In-Place Translation in Software Development
- A Design Science Research Approach

Madelen Hermelin

May 26, 2016
Abstract

Computers have been used for natural language translation since the 1940s. The role for computers to support translation work has expanded since the commercialisation of the Internet in the early 1990s. As an implication of 'crowd sourcing', translation is now supported by users through community translation, meaning that people - users of web sites - actively translate web sites into different languages. In this thesis we will present Babbler - a novel concept for community translation. Following a design science research approach, a software was designed and implemented in an information systems development project. We present our conceptual design and its software implementation, and evaluate it using different techniques, including log analysis, interviews with translators, and an informed argument contrasting our design to other community translation approaches. The evaluation addresses various qualities of the design, including effectiveness, efficiency, reliability, workflow, implementability and performance. Based on our results, we reflect about translations to different target groups (based on demographics such as gender, age and culture), informed by an explorative analysis of translation results in the empirical context. We also show further implications for future design of community translation artefacts, and future research in the area.
Acknowledgements

This is the only place in the thesis that gives me the opportunity to express myself freely. I want to thank my children for creating me such a beautiful life. You give me daily joy! You make me think about life, and how to prioritize. Special thanks to my son for intelligently correcting my texts and statements, and inspired me by making sarcastic remarks about my manuscript. Thanks to my youngest daughter, who daily defines the word "creative". Thanks to my oldest daughter for showing a unique vulnerability and spontaneity. Thanks to you who are unborn, I think I'll love you unconditionally just like I do for my children. No words can explain what I feel for Rebecca, Eleonor and Kevin, but at least I tried <3.

Thanks to my husband, Stefan that gave me wonderful children, thanks for trying to give me more children in the future.

Last but not least, I would like to thank my supervisor for his support and his incredible cleverness.
## Contents

1. **Introduction**  
   1.1 Translation  
   1.2 ISD - Information system Development  
   1.3 Research question and scope  

2. **Knowledge base**  
   2.1 Translation approaches  
   2.2 Existing artefacts  
      2.2.1 Facebook translation  
      2.2.2 TED open translation project  
      2.2.3 Twitter translation  
      2.2.4 Chrome translation  
   2.3 A model for comparison of translation approaches  
      2.3.1 Common issues in Collaborative Translation  
      2.3.2 Comparison Model  

3. **Method**  
   3.1 Design science  
   3.2 Research approach  
      3.2.1 Part I  
      3.2.2 Part II  
   3.3 The design context  
   3.4 Limitations  

4. **Artefact Description**  
   4.1 Design considerations  
   4.2 Workflow for the developer  
   4.3 Workflow for the translator  
   4.4 Workflow for reporting bugs  

5. **Evaluation Part I**  
   5.1 Log analysis  
   5.2 Interview results  
      5.2.1 Effectiveness  
      5.2.2 Efficiency  
      5.2.3 Reliability  
      5.2.4 Workflow  

6. **Evaluation Part II**  
   6.1 Language analysis  
      6.1.1 Some examples of phrases
7 Discussion

7.1 Improvements after evaluation 36
7.2 Relating Babbler to Facebook 36
7.3 Babbler - a point for discussion 38
7.4 Interesting observations from language analysis 39
  7.4.1 Inconsistency 40
  7.4.2 Unclear phrase keys 40
  7.4.3 Variables in phrases 41
  7.4.4 Use of Global keys 41
7.5 Summarizing discussions 42

8 Conclusions 43

8.1 Design considerations 43
8.2 Translation in ISD 44
8.3 Language variation in practice 45
1 Introduction

Once upon a time there was a need for translation into a portal - a group of developers felt a need for following stakeholders wish and implemented a tool inside a research project. After a while a researcher came in and looked through it with her critical eyes and found interesting things that she wanted to share. But before we start to present the story we like you to go back in the history for a better understanding how we even got to the point of a new story.

1.1 Translation

Nowadays web sites and other types of online applications often support multiple languages. Translation into different languages, however, causes challenges of different kinds. Tools and processes are needed to support efficient and accurate translation. When initially reflecting about what language is, I realised that it is a complex term with many meanings. You can talk about language in terms of body language, a specific language (e.g. English), or language as something that evolves over time in a social context. For the purpose of this thesis, we think of a language as "a language is a system of communication used by a particular country or community" (oxford dictionaries, 2015). A language, following this definition, is not bound to a specific country, rather to a community. Communities like Facebook and Twitter seeking not only for a country specific language but inside a community. Following such definition, it is given that those who will use the product are also the ones translating for the community.

Even though computer-supported translation has been discussed since 1949 (Weaver, 1955), computers have supported individual translators even longer. The concept of Translation memory (TM) was used long before machine translation (Garcia, 2009). TM is the idea supporting translators to work more efficiently through the use of technology. Originally, TM was used by individuals for solitary work. When the Internet evolved it became easier for translators to share amongst other translators. Google translate, for example, can be seen as a tool for users to be able to translate their own text - but every time we use the tool we extend the TM for Google Translate. Additionally, since the tool draws from user-generated content, the more people that use the translation tool, the more powerful it will become.

Warren (Weaver, 1955) introduced the idea of using a computer as a translation tool. Even though Weaver introduced the idea it has become both viable and practical, so much so that it has become part of many translators’s jobs. Using crowdsourcing to solve problems is not a new phenomenon (Howe, 2006) the translation of webpages, documents, books have also embraced this technique to allow users and non-professional translators to participate
in the translation process. As we will see later in this thesis, users of online communities like Facebook, Twitter and TED-Talk participate in translation, thus making their translation work and its context available for the global public.

Community translation (we use the term synonymously with collaborative translation throughout the thesis) brings the social dimension into translating. It engages people to contribute to the translation effort. In many cases, community translation is performed by non-experts rather than by professionals. Even though community translation appears to be trivial, it has to deal with several challenges related to business goals, quality control, crowd motivation, role of professionals, parallelism, de-contextualisation and how we can adopt (and further develop) best practices (Désilets and van der Meer, 2011).

1.2 ISD - Information system Development

It is well known that design processes benefit from stakeholder participation (K, 2006; Avison and Wood-Harper, 1990; Mumford, 1995; Sharp et al., 2007; Sjöström, 2010).

We have, during the process of building a translation tool into a portal seen that it engages and supports the ISD. We haven't found any earlier research that explores translation as a part of the ISD. This thesis will investigate and present how letting translation be part of ISD has become an effective strategy to improve the quality of information.

The core of ISD is to improve human activity by designing and implementing software technology (Martin, 2003). For a long time, plan-driven ISD (‘waterfall models’ for ISD) was prevalent in ISD. The waterfall model for ISD was introduced in 1970 by Winston W. Royce (Royce, 1970). Nowadays agile methods have emerged as an alternative development strategy. The agile manifesto is a set of values intended to support developers’ productivity, and adopt their work better to changing circumstances (e.g. requirements). The agile approach - as opposed to the plan-driven - is thus characterised by short development cycles and intensive communication between users and developers. In the agile manifesto they talk about "customer collaboration over contract negotiation" and there is a need to involve customer feedback in a regular and frequent basis (Martin, 2003). In agile methods, users are thus expected to engage in various ways throughout the project.

As a designer during system development, it is important to understand the business domain, i.e. grasping the meaning of concepts in the 'language' spoken by the future users of the software. There are several strategies to engage users in ISD. Krippendorff (2006) summaries several methods that may be employed to facilitate increased understanding between users and developers:

Surveys and structured interviews are arguably the least informative
methods for elucidating stakeholders concepts. Surveys are often sent out to a big population which often does not provide any new and unexpected information. Interviews often tend to make the stakeholder say what they believe that you want to hear.

**Unstructured interviews** lead to a less formalised conversation, but takes more time than structured interviews. This may lead to more in-depth knowledge, but there is also a risk that less stakeholders get to express their views. It is also crucial to be able to ask creative questions that brings something new.

**Focus groups** are based on the idea of having 8-10 stakeholders that talk about a subject, moderated by a designer. It is often hard to see and get a group that represents the community. The focus group may lead the design in unanticipated directions.

**Observational methods** In contrast to interviews, surveys and focus groups, observational methods are focusing on what people do when new artefacts enter their lives. You can use different methods for observation and representing observations, e.g. video, audio, and field notes in writing.

**Protocol analysis** is based on what people say while solving a task. Here you can observe what they say but also get a feel of what they are doing at the same time.

**Ethnography** stems from anthropologists and it is about understanding the whole area. By doing field studies you attempt to understand what is being done as well as the context of action.

