ON INTERNATIONALIZATION (I18N)

Cristian LAKÓ[1]

Abstract

This article examines product and text internationalization within the broader paradigm of GILT (globalization, internationalization, localization, internationalization) processes. It considers perspectives from both the localization industry and academics in translation studies and demonstrates the extent to which theoretical concepts apply to the specificity of website localization and web content translation.

Keywords: internationalization, web content internationalization, text internationalization, GILT processes.

Subsequent to the process of globalization, within the GILT processes, is internationalization (I18n). Schäler defines the term as "[...] the process of designing (or modifying) software so as to enable users to work in the language of their choice (even if the software is not localized) and to isolate the linguistically and culturally dependent parts of an application in preparation for localization." (2009: 158) Software is not the only product that can be internationalized. Dacia Duster, or any other car maker, needs to have its driving wheel on the right-hand side, for instance, if it is to comply with UK driving regulations, and, although there are only two options of placing the driving wheel, it is still about designing the product while minding local rules and regulations. Therefore, internationalization is not only about preparing a product or service for language changes, but also preparing it for seamless integration with the target locales. Inappropriate or partial internationalization, especially if the target culture has already been in contact with similar products or services, can lead to readjustments and extra costs, or even rejection of the products or services. Previous user experience, as part of the reader/user's profile, is an important background factor that can determine how a product or service is perceived. A car with the driving wheel on the left will not sell well in the UK or Japan as it does not comply with what drivers are used to.

On the other hand, in the case of some products, fortunate coincidence helps. For instance, by using keyboard shortcuts such as CTRL+S for saving a document, "S" for "saving" works for the Romanian "salvează" as well. The same happens for CTRL+C ("copy"-"copiază") and CTRL+P ("print"-"printează", increasingly used instead of "tipăreşte" when using a printer). On the other hand, even in English some of the shortcuts do not necessarily indicate the operation to be achieved. "Undo" and "redo" cannot use the "U" and "R" keys as they are used for "underline" and "align right", respectively. On the other hand, the use of CTRL+X for "cutting" is rather iconographic (a welcome tendency especially in the use of gadgets - a triangle used for "play", a square symbol used for "stop", etc.). "Open" and the Romanian "Deschide" use the same

[1] Assistant PhD, "Petru Maior" University of Târgu-Mureş.

combination - CTRL+O. Hence, preparing software for a new market would be rather difficult. In such cases, it is more appropriate to let the user learn the key combinations. The process of preparing a service, in this case keyboard software, can be taken even further. For non-English users you can even choose between using a standard keyboard layout with diacritics and a programmer keyboard layout with or without diacritics. Another argument is the usage of a slightly different keyboard in the UK - an extra £, as compared to the standard US keyboard.

Thus, internationalization is not only about developing a strategy for providing a product in several languages. It also involves preparing a product for local rules and regulations, habits and needs while generating and preserving common communication patterns. Offering various customers from different parts of the world the opportunity to use commonly recognizable icons such as the previously mentioned "play" and "stop" buttons on gadgets represents a plausible solution. Another example is the "Windows" function key on Windows Operating System compatible keyboards, which is now universally known to trigger the Windows menu.

Considering the above remarks, the definition provided by LISA for internationalization found in the MultLingual Magazine encompasses much more accurately the current state of this phenomenon: "... the process of generalizing a product [or service] so that it can handle multiple languages and cultural [and legal] conventions (currency, number separators, dates) without the need for redesign." (LISA cited in Pym 2004a:29) Similarly to globalization, internationalization is rather a managerial decision involving financial and technical aspects.

Pym considers internationalization an "intermediary version". The production of that intermediary version is called "internationalization," and the thing produced is the "internationalized" version." (2014:121) In my opinion, this intermediary version should be based on human-computer interaction findings and any typical locale references be excluded. Pym appropriately names the process of producing an internationalized version delocalization or interlocalization, as all locale traces are to be removed. (2004a:30) Similarly, Cronin speaks of internationalization as a process of accomplishing maximum possible cultural neutrality. (cited in Jiménez-Crespo 2013: 32) From a TS perspective, Pym considers that equivalence between ST, on the one hand, and TTs, on the other, is achieved during the internationalization process: "equivalence [is] created not particularly by one-to- one translation, but by the prior moment of internationalization."(2004a: 31)

Strictly referring to software internationalization, Microsoft uses for the same concept the term localizibility in its documentation for Visual Studio 2005: "An intermediate step prior to localization is testing for localizability. In this step, you ensure that you have separated the application's resources that require translation from the rest of the application's code. "[1]

