Language-specific quality issues in a real world localization process essay

Business

Language-specific quality issues in a real world localization process Master Thesis in Cognitive Science JENNY BORJEL 2007-10-02 ISRN: LIU-KOGVET-D–07/19–SE Abstract This study was made at the localization section at Sony Ericsson Mobile Communications and investigates the part of a localization process where language issues are handled. The purpose of the thesis is to investigate which main factors influence the language quality in a real world localization of a mobile phone. The study is divided into two parts, one initial study and one main study.

In the initial study, the five tasks of the localization process were identified as organization, coordination, technical preparation, creation of source text, and creation of target text. In the main study, the employees' views on language-specific quality issues in the process of localizing mobile phones were mapped out through deep-interviews. The results from the interviews were divided into the categories organization, technical preparation, linguistic issues, and language tools. Organizational prerequisites for achieving high quality were awareness, updated material and structured communication.

The linguistic issues were unambiguous and understandable language, correct language level, standardization of terminology, providing sufficient context to the translators and validation of the target text. Concerning the language tools, both manual and repetitive labor as well as the absence of spell and grammar checkers were two main factors considered to influence the quality of the final text strings in the mobile phone. To contribute to the overall picture of the process, some parts should be further investigated, like the translators' working situation and the end-users' expectations of the

product. To conclude, the most challenging part is determining what the highly subjective term quality means in the context of language, which demands clear guidelines for expected outcomes and especially for whom. Only then can it be determined how it should be done. i Sammanfattning Denna studie utfordes pa Sony Ericsson Mobile Communications och inriktar sig mot den specifika del av lokaliseringsprocessen som hanterar sprakliga problem.

Syftet med uppsatsen ar att i en verklig lokaliseringsprocess undersoka vilka huvudfaktorer som paverkar den sprakliga kvaliteten i den slutgiltiga produkten. I forstudien identifierades de fem huvudsakliga uppgifterna i denna del av processen som organisering, koordinering, teknisk forberedelse, skapande av kalltext och skapande av maltext. I huvudstudien kartlades de anstalldas syn pa sprakliga kvalitetsaspekter i lokaliseringsprocessen genom djupintervjuer. Resultaten delades in i kategorierna organisation, teknisk forberedelse, lingvistiska aspekter och sprakverktyg.

Nodvandiga organisatoriska forutsattningar var: medvetenhet, uppdaterat material och strukturerad kommunikation. Avgorande lingvistiska faktorer var: otvetydigt och forstaeligt sprak, korrekt sprakniva, standardisering av terminologi, tillracklig kontext till oversattarna och validering av maltexten. Vad galler sprakverktygen ansags manuellt och repetitivt arbete samt franvaron av stavnings- och grammatikkontroll vara tva faktorer som paverkar den slutliga kvaliteten hos textstrangarna i en mobiltelefon. Vidare undersokning om oversattarnas arbetssituation och slutanvandarnas

forvantningar pa produkten skulle ge en battre bild av processen som helhet. Den storsta utmaningen ar troligen att faststalla vad begreppet spraklig kvalitet innebar, vilket ar problematiskt och kraver tydliga riktlinjer for vad det ar som ska uppnas och for vem.

Forst darefter kan man tala om hur det ska utforas. ii Acknowledgements Although authored by me, this thesis is not the creation of one person. I want to give my humble thanks to the people supporting me throughout the process. First of all, it wouldn’t even have started if it weren’t for my utors at Sony Ericsson, Peter Warren and Goran Magnusson, who welcomed me and believed in my thesis plan, even with its initial vagueness. I would also like to thank Magnus Merkel, my supervisor at the Department of Computer Science at Linkoping University, who has been supporting my decisions and giving me advice whenever I needed it.

All of you who took part in the interviews and meetings were essential in obtaining the results, both within the sections of UI Localization as well as external actors. Thank you so much for participating in spite of your often busy schedules. Another point of view has been given by my father Ulf Borjel, who has worked as an organizational consultant for several decades, and whose knowledge, comments and insights have been very appreciated, not to say indispensable, by me during this time. I also want to give my appreciation to my companion Magnus Jonsson for his never-ending loving support and my mother Birgitta Borjel, who always knows how to handle my state of mind, whether it’s just stressed out or completely devastated, she puts me back on the right track.

Thank you all! Jenny Borjel September 2007 Lund, Sweden ii Table of Contents Introduction ……

……..

…….

…

…….

……..

…

….

……..

……

…

…..

……..

….

…..

.. – 1 Sony Ericsson Mobile Communications …..

.....

.....

....

.....

.....

.....

.....

....

.....

... – 2 Problem Definition.

....

...

.....

...

.....

.....

.....

....

....

.....

.....

.....

.....

....

.....

.....

.....

. – 4 Delimitations .....

.....

.....

.....

...

.....

....

.....

...

.....

.....

.....

...

.....

...

.....

.....

.....

..... – 4 Outcomes.

.....

.....

...

.....

....

.....

.....

.....

.....

...

.....

.....

.....

.....

....

....

.....

... 5 Scope of the Thesis .

.....

.....

...

.....

.....

.....

....

.....

.....

.....

.....

....