Krippendorff further suggests that each approach has strengths and weaknesses, and that better results are achieved by a *triangulation* of the techniques above, arguing that no single method can provide everything that designers need to know.

### 1.3 Research question and scope

In this thesis we explore community translation and the ways in which it may affect stakeholder participation in ISD.

We will also argue for translation supporting information system development based on design and evaluation. We will look at how to build an artefact to translate webpages. We will evaluate the product with interviews, observations but giving some insight of examples that occurred during the implementation and usage of the product. We investigate other existing tools that can be seen as community translation, collaborative translation or/and crowdsourcing translation systems.
The aim of this thesis is to show how one can build a translation tool in a proper way that supports community translations. We will show how a translation tool can support and be a part of ISD. Since this is a rather broad subject to cover in one thesis the idea is to divide findings and research questions into several smaller questions. Questions that we will consider are:

1. How can we design a translation tool to support community translation?

2. In what way does community translation relate to stakeholder engagement in ISD?

This thesis is an extension and deepening of a paper written by the author and Jonas Sjöström (Sjöström and Hermelin, 2013), it elaborates on the design of the tool itself but also from an organisation aspect. Question 1 takes care of the artefact itself meaning that we want to share our knowledge of how you can build a translation tool for collaborative translation. Question 2 covers that we found several interesting processes that occurred during the use of the translation tool and an found outcomes to have this kind of translation approach into a information system development process. These questions should not be considered like separate parts of the main purpose since they are connected. For example to understand the second question it is also crucial to know how the translation tool was designed.
2  Knowledge base

In this chapter we present other approaches and tools that are used in a similar fashion to the artefact that we are going to present. We will begin by describing several other communities and portals that are using translation tool and then we present an overall picture of related research in the area. First, we outline some common terms and practices related to existing translation approaches in section (2.1). Second, we provide an overview of some well-known Internet-based software companies and their various approaches to translation in section (2.2). Third, in chapter 2.3 we propose a model to characterise and compare translation strategies based on the ideas presented in chapters 2.1 - 2.2

2.1  Translation approaches

Using a computer for translation is not a new phenomenon, but still there are many different approaches on how to manage it. Weaver was the first that introduced the idea of letting the computer translate by statistical techniques (Weaver, 1955). We can also see that collaborative translation and user-generated translation or crowd translation have become important for several reason (Désilets, 2010; O’Hagan, 2011)

Previous research has listed several reasons why it is hard to design translation software.

"Part of the reason why translation is difficult for computers is that translation is just difficult: difficult even for humans." (Arnold, 2003, p. 119)

A translator should be aware that even if they produce a text that is equivalent then they still need to keep the reader interested by making elegant, good translations. Even if its hard to machine translate it’s still a challenge/achievement to effective a fast translation. So an idea is to let the computer do the initial translation and use humans to translate and correct the output.

Several studies have proposed the idea of letting the translation go online and point out the importance of providing features that can support collaborative tasks and improve translations skills. Even if automated translators cannot be compared with professional human translators, we shall see that collaborative translation deliver the same quality. (Kageura et al., 2011)

Machine translation (MT) is the idea of letting the computer translate with or without human assistance. Our aim is not to give you a full, deep descriptions about what it is but an overview. MT tools allows us to translate natural language into another language. Inside MT we can see different kind of approaches since they are more sophisticated than only translate word by word.
Crowd sourcing was first described by Jeff Howe in the article "The Rise of Crowdsourcing" in the magazine Wired (Howe 2006). He states that crowd translation is when you outsource a job into an open call to a (often big) group of undefined people. Crowd translation is when we let a lot of unprofessional people translate. There are a lot of examples and studies that have been made (Munro 2010) (Zaidan and Callison-Burch 2011).

In the definition of crowd sourcing it includes the requirements for those that does the work doesn’t require a worker to be a part of the domain.

Community translation - or collaborative translation - is letting unprofessional people translate. There is an obvious similarity between between crowd and community translation. Even though they’re easy to actually call them interchangeable community and crowd translated there are many shows the differences between them. Community translation, Crowdsourcing translation and collaborative translation share the idea the crowd is better than a few (Howe 2006). Babbler is built on the idea of letting the community translate. Even though there are many ways if describing the same phenomenon, collaborative translation has the potential to change how we translate content. There are sites that are made for the purpose of collecting and sharing good practice or even design patterns that can gain translating processes (Désilets and van der Meer 2011).

Even though it seems easy and trivial to use collaborative translation there is a need to collect some good design patterns on how it can support the process. There was a day-long workshop held in TAUS in October of 2011 and the goal was to generate this kind of patterns, see more on the site www.collaborative-translation-patterns.com.

2.2 Existing artefacts

There are numerous of existing artefacts to support translation in various ways. In this section, we present a set of well-known approaches that relate to our work, including translation strategies of (i) Facebook, (ii) TED talks, (iii) Twitter and (iv) Chrome.

2.2.1 Facebook translation

Our main focus is on Facebook (2.2.1) since it is well-known and has been used on a large scale. Facebook is translated into 75 (October 2011) different languages and they try to motivate users aid the translation process by saying:

"When you contribute with translation you are helping 278 of your friends that might by using Facebook in Swedish"
Facebook translation project started at the end of 2007 and there were many arguments against free translations. There were protests to leave the translation to the translator. \cite{O'Hagan2011}.

Facebook uses several different strategies to translate their platform where they have professional translators but they also use BING for translating user-comments and other user generated content. We will only touch and concern ourselves with their idea of allowing Facebook users to join to help translate the site. Initially the idea was to encourage participation rather than financial gain. The idea was to let user produce a localised website that can fulfill user expectations better than professional translators can. The first site was translated in a week and the next ws achieved in one day using the community \cite{Jimenez2011}. Facebook have for a long time take advantages of community translation, meaning letting the users contribute to the different kinds of language to the platform. They provide a similar product that we propose (Babbler) and they filled in an application for a patent on their work (US patent No. 2009/0198487A1). Babbler and Facebook’s "In-Line" translation tools are similar although they were developed separately and independently from each other. Facebook has about 250,000 translators and they translated 350,000 words into 70 languages in a short time. For some languages it took only 2 days \cite{Desilets2011}.

I will explain and present Facebooks translation process by describing the stages of the translation

**Phase I:** all members of Facebook can apply to be translator. It means that you are going to be assigned to a group where you are gain some rules, can join discussions and get rewarded for being active. As a translator you can choose to see all phrases and take any phrase and translate it. If the phrase has already been translated you can vote it down or up. If you don’t want to work only with translation you can turn on translation and each time you are coming to a phrase that needs to be translated it gets marked and you can right-click on the phrase. No phrase are visible or translated just in time. In the forum you can discuss several issues or interact with other translators. For example a new and highlighted question right now (20 mars) is how they can make the platform more gender neutral. Naturally there are many discussions on the forum by typically those that will affect you are (1) phrases that affect the graphical interface; (2) political discussions about what is correct to say without offending or that words will create irritation among users; (3) reminders on previous decisions made regarding translation and word decisions (4) calls to vote down the wrong phrases that goes against decisions; Facebook has an easy way of filtering and searching for words and you can easily search and vote down phrases. Many of the posts in the forum are about voting down or up phrases.
**Phase II** You only can vote for good or bad translations - this is one of Facebook’s techniques to guarantee the quality on the translations. There is no way of commenting about a specific phrase or translation. It is only in the forum that you can discuss this kind of question - and there are a lot of discussions and opinions. In this step translators can see the phrases in context, meaning that they can see whether if the translations still suit. There have been indications that this process makes it easier for the translators to change small segments of texts instead of a whole bunch of text.

**Phase III** Last step is that Facebook is letting professional translators go through the translations and this is the second quality check Facebook do. They are checking that translators are consistent since it has been indicated that volunteer translations are often inconsistent (Cintas and Sánchez 2006).

Facebook has some rules and FAQ to support translators and their work. They also encourage people to: "It does not hurt to translate to a casual, flowing language, more focused on comprehension than on Sloppy, fun formulations (unless they are substantially creative)" (Facebook 2014). This indicates that they want a language that is closer to the users rather than formal language. Even if there can be a lack of quality there are discussions if community translation is failing on the quality, Jiménez-Crespo and Miguel A are using Facebook as a model of how they can use Facebook’s approach as a method for evaluation of translation studies. Eventhough they can include that Facebook quality evaluation seems to go against many of their principles and guidelines used in Translation studies. They argue that if they want a localised product then it’s not a bad idea of letting a large number of non-professional users contribute with their deep knowledge about the digital genre rather than letting professional translators do it (Jiménez-Crespo 2011).  

