at the source text provided within the context of the website if they so wished.

To register their score, they had to put a ‘1’ in the score that they agreed with the most. Once this one done for one of the localisation tools, they were asked to do the same for the other tool. At no point did the participants know which translation came from which system. Once the evaluations were completed for both localisation tools, the evaluation was finished.

6.7 Conclusion

This chapter has described the design and execution of the evaluations, including the tools used, material provided, and pilot test results. It has also described how the evaluations were carried out in more detail, and has therefore come close to completing the seventh step of the EAGLES 7-step recipe for evaluating systems. The next chapter will present the results of the evaluation, thus completing the EAGLES final step, and will continue by analysing the results.
7 Results and Analysis

This section will present and analyse the results obtained during the study. The structure of this section will follow the evaluation methods table set out in the methodology section: the localised site analysis for Reverso Localize and SYSTRANLinks will be in section 7.1, before moving onto the video analysis in section 7.2. Section 7.3 will analyse the questionnaire responses from each participant, with section 7.4 being dedicated to the machine translation human, BLEU, and quality assessment evaluations. A final section will display a summary of the results, with their scores, for easier comparison.

In all comparison tables, Reverso Localize has been shortened to RL and SYSTRANLinks to SL. It has been decided that all sections analysed here will carry the same importance, because deciding upon a weighting system of this size is a body of work in itself. This will be briefly discussed in the conclusion of this thesis.

7.1 Localised Site Analysis

The analysis of the localised site will be evaluating the accuracy of the localisation tools, by counting localisation errors, omissions, and silences for both Reverso Localize and SYSTRANLinks. Reverso Localize will be examined first.

Reverso Localize displayed two localisation problems, and one that could be argued to be a problem, but which could make the website more accessible from a language point of view. This final ‘problem’ changed all ampersands present in the original website to the written ‘and’ in the localised website (see Figure 32). In her study of Reverso Localize, Peron (2013) mentioned that she had a localisation problem with ampersands, in that the localisation system changed all of them to ‘&amp;’. It may be possible that the developers of Reverso Localize have noticed this problem, and have decided to change the coding of the platform so that all ampersands are changed to ‘and,’ to prevent any future problems. This could, however, have an adverse effect on space; it is possible this needs to be changed so that ampersands are correctly localised.

CONTACT & PLAN D’ACCÈS
CONTACT AND AREA MAP

Figure 32: Original (top) and localised (bottom) versions of ampersand

Another problem was the occurrence of lowercase letters being replaced by capitalised ones in ‘BOAS Swiss Hotels’ on each page. An example of this is shown in...
the ‘Information’ tab as presented in Figure 33 below. For each occurrence of ‘BOAS Swiss Hotels,’ Reverso Localize has capitalised everything to ‘BOAS SWISS HOTELS.’

Figure 33: Original (left) and localised (right) website: capitalisation issue

Finally, there was an issue with the text formatting, in that the localisation system did not format the correct words in bold, as shown in Figure 34.

Figure 34: Original (right) and localised (left) website: text formatting issue

There was, however, only one omission. When the mouse hovers over the contact email address in the original website, a tooltip appears with the same email address. This does not happen in the localised version, as shown in Figure 35 below.

Figure 35: Original (left) and localised (right) tooltips

Figure 35 also shows that Reverso Localize’s system capitalised the ‘I’ in ‘info’ for the contact email address.

As for the silences, there were three main occurrences. The calendar used to reserve rooms was not translated into English, and all buttons save for those labelled ‘Details’ were left in French. Finally, when the link to the ‘Gallery’ page was opened in the localised website, Reverso Localize would revert back to the original French website, and the only way to get back to the localised site was to click the back button on the web browser.
While all these omissions are serious in their own right, in that they could stop potential customers from continuing with their exploration of the website, this final omission is a big issue. From further examination of the Reverso Localize dashboard, it is evident that Reverso Localize did not recognise the Gallery as part of the sitemap and therefore did not extract it for localisation. Because of this, a user hoping to localise this website using Reverso Localize would not be able to fix the problem from within the localisation system.

SYSTRANLinks displayed two localisation problems across the entire site: any sections that were underlined in the original website were not underlined in the localised version, as shown in Figure 36.

![Figure 36: Original (left) and localised (right) underline problem](image)

Continuing with formatting problems, SYSTRANLinks also displayed an issue with text formatting in the localised website, by formatting the wrong text as bold, as shown in Figure 37.

![Figure 37: Original (left) and localised (right) text formatting problem](image)

Unlike Reverso Localize, there were no omissions, and the only silence was that the calendar box was not translated, as with Reverso Localize.

Table 5 shows the problems described above, put together for easier comparison.
Table 5: Localisation systems’ problems, omissions, and silences

<table>
<thead>
<tr>
<th>Localisation Problems</th>
<th>Reverso Localize</th>
<th>Occurs</th>
<th>SYSTRANLinks</th>
<th>Occurs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bold font</td>
<td>4</td>
<td>Bold font</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Capitalisation</td>
<td>2</td>
<td>No underline</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Written ampersand</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Localisation Omissions</th>
<th>Reverso Localize</th>
<th>Occurs</th>
<th>SYSTRANLinks</th>
<th>Occurs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tooltip</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Localisation Silences</th>
<th>Reverso Localize</th>
<th>Occurs</th>
<th>SYSTRANLinks</th>
<th>Occurs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untranslated calendar</td>
<td>1</td>
<td>Untranslated calendar</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Untranslated button</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gallery page</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.2 Video Analysis

The videos of the evaluations carried out by each participant will be analysed to evaluate the operability of the tools, by counting clicks, keys pressed, and tabs open for each task (section 7.2.1), as well as the efficiency of the localisation tools, by analysing the time each tool took to localise the website chosen, and the time taken by each participant to fulfill certain tasks (section 7.2.2).

7.2.1 Operability

The operability of the localisation systems was measured through the number of clicks and key touches during the localisation process, as well as how many tabs it was necessary to have open.

Number of Clicks

Table 6 below shows the number of clicks in the localisation process for each participant for Reverso Localize and SYSTRANLinks, with the average clicks shown in whole numbers at the bottom. The clicks were counted from when the participants logged in to the tools to start the localisation task. The localisation process was separated into three main tasks: creating the project, publishing the website, and linking the original website to the localised website.
A note must first be made. It was not possible for the evaluation participants to publish the localised website with SYSTRANLinks due to security restrictions. There are two methods for publishing a website localised by SYSTRANLinks:

1 Default method

1 Download HTML verification file given by SYSTRANLinks
2 Upload the file to source website
3 Confirm upload by visiting link given by SYSTRANLinks
4 Click on ‘Check’ button

2 Alternative method

1 Insert a meta tag given by SYSTRANLinks into a certain section of the source website’s homepage
2 Click on ‘Check’ button

If using the default method, this would include, at a conservative estimate, at least an extra 10 clicks (downloading and saving the file, going to the source website, uploading the file, going back to SYSTRANLinks to confirm upload, and clicking the ‘Check’ button, as well as clicking between webpages). As it was impossible for the users to do this, clicks could not be counted. It is, however, impossible to deny that the method used by SYSTRANLinks is much more complex than Reverso Localize, and therefore this task is made much less easy to operate, although perhaps more secure.

Therefore the total number of clicks in Table 6 has a different format to what might be expected. In the ‘Total’ column for SYSTRANLinks, there is a * next to the counted total and a ~ next to the estimated total, with the estimated number of clicks for publishing written in italics in the ‘Publishing’ column.

<table>
<thead>
<tr>
<th>User</th>
<th>Total</th>
<th>Project Creation</th>
<th>Publishing</th>
<th>Linking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>SL</td>
<td>RL</td>
<td>SL</td>
</tr>
<tr>
<td>1</td>
<td>31</td>
<td>*25 ~35</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>*23 ~33</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>*29 ~39</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>Average</td>
<td>33</td>
<td>*26 ~36</td>
<td>13</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 6: Total clicks per user and task for Reverso Localize and SYSTRANLinks

When looking at the tools separately, the results do not greatly deviate, although user 2 did have a slightly higher click rate using Reverso Localize and user 3 had a slightly
higher click rate using SYSTRANLinks. On further examination of user 2’s Reverso Localize video this is because they changed much more between the tabs containing Reverso Localize and their email account, belying their confusion at not having received any email notifications as promised by the platform. Although this confusion was also the case with the other two users, they gave up expecting an email quicker than user 2. This notification method must be improved by Reverso Localize, in order to improve the understanding and usability of the platform. On further study of user 3’s SYSTRANLinks video, the slightly higher click rate is due to them making an error with the language selection during the project creation stage.

It can be seen that the number of clicks for project creation was much lower for SYSTRANLinks than for Reverso Localize. This is partially because the steps are laid out in a simpler way for SYSTRANLinks than for Reverso Localize, with only one or two clicks per step. Reverso Localize, on the other hand, has expandable menus that must be opened to select extra information, such as whether the user wants to create a translation memory - without these menus the process would be simpler. The lower click rate for SYSTRANLinks can also be attributed to the fact that some of the steps involved with creating a project in SYSTRANLinks is automated, such as setting the source language of the website to be localised.

The click rate for each tool in the project creation stage was expected to be similar, as the steps involved are similar. This is the case for SYSTRANLinks (with user 3 having a slightly higher click rate due to selecting the wrong language), but user 1 had a much lower click rate when creating a project with Reverso Localize. This is because both users 2 and 3 had problems loading the page after the creation of the project and had to click on the back button in order to reload the page.

The highest number of clicks for all users for both tools is attributed to the task of linking the source website to its localised version. For Reverso Localize, this is because there is no simple way offered to link the websites. SYSTRANLinks’ method was not suitable for the website used in this evaluation, and no instructions were offered if users wanted to link without the language selector display. This was most likely done to encourage users to stay with the language selector, thereby giving SYSTRANLinks’ logo more of a presence on the localised website.

It was expected that the click rate for both tools would match when linking the source and localised websites, as the process is almost exactly the same. On average, SYSTRANLinks required one click more, in order to hide the language selector bar automatically provided. This bar had to be hidden as it only gave two language options, French and English, despite the website also being written in German.
Number of Key Touches

Using BB Flashback Express, it was possible to extract an XML file of all keys pressed, which could be examined in Microsoft Excel. BB Flashback Express separated the keys into two types: ‘Key,’ which can be considered as normal keys such as letters and symbols; and ‘Virtual Key,’ which are keys used for an action, such as deleting, or for commands or shortcuts. Below is a list of the Virtual Keys used by all participants, split into command keys on the left and shortcuts on the right:

Ret 
TAB 
ArwUp 
HOME 
ArwRight 
Del 
ArwDwn 
ArwLft 
Esc 
Ins 
BkSp

It can be seen from this table that the key touches required for both platforms are quite simple, and are not shortcuts or keys that are unknown to most computer users.

Table 7 below shows the keys pressed for both Reverso Localize and SYSTRAN Links, according to the Virtual Key or Key category.

<table>
<thead>
<tr>
<th>User</th>
<th>Keys</th>
<th>Virtual Keys</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>SL</td>
<td>RL</td>
</tr>
<tr>
<td>1</td>
<td>72</td>
<td>92</td>
<td>197</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>79</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>65</td>
<td>215</td>
</tr>
</tbody>
</table>

Table 7: Key touches for Reverso Localize and SYSTRAN Links

It can be clearly seen here that in almost every case, the participants used more Virtual Keys than Keys, with the exception of user 2, who used more Keys than Virtual Keys when using SYSTRAN Links, while favouring the mouse when using Notepad ++. Users 1 and 3, on the other hand, used the arrow and backspace keys a total of 105 and 56 times respectively, almost all of which in Notepad +++, and in high contrast to user 2’s 2 times.
Table 7 also shows that Reverso Localize had more Keys pressed, and, after further examination of the video, this shows that during the project creation phase Reverso Localize needed extra information typed in in order to create the project, such as the project name and details, which SYSTRANLinks did not need.

Further examination of the XML file shows that the three most used Virtual Keys for both Reverso Localize and SYSTRANLinks were [Ctrl]+[C], ArwRight, and ArwDwn, showing that the Virtual Keys were mostly used for linking the source and localised websites together.

Overall, SYSTRANLinks required less key touches than Reverso Localize throughout the localisation process, probably due to the slightly simpler project creation process, as described above.

**Tabs Open**

Reverso Localize required more tabs to be open throughout the localisation process, as shown in Table 8 below. This is because of the email notifications that it sends out, thus requiring an extra tab to be open for the user to check their emails.