Introduction Localizations are all around us. This very text was written using Microsoft Word, localized into a large number of languages, such as Colombian Spanish, Cherokee, Indian English and Swedish. This word processing software has had its visible natural-language strings translated, its models for letters adapted, its alphabetical orders adjusted, its dialogue boxes widened for the changed lengths, its dates reordered, and so on and so forth. Those are all part of the localization process, where localization is used to reach more and more markets including more and more languages. (Pym 2004a) Over the past decades, localization has progressed from being an added effort by some software publishers to a multi-billion dollar professional industry. Localization, web site globalization, language engineering, and software internationalization have all become important issues for companies that want to market and sell their products in international markets. In many cases, localization has proven to be the key factor for international product acceptance and success. (Esselink 2000) The reaching of international markets depends on how effectively the resources are used. Success requires reducing engineering and technical obstacles, and supplying qualified staff with the right mix of linguistic, engineering, and project-management skills. Brooks 2000) In a practical sense, localization is the adaptation and translation of a text (e. g. in a software program) to suit a particular locale, for example Catalan-speaking computer-users in Spain, i. e. particular recipients in a certain culture (Pym 2004a). The more the user must interact with a product, the more localization it will require (Lommel

and Ray 2007). Consequently, the localization of mobile phones is as crucial as it gets, and needless to say, it is often the end-users who become the victims of poor localization (DiFranco 2006). Localization emerged from the computer software and hardware industries, which makes it easy to understand why the U. S. English is often the source language for these types of products or services. It is in fact the source language in approximately three quarters of the language pairs used in localization. (Lommel and Ray 2007) This is also the case within the products of Sony Ericsson Mobile Communications. -1- Sony Ericsson Mobile Communications The process of translating and localizing text in mobile phones has been part of the Ericsson organization since the 1980's. Since 2001 it is part of Sony Ericsson Mobile Communications (SEMC). Today, they localize their products into over 55 different languages. An increased number of languages and text strings have made the localization more difficult. As a leading player on the market, the demands for high quality results are high, manifested in their organization, where localization has its own department. Organization The specific localization department is subordinated the SEMC organization in the following way: President Product Business Units Product Business GroupPortfolio & Platform Planning sector User Interface Design User Interface Localization & Customization department Figure 1: Location of UI Localization department in the SEMC organization. (Source: Internal documents at http://agora. sonyericsson. net) -2- The UI Localization & Customization department is in turn divided into four sections: User Interface Design UI Design Central/Entry UCD Office UI Localization & Customization UI Localization Central/Entry UI Customization UI Localization lab in Manchester

UI Localization & Support Figure 2: Organization chart of UI Design. Source: Internal Documents, no 120120-LXE10858Uen19) All sections are primarily located in Lund, Sweden, except the UI Localization lab in Manchester, England. Central and Entry are two types of hardware platforms which in turn affect the software platform. Central phones are made of hardware from Ericsson and have the software platform OSE (Operating System Enea), as do some Entry phones. However, most of the Entry phones have other types of hardware and software. UI Localization lab in Manchester develops phones with the operating system Symbian, often made of Ericsson hardware platforms. Another translation vendor is used to translate the text string in these phones. The localization department at SEMC has the overall responsibility for text string localization, language tool development, font supply, operator requirements and other language related issues in the process. They also have their own tool to administrate text strings, the LabelTool suite, developed in cooperation with another company – AU Systems. SEMC is also putting resources into developing their term database, TermBase, a trademark purchased from TRADOS. -3- Problem DefinitionTo begin with, a theoretical understanding of the term localization needs to be investigated. The initial study then explores a part of a localization process in the real world, examining what localization means and which key tasks it includes. It is very difficult to aim for high quality if there is not a clear understanding and a common notion about what it is, which is why the main study starts with examining the notion of quality in this context. It continues with a gathering of which factors in the process have the most influence on the quality of the mobile phone text strings. These

include short terms (“ Save”, “ Try again”) as well as longer help texts available to the end-user. The question is where the largest problems are located, where they arise and how they may be resolved. More precisely, the aim is to map out where the largest obstacles are located and how they are affecting and influencing, direct as well as indirect, the quality of the language in the end product. Narrowing it down, the central question is: Which main factors influence the natural language in a mobile phone in a real world localization process? This question is divided into several clarification questions: 0 0 0 0 0 What does localization mean in the real world? Which tasks are included in a localization process? What is the notion of high quality in this context? From whose perspective should the quality be reviewed? Which main factors influence the quality? The questions will be explored in a qualitative case study with expert interviews, aiming for a quality-centred problem inventory of the issues in a localization process in the real world. The first two will be dealt with in Part I of the report, leaving the final three to be explored in Part II. A glossary defining specific terms and expressions can be found in the end of the report. Delimitations 0 This study will only include one case of a localization process, namely the one performed at SEMC. -4- 0 Only a part of a localization process will be studied, namely the part where most of the natural language issues are handled. 0 Not all actors in the process will participate in the deep-interviews, simply those who perform tasks of interest. No end-user quality survey will be conducted. Outcomes The first part of the study will result in a description of a specific part of the localization process at SEMC, its discrete levels and the tasks included in each step. The main study contains a more thorough