2.2.2 TED open translation project

TED talk is a project that has its goal to spread ideas around the world. The idea is to present something in a short video from different kinds of disciplines and cultures. Many of TEDs viewers contacted TED and wanted to help with subtitling and translate videos. TED started the translation project and now they have 104 languages, 15,637 translators, 54,449 translations.

In their guidelines of translation they state:

*Informal over formal Where appropriate, choose informal, colloquial terms over formal or academic ones.*
"Modern over traditional Choose modern terms and phrases over the traditional. Translators should strive to be up-to-date in the topics covered."

"Personal over generic Strive to match the tone and flow of the speaker’s original talk. Rather than produce a word-for-word translation, aim to find the color, energy and "poetry" in the speaker’s organic style and try to emulate it.

As you can see from the above TED prompts are like facebook. Namely that translators should use words that are more close to the community and do not need to be formally correct. To translate from a film is not just about translating it correctly but also being able to translate it with the same style as the one who speaks. Also of relevance is that TED translation was initiated not by themselves but by the people who view the films. This is to create an open and accessible environment which is fun to participate in.

2.2.3 Twitter translation

"An indispensable and vibrant group of more than 350,000 translators use our translation center that Twitter will become available for everyone around the world.

Twitter also allows users to help to translate pages. They use different approach by giving the translator a phrase which is not in its context then the translator provide suggestions of a new phrase. They now have 350,000 translators that are active in the process of making Twitter available for the whole world. They are using a vote system to decide which translated phrase
is going to be used. One difference from Facebook however, is that Twitter extracts the phrase from its context, and you can’t translate on the fly. Its more like you’re sitting for the purpose of working with the translation.

Twitter has two approaches on deciding who is going to translate - one is that they actually ask bi-lingual speaking people to help, or you as a user can apply for being a translator. Twitter has moderators that are providing help to groups of translators. This means that they will be involved in discussions, decisions and policy/guidelines document.

2.2.4 Chrome translation

Chrome has yet another approach to translating. They offer a tool to translate the whole page into your own language. The purpose was to make websites available for the whole world. Chrome uses google’s translation technology. Google translate uses a technique that looks for patterns on already translated pages and tries to decide which is the best translation for the user. This method of solving translation by pattern is called statistical machine translation [Koehn, 2009]. The quality is not always perfect but they state that this is not their focus, the generated page is so that the user can have an idea about what the site is about. Using statistical machine translation the quality is better depending on how many human-translated documents it can find.

2.3 A model for comparison of translation approaches

In order to talk about translations and to make a comparison between the different translation techniques we need to make a categorisation of aspects or characteristics of different kinds of translations approaches. In this section I will present different categories that are used later in the evaluation but also what I have seen during the process of going through existing artefacts and translation approaches. Most of the categories are from a paper where the authors tried to talk about common issues in collaborative translation [Désilets and van der Meer, 2011]. They divided issues into: Business goals, Quality control, Crowd motivation, Role of professionals and parallelism and de-contextualisation. I will summarise each in order to give an understanding of why I choose to expand or why I choose to rename them. After this work I will present my model and try to motivate what it is, why it is important and how it can be used on the existing artefacts presented earlier in this thesis.

2.3.1 Common issues in Collaborative Translation

This is a summarized interpretation of previous research ([Désilets and van der Meer, 2011]) and is presented here for ground work and future comparison.
**Business goals** is about what kind of benefits we can have from collaborative translation. For example cost reductions and so on. But it is important that different kind of flavours of technologies have different kind of benefits.

**Quality control** is about how much we control translation from the top, meaning that we can have no control at all or that we can have professional translators that go through the translations before they get published. It depends on the context in which it will be used.

**Crowd motivation** is about how we motivate our translators and it can depend on what purpose they have. For example Facebook motivates translators by being a part of something while other collaborative translations have emotional motivations (Munro, 2010).

**Role of professionals** is about the role professional translators play in collaborative translation. It is also about how involved professionals are and whether they can be used in another way together with collaborative translations.

**Parallelism and de-contextualisation** many implementations of collaborative translations are made by splitting the work into smaller chunks and splitting them to different members of the community. How you split and to who (also that taking out words from its context) can affect the final result.

### 2.3.2 Comparison Model

Here I will present a new model that is used comparison between collaborative translation and translation approaches. I will motivate each category separately. It is important to understand that it will differ from common issues only because I changed the aspect and purpose of the model. Also I want to open up the model so that can be used for deciding which translation approach fits best.

**Goal and purpose** I created this category to separate between a purpose and a goal. The purpose covers why one would want something, or what will be the benefits of the action. The aim is often more comprehensive during entire project and doesn’t need to be measurable. A goal is more measurable and is a result that needs to be achieved. For example, the purpose of Twitter is to make the portal available to all over the world while a goal is to allow the user to request a language to be implemented if missing. The goal and purpose are means of describing why the artefact was built.
**Translator** I wanted a comparison model that can be used even between different kinds of translation approaches. It is important to know who are the translators? It can also be more clear if we are using some kind of machine translating where the translators are not human at all. It is important that we try to give an idea of what kind of translators we are dealing with and why they are translators. For example, a Facebook user can become a translators only if they are Facebook members. One further example is Google Translate that anyone can use and at the same time contribute to.

**Domain** In many of the translation tools that I looked at it was either closed to members or very open. For example every time we use Google translate we improve Googles translator memory. Even if we don’t take notice, Google will be providing prompts on the same page. This means that Google’s tool is very open for everyone. I want to make distinction between the translation domain and its accessibility.

**Quality control** Quality control is about how we control translators and what they are doing. It also states how important it is to have a correct language. For example using Chromes translation, its only purpose is to make the website understandable while Facebook intention in providing a good language experience. Facebook never lets unedited translations go live without having a professional translator go through the phrases. In Twitter they have moderators that are chosen by Twitter that can both change phrases and are able to suspend users that translates poorly or against the rules.

**Motivation** How motivated are the translators and how can you motivate users to contribute. In order to gain motivation it is important to compare machine translation against collaborative translation. The computer does not need motivation to translate but are still require to motivate people to correct machine translated texts or that we maybe need to tweak a translation algorithm so that it better suits our purpose.

**Role of Professional** This is important and was presented in the previous section as well but we want to make a distinction and elaborate the meaning of "Role of professionals". It can be important in to decide if we should use a professional translator and also know what kind of skills we require on the translators.

**Subdivision and degree of context** During the process of looking at different kind of existing artefacts it became obvious that every artefact use their own way of dividing a context into different kinds of chunks more over one must also consider the aspect of how they present the content. For example Twitter assigns to translators longer phrases
from the context and translates bigger texts; while Facebook assigns shorter phrases and from the context where the phrase is used.

**Numbers of translators** Most of the artefacts that were looked at had a big community with many translators. That is what you could call it a crowd. But in smaller communities or when you have a minor language the crowd becomes smaller and that will be a factor on how you should design the translation tool.

**Opportunity for social interaction** How remote are the translators and how do they engage with the translation process. If we believe that translation is a social activity then we also look how we should support social interaction between translators.

One might think that trying to simplify a complex problem one should only take into account the separate parts. But our goal is not to give you a model that is a recipe of how to succeed with translation. The model is only a description and generalised idea of how we can compare and talk about different kind of approaches. Only making that kind of model will require further research and more testing on different kinds of existing artefacts in translation. But for this thesis I claim that it is enough for describing where my artefact belongs and how it compares to others.
3 Method

This chapter motivates and outlines the research method. Section 3.1 describes the overall method that we used during the thesis. Section 3.2 presents how the work progressed during the project. To get a better understanding of the contribution I have divided the work into two parts. Section 3.3 describes the domain that the artifact has been developed for.

3.1 Design science

Design science is important to IS (information system) research. It is a way of developing innovative artefacts solve real problems. Also it facilitate the extraction of knowledge from the solutions and contributes with solutions for a given class of problems. Design science consists of three major cycles: Design, rigor and relevance (Hevner 2007).

It is easy to argue that there is a relevance cycle, meaning that the artefact is developed in a sophisticated environment. U-Care consist of therapists, researchers and real patients. It also involves collaborating with hospitals and other actors that are involved in patients health care. Both the development, use and evaluation are undertaken in the same environment and consists of real problems.

Rigor comes from the collection of researches from multiple disciplines that contribute to design the platform. Most of these experts have very good knowledge about other platforms. It was also important to keep data about how they worked with the translation tool to gain more information about how they interacted with the artefact. We also monitored phrases that were edited and who made the changes.