Apart from separating the resources that require translation, the main technical aspect involved in the internationalization process is to delimit and separate the localizable parts from the whole. In the case of a website, the developer has to establish

variables - references or, more technically called, Uniform Resource Identifier (URI), both for menus, instruction and help files on the one hand, and for content on the other (text strings and media files). HTML5 standards use specific coding page elements that describe in a more appropriate manner what a certain page element includes - <article> for the main part of a page, <nav> for navigation menu, <section> for a more general type of content, <footer> for the footer of the webpage, and so on. For non-text content like images and animation, no text should be fixed, but dynamic. Texts should be displayed over pictures through dynamic code with the content of the text depending on the user's language and locale, controlled from an XML file or, even more so, database driven so that it allows the usage of an administration interface through which non-programmer employees can operate changes.

From a TS perspective on internationalization, the source text is produced in such a way that it can be straightforwardly transferred into the target languages. This is related to the concept of one-to-many: "This is a term for translation processes that go from an internationalized version to many target-language versions simultaneously." (Pym 2014:120)

Internationalization and translation strategies

In this part I will examine several aspects that have been, more or less, overlooked by the academic community. While these issues are important for generally more efficient GILT processes, there are also numerous aspects that need to be considered from the perspective of various translation study theories.

Thus, in order to maintain the same branding yet adapt the marketing strategy, text source content should be written in such a way as to be easily adapted for any target market. By undertaking this properly at this stage can save time and money, and the company will be able to avoid erroneous translations. For example, by using "plain English" for the source text, while simultaneously avoiding non-standard English, the translators may, with less effort, understand and then transpose the message into the target language. Texts that cannot be generalized, such as rhetorical figures, puns, alliteration, rhyme must be avoided. Contentwhich is difficult to understand is difficult to translate. Furthermore, Oracle.com, while referring to user interface labels (buttons, links, etc.), suggests that at the stage of internationalization it is crucial to avoid jargon, idioms, slang, colloquialisms, abbreviations, Latinate abbreviations (i.e., e.g.), contractions, possessives, capitalization (not all languages use capital letters). In the same article they suggest using the plural instead of the singular or adding both the singular and the plural form, and replacing the slash symbol with "or" and "&" with "and" [2]. Clear and full phrases, simple, active-voice sentences, if the genre allows it, are the most efficient both from the perspective of localization and translation, on the one hand, and from the perspective of marketing on the other. If the message is understood by the translator, the translation as a product will have a higher degree of accuracy and the local market will be more receptive to the message. This is a move towards facilitating the comprehension of

the source text, - that Chesterman (1997:92) incorporates under the "comprehension strategies" paradigm; they refer to analyzing and understanding the source text as compared to the "production strategies" which focus on translation asan output. Most companies require that the translations should be made by translators thatare native in the target language. Consequently, in my opinion, more emphasis should be laid on the comprehension process of the source text, as the translator is often not a native speaker of the source language. Comprehension strategies include various sets of activities such as organizational, reading, text-analysis, terminological issues (which in the case of websites is determined by keyword usage), further information retrieval, consulting experts, etc. It is more beneficial to hire translators specialized in the field of activity required by the client, as most of the previously listed issues are also a matter of the translator's experience in a certain field of human activity. Also, if the text is internationalized efficiently, machine translation can be useful as it works much more capably with texts written in a clear, unambiguous language (in a similar manner to technical texts).

Throughout the internationalization stage the internationalization team/individual has to decide upon the global translation strategy (Bell 1998:188). This is the moment when features such as the style of the text, discourse, genre, and so on, that are to be applied to all localized texts in their respective target markets are decided. For instance, if an e-learning website providing math tutoring in the source language addresses users aged 10 to 12, the vocabularyused should be comprehensible for this age group - simple and clear sentences, active voice, friendly tone – even if it uses terms that are specific to math.

On the other hand, if we are to adopt Roman Jakobson's perspective of the verbal sign (1987: 429) - rendering meaning in the same language, into a different language, or into a nonverbal code - during the internationalization stage of the source text, and more generally of the entire website, there are two processes that are used:

- intralingual translation or rewording, to standardize the source text/content (See previous Oracle.com examples above [2]) and

- rendering message by means of nonverbal code or iconization. Common examples on websites are the use of a house icon for the home or start page, several horizontal lines for opening the menu on a mobile device, an envelope for the contact page, and so on.