examination of the employees' notion of localization issues, such as how to define quality, how to achieve quality in the process, which advantages and disadvantages the tools they use have and if the communication between actors in the process is satisfying. Scope of the ThesisThe overall aim behind this thesis is to deliver suitable and high quality language to the enduser of a Sony Ericsson mobile phone. For this reason, it is important to reach for a common notion of quality within the organization, concerning text and terms which function as bearers of information. The outcome is supposed to contribute to the process with a theoretical background as well as an awareness of where the process stands today and where it is heading. -5- Method All research questions were explored with a qualitative method and within the framework of a case study. Case studies are an advantageous method when the aim is a description of how a process is carried out in the real world or why something is performed in a certain way. (Yin 1984) The approach of the study is everyday practices rather than an idealized account which suggests interviews as a research method (Preece et al 2002). The data was gathered from several sources, illustrated in the thesis progress in the figure below. Process in Theory Literature Internal documentation Results Quality-centered problem inventory of the issues in the real world localization processAnalysis & Comparison Process in Practice Interviews Tool study Informal sessions Figure 3: Thesis progress from data collection to results. Each type of data collection technique generates a certain kind of information. To use different methods is one way to make sure that different perspectives and issues are being found. (Preece et al 2002) Process in Theory As stated earlier, the initial focus of interest is to obtain a notion

about the process in practice, as opposed to the theoretical models. A theoretical background was achieved through localization literature and internal documents. -6-Literature study A thorough review of the available literature within the domain localization was initially made, before embarking on the actual study. The objective was to obtain a profound understanding about the localization process and its characteristics. The established common and principal issues were studied, as well as the quality assurance and quality control standardizations in the localization business. Internal documentation study This method is used when learning about procedures, regulations and standards of a company. It requires no time commitment from users, which is an advantage. It is however important to be aware of the fact that the procedures in the real world may, and probably will, differ from the documented procedures. (Preece et al 2002) This issue has been carefully considered while studying the internal documents at SEMC. The purpose was to get an understanding about the overall procedures and processes, guiding principles and specifications concerning work tasks, tool requirements etc. Even more interesting for this case were the existing style guides, check lists, results from performed surveys and other documents created to ensure quality of the language in the final product. Process in Practice Since one of the fundamental ambitions in this study is to examine a process in reality, it was important to be physically present at the place of work. To conduct a set of interviews without seeing and experiencing the actual work setting might be deceiving. Not because people tend to be dishonest, but because there might be a risk of discrepancy between the perceived process and the actual process (Kvale 1997). The bias that might

arise from this risk is hopefully avoided by combining the interviews with tool studies and informal sessions with the interviewees. One of the contributions of this study was to try and see through this misconception, not to point out people's possible flaws, but to supply awareness and establish a connection between the perception and the reality. To visualize issues to the people involved makes them easier to deal with; the most common reason for confusion and insecurities about a process is to not know why something is happening (Dicander Alexandersson et al 1998). Tool study When conducting a tool analysis one of the most important things to avoid, although frequently neglected, is ad-hoc-testing. This is principally unstructured testing without a clear goal or course of action. To sit down and simply try things is usually thought upon as -7- something time-saving, but in reality this kind of testing always costs more in the long run. Selecting cases at random is not an effective approach to testing tools, which has been shown both in experiments and from experience. (Fewster and Graham 1999) However, since the main purpose of the tool testing in this case is to reproduce the situation in which the employees work, without any deeper usability perspective, only a few functions were distinguished and tested. The tool analysis was made at an early stage in the study and served as a basis for comprehending the discussions made during the interviews in the latter stage of the study. LabelTool To explore and describe a tool like the LabelTool suite has its difficulties. The users perform very different tasks depending on their job situation; to solely collect descriptions from them was not seen as sufficient. For this reason, it was important to study the tool practically. First of all, the LabelTool suite was created to administrate the

increasing number of text strings in the database. The LabelTool suite contains three applications: LabelTool, TranslatorTool and ValidatorTool. The tool makes it possible to reuse text strings from old projects and makes it easier for the users to categorize, weed out, search, filter, validate and in other ways handle the large amount of text strings. A main feature is the out of rangecheck, i. e. if the text exceeds the actual product's display limits. The picture below shows an example of the LabelTool interface. (See page 31 for more information about LabelTool.) -8- Picture 1: Screendump of the LabelTool interface. (www-d-localizer. com)To obtain a basic understanding about the tool's usability, features and capacities, five major functions were selected and briefly tested. These functions were: 0 0 0 0 0 Open a project and choose a stored translation to this project Search for a specific master text, change and validate it Filter the labels on a specific term, chosen from the termbase Import a translation file into a project Export a project to a translation file Input through informal meetings from the actual users established these specific tasks. TermBase Another important source to study is the TermBase, which includes specific SEMC terms. The aim of the TermBase is to define, collect, and standardize terms that should be used in both the mobile phones as well as in the user manuals. The TermBase is under construction but is available for employees at SEMC and external translators, providing source and target terms, descriptions and possible comments to the terms. -9- The purpose is to get an understanding of how the TermBase works in use, test the search functions, and consider how the information is presented in the term entries. Informal sessions To study a process without including the people involved is destined to fail. A process