Design and creation follows a problem-solving approach. This means that it uses an iterative process involving 5 steps: awareness, suggestion, development, evaluation and conclusion.

We decided to apply design science as a method by using the 7 guidelines that were made for assisting researchers, reviewers, editors and readers to understand the requirements for effective design science researchers (Hevner et al. 2004). Hevner et al. argues that each of this guidelines should be addressed in some manner for design science to be complete. However it has been up to the reader and editors to determine how well they are satisfied. I will first present a table with short summaries of the guidelines made by the author of the article. Then I present how our work applies to each of the guidelines.

Guideline 1: Design as an artefact We have built a fully implemented artefact as an instantiation. Since I see the artefact more than the actually instantiation and that tries to solve how we can engage people
Table 1: Design-Science Research Guidelines (Hevner et al., 2004)

<table>
<thead>
<tr>
<th>Guidelines</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guideline 1: Design as an artefact</td>
<td>Design-science research must produce a viable artefact in the form of a construct, a model, a method, or an instantiation</td>
</tr>
<tr>
<td>Guideline 2: Problem Relevance</td>
<td>The objective of design-science research is to develop technology-based solutions to important and relevant business problems.</td>
</tr>
<tr>
<td>Guideline 3: Design Evaluation</td>
<td>The utility, quality and efficacy of a design artefact must be rigorously demonstrated via well-executed evaluation methods.</td>
</tr>
<tr>
<td>Guideline 4: Research Contributions</td>
<td>Effective design-science must provide clear and verifiable contributions in the areas of the design artefacts, design foundations, and/or design methodologies.</td>
</tr>
<tr>
<td>Guideline 5: Research Rigour</td>
<td>Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artefact.</td>
</tr>
<tr>
<td>Guideline 6: Design as a Search Process</td>
<td>The search for an effective artefact requires utilising available means to reach desired ends while satisfying laws in the problem environment.</td>
</tr>
<tr>
<td>Guideline 7: Communication of Research</td>
<td>Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.</td>
</tr>
</tbody>
</table>

and involve users into the ISD with a translation tool. The artefact is built and lessons learnt of the outcome of the design process during the building and use of the artefact.

Guideline 2: Problem Relevance Doing translation and how to support collaborative translation is a big problem by itself. As we saw in the introduction, we have several methods to engage and how get to know our users. The artefact comes from a real organisation that wants to have separate languages depending on what culture that the user is from.

Guideline 3: Design Evaluation Is how we as a researcher must be able to use well-executed evaluation tools. The artefact is implemented and we evaluated the tool by a qualitative approach. This was achieved by having interviews of translators who were most active. While the artefact was built we also collected and logged data that was important for traceability. We collected data for comparison between languages used in the portal.

Guideline 4: Research Contributions In this thesis we will show how the artefact stands against others artefacts in the same domain to illustrate the novelty of our approach thereby exposing the contribution is more than the learning outcome of the instantiation but it is more innovative way to see and implement the artefact into/part of ISD.
Guideline 5: Research Rigor arises from the development of an artifact within a research project, coupled with our awareness of existing tools in the same domain. We have chosen evaluation methods, analysed deliverables, created results by using methods that are common and well-known in the Information system field.

Guideline 6: Design as a Search Process As we will see in this thesis we have adopted an iterative process that bridges the requirements and constraints in the application domain with the solution domain (technical and organisational).

Guideline 7: Communication of Research We focus in this thesis both on a technical view but also on the organisational aspects based on observations of the working artefact. Even if we demonstrate how we can build the tool itself, we are also keen on how it supports the ISD. Focusing on how they interact and change behavioural in the focus of the implemented artefact.

Based on these guidelines; we discuss how they inform our research. This thesis dispositions are created after (Gregor and Hevner 2013) and is a suggestion of how we can contribute with knowledge so that others can understand and position their research project in a sophisticated way.

3.2 Research approach

This chapter will guide you through the work that was undertaken during the thesis. For the reader I will separate my contribution into two parts. This means that I can make a better connection into what part that contribute and answer what kind of question it takes care of.

3.2.1 Part I

Part I of this study was previously presented in a paper (Sjöström and Hermelin 2013). This thesis will go further by investigating differences and similarities of existing translations tool. Motivated by a comment from one of the reviewers, it goes deeper into how it actually supports the ISD. Based on comments from stakeholders we have developed the platform and made more evaluation in the form of one deep interview.

The artefact was developed and rolled out to the stakeholders. Stakeholders were only introduced to the translation API and we did not try to give them guidance in how they should engage in the translation process. The idea of not giving them to much instruction was to see what happened and also how they organised the work. We then performed a log analysis trying to pick those people that frequently translated in the platform and asked them two opened questions:
We are interested in your impressions from the translation process. What worked well, and what did not work well?

If the translation process – the software and/or the way you work with it – should be changed, what would be the most important changes?

The answer was then categorized into 4 kinds of categories Effectiveness, Efficiency, Reliability and Workflow.

Since we want to make a clear connection between research questions it is here that we began making the first steps at answering and evaluation Q1, namely how we can build a tool that supports translators. We both focusing and tries to make a learning outcome both from the technical aspect of the artefact but also the process of translating. This means that we both evaluate interviews but also the observations we have seen and document during the use of the artefact.

3.2.2 Part II

Following the interviews from part I, the changes in the portal we could affect the design. This is shown in the changes that have been made and the discussion we will try to explore what has been observed in meetings, documents demonstrate how this may have affected users. Part II of the thesis is coming from the fact that the stakeholders wanted to have several languages within the portal. For this part we took out each phrase in two different languages and analyzed those phrases that was different between the languages. Our First language was SWETEEN who worked with young people and the second language was SWEADULT that was made for adults. These two languages were the ones that were interesting as the users of these studies felt that it was important to use "a youthful language" to young people and a more "formal and descriptive language" to the elderly. This data was reviewed iteratively and done with a qualitative analysis. My starting points were:

(i) You use more / longer words for the elderly
(ii) You use more personalised calls and texts to the elderly such as using "I", "you" and "here" instead of a direct verbs (eg Send).
(iii) You use a more informal and colloquial language to younger
(iiiii) There is a direct differences in semantics between languages.
(ii iii) For a older domain stakeholders use more specific names than for younger (eg send message instead of send)

Those starting points come from how stakeholders argue why it was important to have different kind of Swedish depending on who they talked to. By analysing the actual translations qualitatively we seek insights about what translators actually did - as a complement to our evaluation part 1, where they expressed their views.
3.3 The design context

The artefact was developed in a research project U-CARE (Uppsala University Psychosocial Care Program) headed in Uppsala University. The artefact is a platform developed for the purpose of giving CBT on internet to patients that have some somatic disorder. In our system, the translation tool was developed and should give stakeholders the opportunity to translate the platform into several languages depending on which patient group that is going to use the platform. Even if the platform can handle several kinds of languages spoken in different countries it can also have different kinds of languages depending on culture, gender and age. In U-Care there are therapists that provide a specific group therapy. this means that there are several kinds of groups,such as young people with cancer or pregnant people that are afraid of giving birth. Even if we have already stated that translation has been a tricky task its important to have a tool that meets the requirements and should be easy to use.

The developer team during the time has involved between 1-5 persons. Even if it wasn’t clearly stated, they worked with agile principles. Every week there was a meeting with the whole developer team and 1-2 persons from each study. The agenda of the meeting was decided in beforehand depending on the results from the week before. Developers work in the same building as the stakeholders, meaning that a close connection is made between those that want the platform and those that implement it. Developers do not have connection with the patients so that every time we referre to stakeholders in this thesis we mean: therapist, health staff or researchers.

U-CARE is for patients but is a research oriented project meaning that the idea is to see if it have a good impact on the health and can complement or replace existing therapies.

3.4 Limitations

This paper promotes further development for to continued the design of translation for optimal use. Since These tools were not the main focus for the research project, only minor changes were made and we have not developed further. In the paper, a qualitative exploratory analysis of two languages, this is only a start and we can only show indications of what we found. This work does provide a basis on which to try to do this during the development stage and thus not encounter the same problems we had.
4 Artefact Description

This chapter describes the implemented artefact. It gives an overview and design considerations before turning into a more technical description of the artefact. Since the processes regarding translation are important for the result Chapter 4.2 and 4.3 presents workflows. During this study we changed the workflow for reporting bugs; therefore we present how the process was from the beginning.