Yet, another aspect that needs to be determined during the internationalization process, is the translation orientations, in other words whether it is necessary to adopt a domesticating (i.e. source-oriented) or a foreignizing (target-oriented) strategy (Venuti, 1998:240), and to what extent either of these approaches should be used. In the world of global e-commerce while the very term of localization suggests "domesticating strategies", the foreignizing approach would still be more suitable if a company decides that potential clients request information on the origin and the story behind the products they intend to buy. However, this may be dependent on the target market. The foreignization-domestication dichotomy, in the case of e-commerce websites, has more to do with the culture-specific terms, references to location and time, relationship between

communicative partners, measuring conventions, formal conventions, text-type and genre conventions, conventional forms of address, salutation formulas, and structural differences in vocabulary, syntax, and supra-segmental features of the two languages. Often foreignization may be applied only to slogans or advertising campaigns. German cars are much appreciated all over the world as reliable (and on the Romanian market even more so, as there is a frequent distrust about national products). On the other hand, a quick search on the Internet shows that Volkswagen's slogan "Das Auto" is used on most international markets as much as the other no less famous previous slogan, "AusLiebezumAutomobil". This can be useful in the case of other German products as well, as potential buyers would infer that any other German technology related product would rise to the same quality as that of German cars. McDonald's and Pepsi, will often use English words in their slogans or advertisements on international markets. Therefore, when the originof the product is important, the foreignizing approach is often used across the world especially in the case of global brands. This strategy may enjoy even a higher degree of success when the target market suffers from the "country of origin" effect (Pucci et al. 2012:155). Mirroring the same idea, in TS terms, I find appropriate Humboldt's definition of translation: "Translation should indeed have a foreign flavour to it, but only to a certain degree; the line beyond which this clearly becomes an error can easily be drawn. As long as one does not feel the foreigness (Fremdheit) yet does feel the foreign (Fremde), a translation has reached its highest goal; but where foreigness appears as such, and more than likely even obscures the foreign, the translator betrays his inadequacy." (cited in Dimitriu 2002:21)

Standardizing the source text does not imply that the localization and (intratextual) translation- for the same language but different cultural values, should not allow puns, rhyme, culture-bound terms and other text specific operations that the localization process involves. If the standardized source text is English, and in most cases it is, localization will involve changes at text level for the US market, the UK or Australia. This means that localization and personalization depends on the marketing strategy. If the language of a product or service provider is not English, English will become the pivot/bridge language. A pivot language in the machine translation terminology is defined as an intermediary language to achieve translation from one language into another. Thus, in order to translate from Romanian into Chinese, one could opt for English as a pivot language – Romanian would be translated first into English and then the English text into Chinese.

Case study

Bitdefender.com is a Romanian antivirus software developer. Under the language option menu there are listed several languages. One could consider that the most delocalized language variant of the website is WorldWide English (form found on bitdefender.com). This English variant can function as the TS for translation or as a pivot language. All the other language options would be considered localized versions.

However, a more in-depth look reveals that choosing from any of the following options: WorldWide English, United States English, or Canada English, displays the same content. By further analyzing the source code and text on the home page for each of the above mentioned options and the other English speaking countries, it is obvious that the differences are only URL related (.com.au, .co.uk). There is no localization involved, they use an all-around standard version.

The standardization of the text is somewhat similar to the process of pre-editing the source text, so that the pre-edited text is easier to be translated by using machine translation. According to Hutchins, "pre-editing of the input, using a controlled language, or restricting the system to a specific subject domain" (2005: 3) will ensure the speed of the translation and localization process and also ensure a higher quality of the output. Similarly to Oracle.com [2], Hutchins (2005: 8) lists several pre-editing procedures and controlled/standardized source text input best-practices which apply to MT and thus to a more efficient website localization process. Cardey et al., referring to MT, speak about controlled language of the input language, which through extrapolation applies to the source text of the website to be localized, during the internationalization process. (2004: 38) Similarly, Pym considers that "internationalization can make a text simpler, reducing surface-level variation through the use of controlled language." (2014:122) A simplified language can be translated more easily and more efficiently through MT. In this way translators could return the translations more rapidly as presumably only the revision of the MT output would be necessary. Pym also considers during the internationalization step expanding the start text so that it allows a multitude of localization possibilities. (2014:121) However, the localized content would bethen the result of an assimilation or gisting translation approachand the text may beperceived by the translators for each of the target texts in various ways. Also, branding at global level may show deficiency in unity.