<table>
<thead>
<tr>
<th>Reverso Localize</th>
<th>SYSTRANLinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source website</td>
<td>Source website</td>
</tr>
<tr>
<td>Dashboard</td>
<td>Platform homepage</td>
</tr>
<tr>
<td>Mirror site</td>
<td>Localised site</td>
</tr>
<tr>
<td>Email page</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Tabs open for Reverso Localize and SYSTRANLinks

SYSTRANLinks has improved operability over Reverso Localize with loading its localised site, as each time a user clicked on the ‘Browse & Edit’ button, the localised website would simply refresh itself if it was already open in a tab. Reverso Localize, on the other hand, would open its mirror site in a new tab every time the link to it was clicked. Peron (2013) also noted a problem with Reverso Localize, in that when a link was opened from an email notification a new tab would open to the Reverso Localize website, even if it was already open in the internet browser. This study was unable to verify this problem due to the lack of email notifications.

**7.2.2 Efficiency**

The efficiency of the platforms was tested in two ways: through the speed of the localisation machines, and through the amount of time it took for each user to complete the evaluation as a whole, as well as key tasks.
The speed of both Reverso Localize and SYSTRANLinks’ localisation machines were timed through the video analysis. In order to localise the source website, both platforms must firstly extract all data from the source website, translate everything into the target language, and then redirect the user to the revision area of the website, after having created a target language website hyperlink. Reverso Localize also claims to check the spelling of the source website, but no evidence was shown of this.

For each localisation tool, the time was calculated for the entire process from the moment each user confirmed the creation of the localisation job, as while Reverso Localize split the process into extraction, translation, and publication, SYSTRANLinks did not. Table 9 shows the results.

<table>
<thead>
<tr>
<th>User</th>
<th>Reverso Localize</th>
<th>SYSTRANLinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01:39</td>
<td>00:56</td>
</tr>
<tr>
<td>2</td>
<td>03:37</td>
<td>01:15</td>
</tr>
<tr>
<td>3</td>
<td>02:13</td>
<td>01:23</td>
</tr>
<tr>
<td>Average</td>
<td>02:29</td>
<td>01:11</td>
</tr>
</tbody>
</table>

Table 9: Localisation times for Reverso Localize and SYSTRANLinks

It can be seen from Table 9 that SYSTRANLinks is significantly faster than Reverso Localize, taking an average of 1 minute and 18 seconds less to extract, translate, and make the localised website available for editing. It can be seen that Reverso Localize took much longer to complete the process for user 2; this is because Reverso Localize experienced some problems loading the revision area, causing user 2 to refresh the screen several times.

The next stage of evaluating the efficiency of the localisation tools was through timing the scenario test carried out. Each video was split into its main stages: introduction, project creation, familiarisation with the localised website, and linking the source and target websites. It must, however, be noted here that the times for each user’s familiarisation with the localised website is displayed for clarification purposes, and will not be commented upon, as it is less to do with how well the tools perform, and more to do with each user’s preferences on how long they took to explore the localised website. Their feelings on the localised websites from both Reverso Localize and SYSTRANLinks will be analysed in the questionnaire analysis in section 7.3.

Each stage was timed using the timer on BB Flashback Express, with the results shown in Tables 10 (for Reverso Localize) and 11 (for SYSTRANLinks).
The tables show that using SYSTRANLinks is more time efficient across all tasks.

It took much less time for each participant to introduce and familiarise themselves to SYSTRANLinks than Reverso Localize. This could be because SYSTRANLinks offers much less information on how the localisation process works, therefore giving users less information to sort through while getting used to the website, although this is not necessarily an advantage.

During the project creation phase, users 2 and 3 took much longer with Reverso Localize than SYSTRANLinks. After further video analysis, this is because of Reverso Localize’s notification system. Both users were confused about the lack of email notifications and the lack of icon notifications, with both checking each notification area several times and user 3 even checking the spam email folder to make sure the email notification had not been filtered by Gmail, although user 2 checked more times than user 3. This clearly shows that if Reverso Localize wishes to continue with its three-fold notification system, it needs to ensure that the system works well to make it more efficient.

Finally, the linking stage was much shorter for SYSTRANLinks. As the process for linking the websites are identical for both platforms, extra video analysis was needed. It was discovered that it took longer to link the websites using Reverso Localize because all of the users glanced through or read the instructions provided by the Reverso Localize blog, with user 1 taking 30 seconds to glance through it, user 2 taking 5 minutes and 22 seconds to read it, and user 3 taking 2 minutes and 54 seconds. SYSTRANLinks does not offer similar instructions: if these Reverso Localize reading times are factored out, the linking times for Reverso Localize and SYSTRANLinks are much more uniform.
7.3 Questionnaire Analysis

After using both localisation platforms, each participant was asked to fill in a questionnaire giving feedback on what they thought of each tool (see Annexes 7 and 8). From the questionnaire, this section will evaluate: the suitability of the tools through the participants’ interaction with them and their views on the mirror site; the accuracy of the localisation system; and finally, the participants’ views on the understandability, learnability, and efficiency of the tools.

Many of the questions the participants answered were closed questions: that is, they were asked to rate opinions or qualities from 1 to 6, or choose a range of options from ‘agree’ to ‘disagree.’ In this case, each category was given a score from 1 to 4, with the most negative category being given the lowest score. These scores were then averaged, which is the questionnaire score that can be seen in the results tables in this section. Where necessary, the corresponding category name will be given along with the average score.

Also where necessary, the questionnaire number will be given to show from which questionnaires the results were taken - for example, RL11 will mean Reverso Localize questionnaire, question 11.

7.3.1 Suitability

The suitability of the systems will be evaluated through the participants’ opinions on their interaction with the tools and on the mirror site provided by each of them, along with publishing and linking the source and localised websites.

Interaction

In the questionnaires, the interaction system was evaluated through the notification system of both Reverso Localize and SYSTRANLinks, with both closed and open questions (RL6, RL7, RL8, RL9; SL6, SL7).

As a reminder, section 4.1 described Reverso Localize’s three-fold notification system: an email notification, a pop up window, and an icon, with SYSTRANLinks making use of pop up windows. Table 12 below shows the results on the efficiency of the notifications systems used by Reverso Localize and SYSTRANLinks (RL6, SL6), with the lowest score being 1 and the highest being 4.
Table 12: Notification system efficiency

Two out of the three evaluators agree with each other, giving Reverso Localize a score of 2 and SYSTRANLinks a score of 3, with user 3 giving Reverso Localize a score of 3 and SYSTRANLinks a score of 4. The difference in the scores, however, does not change, with all evaluators giving SYSTRANLinks one point more over Reverso Localize. This is reflected in the average scores, with Reverso Localize receiving an average score of 2.3, and SYSTRANLinks an average of 3.3.

This is most likely due to the likelihood of the systems working well: SYSTRANLinks’ one notification was far more likely to work efficiently than Reverso Localize’s three-notification system. The Reverso Localize score can perhaps be explained by examining the answers to RL8 and RL9, where all the evaluators stated that it was unnecessary to have all three notification types (email notification, pop up window, and icon), although their choices ranged widely, with user 1 preferring a pop up window only, user 2 preferring a pop up window and the icon, and user 3 preferring the email notification. It must be noted that both user 1 and 2 did not receive email notifications until the day after the evaluation (when the email accounts were checked by the researcher), whereas user 3 did receive an email notification, but not until some time after the localisation of the website was completed.

SYSTRANLinks scores higher for interaction, with suggestions being made that Reverso Localize either ensures that emails reach users in a more timely manner; gives users a choice of notification systems; or changes their notification system so that there are only one or two notification types used throughout the localisation process.

Localised site

As Reverso Localize and SYSTRANLinks have different ways of displaying the localised site, different questions were asked in the two questionnaires (RL13, RL14; SL11, SL12, SL13, SL14). However, the first question for both systems asked whether the localised site was useful, with users being asked to rate the usefulness on a scale of 1 to 6, with 1 being not useful at all and 6 being very useful. Table 13 below shows the results and average scores.
SYSTRANLinks consistently scores higher for all participants in its use of the localised site. All users said in their comments on Reverso Localize’s localised site that the banner advertising the company at the bottom of the site was too intrusive, and that the dashboard option offered by SYSTRANLinks was more flexible in changing the site around. Two of the users noted that the pop up windows to enable post-editing in the Reverso Localize site annoyed them, whereas the SYSTRANLinks post-editing was more discreet.

However, when given the option on what post-editing environment they would prefer, the participants were completely split, with user 1 preferring both the dashboard offered by SYSTRANLinks and the localised site, because they liked being able to change resources as necessary, user 2 preferring the localised site, but admitting the dashboard was more practical, and user 3 preferring the dashboard as it was more advanced and gave more options for changing the localised site as needed.

From these results, it could be possible that the reason SYSTRANLinks’ localised site was preferred was due to the much less intrusive logo advertising its services, combined with the possibility of using the dashboard. Despite SYSTRANLinks being preferred, it must be noted that Reverso Localize did get a better than average score, which shows that its localised site was helpful to the participants. A suggestion to Reverso Localize could be to give users an option on how to edit the localised website.

**Publishing and Linking**

In the questionnaires, publishing and linking the localised site to the source website were evaluated from the answers to RL15, RL16, RL17, RL18, and RL19, and SL15, SL16, and SL17.

On a scale of 1 to 6, with 6 being very difficult, the average score for how difficult the users found the publishing process with Reverso Localize was 3, so it was found averagely difficult. This score is slightly more than Reverso Localize developers would expect, as the tool is advertised as being easy to use from beginning to end.

All three users were very critical of the Reverso Localize blog, which gives instructions on how to link the websites together, with the main criticisms being that

<table>
<thead>
<tr>
<th>User</th>
<th>RL Score</th>
<th>SL Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Average Score</td>
<td>3.7</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Table 13: Localised site results
the blog was clearly not written by a native English speaker, and it assumed users would have more HTML knowledge than a beginner would. The method for confirming administrative and ownership rights over the website (in order to legally publish it) was highly criticised as well, with all participants pointing out that no proof was needed that a person has the right to publish a localised version of the original website, as all that was needed to confirm publishing rights was checking a box.

The method for linking and publishing the localised and original website with SYSTRANLinks was much different than the method used for Reverso Localize, and therefore slightly different questions were asked in the questionnaire. SYSTRANLinks offered what it calls a ‘language selector display,’ which scored an average of 3.34 out of 6. This score was brought down by user 1, who noted that the language selector was only useful if the website to be localised was monolingual at the beginning of the process. Users 1 and 2 noted that the Lake Geneva Hotel website, which was in French and German, would not work with the language selector display as it cut out the option to view the German language part of the website. None of the users felt that they would be able to link the localised and original language websites together easily without the language selector display, with user 1 noting there were no instructions provided, despite SYSTRANLinks giving users the option to not have a language selector display. However, all three users were impressed with the system used to verify that the user had the right to publish the website, noting that it relatively secure, as it asks for proof (by inserting a meta-tag or uploading a verification file).

7.3.2 Accuracy

The accuracy of the localisation systems was evaluated from questions RL10, RL11 and RL12, and SL8, SL9, and SL10. The questions asked the participants’ opinions on how well the localisation tools worked, the number of localisation problems, and extra localisation needed.

On a scale of 1 to 4, where 1 was ‘strongly disagree’ and 4 was ‘strongly agree,’ all of the participants said that they ‘agreed’ (a score of 3) that the Reverso Localize system worked well. When the same question was asked of SYSTRANLinks, users 1 and 2 agreed again that the localisation system worked well, but this time user 3 ‘strongly agreed,’ which slightly brings up the average opinion of SYSTRANLinks over Reverso Localize.

This leads onto the next two questions, where the participants were asked to quantify how many localisation problems they noticed, from the options ‘many,’ ‘quite a lot,’ ‘some,’ ‘not very many,’ and ‘none.’
Here none of the participants agreed with each other when scoring Reverso Localize, with user 1 saying ‘not very many,’ user 2 saying ‘some,’ and user 3 saying ‘many.’ If each category is given a score of 1 to 5, with 5 being given to ‘none’ and 1 being given to ‘many,’ the localisation problems that they users noticed is at an average of ‘some,’ which fits in with the assessment given by the localisation site analysis in section 7.1.

As for SYSTRANLinks, user 1 stayed with the same answer of ‘not very many,’ with user 2 agreeing with that opinion. User 3, however, stayed with the ‘many’ assessment that was also given in the Reverso Localize questionnaire. It is believed that a lot of the errors noticed by user 3 point to the errors made by the machine translation.

Despite this, the assessment given of SYSTRANLinks by the participants is slightly favourable over Reverso Localize, which once again correlates with the localised site analysis.

The final question asked the users to quantify how much extra localisation the localised website would need to be in a publishable state. Each participant was asked to choose between ‘a lot,’ ‘not much,’ and ‘none.’ Users 1 and 2 gave their opinion as ‘not much,’ with user 3 giving their opinion as ‘a lot.’

The same opinion was given for SYSTRANLinks from all three participants concerning the same question: users 1 and 2 believed ‘not much’ localisation was needed, with user 3 believing ‘a lot’ of extra localisation was needed.

Table 14 below shows the Reverso Localize and SYSTRANLinks results. Where necessary, each ordinal answer has been given a score, and the table shows the average score, along with the category that the score relates to.