4.1 Design considerations

A set of design considerations were phrased to guide the initial design of the Babble artefact:

- Make translation as simple as possible (one-click to translate) in the user’s current view
- Allow for context-dependent translation as well as translation of ’global’ phrases
- In a translation situation, provide the user with relevant alternative translations
- Any type of text element, independent of how it is embedded in the page, should be possible to make ’translatable’
- Support translation both for full page requests and for asynchronous requests

Figure 2: Conceptual Model for Babblor

Babbler uses a phrase key value (See Fig 2) – that means that we store phrases into a database and every time a page is called then we translate the page into the right language. There is no storing of the language at the client because we wanted it to be updated as soon as possible. Basically a phrase key is created by the developer and can have several translations provided
by different users. Babbler behaves as follows: if a phrase is translated the
existing phrase will be marked as deleted in the database. This approach
allows us to have old translations available but also keep track of several
different languages in the plattform. Babbler also supports variables within
a phrase, this means that some values doesn’t need to be hardcoded and
can change depending on who is looking at the page. By putting the phrase
enclosed by brackets it finds a value for that variable.

Babbler is dynamic and complex because it can manage two kinds of
situations: Page load-translation and asynchronous translation. There is
an increased need for asynchronous web requests, i.e scripted code on the
client side that makes it possible to send requests to the web server without
reloading the whole web page.

![Babbler translation dynamics](image)

**Figure 3: Babbler translation dynamics**

Fig 3 states that how a web browser requests a page from the server. It
behaves as follows, first the URL is analysed to set the right translation into
text. This is necessary to take care of the possibility of having a situated
translation only on that page or generic translation, meaning that they have
the same translation on all pages. This approach enforces the developer
to use Babbler function in all places that there is text. Both to make the
translation but also to ensure that a phrase is translatable. Developers also
decide on every phrase if it should be translated on all pages or just locally.
The server stores information such as the choosen language in a load session.

Sometimes, text in a web page is encapsulated, e.g. in a dropdown list.
Because of this there was a need for the Babbler API to support different ele-
ment types that need to behave differently depending on whether translation
was activated or not. The normal state is for regular text. The embedded
type is used when phrases are nested within something further on, and there
was a need to construct a helper element to be able to right click and trans-
late on. So workflow for a developer requires a more descriptive explanation.
In table 2 we show the different kind of states.

<table>
<thead>
<tr>
<th>Translation state</th>
<th>Normal</th>
<th>Embedded</th>
<th>Helper</th>
</tr>
</thead>
<tbody>
<tr>
<td>On</td>
<td>Translatable</td>
<td>plain text</td>
<td>Translatable</td>
</tr>
<tr>
<td>Off</td>
<td>Plain text</td>
<td>Plain text</td>
<td>Invisible</td>
</tr>
</tbody>
</table>

Table 2: Translation states and element types

It means that a developer needs to use Babbler functions to make embedded text work properly.

Further on there is a share feature between the languages meaning that every time you right click on a phrase it also shows earlier translations. This was idea allows us to have a culture specific language.

4.2 Workflow for the developer

Babbler is implemented and abstracted by having a class for helping developers create new words. For a better understanding and to demonstrate its simplicity the method is showed below.

```csharp
public class Babbler
{
    public static String Say(String phraseKey,
                                bool allPages = false,
                                BabblereturnType returnType = BabblereturnType.NORMAL)
}
```

Parameter one is the key for the language. Parameter 2 is a boolean that indicates if all phrases with that key should be applied for the whole platform or only at that page. Parameter 3 is the return type that is an enum in the Babbler class:

```csharp
using System;
public class Babbler
{
    public enum BabblereturnType
    {
        NORMAL = 0,
        HELPER = 1,
        EMBEDDED = 2
    }
}
```
The enum is divided into three states, one for plain text and the other for texts that are imbedded. Imbedded text is for buttons or drop menus. If the text was imbedded there was a need for an extra row of code.

```html
1 <%= Babbler.Say("Add observation point", false, Babbler.BabblerReturnType.HELPER) %>
2 </button>
3 <button>
4 <%= Babbler.Say("Add observation point", false, Babbler.BabblerReturnType.EMBEDDED) %>
```

For plain text, one row is sufficient.

```html
1 <%= Babbler.Say("Add observation point") %>
```

The key in the database was generated and processed when a webpage was requested for the first time. When someone translated a phrase it became visible just in time, meaning that the last translation was used. There was no voting system meaning that if someone wanted to change the phrase; changes were visible for all users at the same time.

In some texts there was a need to make the phrase more personalized. For example a text that was more study specific like “In study X you have one more massage”. For personalized cases we used brackets to create variables in the database. For every variable we had to have a dictionary in the database.

Many webpages today use tooltips. Tooltips is a way of communicating what will happen if someone is taking an action. It was hard to use the same approach (left-click) for tooltips so we solved it by having all tooltip-text listed at the top left of the page during translation.

### 4.3 Workflow for the translator

Table 2 shows different kind of translation modes. To be able to translate you needed to set on the translation mode (compare to Facebook’s translation that can be done on the fly). Every phrase in the platform that was translatable was marked with red when the translation mode was on. Translating a phrase was made by right-clicking on a segment of text, a pop-up window was shown and changes were made immediately (compare to Facebook’s voting system).

### 4.4 Workflow for reporting bugs

Because the second evaluation of the artefact will be affected by the fact that we changed the process of reporting errors, we describe the process below. Note that changes were done after the first evaluation. Initially bugs were reported directly by email to developers. Often information about the bugs was missed, information that was important to fix the issue. Problems that often occurred were problems with the use of local or global translations. If
Figure 4: View after right-clicking a phrase

a page was moved to a popup-window and set as local then the URL was changed depending on where they open the window. Users often gathered several translation problems in one word document – that included only a print screen of the translation problem. Often there was a lack of information. One step to solve this problem was to have an icon in the platform. If the user found a translation problem they needed to click on the icon and were able to fill in a report. With the report the platform collected information about the problem (like: username, study name, URL) and took a print screen on the page they were on. The report was transmitted to all developers with an ID, which was able to fix the bug.
5 Evaluation Part I

Evaluation of the software were made by analyse of log data and then interviews with the most frequent users. Translation in U-CARE was self-organised – meaning that there were no rules set for how people should work with translations. The users were domain experts – e.g. psychologists and researchers – that were supposed to work with the U-CARE portal. They were introduced to the translation tool without any particular training in advance. Developers were informed of how they should make views translatable by using the Babbler API.

5.1 Log analysis

All translations were stored in the database which also helped us to keep track of all changes made and when they occurred.

<table>
<thead>
<tr>
<th>User</th>
<th>Count</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donald</td>
<td>1744</td>
<td>Developer</td>
</tr>
<tr>
<td>Claire</td>
<td>226</td>
<td>Research coordinator</td>
</tr>
<tr>
<td>Peter</td>
<td>186</td>
<td>Used by Donald to translate participants’ view</td>
</tr>
<tr>
<td>Thomas</td>
<td>98</td>
<td>Psychologist</td>
</tr>
<tr>
<td>Paula</td>
<td>97</td>
<td>Used by Claire to translate participants’ view</td>
</tr>
<tr>
<td>Audrey</td>
<td>81</td>
<td>Research assistant</td>
</tr>
<tr>
<td>Phil</td>
<td>81</td>
<td>Used by Thomas to translate participants’ view</td>
</tr>
<tr>
<td>David</td>
<td>59</td>
<td>Developer</td>
</tr>
<tr>
<td>Pagona</td>
<td>55</td>
<td>Used by a therapist to translate participants’ view</td>
</tr>
<tr>
<td>Penelope</td>
<td>49</td>
<td>Used by a therapist to translate participants’ view</td>
</tr>
<tr>
<td>Tuula</td>
<td>47</td>
<td>Psychologist</td>
</tr>
<tr>
<td>Tony</td>
<td>39</td>
<td>Psychologist</td>
</tr>
</tbody>
</table>

Table 3: Top 12 translators and their translation count

5.2 Interview results

Question that we asked for the evaluation of the artefact:

- We are interested in your impressions from the translation process. What worked well, and what did not work well?

- If the translation process – the software and/or the way you work with it – should be changed, what would be the most important changes?

We had in total six interviews and the choice of having open questions was to ensure that the users spoke freely about the process and tool. Answers were sorted into four categories using an inductive interpretation process. All the text is translated into English by the author.
5.2.1 Effectiveness

Effectiveness is about the quality and fulfilment of the goal. Generally the respondents mentioned that they were happy and could do the translation smoothly. But in some cases they felt that it was hard to translate if the phrase key was unclear.