materializes in a system with people and tasks; they steer and develop the system. One of the most important issues when improving a process is to make sure the human competence is exploited in an appropriate way. (Johansson 1988) This is why much information about the localization work was gathered through informal sessions, trying to get the most accurate picture as possible of the process. Interviews This method was chosen because of its ability to explore the work situation of the key people on a deep level, assemble their specific subject knowledge, and their opinions about quality issues. An advantage is also the opportunity to explore the reasons behind the respondent's answers with further questioning (Keats 2000). Generally, interviews are classified as structured, unstructured or semi-structured, depending on how strictly the interviewer follows a prepared series of questions (Preece et al 2002). The interviews in this study were conducted in a semi-structured fashion. The reason for this was the importance to both follow a certain series of questions, and to let the interviewees feel free to express themselves if any thoughts emerged during the interview. This was also a way to fulfil the aim of completeness of the study and to guarantee that nothing essential was being neglected because of the possible rigidness of the prepared questions. An interview should not be looked upon as an informal chat, but rather a controlled interaction between a researcher and a respondent. It serves a specific purpose and is created for a specific individual or group of individuals. (Keats 2000) For the interviewee, this human and direct contact with the researcher can make the assignment more agreeable, compared to an impersonal survey or electronic questionnaire. The interaction itself also encourages the respondent to

answer the questions. The researcher may also guide the interviewee to remain focused and give relevant answers. Interviews tend to be conducted face-to-face with one respondent at a time, obtaining one person's individual point of view at a time. (Preece et al 2002) Martin Weisbord, who has great experience in organizational diagnosis, claims that an interview concerning a working process, which he calls a diagnostic interview, can be summed up in these three questions (Johansson 1988): 1. What are you doing here anyway? – 10 – 2. How do you feel about it? 3. What are you going to do about it? This is a method of exploring possible problems in a working process, to be able to make it more efficient (ibid). These are of course simplifications of more profound and deeper-going questions, but they still reflect the essence of the objective. Objective The objective of the interviews is to gather information about how the actual labor is performed in the real world, what perception the interviewees had concerning quality, quality control and customer awareness. The aim is to map out how different factors in the process complicate or facilitate their work in the localization process. Selection of users The respondents must be selected according to certain established criteria, in line with the research plan (Keats 2000). Since a broad range of opinions from several steps in the process wanted to be gathered, eleven people were asked to be interviewed, all with different assignments and perspectives on the localization process. The only requirement was that they were involved in the localization process at SEMC. Two of the interviewees worked outside the company, performing services essential to the process; translation and language tool support. The other respondents had tasks concerning terminology, validation, customization,

language fonts, process management and support. – 11 – Official Title Senior Manager Staff Engineer Engineer Senior Staff Engineer Hourly Worker Task Process management Terminology Terminology Language validation Localization process support (bidirectional languages) Work Experience (in localization business) Company SEMC SEMC SEMC SEMC SEMC 7 years 1, 5 years 5 years 7 years 11 yearsEngineer Development Engineer Senior Staff Engineer Consultant Translator Application manager Language validation Customization validation Language fonts Designer and specification author Translation Development and support of language tool 9 months (5 years in translation) 2, 5 years 23 years 8 years 8 years SEMC SEMC SEMC SEMC Alinoma CyberCom Sweden South Table 1: All participating interviewees. Except for the eleven interviews with key people working with tasks in the core of the process being studied, three additional meetings were arranged to get the view from people working a few steps away from the core activity. Before final approval of the questions, a pilot interview was performed with a localization engineer within the section. Realization There are countless pieces of advices as to how to conduct during an interview to ensure honest and complete answers from the interviewee. Numerous factors may affect the situation negatively; from the concrete location to the more subtle relationship between the actors. (Hager 2001) Several of these factors have been taken into account in the realization of the interviews. The sessions start out following a set of prepared questions, which had been carefully designed to meet the purpose of the problem definition and tested with a pilot interview, which resulted in some rephrasing, before being used in the main study. Since the respondents within the group have different tasks in

the localization process, some of the questions are not relevant in all situations. According to Keats (2000) the relationship between the two participants will quickly worsen if the questions are not seen to be relevant by the respondent. Therefore some of them were cut out, either because they had already been answered, or because they were not relevant to a specific interviewee. A few follow-up questions had been prepared; most of them were posed to clarify or develop previous – 12 – statements by the interviewee. Side-track topics introduced by the respondent were not dismissed unless they were directly irrelevant. (For complete questionnaire, see Appendix) All interviews except one were conducted in the actual work environment of the respondent. The interviewees were therefore assumed to feel safe and confident, thus not holding out on information because of insecurities concerning the immediate setting. All the interviews were recorded with an audio device. Notes were made with the main purpose to facilitate the position of the respondent. The recordings were inspected shortly after the actual interview, a way to structure the data with the session fresh in mind. Analysis For the answers to be correctly analyzed, the questions must be consistent from one interview to another and as free from bias as possible. In interpreting the respondent’s answers, it is important to understand the bias of, for example inconsistency, evasion, inaccuracy, and conceptual difficulty affecting the answers. The answers may also be influenced by the respondents’ emotional states or verbal skills. Needless to say, the interviewer’s verbal and listening skills also have to be taken into account when interpreting the interview data. (Keats 2000) First of all, a summarization of the collected material was made to identify patterns and to