<table>
<thead>
<tr>
<th></th>
<th>Efficiency</th>
<th>Problems</th>
<th>Publishable State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverso Localize</td>
<td>‘Agree’</td>
<td>‘Some’</td>
<td>‘Not much’</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(2.7)</td>
<td>(1.7)</td>
</tr>
<tr>
<td>SYSTRANLinks</td>
<td>‘Agree’</td>
<td>‘Some’</td>
<td>‘Not much’</td>
</tr>
<tr>
<td></td>
<td>(3.3)</td>
<td>(3)</td>
<td>(1.7)</td>
</tr>
</tbody>
</table>

Table 14: Questionnaire results on localisation systems

It can be seen from the table that the participants’ opinion is that SYSTRANLinks’ localisation system is more effective, with a slightly higher score given - although the two platforms’ scores match for how much extra localisation would be needed for the localised sites to be in a publishable state. This last score is probably the same for both systems, because, although Reverso Localize has more localisation issues, most are relatively easy to fix and therefore needs little extra localisation.
7.3.3 Understandability

Due to the nature of the tools, the questions relating to the understandability of Reverso localize and SYSTRANLinks were identical, which therefore makes it easier to compare the systems. There were three main questions used to draw information on the understandability of the tools, centering on their interface, as well as the clarity of information and instructions given (RL1, RL2, RL4; SL1, SL2, SL4).

Reverso Localize consistently scored higher in the interface categories (structure, accessibility, simplicity, and clarity), scoring a total average of 4.34 out of 6 over SYSTRANLink’s 3.33 out of 6. As for the clarity of the information provided by the two tools on how their sites work, Reverso Localize once again scored the highest with all three users scoring 5 out of 6, while SYSTRANLinks split the group (with user 3 scoring 5 and user 1 scoring 2), earning an average score of 3.67.

Perhaps surprisingly, however, SYSTRANLinks was rated highest in the clarity of the directions it provided throughout the localisation process, although marginally so, with an average score of 4.33 out of 6 as opposed to Reverso Localize’s 4 out of 6.

7.3.4 Learnability

Two questions were used to assess the learnability of the tools (RL3, RL5; SL3, SL5). When asked how much effort went into learning to use the tools, Reverso Localize received an average of 2.33 (with 6 being very difficult), and SYSTRANLinks scored 3.33. Reverso Localize was probably seen as easier to learn because of its interface: as mentioned above, it was seen as much clearer and easier to navigate, and the information given on how the tools works was rated much higher. This is in contrast to the SYSTRANLinks, where the website does not give a lot of information of how the tool works.

When asked if they thought they would be able to use the tools without the instructions provided, all users said that they would for both tools, either without any difficulty, or taking some more time to better understand the process, as would be expected given the advertised ‘intuitive’ localisation process of both tools.

7.3.5 Efficiency

The final area evaluated by the questionnaires was the time behaviour of the localisation tools, as asked in RL19 and RL20, and SL18 and SL19.

The results for both tools are quite similar: when asked whether the time taken for the entire process is acceptable, all 3 users ticked ‘agree’ for both tools, except user 3 for
SYSTRANLinks, who ticked ‘strongly agree,’ therefore giving SYSTRANLinks a slight edge over Reverso Localize.

When asked whether the tools could help localisers save time, users 1 and 2 said that they agreed that both tools could help localisers save time, with user 3 saying ‘disagree’ to both, saying that too much post-editing would be required for the tools to really save localisers time.

It can be seen here that the tools are rated almost exactly the same for time efficiency, perhaps because the task of localising a website seems such a big one that it is astonishing to people that it could be done with the help of an online tool in such little time.

### 7.3.6 General Satisfaction

The final questions were related to general thoughts on the tools: RL22, RL23, RL24; SL20, SL21, SL22. For Reverso Localize, none of the participants agreed with each other when asked how useful they felt the tool was for localisers, with user 1 giving Reverso Localize a score of 4 out of 6, user 2 giving a score of 5 out of 6, and user 3 giving a score of 2 out of 6. Despite this low score from user 2, the average score is 3.6, suggesting it is slightly more useful than not.

When asked if they felt Reverso Localize was harmful towards the localisation industry, the participants were once again split, with the average score being 3 out of 6. Despite this, the users were more than happy with their overall experience with Reverso Localize, awarding it an average score of 4.3, suggesting that the lower scores awarded previously may be more to do with ideology than the Reverso Localize experience.

When asked whether SYSTRANLinks was useful for localisers, users 1 and 2 gave the same responses, but user 3 gave a much higher score of 4 out of 6 as opposed to the 2 out of 6 awarded to Reverso Localize, giving SYSTRANLinks an overall usefulness score 4.3 out of 6.

SYSTRANLinks received the same scores as Reverso Localize when asked if it was harmful towards the localisation industry, once again suggesting that it is more of an ideological response than anything based on the experience with SYSTRANLinks. Finally, SYSTRANLinks received an average overall experience score of 4.3 out of 6, the same score as Reverso Localize, suggesting that the tool performed quite well.

### 7.3.7 Questionnaire Summary

Table 15 below shows the average results from the closed questions analysed above. The interaction, localisation system, localised site, and publishing and linking questions
have been averaged in order to give a score for the functionality characteristic, as they are all part of the subcharacteristics detailed in Table 3 in section 5.3. The understandability and learnability scores have been averaged to give a score for the usability characteristic, with the score for time behaviour being part of the efficiency characteristic. General satisfaction has been given its own category.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>RL average score (/6)</th>
<th>SL average score (/6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functionality</td>
<td>2.9</td>
<td>3.7</td>
</tr>
<tr>
<td>Usability</td>
<td>4.6</td>
<td>3.8</td>
</tr>
<tr>
<td>Efficiency</td>
<td>2.8</td>
<td>2.9</td>
</tr>
<tr>
<td>General satisfaction</td>
<td>3.6</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Table 15: Average score for each characteristic

### 7.4 Machine Translation Analysis

As explained in Section 6, the machine translation is to be evaluated by BLEU, human evaluations, and a quality assurance metric. Table 16 below shows a condensed version of the BLEU and human evaluation results, with Annex X showing the results in full.

The analysis of the performance of the machine translations has been split into three sections. The first, section 7.4.1 will look into some of the differences between the machine translations produced by Reverso Localize and SYSTRANLinks, with section 7.4.2 examining the correlation between the human and BLEU evaluation scores. Section 7.4.3 will carry out a quality assurance evaluation using the SAE J2450 Quality Assurance metric.

#### 7.4.1 Human Evaluation and BLEU Scores

The human evaluations scored the machine translations from 1 to 5 on a scale of accuracy, with 1 being 'not accurate at all' and 5 being 'completely accurate.' The BLEU scores were also based on accuracy, but on a scale of 0 to 1, with 1 being the highest score.
Both the human evaluation and BLEU scores agree that SYSTRANLinks’ machine translation scored higher than Reverso Localize. Three sentences have been studied where SYSTRANLinks performed better than Reverso Localize, in an attempt to discover why the system may have performed better overall. The segments chosen are shown below; they were chosen amongst the translations that scored significantly more for SYSTRANLinks than Reverso Localize, with an added criteria being that they were sentences or bullet points rather than one or two words, to make for better analysis.

<table>
<thead>
<tr>
<th>№</th>
<th>Source (FR)</th>
<th>RL Target</th>
<th>SL Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Au bord du lac Léman et à proximité du centre-ville de Genève, des organisations internationales et du centre de congrès Palexpo, le Lake Geneva Hotel***S est idéalement situé pour vos voyages d’affaires et de loisirs.</td>
<td>By the lake Lake Geneva and near the city center of Geneva, international organizations and the center of congress Palexpo, Lake Geneva Hotel ***S was ideally situated for business and leisure trips.</td>
<td>At the edge of the Lake Geneva and the downtown area of Geneva, international organizations and Palexpo convention center, Lake Geneva Hotel ***S is ideally located for your businesses and leisure trips.</td>
</tr>
<tr>
<td>26</td>
<td>De par son emplacement géographique, au cœur de Versoix et sur les rives du lac Léman, le Lake Geneva Hotel***S vous offre un environnement calme et agréable tout en étant parfaitement desservi par les moyens de transports.</td>
<td>Due to its geographical location, at the heart of Versoix and on the banks of the Lake Léman, Lake Geneva Hotel ***S offers itself a quiet and pleasant environment to you while being perfectly served by the means of transportation.</td>
<td>From its geographical site, in the middle of Versoix and on the Lake Geneva lakeshores, Lake Geneva Hotel ***S offers to you a calm and pleasant environment while being served perfectly by the means of transport.</td>
</tr>
</tbody>
</table>
### Table 17: Sentences to be evaluated

In the Reverso Localize column, some text has been highlighted in red, with the corresponding source and SYSTRANLinks text in bold. This is the text that is interesting from a translation comparison point of view, as it is what could have given Reverso Localize the lower score when compared to SYSTRANLinks. These sections in red are analysed below:

- **lac Léman/ lake Lake Geneva/ Lake Geneva**: it can be seen here, and throughout the website, that ‘lac Léman’ triggers Reverso Localize to produce a translation with two ‘lakes’. The reason for this is somewhat unclear, but what is clear is that the translation is not as accurate as it could be.

- **centre de congrès Palexpo/ center of congress Palexpo/ Palexpo convention centre**: this is a literal translation of the French for Reverso Localize, resulting in a stilted translation which does not flow in the same way that the more natural SYSTRANLinks translation offers.

- **est idéalement situé/ was ideally situated/ is ideally located**: it seems here that the Reverso Localize machine translation system has been programmed to translate ‘est situé’ as a verb that conjugates with ‘être’ in the past tense, therefore producing ‘was’ instead of ‘is,’ and changing the meaning of the text.

- **vous offre/ offers itself... to you/ offers to you**: here the indirect object used in the original French has been poorly translated into a reflexive verb and indirect object. This once again reads poorly, and although the translation offered by SYSTRANLinks is by no means perfect, it is more accurate than the translation offered by Reverso Localize.

- **desservi/ harmed/ served**: it is odd here that Reverso Localize uses the wrong word in this translation, therefore changing the sentence completely and making it much less accurate.

- **à moins de/ unless/ with less than**: as with the example above, this is a poor word choice for the translation of ‘à moins de,’ therefore making the translation

<table>
<thead>
<tr>
<th>No</th>
<th>Source (FR)</th>
<th>RL Target</th>
<th>SL Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>De plus, l’aérop ort international de Genèv e (GVA) se trouve à moins de 10 kilomètres.</td>
<td>Furthermore, the international airport of Geneva (GVA) is unless 10 kilometers.</td>
<td>Moreover, the international airport of Geneva (GVA) is with less than 10 kilometers.</td>
</tr>
</tbody>
</table>
less accurate. Once again, the SYSTRANLinks can be improved, but it is more accurate than the Reverso Localize translation.

In these examples, there were no huge errors made by the Reverso Localize translation system; however, if small mistakes like this were made across the 61 segments, this would explain why it performed slightly worse than SYSTRANLinks overall.

### 7.4.2 Correlation

Generally, as shown by Coughlin (2003), there is a correlation between BLEU and human evaluation scores. Despite this general correlation, there will always be some differences in the main scores between the human and BLEU evaluations.

In this evaluation, the differences have arisen particularly with the average scores given and the differences in the lowest scores awarded. The average human scores for Reverso Localize and SYSTRANLinks were 3.44 and 3.51 out of 5, giving an accuracy score of more than 60%. However, the BLEU scores given were at less than half of the human evaluation accuracy scores, with Reverso Localize receiving a score of 0.2624 out of 1 (26%) and SYSTRANLinks a score of 0.2934 out of 1 (29%). Another difference is in the lowest score awarded to each system, with both receiving a lowest score of 1 from the human evaluators, but the lowest blue score for SYSTRANLinks was 23.6% higher than for Reverso Localize. These differences mean that the question of whether the BLEU and human evaluation scores correlate must be asked.

Table 18 below shows the Pearson correlation between the BLEU and human evaluation scores for each machine translation system, with Annex X showing the correlations on a graph.

<table>
<thead>
<tr>
<th></th>
<th>Reverso Localize</th>
<th>SYSTRANLinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.56</td>
<td>0.54</td>
</tr>
</tbody>
</table>

**Table 18: Human and BLEU Pearson correlation**

The BLEU and human evaluation scores received a moderate correlation for both Reverso Localize and SYSTRANLinks’ machine translation systems. This means that in some cases there was no agreement at all, as can be seen in the graphs in Annex X. The three sentences that have correlated the least have been chosen to be examined as to why they might not correlate: segments 47, 51, and 49.

In these cases the scores given to Reverso Localize and SYSTRANLinks are completely inverted: where the human evaluations scored SYSTRANLinks lower for
segments 47 and 51 and higher for segment 49, the BLEU scores were the opposite. Below is a table with the segments, as well as the reference translation given to Asiya, the system used to calculate the BLEU score.