_The original text [phrase key] can be tricky the thing that you see before someone provided a translation. This has sometimes been incomprehensible. In those cases you don’t know how to translate, but I guess the developers follow a standard we don’t know._ (Tony)

_It can sometimes be hard to translate, you don’t know the meaning of the text. Example: “Mark as read” and “mark as archived” what does that mean?_ (Claire)

_Sometimes you are a bit unsure if the last translation is really the best._ (Thomas)

Several respondents stated that they missed the opportunity to see how the key would have been translated into other languages. It was a good learning experience for the developers to understand that they must be more careful of how they chose the phrase key. Another thing that also came up was that sometimes some keys disappeared and we realised that several phrases were empty.

5.2.2 Efficiency

Efficiency here is defined as the amount of work required to perform translation. By the following statements we can see that translation work was efficient and easy.

_In general, I think that the [translation] tool worked smoothly. I haven’t encountered any bugs or misunderstandings. It’s a very good idea that users can translate on the fly._ (Tony)

Even if many were happy with the tool they suggested a lot of further improvements. For example for some it was sometimes hard to find the phrases and felt that they lacked the support to decide how they would translate. Some of them pointed out that they really wanted to have some kind of function were they can make notes. Another factor that affected the automation of the process is that the translation tool was unable to spellcheck the word.

_You can often go to the same submenu from different places._ Then you
have to change [translate] every submenu separately. It is hard to find all the places before you are used to it [the translation tool] (Audrey)

[...] You often miss translation alternatives. I use a thesaurus while translating. As usual you wish the software could do more work for you, e.g. a search function for text that needs translation. It is easy to miss things, e.g. [...] when you think you translated something [but it isn’t reflected across pages as anticipated]. Maybe you would want to return to things you noted [...] The software does not appear to have any automatic spelling correction, which would reduce simple mistakes. I also miss an undo button (Audrey)

Hard to say who translated. During translation, I am unsure if the translation will only be visible for a psychologist or, for instance, for health staff. It should be more explicit where the translation is visible, and for whom (Claire)

5.2.3 Reliability

How well the users could rely on the tool. There were concerns regarding the reliability of the translation tool and we can see that two factors: (1) Human error by developers when building the pages (2) problems attached to software development in a heterogeneous technical environment - mean the use of browser-side scripting and problems with incompatibilities between web browsers. We focus now on human error. If there was a non-use of Babbler API when creating or updating of pages, user experiences were as follow:

Everything hasn’t been translatable. (Thomas)

Perhaps it is not related to Babbler, but it is a bit frustrating when you see a text in English that is not directly translatable. (Tony)

As a user, one can be frustrated when it doesn’t work. For example if it [text] is pink but still not translatable, then you have to disturb the developers. (Claire)

5.2.4 Workflow

Workflow concerns the collaboration quality in the translation work. The respondents spontaneously brought up the workflow in the translation process, including many cues about a need for a structured workflow. Most of the comments were related to uncertainty about the quality and status of the translation process.
It is unclear who decides what translation to use when there are multiple suggestions. [..] In a way I’m happy that we didn’t get any education, or that we didn’t have eternal discussions about how things should be done. One could get started on the fly. [..] Maybe you should be able to vote for translation so that the most popular one wins (Thomas)

[..] Now we had to make a print screen and send it to the developers. [..] but maybe some solution that made it possible to highlight [..] phrases directly in the software. There is a need for better logistics. Now I couldn’t notice if a word had been made translatable until I accidentally checked it later (Tony)

[..] it feels like [the developers] have many important things to do, so you don’t want to disturb. [..] we need to find a workflow for translation. Sometimes I don’t know who translated. [..] Everybody can translate, different concepts are used [..] (Claire)
6 Evaluation Part II

In this chapter we will analyze findings we saw when we wanted to have a culture specific language.

6.1 Language analysis

The flexibility of having their own languages in each study was experienced as important by users. It was mentioned on several occasions that it is important for the group to feel comfortable in the language - important to use the concepts that users use. Both studies use Swedish, but still there was a desire to have different languages for elderly and teenagers. Therefore, there is a potential to investigate if we could see if the languages were different from each other. To investigate whether there were differences, we took out all the phrases from each language. Then we filtered those phrases (with the same key) that were different from each other. We choose to focus on the 'language SWETEEN (Swedish for teenagers) and SWEADULT (Swedish for adults). Each line of the results was reviewed and we tried to find patterns. To provide a better insight into the size of the data, there were 2155 phrases for each language, 380 of these differed. SWETEEN had 11,681 characters, whereas SWEADULT had 11,299 characters.

What we thought we could find during the analysis (as we already mentioned in chapter 3.2) were:

1. You use more / longer words for elderly
2. You use more personalized calls and texts to the elderly such as using "I", "you" and "here" instead of a direct verbs (eg Send).
3. You use a more informal and colloquial language to younger
4. There is a direct differences in semantics between languages.
5. You use more specific for elderly than for younger (eg send message instead of send)

6.1.1 Some examples of phrases

Table 4 shows some examples of phrases and gives an idea of what kind of differences we’re talking about. The evaluation did not show that there are any major differences between the languages. Small differences that we can see are not an active choice for that target group, they happened spontaneously. Translators use similar expressions in both languages. Both languages vary between using shorter or longer descriptions, however, they are not consistent in the language. There are no indications that any of the languages use more formal or non-informal language. The only things that really differ were in the concept for a function which SWETEEN called ‘diary’ and SWEADULT called it ’logbook’. This was already known since the concept was discussed at a meeting (more on that in Chapter 7.3). Another
finding was that SWEADULT did not use the word 'logbook' consistently. This is something that has been discussed in the context of collaborative translating. Translators find it difficult to produce translations that are consistent [Cintas and Sánchez 2006]. We also found that the translators used 6 different synonyms to say that a questionnaire was answered. Like: (1) submitted (2) answered (3) responded (4) sent in (5) done (6) finished. All the words above were used interchangeably and none of the languages were consistent.

Some of the phrases were totally different. When there was no posts in the diary SWETEEN showed "the diary is blank" while SWEADULT referred to "Here you can see your posts if there have been any". Many of these differences were context-sensitive phrases, meaning that if you saw it in its context you understood why the differences occurred. Other context-sensitive phrases where that in some cases translators translate “view post” while in the other language only wrote “show”. Even if this is nothing new or surprising it can indicate that translation in its context can be important.

What we saw in the previous chapter (Chapter 5) stakeholders suggested improvements to have the chance to mark a word so they can come back to it. This became obvious when we found 8 places where the key was translated into three question marks (???) . This can also indicate that translators found it hard to understand certain keys.
<table>
<thead>
<tr>
<th>Nr</th>
<th>Phrase Key</th>
<th>SWETEEN</th>
<th>SWEADULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>USER-PARTICIPANT ONEACTIVEMODULE</td>
<td>A module in progress.</td>
<td>Active modules:</td>
</tr>
<tr>
<td>2</td>
<td>MONITOR MODERATION G?TOFORUM</td>
<td>Go to forum</td>
<td>Show</td>
</tr>
<tr>
<td>3</td>
<td>MESSAGE CONVERSATION WRITE MESSAGE TO USERNAME</td>
<td>Send message to Username</td>
<td>Write message to Username</td>
</tr>
<tr>
<td>4</td>
<td>MESSAGE REPLYTITLE</td>
<td>Reply to message</td>
<td>Send message</td>
</tr>
<tr>
<td>5</td>
<td>USER OTHERUSER YOUMANOT-COMPLETEDANY-MODULE</td>
<td>You have not completed module.</td>
<td>The participant has yet not complete one module.</td>
</tr>
<tr>
<td>6</td>
<td>SIDEMENUITEM DIARY ADDDIARY</td>
<td>Write in the diary</td>
<td>Write in the log book</td>
</tr>
<tr>
<td>7</td>
<td>YOU CAN TAG ANY USER(S) IN YOUR NOTE TOMAKE THEREMINDER APPEAR WHEN YOU VIEW THAT USER’S PROFILE.</td>
<td>???</td>
<td>You can also choose to add a person of note. The note is then displayed when you enter the profile page of the person you have chosen to add.</td>
</tr>
<tr>
<td>8</td>
<td>DIARY VIEWALL-DIARY DIARYISEMPTY</td>
<td>Here you can read all of your diary entries.</td>
<td>Nothing is written in the logbook.</td>
</tr>
<tr>
<td>9</td>
<td>PATIENT VIEW ONEACTIVEMODULE</td>
<td>An active module - you can choose to activate a new module.</td>
<td>Active modules:</td>
</tr>
</tbody>
</table>

Table 4: Examples of phrases.