While MT is of use to the translation process, oversimplified language of the source text will require extra effort on the part of the translator if the text is to be perceived as appropriate and felt as created in the target language. What makes a language and culture specific are precisely the rhetorical figures, puns, alliterations, rhymes etc. Under these circumstances, what would be the most appropriate approach? An attempt to answer this question through a conceptual analysis is made in Lakó (2014:82-86) and Lakó (2014:229-237).

One last important issue to be elucidated during the internationalization process is to study how the information on the webpages is accessed and read by each of the target nations. There have been numerous studies regarding user interaction with the online medium. One ofthe most important research groups in this area is the nngroup.com, led by Jakob Nielsen. They have hundreds of studies regarding usability, user interface and user interaction. Some of the these findings must be considered during the internationalization process. They relate to reading patterns, content prioritization, types of e-commerce web-users, international usability (effectiveness of user interface used in a

target market different than the one in which user interface was created) (Nielsen 1996), localized website testing, level of Internet maturity among various nations (rapidly leveling over the last years), and many others.

Regarding internationalization, Nielsen found in one of his studies (Nielsen 2011) that, in general, the basics of the interaction with website interface are the same all over the world and founded on human-computer interaction findings, regardless of the country of origin or of their language and cultural background.

Reading from right to left, in the case of Arabic or Hebrew, is only a matter of direction, as the users will still focus their attention more on the beginning of the message, hence, it is a mirrored F-pattern (Nielsen 2006).

Internationalization and text length

The internationalization process stage of the source text in web localization must also consider the various character sets necessary for displaying the multitude of international languages: variants and extension of the Basic Latin alphabet, Cyrillic, Japanese, Chinese, Arabic, Greek, etc. While there are similarities with the keyboard localization issues previously mentioned, there is a necessity at this stage that all possible languages should be accounted for.

In my opinion, very often, even pictures need to be changed depending on the target locale, but for such a case a URI (Uniform Resource Identifier) needs to be set from this stage. For instance, it would look inappropriate to use black people to advertise insurance services, buying electronics or employ them to be part of contextual pictures on a website on the East- European market, where black populations are insignificant. It would not convey consistency with the beliefs and common knowledge of the potential clients; it would appear peculiar, orUS-like. Thus, knowing that a picture must be kept dynamic within a design is as important during the internationalization stage as separating text strings from code. Very often, local celebrities are used to convey to the audience "I am like you and I use it. Follow my example!". Consequently, the spot for an image or a video on a webpage should also be considered from the perspective of the internationalization process as it will be significantly easier to change and test various non-linguistic elements when reaching the localization process.

In website localization web design and the websites' layout must be consistent across the various target languages. There have been numerous studies in Machine Translation and Comparative Corpora showing that the target text (TT) is often longer than the English source text (ST). The length of the text can be measured either by the number of words or by the number of characters. A straightforward free online tool to test the length of ST and TT is Side-by-Side SEO Comparison Tool [3]. For instance, one can compare the text length of the English and Romanian versions of two pages from the EU official website (europa.eu). If we consider languages other than English, then the TT can be longer or shorter. However, in the case of menus and other buttons, there is often a limited space for the text, and the target language may use certain corresponding terms,

from several competing synonyms that are longer than the available space. Synonyms may be used but they may not in all cases be perceived as natural replacements. This is called Text expansion and is defined by Roger Chriss as "an increase in the length of the target text as compared to the source text. […] but must be given due consideration by graphic artists and desktop publishers who want to use the same format or templates for both the source text and translated text. Similarly, text expansion must be taken into consideration when translating software. Since dialog boxes and windows may have to be resized to accommodate the translated text." (2006:202) It is also worthwhile mentioning that, conversely, there is also a possibility for the TT to be shorter for certain language pairs. High-context cultures require less text. It would be useful to conduct a research on verbosity with regards to the target market. If, for example, a low- context culture would require producing 500-word-long textual content, in the case of a high- context culture 300 would suffice. Also, it would be useful to determine if the difference in text length is compensated for with images or other non-textual elements.

Nielsen (2011) also reveals some important data relating to the "degree of verbosity" between languages. Some languages are seemingly more verbose than others. Arabic and Germanwould need more space to display, and this is not a matter of translation but a cultural issue, which means that the translator should be a native speaker of the target language and use a lengthier and wordier text in a natural manner. This might imply copywriting skills as well. There is also a culture specific issue linked to textual length, as there are two main types of communication which also influence the length of text but from a different perspective (cf. Hall 1976: 105-116): high-context (HC) and low-context (LC) cultures. LC cultures use clear, direct and unambiguous communication as well as an argumentative style, whereas HC cultures require less wording but more context. German, for instance is an LC culture, so there is no doubt it requires more display space on webpages. More details regarding the cultural impact on communication, translation and localization in later chapters.