<table>
<thead>
<tr>
<th>Source</th>
<th>Reference</th>
<th>RL Target</th>
<th>SL Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Au bord du Lac Léman</td>
<td>On the shores of Lake Geneva</td>
<td>By the Lake Lake Geneva</td>
<td>At the edge of Lake Geneva</td>
</tr>
<tr>
<td>Carte gratuite des transports publics pour le Canton de Genève (durée du séjour)</td>
<td>Free pass for all kinds of public transport in the Canton of Geneva for the duration of your stay</td>
<td>Free card of public transports for the Canton of Geneva (duration of the stay)</td>
<td>Free map of public transport for the Canton of Geneva (lasted of the stay)</td>
</tr>
<tr>
<td>A quelques mètres des transports publics et des commerces</td>
<td>Close to public transport connections and local businesses</td>
<td>A few meters away from public transports and businesses</td>
<td>With a few meters of public transport and trade</td>
</tr>
</tbody>
</table>

Table 19: Non-correlated BLEU and human evaluation sentence scores

Here the reason for the lack of correlation seems to be the n-gram system that the BLEU evaluation system uses to score the segments. It can be seen in the last two segments that the human evaluators gave Reverso Localize a higher score in the first segment because of the accuracy of ‘card’ and ‘duration,’ and a higher score in the second one because of ‘a few meters.’ However, the BLEU score penalises these two segments in the n-gram scores, such as with a possible 2-gram score for ‘public transport’ as the machine translation system for Reverso Localize added an ‘s’ to ‘transport’ where the reference translation and SYSTRANLinks translation did not have one.

In the first segment, SYSTRANLinks scored higher in the human evaluation but lower in the BLEU evaluation. It scored higher in the human evaluation because of the use of only one ‘Lake,’ but because BLEU uses an entire reference corpus when scoring the translation to be evaluated, it is possible that Reverso Localize’s more than one ‘lake’ gave it a higher n-gram score, therefore giving it a higher overall score by the BLEU system.

7.4.3 SAE J2450 Quality Assurance

The BLEU and human evaluations give a general overview of the quality of the translations provided by Reverso Localize and SYSTRANLinks, but they do not give
information on the type and severity of the errors made. The SAE J2450 Quality Assurance Model is one of many models that does give information on this, by separating errors made by translation systems into seven categories (see section 3.3.1), with each category having a different weighting according to whether the error is minor or serious.

In this way, the translations by Reverso Localize (Figure 38) and SYSTRANLinks (Figure 39) were assessed, with full results shown in Annex XI and calculations in Annex XII:

![Figure 38: Reverso Localize translation errors](image)

Here it can be seen that the majority of the errors made by Reverso Localize were for either the wrong term or word structure and agreement error. There were many lexical errors made by Reverso Localize, such as translating “Restauration” as “Restoration,” where it should typically be “Restaurant,” “Food and Drink,” or “Catering.” 9 out of the 15 errors made were serious, where the word used in the translation could not in any way be interpreted as an equivalent of the source text, such as translating “Junior suites” as “Junior consequences.”

There were 12 minor and 3 serious errors for word structure and agreement, particularly with using the plural form of words when it should be the singular, or with using a noun when an adverb is needed, such as translating “avec soin et élegance” as “carefully and elegance.”

The next biggest error type made was syntactical, with 9 minor errors being made. These errors were all noted because the word order was not quite right, for example, in the phrase “the hotel welcomes you for a stay any comfort of high quality.” The word order is serious enough to warrant an error, but as it is understandable from a native English language perspective, it was considered a minor error.
There were 6 miscellaneous errors, 2 of which were serious. One of these was related to Reverso Localize translating the language code “DE” as the literal “OF,” with other errors being noted for the use of “lake lake” instead of “lake,” and “Route de Suisse 77” being translated as “Road of Switzerland 77.”

Finally, there was 1 minor punctuation error, with punctuation missing in a list of places where the hotel is nearby.

Two positives must be noticed for Reverso Localize: there were no omission or misspelling errors noted. The weighted scores for each category were added together and divided by the number of words in the source text (see Annex XII), to give an overall document score of 0.334.

Below, Figure 39 shows the results for SYSTRANLinks’ quality assessment. It can be seen how it differs from Reverso Localize.

![Figure 39: SYSTRANLinks translation errors](image)

Here it is clearly the term errors that make up the majority of the errors, with 6 minor and 10 serious errors being recorded, just one more error than Reverso Localize. The term errors made by SYSTRANLinks were quite similar to Reverso Localize, with the serious error of “Point presse” being translated as “Not press.”

Unlike Reverso Localize, syntactic errors was the category where next highest errors were recorded, with 10 minor and 3 serious errors. The serious errors were recorded as such when the target text was near-incomprehensible, for example, with the target sentence “the 103 rooms of the establishment invite you to relieving.”

The third biggest error type was word structure and agreement error, with 8 minor errors and 1 serious error. As with Reverso Localize, the minor errors were due to the the incorrect plural form being used instead of the singular, such as using “lakeshores” instead of “lakeshore.”
Finally, there were 2 serious miscellaneous errors, both of which being noted as SYSTRANLinks did not recognise the language codes “DE” and “EN,” instead translating them as “OF” and “IN.” There was one minor punctuation error made in the same segment as with Reverso Localize, where there was poor punctuation in a list of places.

As with Reverso Localize, SYSTRANLinks did not display any omissions or misspellings. When the weighted scores were added together and divided by the number of words in the source text (see Annex XII), it was given an overall document score of 0.332. As with the BLEU and human evaluations of the translations, this score is marginally higher than Reverso Localize’s.

This is interesting from an analytical point of view, as SYSTRANLinks in fact received a higher weighted score for serious errors, receiving a score of 72 as opposed to Reverso Localize’s 63, as well as receiving higher error marks for the most serious error categories (wrong term and syntactic error). SYSTRANLinks, did, however, receive a lower minor error score, with 49 points to Reverso Localize’s 59.

Despite this, the total of the weighted scores was incredibly close, with Reverso Localize receiving a score of 122 and SYSTRANLinks receiving a score of 121. It must be noted that if only the serious errors were to be taken into account, SYSTRANLinks’ machine translation was not as accurate as Reverso Localize’s, in terms of word choice and style (syntactic choices).
7.5 Summary

It can be seen from Table 20 that two out of the three characteristics’ results correlate. In the functionality characteristic, all three evaluation methods (questionnaire, site analysis, and machine translation) agree that SYSTRANLinks was more effective than Reverso Localize, receiving a higher average score in the questionnaire, as well as a higher BLEU, human evaluation, and quality score, and displaying fewer localisation issues. For the questionnaire, SYSTRANLinks received an average of 3.7 over Reverso Localize’s 2.9, mostly because of its success in the user opinion of its localised site. It displayed fewer localisation issues, too, although it must be pointed out that if the number of localisation errors were to be taken into account, Reverso Localize would have outperformed SYSTRANLinks, as SYSTRANLinks displayed 30 instances of the ‘no underline’ error mentioned. The machine translation evaluation scores were very close, with SYSTRANLinks performing slightly better, mostly due to small errors made on Reverso Localize’s part.

The efficiency characteristic also showed a correlation between the two evaluation methods (questionnaire and video analysis), with SYSTRANLinks taking less time than Reverso Localize. In the video analysis, the times for linking the source and localised websites were fairly similar, and more or less disregarded, as the process was almost exactly the same. However, for creating the project, SYSTRANLinks’ average time was lower than Reverso Localize, and it was more efficient in the time taken to localise the website. These slight differences are reflected in the questionnaire scores, with SYSTRANLinks receiving a score of 2.9 to Reverso Localize’s 2.8, although it is surprising there is not more of a difference between them.

However, there was no correlation between the two evaluation methods for the usability characteristic (questionnaire and video analysis). If total clicks and key touches were to be taken as one entity, using SYSTRANLinks meant that users had to use the keyboard and mouse 30% less than when using Reverso Localize, and had one tab less open throughout the process. However, Reverso Localize performed much better in the questionnaire, with an average score of 4.6 to SYSTRANLinks’ 3.8. This may be an indication that while the information and directions provided by Reverso Localize are superior to those provided by SYSTRANLinks, the SYSTRANLinks interface is more intuitive than Reverso Localize’s.

Table 20 has condensed the results described above to make for slightly easier comparison. It has been set out following the EAGLES criteria described in section 5, with the evaluation methods being split by questionnaire or video analysis, with a section for the machine translation evaluation added.

All scores given (except the localised site analysis and the quality assurance score)
are an average of the three participants’ data, whether that be from the video analysis or the questionnaire scores. The learnability score has been inversed so that it fits in with the other scores, in that the higher score is the better score (see page 82). In the case of the quality assurance score, the lowest score is the better score. Where a questionnaire was used, the average scores given in Table 15 will be written in bold.

Each evaluation method will be highlighted in either green or red, depending on which tool received a higher score: green for Reverso Localize, and red for SYSTRANLinks.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Evaluation</th>
<th>RL Results</th>
<th>SL Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functionality</strong></td>
<td><strong>Questionnaire /6</strong></td>
<td>Interaction: 2.3</td>
<td>Interaction: 3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Localisation system: 2.4</td>
<td>Localisation system: 2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Localised site: 3.7</td>
<td>Localised site: 5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Publishing &amp; linking: 3</td>
<td>Publishing &amp; linking: 3.3</td>
</tr>
<tr>
<td></td>
<td><strong>RL: 2.9; SL: 3.7</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Site analysis</strong></td>
<td>Localisation system:</td>
<td>Localisation system:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.5 problems*</td>
<td>2 problems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 omission</td>
<td>0 omissions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 silences</td>
<td>1 silence</td>
</tr>
<tr>
<td></td>
<td><strong>Machine translation</strong></td>
<td>BLEU: 0.26/1</td>
<td>BLEU: 0.29/1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human: 3.4/5</td>
<td>Human: 3.5/5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality assurance: 0.334</td>
<td>Quality assurance: 0.332</td>
</tr>
<tr>
<td><strong>Usability</strong></td>
<td><strong>Questionnaire /6</strong></td>
<td>Understandability: 4.4</td>
<td>Understandability: 3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learnability: 4.7</td>
<td>Learnability: 3.7</td>
</tr>
<tr>
<td></td>
<td><strong>RL: 4.6; SL: 3.8</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Video analysis</strong></td>
<td>Total clicks: 33 (most for linking)</td>
<td>Total clicks: 26 (+10 for publishing)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total key touches: 233</td>
<td>Total key touches: 154</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total tabs open: 4</td>
<td>Total tabs open: 3</td>
</tr>
<tr>
<td><strong>Efficiency</strong></td>
<td><strong>Questionnaire /6</strong></td>
<td>Time behaviour: 2.8</td>
<td>Time behaviour: 2.9</td>
</tr>
<tr>
<td></td>
<td><strong>RL: 2.8; SL: 2.9</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Video analysis</strong></td>
<td>System localisation time:</td>
<td>System localisation time:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>02:29</td>
<td>01:11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human localisation time:</td>
<td>Human localisation time:</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>project creation:</strong> 05:05</td>
<td><strong>project creation:</strong> 02:14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>linking: 12:34</td>
<td>linking: 08:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(minus reading the blog: 09:19)</td>
<td></td>
</tr>
</tbody>
</table>

*RL has 2.5 problems instead of 3 due to the ‘nuance’ of the ampersand ‘error’

**Table 20: Results in brief**
Conclusion

The objective of this thesis was to compare the two online localisation tools Reverso Localize and SYSTRANLinks, in order to see to what extent these free localisation tools satisfy end users, to give recommendations, and to ascertain which is more efficient and effective for end users. In order to do this, this thesis first of all described the concepts of localisation and machine translation, in order to give a theoretical background to the practical evaluations carried out. How the two systems work was then described, with details of their similarities and differences. Afterwards, a methodology was set out using the EAGLES 7-step method, which helped define the characteristics to be evaluated for each localisation tool. Still using the EAGLES method, the design of the evaluation was set, before carrying out the pilot test and evaluations using various methods, such as video analysis and questionnaires. The results from these tests were then analysed, before being summarised.

Many conclusions can be drawn about these two localisation tools. On a grand scale, both tools successfully completed the task assigned to them: all three of the non-professional participants managed to localise the Lake Geneva Hotel website, both within less than 45 minutes. However, these localisations were done with a certain level of success, with some functions performing better than others, and one tool performing better than the other in some cases.

This conclusion will base itself on the EAGLES characteristics: functionality, usability, and efficiency. For functionality, SYSTRANLinks scored higher in every category, particularly for the scores relating to its localised site and localisation system.

It must also be pointed out, however, that while SYSTRANLinks was rated higher than Reverso Localize, this was more due to mistakes made by Reverso Localize than a superior system by SYSTRANLinks. A suggestion to Reverso Localize would therefore be to simplify its notification system, or ensure that notifications are received in a more timely manner.

Despite SYSTRANLinks receiving higher accuracy scores in the machine translation evaluation, the scores were still close enough to be almost identical, from all three of the evaluations carried out, leading to the conclusion that while the localisation system and localised site were more appreciated and worked better during the evaluations, the written content was not necessarily translated to a higher quality by the SYSTRANLinks machine translation system.

