

Machine Translation and Technicalities of Website Localization

Hamza Ethelb

Faculty of Languages, University of Tripoli
Tripoli, Libya

Abstract

Machine translation tools are widely used by companies. The tools are on an increasing demand. Translators need to equip themselves with the knowledge and the mastering of these tools. This study explores two machine translation tools involved in website localization. These tools are Alchemy Catalyst and Trados Tageditor. The study adopts an evaluative methodology to shed light on the intricacies and technicalities of these two localization tools. It discusses some of the cultural issues that localizers come across in the process of localization. In addition, it delves into the technical issues, mainly focusing on localizing into Arabic with a special focus on string, text, lexis, and orthography. The study concludes that the process of localization requires teamwork and involvement of computer engineers, and both localization tools are valuable in achieving a localization task.

Keywords: Alchemy Catalyst, Arabic, localization, machine translation, Trados

Cites as: Ethelb, H. (2019). Machine Translation and Technicalities of Website Localization. *Arab World English Journal for Translation & Literary Studies*, 3 (3) 107- 112.
DOI: <http://dx.doi.org/10.24093/awejtls/vol3no3.7>

Introduction

Machine translation (MT) and localization are interrelated industries. The developments in the area of machine translation will ultimately reflect positively on the industry of localization and website localization. This is because the localization companies have an increasing interest in offering regulated services in their competitive markets. Machine translation has its sister tool called translation memory (TM). With the development of technology of translation, they are "becoming the one thing" (Pym, 2013, p. 61). Machine translation programs and software work through translation memories to provide consistent translations. Pym (2013) explains that companies have their translation databases that have been built through time with translation memory applications. Due to its productivity and consistency, TM opens the door for companies to work on translation projects of large scale. This is what has made the localization industry flourishes – the usefulness of TM in producing fast translations with consistency.

Website localization widens the scope of communication among cultures and opens new doors for businesses for some commercial websites. According to Pierini (2007, p. 85), there are three types of web localization: informational, promotional, and commercial. Although some web pages may be restricted to only one of these three kinds, other web pages may share more than one of these features. Pierini (2007, p. 86) further elaborates that informational websites are those sites which provide information on particular topics such as tourism. Promotional websites are those which promote specific products: an advertisement on a particular destination in the field of tourism, for example. Commercial sites, on the other hand, are likely to be related to market and business and customer services via the Internet.

Internet surfing is changing, and there are many languages found to serve the needs of the various locales. Such surfing is not restricted to English and those who are interested in the economy and markets will do their best to promote their products on the Net, making them available and accessible by using the language of the target audience (Maroto & Bortoli, 2001). One way of achieving this goal is to *go local* and translate or adapt the web page using a process known as localization.

Localization is a process that is carried out on a particular product (website page) to translate and adapt the content and presentation of that website page. Content includes text and style, while presentation consists of the "graphical" and "technical components" (Maroto & Bortoli, 2001, p. 4). Therefore, it is crucial for the translator (localizer) to have a technical expertise in some computer programs.

This paper is intended to make a critical appraisal of some of the localizing tools and to critically discuss the technical issues involved in the process of localizing websites, using Alchemy Catalyst and Trados Tageditor. It will start with a definition of the term *localization* and will then compare it with *internationalization*. Moreover, some localization concepts will be presented to link them with the intended web site for localization. However, the main discussion will be concerned with the technical issues that occur during the localization operations.

Research Context

This paper sheds light on the technical issues that arise in the process of localizing websites. It adopts an evaluative method to evaluate two localizing tools: Alchemy Catalyst and Trados TagEditor in order to check their effectiveness and which one can be more attractive to users. It is not the intention of the study to examine the translation accuracy of the machine. This is because it is a translation memory, which means the input will have a direct effect on the output. Thus, the evaluation of the tools does not mean which one offers the best translation, nor to compare it to human translation. Instead, it highlights the technicalities of both systems in terms of their usability, applicability, and efficacy.