get a general view of the responses. This also made it easier to compare the answers in a standardized format. Categorizations of the data were created, following the results from the initial study. This way, frequently raised issues from the interviewees could be identified and revealed. After this preparation, the information in each category could be interpreted in a controlled way to form the results. (Kvale 1997) Outline of the Report Part I of the report offers a theoretical background concerning the concept of localization and its process in theory. The results in this section present how practitioners within the area choose to define localization and, consequently, which specific localization tasks are handled in this real world process. This is the initial study. Part II of the report opens with a background covering localization quality issues established on a theoretical level, followed by the results of the deep-interviews on the same topic. This is the main study. Each theoretical section is completed with a summary and each result section is followed by a conclusion of the questions at issue, defined above under Problem Definition. 13 – PART I – Localization Process – 14 – Perspectives on Localization The localization industry has been expanding considerably over the past 15 years. In spite of this, the notion about the concept remains within a secluded area of practitioners and their clients. Because of the wide perception of the term, there is no consensus as to what it represents exactly. (Dunne 2006) But localization can not be fully or correctly understood without being contextualized in reference to a number of interdependent processes. These are referred to as a group by the acronym GILT – Globalization, Internationalization, Localization and Translation. (Dunne 2006) One common way to explain the relationship

between these terms is that, in order to globalize, an internationalization of the product must be implemented to make it general, and then the adaptation (localization) to specific locales is possible. Although translation is an essential part of the process, these terms rarely appear in the area of linguistics and they are still far from standardized in the business. (Pym 2004a) TranslationEven though translation could be seen as a general adaptation of anything, as long as it concerns a source culture and a target culture, translation is within the GILT framework seen as just a part of localization. The terms are intertwined in the process and there are evident connections between them. Translation and localization are for example both most visible when neglected and when containing shortages (Fry and Lommel 2003). But localization generally also attends to important non-textual components of products or services in addition to translation (Lommel and Ray 2007). A product or service is well-localized when its origin is not visible, e. g. the material should not be perceived by the user as localized (Esselink 2000; Fry and Lommel 2003). The same principle is valid for translations which are done according to dynamic equivalence. They should function according to the premises of the target locale and the reaction of the perceiver should carefully be taken into account. The content should have the same effect on both the user in the source locale as well as the user in the target locale. In translation studies this is called dynamic equivalence, as opposed to formal equivalence. Ingo 1999) – 15 – However, the old restricted view of translation has been seriously challenged. Today's software companies who translate their products and services often choose to distinguish themselves from the traditional translation business by "

localizing" instead of " translating". (Fry and Lommel 2003) Some think there has been a shift to a more market-driven translation theory which means specific readerships and client-defined purposes are being emphasized and cultural problems are seen as more important than questions of lexis and grammar (Pym 2004a). From a localization perspective, translation means to, on a natural language level, replace the source text string with a target text string, and this is considered to be a relatively small part of the localization process (Fry and Lommel 2003). Localization also generally has stronger emphasis on translation tools and technologies, compared to the traditional translation industry (Esselink 2000). The customization from one market to another is less a transfer from one language to another and more a global adaptation into several markets. GlobalizationLocalization is an integral part of globalization, and without it, other globalization efforts are ineffective. (Lommel and Ray 2007, page 11) The somewhat wide term of globalization includes the business issues and decisions associated with taking a product global. It involves integrating both internationalization and localization throughout a company, as well as having an appropriate sale and support strategy. The term contains thus the change of business and processes to support customers around the world, regardless of which language they prefer, which country they live in, or which culture they belong to. Lommel and Ray 2007; Pym 2004a) Globalization has become more and more of an irreversible process if a company seriously wants to expand. It is no longer an option to not globalize but there is, however, an option to globalize responsibly and in a way that provides value to the customers and pays respects to their local languages and their cultures.

(Lommel and Ray 2007) It can be argued that the globalization helps to restore economic inequalities through letting speakers of less common languages use products and services which otherwise would not be available to them. Rather than imposing English as the global language, it indicates a motivation to form a multilingual world. Through localization a world without exclusion could be possible. (Fry and Lommel 2003) Another definition is to view globalization as the process of separating localizable content from source code and assuring that a product has capacity for foreign-language data (Brooks – 16 – 2000; DiFranco 2006). This means that the goal of globalization is to eliminate the need to completely re-engineer the product for multiple languages simultaneously (DiFranco 2006, page 59). This is however what many others choose to call internationalization, but perhaps the term “ delocalization” is more accurate for what has come to be called internationalization. In reality it has nothing to do with nations and borders, but with locales. The term “ interlocalization” could be even more appropriate, since it would accurately represent the preparation of a product for many features. (Pym 2004b) Nevertheless, the most common term, internationalization, will be used for now. InternationalizationToday’s business climate, with all the more focus on globalization, has made it ever more important to internationalize products. This process includes the technical preparation of a product to be localized. (Fry and Lommel 2003) On a technical level, this means the product has been designed and built to be easily adapted to a specific market after the engineering phase and that it does not need redesigning or reengineering at the time of localization (Lommel and Ray 2007). At an early stage, target language fonts should be