Under the usability characteristic there were mixed results. Reverso Localize scored much higher than SYSTRANLinks for learnability and understandability, most likely
because of the quality of information provided by Reverso Localize on how the site and tool functions.

With the video analysis for the usability characteristic, if publishing the websites are to be taken into account, Reverso Localize was more effective, although it must be pointed out that publishing a localised website with Reverso Localize is less secure. However, for the number of key touches, SYSTRANLinks displayed less keys being pressed, mostly because less information was needed during the project creation stage. Finally, SYSTRANLinks was also more usable with regards to the number of tabs open, as its notification system did not include emails, and therefore it was necessary to have the user email account open in a different tab.

Finally, under the efficiency category, SYSTRANLinks was more effective across the board. The participants scored it slightly higher on the time behaviour scale, which was mostly affected by user 2’s slightly negative experience with Reverso Localize, although despite this two out of the three participants did agree that both tools would help save time when localising a website.

From the video analysis, SYSTRANLinks took much less time to localise the website than Reverso Localize, with Reverso Localize being even less efficient when considering that it did not load during the localisation process for one of the participants. Project creation was also longer for Reverso Localize, as more information was wanted by the system during this stage, although the times for linking the source and localised website were similar for both localisation tools. However, the participants found reading the Reverso Localize blog useless, for reasons such as poor language and unrelatability to the particular scenario used for the evaluation. Despite this, the efficiency of both tools could be argued to be high, as the entire process for both tools took less than 45 minutes, which, as mentioned above, all participants agreed was acceptable.

If a suggestion were to be made concerning which tool would should be recommended for use, it would be SYSTRANLinks which should be suggested, on the basis of the results shown, as it was preferred in the majority of the characteristics evaluated. However, both tools could be improved: in particular, they both displayed some errors when localising the website, which for a non-professional, could be an issue.

Both tools also need to improve how the source and localised websites are linked together: Reverso Localize, by giving more understandable instructions, and SYSTRANLinks, by providing instructions if a user decides not to use the language selector display. In general as well, SYSTRANLinks needs to provide more instructions and information, as it was the one area (understandability and learnability) where it performed unfavourably in contrast to Reverso Localize.

Despite these complaints, the fact that both tools are able to localise a website to
the extent that they did is highly impressive, as it gives more control to the client when localising their website. A particularly good factor to both the tools is the time taken to localise the website; a project which could take weeks is condensed into a matter of hours. However, the machine translation systems do need to be improved, to make the job easier for those who do not know their target language and to make the process truly autonomous.

**Limits and Further Work**

There were several limits to this thesis. Firstly, the number of participants was minimal: as this is a master’s thesis, it was not possible to get a large number of people to try the two tools. Only one domain was tested, and only one language pair was selected, but this was all that could done in the time frame set. The webpage chosen to evaluate the machine translation systems was also relatively small, at less than one tenth of the overall website word count, which could have had an adverse effect on the results.

Another limit is the tools themselves: a cursory web search shows that a lot of online localisation tools mostly specialise in software localisation, with little thought to web localisation. It is therefore very difficult at the moment to truly evaluate these tools, as this thesis has had to take its evaluation methods from a variety of sources for different types of tool, as both Reverso Localize and SYSTRANLinks use many types of tool in the website localisation process.

It would be particularly interesting in the future to attempt to creating a model which specifically evaluates tools such as these ones, where all functions are incorporated. These would have been particularly useful in the results and analysis section of this thesis, where there were many struggles in bringing the results together for analysis. As can be seen, it was decided that all scores and all characteristics had an equal weighting; in an evaluation model specifically assigned to free and online localisation tools, it would be a good idea to assign weightings depending on how important the category is to the general functioning of the localisation tool.

What is interesting about this evaluation is that the section focussing on Reverso Localize can be compared to an evaluation done by Cristina Peron a year ago, in order to see what changes have occurred in that time - for example, the spell check has more or less disappeared, and languages have been added or removed. It would be interesting to see in the future how much more Reverso Localize has changed, and if SYSTRANLinks will, too.
Bibliography


98


Torres del Rey, J. and Rodríguez V. de Aldana, E. (2014), Dynamic Web Localisation, in Localisation et gestion de projet, not published.


Annex I: Participant Information

INFORMATION FOR PARTICIPANTS

Comparative evaluation of localisation tools: Reverso Localize and SYSTRANLinks

OBJECTIVES OF THE EVALUATION
This evaluation, as part of a master’s thesis, is being performed to assess and compare various functions of the online localisation platforms Reverso Localize and SYSTRANLinks, in order to evaluate their impact on localisers. You are either being asked to carry out a series of localisation tasks with these two tools and then evaluate their performance, or to compare machine translations of web page text by these two tools.

LENGTH OF EVALUATION
The evaluation should take no longer than an hour and a half in the case of carrying out the localisation tasks, and no more than an hour in the case of evaluating the machine translations.

RESULTS
The results of this work should be available by the end of August 2014.

BENEFITS AND RISKS
This evaluation will lead you to become better acquainted with two very large machine translation companies, and in particular their localisation platforms and their raw machine translation output. There are no known risks to taking part in this evaluation. If at any time you are unhappy with the process, you may leave.

DATA PROTECTION
You will be asked for some information, such as your name, your age, and level of education which will then be linked to a participant number throughout the evaluation. This information will not be given to any third party, and any results coming from this evaluation will be anonymous in the thesis.

CONTACT DETAILS
For any other information, contact Charlotte Gray using the email address charlotte.gray@etu.unige.ch
Annex II: Consent Form

CONSENT FORM

EVALUATION TITLE
Comparative evaluation of localisation tools: Reverso Localize and SYSTRANLinks

PERSON CARRYING OUT RESEARCH
Charlotte Gray, Masters in Translation in the Department of Multilingual Information Processing

RESPONSIBLE FACULTY MEMBER
Pierrette Bouillon

DECLARATION
I agree to the following and am ready to take part in this study:

- I have read the information provided and understand the study’s objectives
- I have asked for clarification on matters that concern me and have received satisfactory answers from the person carrying out the research
- I understand the conditions of participation
- I understand that the results of the study carried out today may be used by the person carrying out the research, that my participation is voluntary and I may leave at any time

PARTICIPANT NAME

PARTICIPANT SIGNATURE

RESEARCHER SIGNATURE

DATE
Annex III: Participant Questionnaire

User no:

PARTICIPANT QUESTIONS

Please fill in the form below. Questions with (...) next to them indicate that you can choose more than one response.

SECTION 1 - PERSONAL INFORMATION

Name: ________________________________
Surname: ________________________________

Gender

□ Male
□ Female

Age

________________________

Native language (...)

□ English
□ French
□ German
□ Spanish
□ Italian
□ Chinese
□ Arabic
□ Other ______________________

1Inspired by questionnaire in Peron (2013)
SECTION 2 - EDUCATION

Highest level of education completed

☐ None
☐ Qualifications aged 16 (GCSEs)
☐ Qualifications aged 17 (AS levels)
☐ Qualifications aged 18 (High school diploma, A-levels, IB)
☐ Apprenticeship/NVQ
☐ Bachelor’s degree
☐ Master’s degree
☐ Doctorate
☐ Other ________________

Current education

☐ None
☐ Qualifications aged 16 (GCSEs)
☐ Qualifications aged 17 (AS levels)
☐ Qualifications aged 18 (High school diploma, A-levels, IB)
☐ Apprenticeship/NVQ
☐ Bachelor’s degree
☐ Master’s degree
☐ Doctorate
☐ Other ________________

Knowledge of French language

☐ None
☐ Less than 2 years’ study
☐ 2-5 years’ study
☐ 5-10 years’ study
☐ More than 10 years’ study
☐ Bilingual with native language
☐ Native language
☐ Other ________________
SECTION 3 - COMPUTER EXPERIENCE

1. Do you own a computer or a similar device?
   □ Yes
   □ No

2. If yes, which type? (...)
   □ Desktop
   □ Laptop
   □ Tablet
   □ Other ______________________

3. How often do you use a computer device?
   □ Hardly ever
   □ Once a month
   □ Once every couple of weeks
   □ Once a week
   □ 2-4 days a week
   □ 5-6 days a week
   □ Every day

4. How many hours a day do you use a computer device?
   □ Less than an hour a day
   □ 1 hour a day
   □ 2 hours a day
   □ 3 hours a day
   □ 4 hours a day
   □ More than 4 hours a day

5. For what reason do you use a computer device? (...)
   □ Professional work
   □ Education
   □ Leisure (shopping, communication, gaming)
   □ Other ______________________
6. Which operating system(s) do you know? (...)

☐ Windows
☐ Mac
☐ Linux
☐ Other _____________________

7. Which operating system(s) do you use on a daily basis? (...)

☐ Windows
☐ Mac
☐ Linux
☐ Other _____________________

SECTION 4 - WEBSITE KNOWLEDGE

8. What knowledge do you have of HTML?

☐ Expert
☐ Proficient
☐ Competent
☐ Advanced beginner
☐ Novice

9. Have you already worked with source codes of webpages?

☐ Yes
☐ No

10. Have you worked with any content management systems (CMS)?

☐ Yes
☐ No

11. If yes, which one(s)? (...)

☐ WordPress
☐ Drupal
☐ Joomla!
☐ TYPO3
☐ Other _____________________
12. If you were to have a website, would you think of using a localisation platform to make it multilingual?

☐ Yes
☐ No
Annex IV: Reverso Localize Instructions\[^2\]  

**REVERSO LOCALIZE INSTRUCTIONS**

These instructions are part of the evaluation process of two online localisation platforms, Reverso Localize and SYSTRANLinks, which you will be using to localise a hotel website. The instructions you are reading now are for Reverso Localize. Read the instructions below first before starting. Don’t hesitate to ask if you have any problems.

What you do will be recorded with BB Flashback, a piece of software which records what happens on the screen and extracts data which can be analysed later. No personal information will be recorded.

These instructions are also available on the desktop (ReversoInstructions.pdf) if you prefer to read them on the screen. Please make sure you read each instruction completely before following it.

**METHOD**

1. Open Google Chrome, go to www.gmail.com, and sign in with the email address and password given to you. Make sure this stays open throughout the evaluation. Afterwards, open a new tab and go to the following website:

   http://www.lakegenevahotel.ch

   *This is the website that has been chosen to be localised. As you can see, it is a website for a hotel on Lake Geneva, which is in French, English and German. Take a couple of minutes to have a look around – get yourself used to the way it works and looks.*

2. Open a new tab in Google Chrome and go to the following website. Make sure that the Lake Geneva Hotel website is still open.


   *This is the website for Reverso Localize, a free online localisation platform, and one of the pieces of software which will be evaluated.*

   *Make sure that you stay on the English language version throughout the entire evaluation. Have a look around the website to get an idea of what it does, or watch the information video, which you can get to from the homepage.*

\[^2\]Inspired by questionnaire in Peron (2013)
3 At the top of the page, click on Login. Enter the email address and password that was given to you, and keep the Keep me logged in box ticked. Click on the Login button.

You will now be in your personal Reverso Localize account area.

4 Click on the Create a new job button, and click on Mirror Websites in the dropdown list.

5 Under Project Settings, put in the URL of the Lake Geneva Hotel website (http://www.lakegenevahotel.ch) in the Website address field.

Leave Description of the website blank.

A name will have appeared automatically under Project Name. Delete it and write ‘evaluationx,” where “x” is the user number given to you (the same number as in the Gmail account created for you). Set the Domain as “General.”

6 Under Select Languages, select the Source Language as French and the Target Language as English.

Click on the plus sign next to Show advanced options, and make sure that both Use translation memory if available and Use dictionary if available are selected.
7 Click on Create. A message will pop up telling you that the job is being processed. Read it, make a note of what it says, and then close it.

8 Wait until the system lets you know that the initial translation has been performed automatically. It should let you know in three ways: through a pop-up screen, through an email notification, and through the information icon on the top right-hand corner, next to Logout. Check each of these three—did you get the notifications as expected?

*Make a note of how long you think you had to wait. Did it take as long as you thought it would? Do you feel that the software has been notifying you of the process as it should be?*

9 Make sure that you are in the Dashboard>Job Details section.

In the Revision section, click on the Revise button. The localised version of the website should be opened in a new tab.

*This website is what Reverso Localize calls a Mirror site. It is a copy of the original website, with the same structure and interface. The difference is that the French pages have been translated into English, and you will not be able to access the German and original English pages, as they were ignored by the software during the localisation process.*

*Here it is possible to post-edit the pages, which you do not need to do.*

10 Have a look at the localised website, and compare it with the original French website, which you should have open in another tab. If not, open a new tab and put in the website's address (http://www.lakegenevahotel.ch).