The study discusses the technicalities of localizing websites into Arabic, along with some of the cultural constraints that appear during the process. The paper handles the steps of carrying out a localization task and mainly focuses on issues pertinent to string, text, lexis, and orthography, along with page formatting, graphics and the tone of the final product.

Localization and Web Localization

The operation of localization is a complex process which involves different factors such as content, communication tone, graphical and scripting components, and page formatting. These web technicalities have to be translated and adapted in a suitable and functional fashion. Pym (2004) defines localization as "taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold" (p. 29). Localization requires teamwork involving translators and computer and programming engineers, as well as a text revision editor, proofreaders and, in some situations, marketing consultants.

Internationalization, on the other hand, is the opposite of localization, and is the process of making a product universal. Pym (2004) defines it as "the process generalizing a product so that it can handle multiple languages and cultural conventions without the need for redesign. Internationalization takes place at the level of program design [in the case of software] and document development" (Pym, 2004, p. 29).

Localization is carried out for several reasons, such as cultural and linguistic restrictions with respect to symbols or images (Pierini, 2007). For example, when the website is technical, the focus needs to be on the technical issues, rather than cultural ones. Merino (2006) argues that the localization process should take into account the target audience of the intended web translation. It is advisable that the audience perceives the product without hindrance. In other words, high quality and efficiency of the localization process should be conformed for the product to be successful.

Translation Software

Alchemy Catalyst is advanced and innovative software that offers achievable solutions for technological obstacles in localization. It also enhances the productivity, quality, and efficiency, as well as the effectiveness of automated translation. The software itself offers compliant solutions for the users of the software, with tutorials and instruction services that facilitate the work of the

localizers and help them to meet the challenges that might occur. It is possible to export what you have translated on Catalyst to other software such as Trados. The translation will then be saved in memory and will provide matches for the rest of the translation required.

On the other hand, Trados is other software that used for translation and website localization. In Trados, a text is not stored as a whole, but rather in segments. The segment or unit usually resembles a sentence. The localization function in Trados is capable of adapting a product – website – to another one that suitable for the market. In addition to translating texts, it can handle graphics and cultural nuances to target the local intended market (Austermuhl, 2014).

The localization Process

Some websites are not complicated and do not contain the features of dynamic webpages, such as texts embedded in graphics, video, or scripts. However, they may contain strings, texts, and links leading to other pages. The first step in localizing a webpage is to save the page in HTML (Hypertext Markup Language) format in the documents file. This will make it possible for Catalyst or Tageditor to allow localization. What follows then is that the webpage needs to be inserted into the localizing tools and will then be ready for working on.

In a particular case of translating a web page into Arabic, one has to be careful with the dialects of the Arabic language (Wilmsen & Youssef, 2009). As the webpage targets an audience, all of whom speak Arabic, it should, therefore, be localized with standard Arabic. However, in other situations of localization such as forums and so on, dialects may be considered when localizing. Likewise, the pictures may remain the same as in case they convey no specific cultural content.

1- Strings

Both localization tools allow translation of the web content by strings. Hughes (2002) defines string as "a sequence of zero or more characters from Unicode character set" (p. 318). In other words, a string is a series of characters usually recognized as a single unit. It may include letters, digits, symbols, a phrase, or a sentence. Each of the strings is translated in turn, usually taking the form of a link which leads to another page. To place the page under the link, the full link will have to be copied with the tags and then pasted onto the translation bar, in order to be rendered into the target language. To activate the link in the target language, the localizer has to work on the language between the tags, keeping the tags as they are, as they contain all the commands and the programming language.

2- Texts

According to Jimenez-Crespo (2011), a web localized text should preserve source text features in terms of macro and microstructure. Some texts in web pages are not generally intended for experts in the field that the website addresses, but rather to the public who have a reasonable interest in particular knowledge and eager to learn new information about specific concepts. Therefore, the nature of the texts can be very flexible, and the localization is achievable working with either tool: Catalyst or Tageditor and the outline (whether macro or micro) of the texts were maintained.