set up and icons should be designed to be interchangeable (Pym 005). On a text level, internationalization means to simplify the source material and remove locale-specific content to reduce the problems later in the process (LISA 2004; Pym 2004a). Although the internationalization problems which could arise have serious consequences, they are usually “ simple” mistakes. For example, when English commands are replaced with translated versions, the number of spaces set aside in the code for the command (i. e. the size of the string buffer) may be insufficient to handle the translated command. This is easy to attend to, but since hundreds of people may work on a piece of software it is almost inevitable that someone makes a mistake. (Brooks 2000) Take the date 11. 12. 03 for example, which has different interpretations in different locales. It is the 11th of December 2003 in British English, November 12th 2003 in American English, and December 3rd 2011 in Swedish. In order to internationalize either of these versions, it is enough to move a level up and generalize with a variable like DATE FIELD. That field can then be formatted in each specific locale. (Pym 2004a) Without proper internationalization, the localization step may in the end be an expensive ordeal (ibid). – 17 – Localization According to the Localization Industry Standards Association localization involves the following issues (Lommel and Ray 2007): 0 0 0 0 Linguistic Physical Business and cultural Technical Aside from linguistic tasks like translation of interfaces, the physical constraints may evoke a form of localization. It can be explained by a car example: in countries with left-hand traffic, the driver seat and steering wheel need to be physically adjusted in the production of the vehicles. The cultural issues include aspects like icons, colours, abbreviations and address formats which

are not used in the same way in all cultures. A fourth characteristic is the technical aspects of localization. (Fry 2003; Pym 2005) These issues often coincide with each other. An example of both a technical and cultural issue is a US produced car racing computer game which needs to be adapted to function on other markets. If “ miles per hour” is to e translated into the Swedish rate of motion (“ kilometer i timmen”), the translation has consequences for both the graphical representation of the speedometer on the screen as well as the source code which needs to be changed, hence the need for a localization process. In short, localization is about combining language and technology to produce a product that can cross cultural and language barriers. (Esselink 2003, referred to in Dunne 2006, page 1) Localization projects include several tasks, such as project management, software engineering, testing, and desktop publishing (Esselink 2000). RolesThe key actor in the process is the manager, serving as a central point of communication for people involved in a localization project (Esselink 2000). In addition, the localization market with its broad range invites several types of actors (Lommel and Ray 2007): 0 0 Clients, creating products and services Localization service providers, providing local-language engineering and linguistic services 0 Consultants of various types, providing expert services – 18 – 0 0 Tool developers, providing technology tools make the process more efficient Academics, providing research in localization-related topics and skillsKey actors in the actual process are executives, developers, authors, translators and foreign language educators. These different areas of expertise require a significant amount of people and a process management that can handle the complexity of the process.

(Dunne 2006) Tasks Localization does not consist of a discrete process or a defined set of tasks, but rather represents a focal point in the corporate matrix at which various business units, objectives, and processes intersect. (Dunne 2006, age 2) However, some typical steps for a localization project sequence have been identified (Esselink 2000): 0 0 0 0 0 0 0 Scheduling and budgeting Analysis of source material Identification and setup of both source and target language terminology Translation of software Engineering and testing of software and online help Processing updates Product quality assurance and delivery These basic steps are not applicable to all projects, but are to be seen as general goals in the localization chain. The workload, duration and productivity levels are influenced by many factors, for example the complexity and size of the source material, the quality of the source material and the tools used. Language tools like translation memories often have a positive effect on both efficiency and productivity in the process, since for new components translated, or as additional products are localized, the larger the database will get, and the more matches will be found automatically. The efficiency also depends on the target language; for Asian languages the process will be longer, because of the complexity of entering and processing Asian characters. (Esselink 2000) Language Tools There are several language tools to use as support in the localization process, with the overall aim to produce better results quicker. Localization activities are often hunted by tight – 19 – deadlines as a result of being last in the process. This time constraint together with its large volumes of material and complicated management means that the industry has always been greatly dependent on language technology tools. (Fry and Lommel 2003) The prediction that

language technologies would replace skilled human translators entirely was made in the early 1950s (Fry and Lommel 2003), when the computational linguistics discipline was progressing rapidly and the faith among researchers was strong, especially in machine translation and computerized tools for translators (Hutchins 1986). In spite of this progress, the machines continued to produce useless results, and the optimism soon diminished. Martin Kay, an esteemed professor in linguistics at Stanford University, wrote in 1980 that even if many tasks in translation are mechanic and repetitive, it is not possible to mechanize the things that are non-mechanical, that is to say the things whose structure we do not yet understand. An example is the human activity and cognitive mystery that takes place when a translation from one natural language to another is made. Kay 1980) It is not a realistic assumption that technology will completely replace human tasks. However the tools can be, with the right design, capable of supporting both the activities of localization and the localization process. According to LISA (the Localization Industry Standards Association) these linguistic tools traditionally have been developed in-house, much because of the requirements of rapid updates and substantial customization to satisfy the individual needs of an organization or its clients. Statistics show that almost 70% of clients and nearly 50% of translation service providers still do some system development internally and these tools have a tendency to be more complex and considerably larger than other tools used by companies who do not have internal development. (Fry and Lommel 2003) One problem is that software publishers have lacked to structure their work and create documentation in a way that would have made the use of translation tools