*Click on the links at the top of the website and make sure that they all go to where they should. Make a note if you see any localisation problems (for example, pictures missing, links not working, or text missing or not translated).*

*If clicking on a word opens the Post-edit text window, close it without changing anything.*

*What do you think of the translation? What kind of general improvements would you suggest?*

*As this study is only evaluating the localisation process, you will now continue as if you had post-edited the translation.*

11 When you have finished, go back to Reverso Localize and Dashboard>Job Details.

12 Go down to the Publication section and click on Publish.
A window called *Publication of the translated website* will open. Make sure that you check *Right to publish* (Mandatory), and leave the other options blank.

**Right to publish** is Reverso Localize’s way of confirming that the person using the website for localisation purposes has the administrative rights to do so. What do you think of this method of confirming ownership?

13 Click on *Publish*. Make a note of how long it takes for the publication to finish.

14 Once Reverso Localize has finished publishing the website, you will see in the *Publication* section that the status is set to *Online*. It is now time to link the English version of the website to the French original. At the bottom of the *Publication* section, there is a link saying *See how to have your visitors switch from one language to another!* Click on it and it will direct you to a Reverso Localize blog post.

Read the blog article that the link directs you to. What do you think of it? If you followed the instructions written there, do you think you would be able to link the original website and the localised website together?

15 Go back to the original Lake Geneva website (http://www.lakegenevahotel.ch), making sure that you are on the homepage (i.e. so that only “www.lakegenevahotel.ch” is visible in the address bar).

You are now going to go through the steps that will link the original website with the localised website. You will not be following the steps given in the blog post above as they do not quite fit the needs of localising this particular website.

Download the source file of the website. To do this, right click anywhere on the page, then go to *Save As*, and choose the Desktop as the location to save the file. Make sure that the type of file is set as *Webpage, Complete*. Keep the file name the same as the one automatically entered in the *File name* box. Click on *Save.*
16 Now go to Desktop>Notepad++. Click on File>Open>Desktop and select the document you just saved, which is called “Accueil Hôtel Lake Geneva.” Make sure you select the Google Chrome icon, not the folder with the same name. Click Open. The source code of the hotel’s website homepage should open.

17 What you will be doing now is replacing the official English version of the website with the Reverso Localize version. Using the search function (Ctrl+F), search “EN”, making sure you check the Match whole word only, Match case, and Wrap around options.
You should find “EN” in the following text:

```
<a href="http://www.lakegenevahotel.ch/en/">EN</a></li>
```

Replace the `http://www.lakegenevahotel.ch/en/` link between the inverted commas with the link given to you by Reverso Localize in the *Publication* section of Reverso Localize (next to the United Kingdom flag). The final text should look like this:

```
<a href="ReversoLocalizeLink">EN</a></li>
```

18 Save the changes by using Ctrl+S or going to *File*>*Save*, the close Notepad++ and go back to the desktop.

19 Open the same file “Accueil Hôtel Lake Geneva.htm” (the one with the Google Chrome icon). This time it should open using Google Chrome. If everything has worked correctly, you should be able to see Reverso Localize’s English version of the website instead of the original website.

*Have a look at the website as it is. What do you think? Would you change or improve anything?*

*The evaluation is now finished. Please fill in the questionnaire that has been given to you.*
Annex V: SYSTRANLinks Instructions

SYSTRANLinks INSTRUCTIONS

These instructions are part of the evaluation process of two online localisation platforms, Reverso Localize and SYSTRANLinks, which you will be using to localise a hotel website. The instructions you are reading now are for SYSTRANLinks. Read the instructions below first before starting. Don’t hesitate to ask if you have any problems.

What you do on the screen will be recorded with BB Flashback, a piece of software which records what happens on the screen and extracts data which can be analysed later. No personal information will be recorded.

These instructions are also available on the desktop (SystranInstructions.pdf) if you prefer to read them on the screen. Please make sure you read each instruction completely before following it.

METHOD

1 Open Google Chrome and go to the following website:

http://www.lakegenevahotel.ch

This is the website that has been chosen to be localised. As you can see, it is a website for a hotel on Lake Geneva, which is in French, English and German. Take a couple of minutes to have a look around – get yourself used to the way it works and looks.

2 Open a new tab in Google Chrome and go to the following website. Make sure that the Lake Geneva Hotel website is still open.

http://systranlinks.com

This is the website for SYSTRANLinks, a free online localisation platform, and one of the pieces of software which will be evaluated.

Make sure that you stay on the English language version throughout the entire evaluation. Have a look around the website to get an idea of what it does.

3 At the top of the page, click on Login. Enter the email address and password that was given to you, and make sure that the Stay connected box is ticked. Click on Sign in.

You will now be in your personal SYSTRANLinks account area.
4 Click on the Add a website box. A window called Add a website will appear.

In the first section, Your Website, enter the link for the Lake Geneva Hotel (http://www.lakegenevahotel.ch) in the URL field, and click on Check. The URL field should turn green.

Make sure that the Source Language field, which has been filled in automatically, has been changed correctly to French. Click on Next.

5 The next step is called Translation Settings. Set the target language as English, and make sure that the Default content for translated websites is set as Machine Translation, and that Community Feedback is set as None. Click on Next.
6 The final step is Display Mode. Under Language selector theme, choose Grey.

7 Click on Create. A message will come up telling you that the job is being processed, and should then be replaced by a window saying that the website has been localised. Once you have read the message, click on Browse & Edit your localized website. 

Make a note of how long you think you had to wait. Did it take as long as you thought it would? Do you feel that the software has been notifying you of the process as it should be?

8 Another tab should open in Google Chrome with the localised website.

This is a copy of the original website, with the same structure and interface. The difference is that the French pages have been translated into English, and there is a language selector in the top right hand corner of the website.

Here it is possible to post-edit the pages, which you do not need to do.

IMPORTANT! Make sure that you keep the localised website open, as you will be needing its address later for linking it to the original website. The address should be “sledit.en.www.lakegenevahotel.ch.systranlinks.net”.
9 Have a look at the localised website, and compare it with the original French website, which you should have open in another tab. If not, open a new tab and put in the website’s address (http://www.lakegenevahotel.ch).

Click on the links at the top of the website and make sure that they all go to where they should. Make a note if you see any localisation problems (for example, pictures missing, links not working, or text missing or not translated).

If clicking on a word opens an editing window, close it without changing anything.

What do you think of the translation? What kind of general improvements would you suggest?

10 SYSTRANLinks also allows users to look through the website from a dashboard, as shown below.

To get to the dashboard, click on the tab called SYSTRANLinks - Pages.

Have a look through the dashboard, especially at the Content and Settings section. Do you think you would be able to edit the localised website through the dashboard? Which interface would you prefer – the dashboard or the localised website?

I would like to draw your attention to the Verification Status box in the Settings > Configuration section. Click on Verify and have a look at the two methods that SYSTRANLinks uses as a way of confirming that the person using
the website for localisation purposes has the administrative rights to do so. What do you think of this method of confirming ownership?

As this study is only evaluating the localisation process, you will now continue as if you had post-edited the translation.

11 Click back on the tab for the localised website. You will notice that choosing to use a language selector display has led to the German part of the website becoming invisible. You will need to fix this by removing the language selector display and then linking the source and target websites manually.

To remove the language selector display, go back to the dashboard, and go to Display Mode. Options to configure the language selector display will be there. Click on the blue button near the top which says Grey, and change it to None. Click on Apply at the bottom of the screen.

If you now refresh the localised website, you will notice that the German part of the website is available again.

12 Go back to the original Lake Geneva website (http://www.lakegenevahotel.ch), making sure that you are on the homepage (i.e. so that only “www.lakegenevahotel.ch” is visible in the address bar).

You are now going to go through the steps that will link the original website with the localised website.

Download the source file of the website. To do this, right click anywhere on the page, then go to Save As, and choose the Desktop as the location to save the file.
Make sure that the type of file is set as *Webpage, Complete*. Keep the file name the same as the one automatically entered in the *File name* box. Click on *Save*.

13 Now go to *Desktop>Notepad++*. Click on *File>Open>Desktop* and select the document you just saved, which is called “Accueil Hôtel Lake Geneva.” Make sure you select the Google Chrome icon, not the folder with the same name. Click *Open*. The source code of the hotel’s website homepage should open.

14 What you will be doing now is replacing the official English version of the website with the SYSTRANLinks version. Using the search function (Ctrl+F), search “EN”, making sure you check the *Match whole word only*, *Match case*, and *Wrap around* options.
You should find “EN” in the following text:

\[<a href="http://www.lakegenevahotel.ch/en/">EN</a></li>\]

Replace the \[http://www.lakegenevahotel.ch/en/\] link between the inverted commas with the localised website link given to you by SYSTRANLinks. You should find this link in the address bar of the localised website. The final text should look like this:

\[<a href="SYSTRANLinks link">EN</a></li>\]

15 Save the changes using Ctrl+S or going to File>Save, close Notepad++ and go back to the desktop.

16 Open the file “Accueil Hôtel Lake Geneva.htm” (the one with the Google Chrome icon). This time it should open using Google Chrome. If everything has worked correctly, you should be able to see SYSTRANLinks’ English version of the website instead of the original website.

Have a look at the website as it is. What do you think? Would you change or improve anything?

The evaluation is now finished. Please fill in the questionnaire that has been given to you.
Annex VI: Machine Translation Evaluation Instructions

MACHINE TRANSLATION EVALUATION INSTRUCTIONS

These instructions are part of the evaluation process of two online localisation platforms, Reverso Localize and SYSTRANLinks, which are being used to localise a hotel website. The instructions you are reading now are to give you guidance for evaluating Reverso Localize’s and SYSTRANLinks’ raw machine translation output of a webpage from the hotel website.

The evaluation data collected in this process will be used later for analysis; however, no personal information will be recorded or published.

Please read the instructions first before starting. They are also available on the desktop (MTInstructions.pdf) if you prefer to read them on the screen. Don’t hesitate to ask if you have any problems.

OBJECTIVE

A webpage has been translated by the machine translation systems from Reverso Localize and SYSTRANLinks.

You are being asked to rate each translated segment according to its accuracy.

For this evaluation, accuracy is defined as to what extent the meaning of the source language segment is kept in the target language segment, depending on the context of the source webpage.

You will be scoring the accuracy of each translated segment on a scale of 1 to 5, where 1 is not at all accurate and 5 is completely accurate.

METHOD

1 Before starting the evaluation, it is strongly advised that you look at the source text for the translations, in order to put the source text in context. It can be found at the address below:

   http://www.lakegenevalhotel.ch/accueil/

2 Next, go to the desktop and open the Microsoft Excel 2010 file called MT Evaluation.xlsx. The spreadsheet should look like the image below.
Note the colours on the image above. The text outlined in purple is the source text, and the yellow is the target text, with each cell being a translation segment.

The area outlined in green is where you will be scoring each segment.

At the bottom of the image you will see two areas circled in red and blue. The worksheet in red is where the translations and evaluation area for system one are stored, and the worksheet in blue is where the translations and evaluation area for system two are stored. **Make sure you complete both worksheets.**

You will not be told which system is Reverso Localize or which system is SYSTRANLinks.

3 As mentioned in the objectives above, you are being asked to evaluate each translated segment from the point of view of **accuracy**, on a scale of 1 to 5, where 1 is **not at all accurate** and 5 is **completely accurate**.

For each translation segment, put a 1 under the score that you judge to be best, as shown below. Make sure that you only put one score for each translation segment.

4 Once completed, go to File>Save As, then save the file under the format MTEvaluationUserx, where x is the user number given to you at the beginning of the evaluation.

The evaluation is now finished.
Annex VII: Reverso Localize Questionnaire

User n°:

QUESTIONNAIRE - REVERSO LOCALIZE

Please answer the questions below concerning the test carried out on Reverso Localize. There are no compulsory questions.

SECTION 1 - INTERFACE AND USABILITY

1 On a scale of 1 to 6, with 6 being the best, rate each of these elements concerning Reverso Localize’s interface.

<table>
<thead>
<tr>
<th>Structure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navigability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simplicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2 On a scale of 1 to 6, with 6 being very clear, how clear was the information provided by Reverso Localize on how the site works (video, text, information windows)?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 On a scale of 1 to 6, with 6 being the most, how much effort went into learning to use Reverso Localize?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 On a scale of 1 to 6, with 6 being very clear, how clear were the directions provided by Reverso Localize throughout the process?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inspired by questionnaire in Peron (2013)
5 Do you think you would have been able to use Reverso Localize without the instructions provided by the creator of the test?

☐ Yes
☐ Yes - but it would take more time
☐ Yes - but I would do some things differently
☐ No

If you wish, use this space to add more details.

________________________________________

________________________________________

________________________________________

____________

SECTION 2 - FUNCTIONS

Interaction

6 Did the notification system (icon at the top of the dashboard, email notifications, pop up window) work as intended?

☐ Yes - I received all notifications when I should have
☐ Yes - I received all notifications, but a while after I should have been notified
☐ No - I only received notifications via one or two methods
☐ No - I received no notifications

7 Do you agree that the notification system used by Reverso Localize is useful?

☐ Strongly agree
☐ Agree
☐ Disagree
☐ Strongly disagree

8 Tick which of the three types of notification you feel are necessary and which you would like to see throughout the localisation process (you can tick more than one).