7 Discussion

To clarify and simplify for the reader this chapter is divided into several sub chapters. Section 7.1 handles and sorts out the changes that were made after the first evaluation of Babbler - This can be seen as an attempt to show how one can refine and improve an artefact like ours. Section 7.2 compares and sorts out how Babbler differs from similar artefacts. We have chosen to compare Babbler against Facebooks inline-translation because they share many properties that may be of interest. Section 7.3 presents further discoveries made during the analysis of the artefact. This section is important to show how the artefact developed into something that the developers did not originally expect. In the penultimate section 7.4 we list the discoveries made.
during the analysis of phrases. Section 7.5 contains a summarized discussion of our findings.

### 7.1 Improvements after evaluation

From the very start of the project, there were no well-developed processes for managing misreporting. The reasons may be many, but there was a closeness to the developers and the project was not in service for the patients. Therefore, known bugs were reported by email and meetings. These bug reports were then transferred into a word document to be handled by the developers. One of the biggest problems that arose was that users had difficulty providing enough information to fix the errors. For the translation errors you only got a picture - additional information such as time, the role and the URL was missing. If the stakeholder was unable to do a print screen you only got the phrase that was problematic. We have seen in the interviews that stakeholders mention this as a workflow problem. Reporting in this way was ineffective but also crucial since sensitive data should not be transported by mail.

As the project grew in size and more studies started there was a need for processes and support to manage and report bugs. An icon on the left on the screen was implemented and stakeholders only had to click to create a bug report. The platform took a print screen and collected information that was valuable for the developers. Stakeholders could add more comments to the report and send it to the developers. This feature solved three issues regarding (1) workflow, (2) correct information, (3) sending sensitive data. This feature was suitable for translation problems but worked even for other bugs. From a developers perspective – translation problems was now both easy to find and solve.

Improvements were also made to support users in their translation work. For example, introduced spell check and the ability to view previous translations between the languages.

### 7.2 Relating Babbler to Facebook

Facebook has a similar approach to Babbler but was developed in parallel and this has not affected Babblers design. Even do there are similarities they differ in such way that it’s not possible to only apply knowledge from Facebook on Babbler. Facebook uses a big crowd for translation. They are doing translation, they are motivated and voluntary. In our project we have been a small community that were obliged to participate. By this we mean that it was part of theirs job.

Another factor that may have influenced the outcome of the various artefacts is the proximity between the translators. By proximity I mean that Facebooks translators can only communicate by forums and chat. In our
Regarding quality Facebook needed to have a professional translator that checked the translation before publishing since they had incidents were people started to destroy translations for fun. In our project this isn’t needed since there is a stronger sense of engagement for their product. But as we saw in the interviews, there were concerns about the quality so they actually had a person that went through the entire platform before publication, but still the phrase was visible just-in-time. There are no rules in U-Cares project that required you to be a good translator and that you could see that after a while it engaged people to contribute to a better language.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Facebook</th>
<th>Babbler</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goal and purpose</strong></td>
<td>To make a more specific language to the target group</td>
<td>To make a more specific language to the target group and be able to have culture-specific language</td>
</tr>
<tr>
<td><strong>Translator</strong></td>
<td>Human</td>
<td>Human</td>
</tr>
<tr>
<td><strong>Domain</strong></td>
<td>Members open</td>
<td>Staff closed</td>
</tr>
<tr>
<td><strong>Quality control</strong></td>
<td>Voting and professionals</td>
<td>None</td>
</tr>
<tr>
<td><strong>Motivation</strong></td>
<td>High score, be a part of somethign bigger</td>
<td>Contribute to your own product</td>
</tr>
<tr>
<td><strong>Role of professional</strong></td>
<td>Hired translators</td>
<td>None</td>
</tr>
<tr>
<td><strong>Subdivision and degree of context</strong></td>
<td>In context and out of context</td>
<td>In context and Global phrases</td>
</tr>
<tr>
<td><strong>Numbers of translators</strong></td>
<td>About 250,000</td>
<td>About 10</td>
</tr>
<tr>
<td><strong>Opportunity for social interaction</strong></td>
<td>Online by forum</td>
<td>In real life communication</td>
</tr>
</tbody>
</table>

Table 5: Differences and similarity Babbler VS Facebook

After completing the analysis, we realized that users began to use translation as a way to fill the platform with information. For example in table 6 we can see that the key became something totally different. So the translation tool changed meaning from only translation to use the same feature as a Content management system (CMS). As a contrast to a phrase that totally changed table 7 gives an example of a pure translation. Even if we didn’t want them to use this as a CMS, we realized that this right-click feature was easy for them to understand and use. Therefore in some cases when they asked for study specific text we used Babbler as a CMS. This is perhaps not interesting in itself; however, from a design perspective it can be a knowledge that can be useful to know about. We chose to use the fact and they let them continue with this behavior, but if you implement similar functionality you should be aware that it can happen.
Here you can keep notes on what you want, it may involve thoughts and feelings about your disease and treatments. You can at every single post choose if you want it to be private, or if other participants will get to read what you write.

Table 6: Babbler used as an CMS

Below are everyone’s programs (interventions) note that an intervention posted can not be edited.

Table 7: Babbler used for translation

### 7.3 Babbler - a point for discussion

During a translation process there is a need for communication between translators. In U-Care most communication occurred daily in the workplace. Facebook translators needed other ways to create a way for users to communicate – in this case they use a forum. Facebook translators were diligent in communicating and they often have excited discussions about what was right or not. Something that often occurred in Facebook was that phrases became too long and messed up the GUI. They solved this by discussing how they could rename things in a shorter way. However, in U-Care there was never any indication of this and due to the proximity with the developers, translators received direct feedback if their translation caused problems in the GUI.

Other interesting discussions were about what was meant by certain words. We noticed while analyzing translations on Facebook that they had deep discussions about the Swedish word "hen". “Hen” is a gender-neutral word that you can use instead of "he or she" and has been a hot topic in the Swedish media. There were protests against the use of the word since it can upset users and that Sweden wasn’t “ready” for the change. This kind of discussions happened in U-Care but during meetings. If they had a word that was translated back and forth or that a phrase was unclear they decided to hold a meeting to resolve the issue. Table 8 gives an example about a
The function was that the patients were supposed to write small texts with optional subjects. Patients text can be private or public. Even if this specific word by itself is not so important for this thesis, it shows that they are working in the same way as in Facebook. We have also seen that it actually enables the discussion that they can change without any involvement of a developer if they want to change. This is also important for the domain language.

7.4 Intresting observations from language analysis

In this section we describe out findings observed during the data analysis of phrases.
Table 8: Design rationale

<table>
<thead>
<tr>
<th>Suggested word</th>
<th>Design Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log</td>
<td>Could be perceived as patients to log their disease more than that they were free to write things that they themselves found important</td>
</tr>
<tr>
<td>Diary</td>
<td>Diary could in some audiences feel as quite childish and that it felt more like you would write every day as part of the treatment</td>
</tr>
<tr>
<td>Blog</td>
<td>Blog could be perceived as something more public than it needed to be. One can choose which patient to have his journal published for those who were part of the same study, but the blog was able to get patients to believe that it was always the case.</td>
</tr>
</tbody>
</table>

7.4.1 Inconsistency

In the evaluation we found that translators in U-CARE are not consistent while they translate. For patients that are using the platform this can be confusing. Problems arise by having many different words for submitting a survey. One way of solving this could be to use more global keys – but it’s always a risk since translators can lose some flexibility when they translate. When I mean that they are losing flexibility I say that translators sometimes want to be more or less specific on different pages. Sometimes you only need to have “saved” but if the GUI is unclear (for any reason) you maybe need to state even what they save (Save your post). Further on there are also different kinds of opinions If you will use more personalized invitations like “Save your work”, “I agree”. Users are more keen in having a flexibility rather than consistenty. Solving this problem can be tricky but can be reduced if we create better conditions. Better conditions we suggest are (1) searchable phrases (2) marking/save words (3) policy documents.

7.4.2 Unclear phrase keys

Stakeholders that were interviewed stated that it sometimes was difficult to understand the phrase keys – this was also confirmed since we found phrases that were translated into three question marks “???”. This shows that there is a need for developers to be trained or have standards of how the key should be set. The key was generated of the initial phrase but was converted to upper case characters and merged with underscore (“Fill all fields” became FILL_ALL_FIELDS). Developers chose (without restrictions) the key. Often, they set the key as short as possible to avoid having long lines of code. However, it was problematic because it must be clear to the translators what was meant. This was more difficult than we thought. Those interviews were
made before we made some changes regarding reporting bugs, meaning that if they are unable to understand the keys today they would probably report it as a bug.