easier and more cost-effective. Since the multilingual business is increasing and the software publishers are getting more aware of the advantages of high quality localized products, they are now working harder to find ways to reduce their costs and solve their localization problems. (Esselink 2000) Industry experts predict further expansion in this field, much because there is a need for smoothing the process of the integration of stand-alone technologies. Fry and Lommel 2003) What has happened so far is that language technologies have made some of the most monotonous and repetitive tasks automatic which has resulted in, among other things, translation volumes that otherwise would be impossible to provide (Lommel and Ray 2007). Another great advantage in using language tools is that they offer the ability to recall and reuse pre-existing translations, which in turn leads to major cost savings, improved language consistency and higher productivity (Bass 2006). – 20 -Esselink (2000) divides the linguistic resources (or “ translation tools” as he denotes them) into three different types: 0 0 0 Translation memory tools Terminology tools Software localization tools The first two are often combined in a tool set for the translation of documentation and other large pieces of text, called “ Localization Workbenches” by the LISA and referring to tools which combine translation technologies together into a single application. (Lommel and Ray 2007) Software localization tools are used to translate, localize and test software user interfaces, i. e. dialog boxes, menus and messages. Most of these tools combine resource editing, translation memory, reusing, validation, and spell checking functionality. (Esselink 2000) DiFranco (2006) defines the term “ tool” as a reference to any specific piece of software used

in the localization process. He divides the tools into two different categories: CAT (i. e. computer-assisted translation) tools and localization tools. CAT tools are widely used because of the frequency of repetitive content, such as text strings in user interfaces and manuals, that can be reused throughout a product, or a platform of similar products. They support the reuse and recycling of pre-existing translations during the update or renewal of old material. They also ensure the consistency of the localized content. A localization tool includes all the functions of a CAT tool but it also supports the solution of basic technical issues. For example, it supports the resizing of dialog boxes and their controls to have capacity for expansion in languages such as Finnish, German or Indic. Since these languages typically take up 40% more space than English, the localization tool ensures the correct display of all the strings in the localized interface. ibid) All in all, there are more than one way of categorizing the tools for localization that exist on today's market. In this paper, the term " language tool" will refer to all software application such as terminology databases, translation memories and localization tools. Localization Tools A comparison between localization tools and common resource editors shows that the latter have neither statistics, spell checking or validation functionality. Also, each new update must be translated from scratch, since resource editors are without any connection to translation memories. Moreover, localization tools prevent unintentional changes or deletions of interface items or encoding and they contain markers indicating which translated strings have been recycled from previous version, which strings need review and which strings are – 21 – signed-off. Some localization tools also support pseudo-translation tasks, an

important feature to solve problems related to both language and technology. (Esselink 2000) Terminology Management Terminology management is essential for maintaining consistent terminology in a product and across different products and releases (Esselink 2000). During translation and editing, the terminological equivalents can be automatically looked up in a term database which reduces time spent on repetitive tasks. Furthermore, it will save costs through increased consistency among products. (Esselink 2000; Lommel and Ray 2007) The term database TRADOS MULTITERM is an application that allows the user to create, manage and present terminology. Apart from entering equivalents in multiple languages, the user can also enter a variety of user-defined text and attribute fields such as definitions, context, grammatical information, and images. The information can be commented, imported and exported, and shared over a network. (Esselink 2000) Picture 2: An example of the interface of TRADOS MULTITERM database. In working with a terminology database, it is important to be able to have access to information about the situational context for each term. It is often advantageous to know whether a term is a menu item, a dialog box option, or a dialog box title. If the contextual – 22 – reference is unclear, it is recommended to add this type of information manually. (Esselink 2000) SummaryIn sum, localization is more about the processes by which products and services are adapted than about specific tasks, which can vary between different projects (Dunne 2006). To globalize products and develop a completely language-neutral development process is difficult. The technology is constantly developing, and the dynamic natural languages will never stop evolving, which means there will always be

plenty of opportunities for new globalization issues to arise. (DiFranco 2006) The four different concepts presented above are indeed in close relations to each other, but they have different functions. To summarize, the definition of the terms are henceforth perceived in the following way: 0 0 0 0 Globalization Internationalization Localization Translation » » » » concept design process implementation This means localization could be a superordinate to internationalization, since the latter is about the concrete design and the former about the process in which the design is made. In any case, localization is not per se something that comes after internationalization or a subordinate to it. Globalization is more on a theoretical level, and translation is more independent – you can translate without globalize or without internationalizing, although translations are greatly affected by the awareness of the other three concepts, especially in the software localization business. Although difficult to define specific tasks in the process of localization, some general objectives are important to implement: analysis of source material, identification of source and target terminology, translation, technical preparation, product quality assurance and product delivery. The use of language tools such as terminology databases and localization tools, which can handle often large amounts of data, usually have a positive effect on both efficiency and productivity, especially when localizing into many markets. The following results of the case study will present one part of a localization process in the real world, its actors and tasks as well as the key tools involved. – 23 – Process in Practice This specific part of the localization process is currently situated at Sony Ericsson Mobile Communications in the section UI Localization & Support,