<table>
<thead>
<tr>
<th>Email notification</th>
<th>Pop up window</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

124
9 Is there anything you would change about the notification system?

Localisation

10 Do you agree that the localisation system worked well?
□ Strongly agree
□ Agree
□ Disagree
□ Strongly disagree

11 How many localisation errors did you notice?
□ Many
□ Quite a lot
□ Some
□ Not very many
□ None

12 How much extra localisation do you feel the site would need to be in a publishable state?
□ A lot
□ Not much
□ None

Mirror site

13 On a scale of 1 to 6, where 6 is very useful, did you think the mirror site was useful?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
14 What do you think of the banner advertising Reverso Localize at the bottom of each page of the localised website?

Publishing

15 On a scale of 1 to 6, with 6 being very difficult, how difficult did you find publishing the website?

16 Do you think you would have been able to link the original website and localised website together with just the instructions from the Reverso Localize blog?

☐ Yes
☐ Yes - but it would take more time
☐ Yes - but I would do some things differently
☐ No

17 What do you think of the blog post?

18 As you have seen, linking the original website and the localised website involves going to the source code of the website and changing some HTML. What do you think of this method?
You will have noticed during the process that before publishing you must confirm that you have the right to publish a localised version of the website. What do you think of this method of confirming ownership?

SECTION 3 - TIME

Do you agree that the time taken for the entire process is acceptable?

- Strongly agree
- Agree
- Disagree
- Strongly disagree

Do you agree that this tool can help localisers save time?

- Strongly agree
- Agree
- Disagree
- Strongly disagree

SECTION 4 - FINAL THOUGHTS

On a scale of 1 to 6, with 6 being very useful, do you think that this tool is useful for localisers?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you think that this tool is useful or harmful to the localisation industry? Tick a place on the scale.

Useful

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>

Harmful

|   |   |   |   |
Why?

______________________________

______________________________

______________________________

______________________________

24 On a scale of 1 to 6, with 6 being excellent, rate your overall experience with Reverso Localize.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Feel free to add any other comments here.

______________________________

______________________________

______________________________

______________________________

Thank you for completing this questionnaire
Annex VIII: SYSTRANLinks Questionnaire

User n°:

QUESTIONNAIRE - SYSTRANLinks

Please answer the questions below concerning the test carried out on SYSTRANLinks. There are no compulsory questions.

SECTION 1 - INTERFACE AND USABILITY

1 On a scale of 1 to 6, with 6 being the best, rate each of these elements concerning SYSTRANLinks’ interface.

<table>
<thead>
<tr>
<th>Structure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Navigability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simplicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2 On a scale of 1 to 6, with 6 being very clear, how clear was the information provided by SYSTRANLinks on how the site works (text, information windows)?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

3 On a scale of 1 to 6, with 6 being the most, how much effort went into learning to use SYSTRANLinks?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

4 On a scale of 1 to 6, with 6 being very clear, how clear were the directions provided by SYSTRANLinks throughout the process?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

*Inspired by questionnaire in Peron (2013)*
5 Do you think you would have been able to use SYSTRANLinks without the instructions provided by the creator of the test?

- Yes
- Yes - but it would take more time
- Yes - but I would do some things differently
- No

If you wish, use this space to add more details.

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

SECTION 2 - FUNCTIONS

Interaction

6 Do you agree that SYSTRANLinks notified you of what was happening during the localisation process?

- Strongly agree
- Agree
- Disagree
- Strongly disagree

7 Is there anything you would change about or add to the notification system?

__________________________________________________________________________

__________________________________________________________________________

__________________________________________________________________________

Localisation

8 Do you agree that the localisation system worked well?

- Strongly agree
- Agree
- Disagree
- Strongly disagree
9 How many localisation errors did you notice?
   - Many
   - Quite a lot
   - Some
   - Not very many
   - None

10 How much extra localisation do you feel the site would need to be in a publishable state?
   - A lot
   - Not much
   - None

Localised site

11 On a scale of 1 to 6, with 6 being very useful, did you think being able to see the localised site was useful?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12 On a scale of 1 to 6, with 6 being definitely necessary, how necessary do you feel it is to have two options for editing the localised site (editing via the localised site, or editing via the dashboard)?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

13 Which would you prefer: editing by using the dashboard or the localised site?
   - Dashboard
   - Localised site
   - Both

   Why?

   ______________________________________________________
   ______________________________________________________
   ______________________________________________________
   ______________________________________________________
   ______________________________________________________
   ______________________________________________________
14 What do you think of the SYSTRANLinks logo on the top of each page of the localised website?

Linking the website

15 On a scale of 1 to 6, with 6 being very useful, how useful do you think SYSTRANLinks' language selector display is?

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

If you wish, use this space to add more details.

16 As you have seen in SYSTRANLinks, it is impossible to use the language selector display if the website is already multilingual, and that linking the pages through the source code is necessary. Do you think you would have been able to do this without the instructions given to you?

- Yes
- Yes - but it would take more time
- Yes - but I would do some things differently
- No

If you wish, use this space to add more details.
17 You will have noticed during the process that before the localised website goes live you must confirm that you have administration rights to the original website, and therefore have the right to publish a localised version. There are two methods to do this: the first, by uploading a verification file to the original website’s homepage, and the second, by inserting a meta tag into the website’s homepage. What do you think of this method of confirming ownership?

SECTION 3 - TIME

18 Do you agree that the time taken for the entire process is acceptable?

□ Strongly agree
□ Agree
□ Disagree
□ Strongly disagree

19 Do you agree that this tool can help localisers save time?

□ Strongly agree
□ Agree
□ Disagree
□ Strongly disagree

SECTION 4 - FINAL THOUGHTS

20 On a scale of 1 to 6, with 6 being very useful, do you think that this tool is useful for localisers?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

21 Do you think that this tool is useful or harmful to the localisation industry? Tick a place on the scale.

<table>
<thead>
<tr>
<th>Useful</th>
<th>Harmful</th>
</tr>
</thead>
</table>
Why?

On a scale of 1 to 6, with 6 being excellent, rate your overall experience with SYSTRANLinks.

22 On a scale of 1 to 6, with 6 being excellent, rate your overall experience with SYSTRANLinks.

Feel free to add any other comments here.