With the ever increasing usage of mobile devices, apart from connection speeds, it is very important to provide content laid out properly according to the screen resolution and size of the device used. Even the buttons used for the menu must be resized to be pressed easily using the tip of the index finger on a 4.52" as compared to pressing buttons of the same menu using the mouse pointer on a 24" desktop monitor. Fortunately, this can be automated if so desired. Through a programming framework (such as Bootstrap - http://getbootstrap.com/) screen resolution, menu position and content layout are adjusted accordingly, maintaining a user-friendly experience across as many screen sizes as possible. This kind of option should be considered as one of the technical aspects of localization. In technical terms, it means applying a responsive design. Responsive design adjusts content layout according to the device on which the information is displayed. Thus, apart from dealing with source content that needs to be displayed correctly on a multitude of devices, textual length in the target language must also be considered during

the stage of internationalization. Thorough planning at this level can save time and money.

Bibliography

Bell, Roger T. (1998). "Psychological / cognitive approaches" in M. Baker (ed.), Routledge Encyclopedia of Translation Studies. London & New York: Routledge, pp. 185-189

Cardey, Sylviane, Peter Greenfieldm, Xiaohong Wu, (2004) "Designing a Controlled Language for the Machine Translation of Medical Protocols: The Case of English to Chinese", in Robert E. Frederking Kathryn B. Taylor (Eds.) Machine Translation: From Real Users to Research, Berlin-Heidelberg: Springer, pp. 37-47

Chesterman, Andrew (1997) Memes of Translation: The Spread of Ideas in Translation Theory. Amsterdam: John Benjamins

Chriss, Roger (2006) Translation as a Profession Morrisville, NC: Lulu Press

Dimitriu, Rodica (2002) Theories and Practice of Translation.Iaşi: Institutul European

Hall, Edward T. (1976) Beyond Culture, New York: Random House

Hutchins, John (2005) "Current commercial machine translation systems and computer-based translation tools: system types and their uses", International Journal of Translation, 17, pp. 5-38. Retrieved from http://www.hutchinsweb.me.uk/IJT-2005.pdf

Jakobson, Roman (1992). "On Linguistic Aspects of Translation" in John Biguenet and Rainer Schulte (eds.). Theories of Translation: An Anthology of Essays from Dryden to Derrida. Chicago and London: The University of Chicago Press, pp. 144-151

Jiménez-Crespo, Miguel A. (2013) Translation and Web Localization London & New York: Routledge

Lakó, Cristian (2014), Localizing Websites: Shifting Focus onto the End-User (Doctoral Dissertation). Available at http://cristianlako.ro/LAKO_PhD_Localizing_Websites_Shifting_Focus_onto_the_End-User.pdf

Nielsen, Jakob (2006, April 17) "F-Shaped Pattern For Reading Web Content", Nielsen Norman Group, retrieved from http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

Nielsen, Jakob (2011, June 6) "International Usability: Big Stuff the Same, Details Differ", Nielsen Norman Group, retrieved from http://www.nngroup.com/articles/international-usability-details-differ/

Pym, Anthony (2004a), The Moving Text. Localization, Translation, and Distribution, Amsterdam, Philadelphia: John Benjamins

Pym, Anthony (2014) Exploring Translation Theories, 2nd edition London & New York: Routledge

Pucci, Tommaso, Christian Simoni and Lorenzo Zanni (2012) "Country of origin effect, brand image and retail management for the exploitation of 'Made in Italy' in China" in

Giuseppe Bertoli and Riccardo Resciniti (Eds.) International Marketing and the Country of Origin Effect. Cheltenham, UK & Northampton, MA: Edward Elgar Publishing
Schäler, Reinhard (2009), "Localization" in Mona Baker and Gabriela Saldanha (eds.) Routledge Encyclopedia of Translation Studies. New York: Routledge, 2nd edition, pp. 157-161
Venuti, L. (1998). "Strategies of translation".In M. Baker (ed.), Encyclopedia of translation studies. London & New York: Routledge, pp. 240-244

Internet sources

[1] http://msdn.microsoft.com/en-us/library/h6270d0z.aspx
[2] http://download.oracle.com/tech/blaf/specs/ssnls.html
[3] http://www.internetmarketingninjas.com/seo-tools/seo-compare/