7.4.3 Variables in phrases

Another problem that arose in connection with the use of variables was that it was hard to understand when you could use variables or not. It often happened that they were trying to use variables where the developer had not defined a variable, and the translator did not understand at all what happened. Although, it was pretty easy to fix by teaching the user how it worked. However, if the original phrase was a variable and translators did not use it so it disappeared and only a developer could discover the mistake. Although I do not intend to implement more in the translation tool, I wanted to give an idea of how this can be solved in a simple way. Suppose that we would have a box in the dialog box as if it were shown a list of variables, and in it was the clickable boxes, and when the user clicks so they are added to the text. Namely, that it is easy to make mistakes by misspelling or it may indicate that the variables are not defined.

Figure 6: Potential solution of problems with variables

7.4.4 Use of Global keys

Babbler was able to handle variables in phrases. This means that you can have a phrase that changed content. Something that happened several times was that the translators had difficulty understanding when they could use variables. A variable was reachable only if the developers implemented variables. For the translator they needed to use brackets and a key for that variable. For those that are aware of using variables, it’s a simple concept
but for those with no technical background it was hard to understand - even for those that we trained. If a translator choose not to use the variable then it was even harder, since the variable became invisible. In those cases the developer needed to take out the variable again. To solve this problem I proposed a solution where each variable is a box that you click into the phrase. Through this, you remove the risk of having translators that misspell the key or variables which disappear.

7.5 Summarizing discussions

We can see from the stakeholder evaluation of the general experience that they use and talk about the artefacts in a good way. We have made improvements after the evaluation and have been taking care of many issues related to the artifact and the process. We can see in several ways that it engages people to participate to the domain language. Even if we can’t find any differences between languages we can see that they feel that it’s important and are motivated to translate. Maybe flexibility and ability are factors that made it successful. While they translate stakeholders learn more about the platform and test it in the same time. Translation during ISD is a form of education. Through testing and refining the process becomes that of creating a domain language.
8 Conclusions

Since this thesis presents different aspects of a major work, I choose to divide this chapter into three different sections depending on what questions it considers. Research questions:

1. How can we design a translation tool to support community translation?

2. In what ways does community translation relate to stakeholder engagement in ISD?

Section 8.1 discusses the first question and talks about design considerations regarding the artefact itself. Section 8.2 will take care of how a translation tool can support and be a part of ISD. Section 8.3 is a reflection on how translators actually worked with translation given the option to create different language variations within a language (i.e. Swedish for teens v. Swedish for adults). In doing so, we provide further insight into how community translation may affect users during ISD. This thesis demonstrates how we can use ideas from community translation in the ISD context. We demonstrate and discuss that translation has a good impact on the ISD by increasing the involvement of users in an ISD. In the DSR knowledge contribution framework (Gregor and Hevner 2013), we characterise our work as an exaptation of knowledge appropriated in a new context. Babbler, a tool for collaborative in-context translation, has been presented and assessed through (i) log analysis that prove that Babbler has been effectively used the tool and (ii) interviews that discuss various qualities of the toll and conditions that need be fulfilled for successful translation.

8.1 Design considerations

Even if we can argue that in-place translation is both easy for developers and translators then is still a need to train developers on how they should set the phrase key and when one should use global or local keys (meaning global translation or local translation). We have also shown that providing translation facilities in ISD has created natural discussions of surrounding concepts and language. If we should suggest specific improvements that have a practical impact on how well a translation tool can be implemented and used, it is primarily a need to reduce human errors. As we have shown, there have been some problems regarding translation caused by the developers. There appears to be a need to train developers to be consistent in showing phrases as Babbles, and develop conventions for phrase-keys. In addition, it makes sense to maintain a shared document with agreed-upon definitions of terms.
including ‘global’ phrase-keys for developers. Although we make proposals to reduce the errors it is only possible to minimize the number of errors, so it is equally important to support the process and report when things go wrong. There is a need for better support for translators. There were several people who complained about (1) they did not know when the phrase was translated globally or locally; (2) it should be easier to see how others translated; (3) there should be integrated spelling correction;

<table>
<thead>
<tr>
<th>Quality</th>
<th>Evaluation results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementability of concept</td>
<td>The in-place translation concept is implementable, as shown through the proof-of-concept (the Bablcer software)</td>
</tr>
<tr>
<td>Software performance</td>
<td>Translators do not mention performance, which indicates that software performance is sufficient. There is a need for more performance evaluation before using the software in a more transaction intensive context.</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>All text elements in the U-CARE portal were translated by users. However, translators expressed uncertainty regarding the quality of translations.</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Translators’ comments indicate that the Bablcer software through its simple design - was efficient to use. There are improvement opportunities, such as better decision support, e.g. spelling correction and more visible info about alternative translations and translation scope</td>
</tr>
<tr>
<td>Software Reliability</td>
<td>The software as such has reached a stable state, but the design concept is prone to human error by developers</td>
</tr>
<tr>
<td>Workflow</td>
<td>The translators self-organized to solve the translation work, but translators express a need for more structured workflow and decision-making</td>
</tr>
</tbody>
</table>

Table 9: Evaluation results

8.2 Translation in ISD

By extending an information system development process with collaborating or community translation we have seen hat it creates a strong of positive desire to translate by users. We have shown that collaborative translation is used and has been of interest for a long time within Facebook, Twitter and Ted talk. We have shown that you don’t need to have a big crowd. Just by seeing that the companies create an individual role that only works with the translation, indicates that it understood the value of the translation
and to trust the company’s resources to have the skills required to make a good translation. Since we noticed during the process of bringing translation inside an ISD that there were no framework in place that helped us to measure the engagement during the development. So to quantify our approach we needed to discover how much time users spent on translation or observation. For now I can only show you examples that they were working with t and also show some issues connected. Many of the discussion occurred in the meetings but also during lunch-time so a full observation must be made to see if we can use a translation process as a part of ISD – or at least say how much it affects ISD. For developers it was easier to develop and to name things in the portal since we knew that they checked and translate the phrase. Since we can see that they changed a lot from the beginning phrase we concludes that it saved us a lot of time. Even if we don’t discuss so deep in this thesis I will say that a translation tool that uses in-place translation is a natural tool for the users to go through the system, both contributing with testing (since they need to set up and try functionality for reaching some of the pages) but also for training the users. While they translated they also found a lot of bugs or issues that may not be marked until later if they did not translated - we are confident that the translation tool is a good way to involve users, also in testing and training for the stakeholders. But we need more research about that to see if they find more bugs during the translation.

Our findings show that the idea of crowd translation may prove useful in ISD. We believe that this research may be useful to other researchers interested in software process improvement and stakeholder-centric design.

8.3 Language variation in practice

As we saw in the previous section, there was no indication that there no practical difference between the languages in the platform. We created the possibility that users felt it was important but that this does not appear in the final result. In contrast, we found some indications that the tool and their process of working with the tool need to be improved in order to create a consistent language on the portal.

So is this a waste? It costs some time to have several different languages in the portal and it costs in performance when you translate to the appropriate language with each download of a page. It has also been costly to create help manuals, I found that users really have understand that it was important to adapt the language to the right audience. Ownership of a language is a strong motivation and hence that it created a strong connection to contribute to "their" language. Viewing it on Facebook and their comments that they find translators effective, as they provide an important knowledge to other facebook users and that there is also a politics Semitic dimension to use (or not use) new words that drive society forward. Since Babbler began with the desire of users interested to see what would happen bu introducing
a translation tool in a group who did not want it from the beginning. It is my belief that given the problems our translators at times had, perhaps the motivation and the final result was affected significantly unless there was a belief that it was important to adapt the language given the target audience who will use the product. We can see that Twitter, TED Talk and Facebook all have in common that they think it is important that the language is not too informal, but it should be just the language users speak. Hence, we can mention that it is important to have a goal-directed language but cannot conclude that it is important to build in support for adapting our example, namely a version for adolescents and one for older people. However, we are using language analysis so that they intervene to provide short exhortations such as "send" to use more expressive commands as saying "send message to X" - this can be a motivation and an opportunity to make users feel that they can affect how the system looks and operates, and a reason enough to build in support for multiple languages in the same language. It does, minimize the number of obstacles and build support for both to report bugs, but also that one should help users to be able to get a consistent language.
References