subordinated to UI Localization & Customization in the organization. All of the results presented below are either data from internal documents or key people within the organization, if nothing else is stated. Defining Localization The interviewees have quite a unison perspective on localization; it involves not just translation of natural language text strings but also technical functionality, for example the implementation of bidirectional languages. A third factor is the cultural which affects applications like the calendar, the clock, measuring systems and other applications. The technical and cultural localizations are prepared before the actual translation. One interviewee said: “ Localization is the converting of information so the local target culture may understand and comprehend it the same way as the source culture. Translation is the converting of text strings which has to be grammatically correct, but it doesn’t necessarily mean it will be understood in the same way in the target culture. Localization is an adaptation – the language is the core. Another respondent pointed out that localization is about being aware of factors like dialects, fonts and language styles which are different from one language to another and which must be taken into account during the process. The awareness of the importance of localization has during the past years developed into a complex process with many roles and tasks involved. The process is directed by heartbeats (often called platforms), which put the technical restraints on the physical products. A project is divided into milestones which each have specific task deadlines. Every platform triggers a new mother text, i. e. general text strings used in all products and every mother text generates daughter texts, i. e. specific text strings used in certain applications. These English source texts are all called master texts.

Every master text has a description to provide context to the text strings. A master text instance can contain everything from one single term (“ Save”) to a longer help text with several sentences. Source text will be used when referring to both master text and descriptions. They are grouped into applications, for example Contacts or E-mail client, which are called modules in the language tool. The LabelTool suite functions as a core – 24 – through the whole process, being opened and closed to certain actors during a certain amount of time in the process. Tasks For each platform a new project with organizing, coordinating, preparing as well as creating both source and target text is required. The following tasks and roles have been distinguished on the basis of the data analysis. Each role may include several people. Task Organizing Coordinating Role Process manager Function group leader Coordinator Technical preparation Text parameter setup Font supplier Creating source text Source text author Source text validator Terminology specialist Customization validator Creating target text Translator Target text validator Function Supervise the process Coordinate the function group Create project and handle deadlines Set text constraints Provide correct fonts Write source text suggestion Edit source text suggestion Define and determine terms Adapt source text to operator demands Translate from source to target text Review market specific issues in target textTable 2: Tasks and functions of roles significant to the process. Every task except for the actual translation is performed within the same sector at SEMC. Some fonts may also be ordered from an external business partner. All but the text parameter setup, source text author, translator and target text validator work within the same section in the organization. Organizing The process

manager has the role of supervising the process, setting up a localization team, making sure deadlines, development and project plans are followed and reports to the superordinate step in the organization. The task is also to lead the sections in the right direction, according to the company's overall goals, and spread awareness about the importance of localization issues. – 25 – Coordinating The Function Group Leader (FGL) operates as the overall coordinator between the steps both within and outside the section, makes sure the time plan for each project is created and updates the project regularly. This role also includes requesting project resources from the management, such as labour required for each specific project. The function group is dynamic and the members change as the project advance. Different points demand different expertise, one of these points is the source language creation. The FGL is primarily responsible for the tasks undertaken by the function group. Besides the FGL, each project has an assigned coordinator who handles issues concerning the planning of deadlines and communication with other sections within the company. The coordinator is in charge of the database, import and exports files and handles and delegates bug reports, both from internal (e. g. software implementation) and external sources (e. g. endusers). The task includes LabelTool coordination, i. e. pening and closing it for specific users. The coordinator also delivers files to the translation company and handles queries from them. The coordinator's delivery consists of files with translated material, which must be ready for the product's user manual handoff, operator acceptance and software implementation. Technical preparation Introducing natural language text in a digital product like a mobile phone, with a limited display area, requires

careful technical preparation. The goal of the technical preparations is to arrange all necessary prerequisites for every languages of interest (i. e. ultilingual support) which includes font setup with characters according to UNICODE standard. The technical preparation also ensures that all characters are correctly displayed in the tools used. Two main tasks are the setup of text parameters and the fonts. Text parameter setup The text parameter is an obligation and prerequisite for each master text instance in a LabelTool project. It contains a restriction of length and height for the text string. In old mobile phone displays, the limit was easier to set since each character had its own square. Today it depends on fonts, which are proportional, e. g. ne text can be 88 pixels long and 10 pixels high. It also depends on the languages and the size of the current product display. The text parameter is a way to control and ensure that the text strings do not exceed the display limits in the end-product, to avoid toggling and disruptions. One text parameter can be connected to several master texts. – 26 – There are presently 80 general (including common labels, e. g. yes/no) text parameters to choose from, and 56 application-specific (including detailed labels, e. g. “ Oppna linsskyddet”). The general text parameters apply to all languages, both source and target. Any adjustments due to specific target language features are handled later in the process. Font supplier For each new language, new characters might need to be put into the font file connected to LabelTool. The proper displaying, i. e. the rendering, of the text strings requires the necessary fonts for all languages. Certain font types are outsourced. Without fonts, the next step in the process would be futile. Creating source text The creation of the source text

involves not only the actual writing, but also the validation, customization and terminology work which will be explained below. Figure 4 shows the author’s interpretation of the work flow. The source text language is Engli