Thank you for completing this questionnaire.
### Annex IX: Translation Segments

<table>
<thead>
<tr>
<th>No</th>
<th>Source Text (FR)</th>
<th>RL Target Text (EN)</th>
<th>SL Target Text (EN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACCUEIL</td>
<td>WELCOME</td>
<td>HOME</td>
</tr>
<tr>
<td>2</td>
<td>HÉBERGEMENT</td>
<td>ACCOMMODATION</td>
<td>ACCOMMODATION</td>
</tr>
<tr>
<td>3</td>
<td>Classiques</td>
<td>Classics</td>
<td>Classics</td>
</tr>
<tr>
<td>4</td>
<td>Supérieures</td>
<td>Superiors</td>
<td>Higher</td>
</tr>
<tr>
<td>5</td>
<td>Deluxes</td>
<td>Deluxes</td>
<td>Deluxes</td>
</tr>
<tr>
<td>6</td>
<td>Junior suites</td>
<td>Junior consequences</td>
<td>Junior continuations</td>
</tr>
<tr>
<td>7</td>
<td>RESTAURATION</td>
<td>THE RESTORATION</td>
<td>RESTORATION</td>
</tr>
<tr>
<td>8</td>
<td>Restaurant</td>
<td>Restaurant</td>
<td>Restaurant</td>
</tr>
<tr>
<td>9</td>
<td>Lounge-bar</td>
<td>Lounge bar-bar</td>
<td>Lounge-bar</td>
</tr>
<tr>
<td>10</td>
<td>ACTIVITÉS</td>
<td>ACTIVITIES</td>
<td>ACTIVITIES</td>
</tr>
<tr>
<td>11</td>
<td>INFORMATIONS</td>
<td>INFORMATION</td>
<td>INFORMATION</td>
</tr>
<tr>
<td>12</td>
<td>Contact &amp; plan d’accès</td>
<td>Contact and area map</td>
<td>Contact &amp; access map</td>
</tr>
<tr>
<td>13</td>
<td>Point presse</td>
<td>Press conference</td>
<td>Not press</td>
</tr>
<tr>
<td>14</td>
<td>Evènements</td>
<td>Events</td>
<td>Events</td>
</tr>
<tr>
<td>15</td>
<td>Newsletter</td>
<td>Newsletter</td>
<td>Newsletter</td>
</tr>
<tr>
<td>16</td>
<td>BOAS Swiss Hotels</td>
<td>BOAS SWISS HOTELS</td>
<td>BOAS Swiss Hotels</td>
</tr>
<tr>
<td>17</td>
<td>FR</td>
<td>FR</td>
<td>FR</td>
</tr>
<tr>
<td>18</td>
<td>DE</td>
<td>OF</td>
<td>OF</td>
</tr>
<tr>
<td>19</td>
<td>EN</td>
<td>EN</td>
<td>IN</td>
</tr>
<tr>
<td>20</td>
<td>LAKE GENEVA HOTEL - 3*SUPÉRIEUR</td>
<td>LAKE GENEVA HOTEL - 3*SUPÉRIEUR</td>
<td>LAKE GENEVA HOTEL - 3*SUPERIOR</td>
</tr>
<tr>
<td>21</td>
<td>Au bord du lac Léman et à proximité du centre-ville de Genève, des organisations internationales et du centre de congrès Palexpo, le Lake Geneva Hotel***S est idéalement situé pour vos voyages d’affaires et de loisirs.</td>
<td>By the lake Lake Geneva and near the city center of Geneva, international organizations and the center of congress Palexpo, Lake Geneva Hotel *** was ideally situated for business and leisure trips</td>
<td>At the edge of the Lake Geneva and the downtown area of Geneva, international organizations and Palexpo convention center, Lake Geneva Hotel ***S is ideally located for your businesses and leisure trips.</td>
</tr>
<tr>
<td>№</td>
<td>Source Text (FR)</td>
<td>RL Target Text (EN)</td>
<td>SL Target Text (EN)</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>22</td>
<td>Dans un style contemporain aux lignes épurées, l’hôtel vous accueille pour un séjour tout confort de haute qualité.</td>
<td>In a contemporary style in clean lines, the hotel welcomes you for a stay any comfort of high quality.</td>
<td>In a contemporary style with the purified lines, the hotel accommodates you for a stay any comfort of top quality.</td>
</tr>
<tr>
<td>23</td>
<td>Symbo lisant et matérialisant la rencontre entre l’eau et la terre, le lobby s’ouvre sur un espace lounge-bar convivial pour vos moments de détente.</td>
<td>Symbolizing and realizing the meeting between the water and the earth, the lobby opens on a space friendly lounge bar-bar for your moments of relaxation.</td>
<td>Symbolizing and materializing the meeting between water and the ground, the lobby opens on a space convivial lounge-bar for your moments of relaxation</td>
</tr>
<tr>
<td>24</td>
<td>Bénéfici ant d’un cadre d’exception, le restaurant cosmopolite et sa terrasse surélevée vous garantissent une vue plongeante et imprenable sur le lac Léman.</td>
<td>Benefiting from a frame of exception, the cosmopolitan restaurant and its heightened terrace guarantee you a plunging and impregnable view on the Lake Léman.</td>
<td>Profiting from a framework of exception, the cosmopolitan restaurant and its elevated terrace guarantee a view from above and impregnable to you on the Lake Geneva</td>
</tr>
<tr>
<td>25</td>
<td>Aménagées avec soin et élégance, les 103 chambres d’établissement vous invitent à la relaxation.</td>
<td>Fitted out carefully and elegance, 103 rooms of the establishment invite you in the relaxation.</td>
<td>Arranged with care and elegance, the 103 rooms of the establishment invite you to relieving</td>
</tr>
<tr>
<td>No</td>
<td>Source Text (FR)</td>
<td>RL Target Text (EN)</td>
<td>SL Target Text (EN)</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>26</td>
<td>De par son emplacement géographique, au cœur de Versoix et sur les rives du lac Léman, le Lake Geneva Hotel*** vous offre un environnement calme et agréable tout en étant parfaitement desservi par les moyens de transports.</td>
<td>Due to its geographical location, at the heart of Versoix and on the banks of the Lake Léman, Lake Geneva Hotel *** offers itself a quiet and pleasant environment to you while being perfectly harmed by the means of transportation.</td>
<td>From its geographical site, in the middle of Versoix and on the Lake Geneva shores, Lake Geneva Hotel *** offers to you a calm and pleasant environment while being served perfectly by the means of transport.</td>
</tr>
<tr>
<td>27</td>
<td>Il est accessible en voiture (parking intérieur à disposition), en train (gare de Versoix), en bus (ligne V) et par le lac (port de Versoix-Bourg).</td>
<td>He is accessible by car (internal parking lot at arrangement), by train (station of Versoix), by bus (line V) and by the lake (port of Versoix-Village).</td>
<td>It is accessible conveys some (interior carpark at disposal), in train (station of Versoix), in bus (line V) and by the lake (wearing of Versoix-Borough).</td>
</tr>
<tr>
<td>28</td>
<td>De plus, l’aéroport international de Genève (GVA) se trouve à moins de 10 kilomètres.</td>
<td>Furthermore, the international airport of Geneva (GVA) is unless 10 kilometers.</td>
<td>Moreover, the international airport of Geneva (GVA) is with less than 10 kilometers.</td>
</tr>
<tr>
<td>29</td>
<td>RÉSERVATION</td>
<td>RESERVATION</td>
<td>RESERVATION</td>
</tr>
<tr>
<td>30</td>
<td>Arrivée</td>
<td>Arrival</td>
<td>Arrival</td>
</tr>
<tr>
<td>31</td>
<td>Nuits</td>
<td>Nights</td>
<td>Nights</td>
</tr>
<tr>
<td>32</td>
<td>Personnes</td>
<td>People</td>
<td>Peoples</td>
</tr>
<tr>
<td>33</td>
<td>Réserver</td>
<td>Réserver</td>
<td>Book</td>
</tr>
<tr>
<td>34</td>
<td>HÔTEL</td>
<td>HOTEL</td>
<td>HOTEL</td>
</tr>
<tr>
<td>35</td>
<td>103 chambres (classiques, supérieures, deluxes et junior suites)</td>
<td>103 rooms (classics, superiors, deluxes and junior consequences)</td>
<td>103 rooms (classical, higher, deluxes and junior continuations)</td>
</tr>
<tr>
<td>36</td>
<td>Restaurant avec terrasse surélevée et vue sur le lac Léman</td>
<td>Restoring with terrace heightened and seen on the Lake Léman</td>
<td>Restoring with elevated terrace and Léman lake view</td>
</tr>
<tr>
<td>37</td>
<td>Possibilité de salles de conférence</td>
<td>Possibility of meeting rooms</td>
<td>Possibility of conference rooms</td>
</tr>
<tr>
<td>38</td>
<td>Parking intérieur</td>
<td>Internal parking lot</td>
<td>Interior carpark</td>
</tr>
<tr>
<td>N°</td>
<td>Source Text (FR)</td>
<td>RL Target Text (EN)</td>
<td>SL Target Text (EN)</td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>39</td>
<td>Accès aux personnes à mobilité réduite</td>
<td>Access to the people with reduced mobility</td>
<td>Access to the people with reduced mobility</td>
</tr>
<tr>
<td>40</td>
<td>SERVICES</td>
<td>SERVICES</td>
<td>SERVICES</td>
</tr>
<tr>
<td>41</td>
<td>Réception 24/24</td>
<td>Reception 24/24</td>
<td>Reception 24/24</td>
</tr>
<tr>
<td>42</td>
<td>Accès internet et wifi gratuits</td>
<td>Free internet and WiFi access</td>
<td>Free Internet accesses and wifi</td>
</tr>
<tr>
<td>43</td>
<td>Presse du jour</td>
<td>Day press</td>
<td>Press of the day</td>
</tr>
<tr>
<td>44</td>
<td>Blanchisserie</td>
<td>Laundry</td>
<td>Cleaning industry</td>
</tr>
<tr>
<td>45</td>
<td>Système de confort climatique</td>
<td>System of climatic comfort</td>
<td>Climatic system of comfort</td>
</tr>
<tr>
<td>46</td>
<td>Navette aérport: le service sera disponible dans les semaines à venir</td>
<td>Shuttle airport: the service will be available in the next few weeks</td>
<td>Airport shuttle: the service will be available in the weeks to come</td>
</tr>
<tr>
<td>47</td>
<td>Carte gratuite des transports publics pour le Canton de Genève (durée du séjour)</td>
<td>Free card of public transports for the Canton of Geneva (duration of the stay)</td>
<td>Free map of public transport for the Canton of Geneva (lasted of the stay)</td>
</tr>
<tr>
<td>48</td>
<td>SITUATION</td>
<td>SITUATION</td>
<td>SITUATION</td>
</tr>
<tr>
<td>49</td>
<td>Au bord du Lac Léman</td>
<td>By the Lake Lake Geneva</td>
<td>At the edge of the Lake Geneva</td>
</tr>
<tr>
<td>50</td>
<td>Au cœur de Versoix</td>
<td>At the heart of Versoix</td>
<td>In the middle of Versoix</td>
</tr>
<tr>
<td>51</td>
<td>A quelques mètres des transports publics et des commerces</td>
<td>A few meters away from public transports and businesses</td>
<td>With a few meters of public transport and trade</td>
</tr>
<tr>
<td>52</td>
<td>A seulement 10 minutes du centre-ville de Genève</td>
<td>In only 10 minutes of the city center of Geneva</td>
<td>At only 10 minutes of the downtown area of Geneva</td>
</tr>
<tr>
<td>53</td>
<td>A moins de 10 km de l’aéroport international de Genève</td>
<td>Unless 10 km of the international airport of Geneva</td>
<td>With less than 10 km of the international airport of Geneva</td>
</tr>
<tr>
<td>54</td>
<td>Lake Geneva Hotel</td>
<td>Lake Geneva Hotel</td>
<td>Lake Geneva Hotel</td>
</tr>
<tr>
<td>55</td>
<td>Route de Suisse 77</td>
<td>Road of Switzerland 77</td>
<td>Route de Suisse 77</td>
</tr>
<tr>
<td>57</td>
<td>Accès piétons: Ch. Jean-Baptiste Vandelle</td>
<td>Pedestrian accesses: HP. Jean-Baptiste Vandelle</td>
<td>Access pedestrians: CH. Jean-Baptiste Vandelle</td>
</tr>
<tr>
<td>№</td>
<td>Source Text (FR)</td>
<td>RL Target Text (EN)</td>
<td>SL Target Text (EN)</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>58</td>
<td>+41 (0)22 907 81 11</td>
<td>+41 (0)22 907 81 11</td>
<td>+41 (0)22 907 81 11</td>
</tr>
<tr>
<td>59</td>
<td>+41 (0)22 907 81 19</td>
<td>+41 (0)22 907 81 19</td>
<td>+41 (0)22 907 81 19</td>
</tr>
<tr>
<td>60</td>
<td><a href="mailto:info@lakegenevahotel.ch">info@lakegenevahotel.ch</a></td>
<td><a href="mailto:Info@lakegenevahotel.ch">Info@lakegenevahotel.ch</a></td>
<td><a href="mailto:info@lakegenevahotel.ch">info@lakegenevahotel.ch</a></td>
</tr>
<tr>
<td>61</td>
<td>Groupe hotelier suisse</td>
<td>Swiss hotel group</td>
<td>Swiss hotel group</td>
</tr>
</tbody>
</table>
Annex X: Machine Translation Evaluation Results

Human Evaluation Average Scores - Segments 1 to 20

Human Evaluation Average Scores - Segments 21 to 40
Human Evaluation Average Scores - Segments 41 to 61

BLEU Evaluation Scores - Segments 1 to 20
BLEU Evaluation Scores - Segments 21 to 40

BLEU Evaluation Scores - Segments 41 to 61
Correlation Between Human and BLEU Evaluation Scores - Reverso Localize

Correlation Between Human and BLEU Evaluation Scores - SYSTRANLinks
Annex XI: Raw Machine Translation Faults

The segment numbers in the table below refer to the numbers of each translation segment as set out in Annex IX.

<table>
<thead>
<tr>
<th>Segment n°</th>
<th>Reverso Localize fault</th>
<th>SYSTRANLinks fault</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>Minor term</td>
</tr>
<tr>
<td>3</td>
<td>Minor word structure</td>
<td>Minor word structure</td>
</tr>
<tr>
<td>4</td>
<td>Minor word structure</td>
<td>Minor term</td>
</tr>
<tr>
<td>5</td>
<td>Minor word structure</td>
<td>Minor word structure</td>
</tr>
<tr>
<td>6</td>
<td>Minor word structure / Serious term</td>
<td>Minor word structure / Serious term</td>
</tr>
<tr>
<td>7</td>
<td>Minor word structure / Serious term</td>
<td>Serious term</td>
</tr>
<tr>
<td>9</td>
<td>Serious term</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>Serious term</td>
<td>Serious term</td>
</tr>
<tr>
<td>18</td>
<td>Serious misc.</td>
<td>Serious misc.</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>Serious misc.</td>
</tr>
<tr>
<td>20</td>
<td>Minor misc.</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>Minor misc. / Minor punctuation / Serious word structure</td>
<td>Minor punctuation / Minor word structure</td>
</tr>
<tr>
<td>22</td>
<td>Minor syntactic error</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>23</td>
<td>Minor syntactic error / Serious term</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>24</td>
<td>Minor syntactic error / Minor term</td>
<td>Serious syntactic error</td>
</tr>
<tr>
<td>25</td>
<td>Minor word structure / Minor syntactic error</td>
<td>Serious syntactic error</td>
</tr>
<tr>
<td>26</td>
<td>Minor term / Minor syntactic error / Serious term</td>
<td>Minor syntactic error / Minor word structure error</td>
</tr>
<tr>
<td>27</td>
<td>Minor syntactic error / Serious word structure</td>
<td>Serious syntactic error / Serious term</td>
</tr>
<tr>
<td>28</td>
<td>Minor syntactic error</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>33</td>
<td>Serious term</td>
<td>-</td>
</tr>
<tr>
<td>35</td>
<td>4 minor word structures / Serious term</td>
<td>2 minor word structures / Minor term / Serious term</td>
</tr>
<tr>
<td>36</td>
<td>Minor syntactic error / Minor term / Serious term</td>
<td>Serious term</td>
</tr>
<tr>
<td>37</td>
<td>-</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>39</td>
<td>Serious word structure</td>
<td>Serious word structure</td>
</tr>
<tr>
<td>Segment n°</td>
<td>Reverso Localize fault</td>
<td>SYSTRANLinks fault</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>41</td>
<td>Minor syntactic error</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>42</td>
<td>-</td>
<td>Minor word structure</td>
</tr>
<tr>
<td>43</td>
<td>Minor term</td>
<td>-</td>
</tr>
<tr>
<td>44</td>
<td>-</td>
<td>Serious term</td>
</tr>
<tr>
<td>45</td>
<td>Serious term</td>
<td>Serious term</td>
</tr>
<tr>
<td>46</td>
<td>Minor syntactic error</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>Minor syntactic error</td>
<td>Minor term/ Minor syntactic error/ Serious term</td>
</tr>
<tr>
<td>48</td>
<td>Serious term</td>
<td>Serious term</td>
</tr>
<tr>
<td>49</td>
<td>Minor misc.</td>
<td>Minor word structure</td>
</tr>
<tr>
<td>51</td>
<td>Minor word structure</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>52</td>
<td>Minor syntactic error</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>53</td>
<td>Minor term</td>
<td>Minor term</td>
</tr>
<tr>
<td>55</td>
<td>Serious misc.</td>
<td>-</td>
</tr>
<tr>
<td>57</td>
<td>Minor term/ Minor word structure</td>
<td>Minor syntactic error</td>
</tr>
<tr>
<td>60</td>
<td>Minor misc.</td>
<td>-</td>
</tr>
</tbody>
</table>
Annex XII: SAE J2450 QA Model

SAE J2450 QA Model for RL translation of Lake Geneva Hotel's 'Accueil' webpage (365 words)

<table>
<thead>
<tr>
<th>Category</th>
<th>Minor</th>
<th>Serious</th>
<th>Weighted Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrong term</td>
<td>6</td>
<td>9</td>
<td>57</td>
</tr>
<tr>
<td>Syntactic error</td>
<td>9</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Omission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word structure/agreement error</td>
<td>12</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>Misspelling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punctuation error</td>
<td>1</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Miscellaneous error</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>59</td>
<td>63</td>
<td>122</td>
</tr>
</tbody>
</table>

Overall Document Weighted Score: 0.334

(sum of weighted scores divided by number of words in source document)

SAE J2450 QA Model for SL translation of Lake Geneva Hotel's 'Accueil' webpage (365 words)

<table>
<thead>
<tr>
<th>Category</th>
<th>Minor</th>
<th>Serious</th>
<th>Weighted Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrong term</td>
<td>6</td>
<td>10</td>
<td>62</td>
</tr>
<tr>
<td>Syntactic error</td>
<td>10</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>Omission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word structure/agreement error</td>
<td>8</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Misspelling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punctuation error</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Miscellaneous error</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>49</td>
<td>72</td>
<td>121</td>
</tr>
</tbody>
</table>

Overall Document Weighted Score: 0.332

(sum of weighted scores divided by number of words in source document)