3- Lexis and Orthography

The lexical terms in the original texts of webpages can generally have equivalence in the target language. However, some very technical terms have no equivalence. Two strategies can be applied in this respect: transliteration and adhering to the same spelling. In fact, Jimenez-Crespo (2011) argues that many of what he calls "lexical and syntactic anglicisms" appears to be maintained in texts localized into other languages (p. 17). Therefore, Phrases such as file names and abbreviations should be left in the original spelling. This strategy is followed as many technical terms are known to the audience in the foreign orthography and many feel comfortable with them, rather than foreignizing the terms. For example, 'exe file' is best written in this form rather than distorting the shape into another language in which it may appear difficult to the mind or the eye of a programmer or computer user. More importantly, this type of file name is well-known to the target audience. Both Catalyst and Tageditor managed to deal with the tags which were to be translated.

In the case of Arabic, one serious problem that can difficult to overcome is the orthography and the left-right issue in writing between Arabic and English. It is not possible to re-arrange the overall appearance of the webpage from right to left, but that would not be an obstacle to the audience, as each chunk or phrase of a sentence is readable and well-presented.

Conclusion

Localization is a very complicated process that requires teamwork as the localizer may face some programming issues which may necessitate the assistance of a computer engineer or programmer language (Pierini, 2007, p. 96). It is also not a cheap process and may, indeed, be highly expensive in some cases. Both localization tools have the advantage of presenting the target text and source text of the webpage on the same page of the tool. Catalyst uses WYSIWYG (What You See Is What You Get) mode and appears to be more practical text as it allows immediate editing. It also features the Comparison Expert and Translation Memories, in which project files such as TTK (Translation Toolkit) are stored and made available for future work. On the other hand, Tageditor is also adequate in translating web pages, and it saves the work as 'save target as'. By choosing this, the page is saved and can be viewed as a web page. Moreover, the page can be saved in an 'index.html' form for future work. Some technical issues have been highlighted in the work of translating webpages into Arabic by comparing two localizing tools.

About the Author:

Hamza Ethelb graduated from University of Tripoli in 2006. He obtained his master's degree from Heriot Watt University, Edinburgh, in 2011 with a concentration in Computer-Assisted Translation Tools. He has received his PhD in Translation Studies from the University of Glasgow. His research interests are translation teaching, translation profession, media, news translation and CATT. <https://orcid.org/0000-0002-6075-255X>

References

- Austermuhl, F. (2014). *Electronic Tools for Translators*. London & New York: Routledge.
- Hughes, D. (2002). *Fundamentals of Computer Sciences Using Java*. UK: Jones & Bartlett Publishers, Inc.
- Jimenez-Crespo, M. A. (2011). To Adapt or not to Adapt in Web Localization: a Contrastive Genre-based Study of Original and Localized Sections in Corporate Websites. *Journal of Specialized Translation*, January, (15), 2-27.
- Maroto, J & Bortoli, M. (2001). Web Site Localization. *Proceedings of European Languages and the implementation of Communication and Information Technologies (Elicit)*, University of Paisley, 1-12.
- Merino, M. B. (2006). On the Translation of Video Games. *Journal of Specialized Translation*, July, (6), 22-36.
- Pierini, P. (2007). Quality in Web Translation: An Investigation into UK and Italian Tourism Web Sites. *Journal of Specialized Translation*, (8), 85-103
- Pym, A. (2013). Translation Technology and Translation Theory: Dialogue with Anthony Pym on Localization. *Chinese Translators Journal*, (4), 61-64.
- Pym, A. (2004). *The Moving Text: Localization, Translation and Distribution*. Amsterdam: John Benjamins Co.
- Wilmsen, D. & Youssef, R. O. (2009). Regional Standards and Local Route in Adoption Techniques for Specialised Technologies in the Dialects of Written Arabic. *Journal of Specialized Translation*. January, (11), 191-210