

## THE WEB LOCALIZATION PROCESS

The web localization process involves a series of steps (Gouadec 2007: 40–43) in which a wide range of professionals can collaborate, such as localization engineers, managers, terminologists, QA operators and localizers-translators. In the industry setting the notion of process differs from that of workflow in that the latter is automated. Normally, industry descriptions of processes and workflows can be considered as prototypes within an organization with large resources devoted to web localization. However, in reality this process varies considerably depending on the nature of the project (marketing website, web-based application, e-commerce site, etc.), the technologies involved, the resources available or even the type of translation procedure. Currently, four main distinct web localization processes can be identified: large corporations, medium and small ones, volunteer-crowdsourcing (see Chapter 9),

and individual localization in which a single agent performs all the tasks. As an example, crowdsourcing approaches require careful planning but the overall configuration of the rest of the project would be quite different from the list below (DePalma and Kelly 2011). This is the workflow model for online volunteer translations by the Center for Next Generation Localization, in which the chunking of texts to accommodate large numbers of volunteers plays a key role. The most commonly described web localization process found in the industry can be subdivided into:

(a) Initial project preparation, project acceptance (performed by localization managers, engineers)

1. Definition of the scope of localization project with clients. Scope requirement collection (including

whether or not the site has been properly internationalized, so that l10n can begin).

2. Setting up the localization environment and managing the process (Dunne and Dunne 2011).
3. Retrieving the site contents and architecture.
4. Analysing the website functions and operation. Website analysis for errors or functionality problems (broken links, missing graphics, lack of uniformity, wrong addresses, etc.).
5. Analysis of third-party components, such as shopping carts, e-commerce platforms, etc.
6. Identifying the adaptations required for the target market according to the localization level. Organizing work specifications and

- the overall planning of the project.
7. Breaking down the website into the different components.
  8. Depending on the localization level, the future architecture of the localized site is defined. Often, not all of the source website is localized (Jiménez-Crespo 2012a). Also, the international site structure might be defined, including domain choices.
  9. The content that will be translated is identified, processed and analysed for estimates. This is often done with automatic systems. Text can be extracted or translated directly on the HTML structure.
  10. The overall distribution of tasks and time estimates is carried out by localization managers (Dunne and Dunne 2011).

11. The localization pack is created with all the necessary files (images, specifications, instructions, proprietary software, etc.), instructions and deliverables.
12. A mirror or clone of the future website structure is created. Folders and files are set up for each new language, and the source files are transferred under their new (local) names.
13. Creating the global gateway. Links are readdressed into the HTML or JavaScript or other files.
14. Similar components are assembled into homogeneous packs consisting of: text, scripts, frames, bars, 'pop-ups' (which appear only in the source for the pages – as tool tips and legends), titles of web pages, sound files, images and

images with embedded texts, and other types (runtimes, databases, etc.).

15. A testing plan is created for the localized website.
16. All the components are set up to be sent to translators (tagging, conversions, formatting, setting up into localization environment, etc.).

(b) Performed by localization specialists or freelance translators:

17. A glossary or termbase with the essential terminology and phraseology is prepared. This step can be assigned to a specialized terminologist (Karsch 2009).

18. The textual components of the websites are localized, including

videos, presentations, etc. This represents the most significant component of localization. As LISA indicated, processing texts is 'the bulk of the localization process' (LISA 2007: 11).

19. Any component that has to be made from scratch or fully adapted for the target locale and is not present in the source website is created and tested. (This is often prepared by the requester of the localization project. A localization or internationalization engineer might be subcontracted to handle the coding work.)

20. All graphics are analysed, adapted as necessary and reprocessed.

21. A new set of keywords and/or description might be prepared to allow

for properly indexing the site in the target locale. This is often referred to as Search Engine Marketing (SEM) and requires a different set of agents.

(c) QA and Integration. Performed by QA specialists, engineers or localization specialists

22. The translation is proofread and checked.

23. The localized components, such as translated text, are reintegrated into the website structure, changing the links (including links to images) within the localized files.

24. All components that were created or fully adapted are integrated and functional quality tested.

25. Functional quality tests are performed.

26. Cosmetic testing is performed.

27. Staged Quality Control. Websites are 'staged' for testing before they go live. Functionality tests are carried out in as many formats and screen and navigator configurations as possible.

28. Any changes are made, and these changes are confirmed and documented.

29. A Web-ready version is created.

30. Cultural acceptability and efficiency is tested. Often an in-country review is performed.

31. Online Quality Control and validation are carried out.

32. Delivery of website to client/posting online.

As seen in this list, the number of agents involved in the process can vary, from a single

person responsible for the entire process to a multiplicity of agents in large organizations: business managers, localization managers, localization engineers, terminologists, localizers, QA operators, freelance translators, etc.

## **Web Localization and Cultural Adaptation**

As we saw in [Chapter 1](#), the critical role of culture and cultural adaptations has been brought to the fore since the early days of localization (Esselink 2001; LISA 2007: 14). This type of adaptation that often appears in localization discourse was nothing new to TS (i.e. Katan 2009), particularly since the emergence of communicative or target-oriented approaches in translation theory. Originally, the emphasis on

cultural adaptations revolved around specific basic issues such as colours, icons/graphics, perception, dates, number and measurement formats, etc. –types of adaptation shared by many other translation types, such as advertising or technical translations. Many other culture-dependent issues were never mentioned explicitly, such as textual structure (Neubert and Shreve 1992), pragmatic differences or genre-specific conventions (Jiménez-Crespo 2009a). This initial emphasis on cultural adaptation soon led to one of the most interesting contradictions in the localization industry, because cost-efficiency considerations favoured a trend running counter to cultural adaptation: the internationalization discourse seeks to neutralize culture-specific features so as to make localization easier, often mentioning the goal of achieving the maximum possible cultural neutrality (Cronin 2003: 18). This has also been referred to as ‘reverse localization’ (Schäler 2008c).

Apart from prescriptive and practical publications, this cultural dimension has been the object of a number of studies both from TS (McDonough 2006a; Tercedor 2005; Schäler 2002) and international marketing perspectives (Singh and Pereira 2005; Singh *et al.* 2004). Scholars have indicated that the goal of the cultural adaptation is not to ‘mislead’ the user into believing that the website is a local production, but rather, to perceive that the company is conscious and respectful of the receiving culture (McDonough 2006a; Yunker 2003: 18). According to the pragmatic- textual and cognitive perspective of Tercedor (2005: 153), four cultural elements are the subject of adaptation:

- Linguistic-textual aspects, such as intertextuality, register or macrotext
- Visual-iconic aspects
- Technical aspects
- Cognitive aspects, such as navigation,

metaphors, mental models or interaction.

These elements relate to a range of culture-determined issues, from cognitive aspects to the visual-iconic ones that are the most commonly found in industry literature. Empirical studies on cultural adaptations have shown that higher degrees of adaptation relate improvements in navigation, interaction and rating of websites. This is because a culturally adapted site requires a lower cognitive effort, and the interaction environment is more efficient and clear (Singh *et al.* 2004).

The role of culture in web localization has been studied from several perspectives. The approach with the highest impact is the international marketing approach of Singh and Pereira. These scholars have extensively researched the role of culture in localization using the dimensions of anthropological psychologist Hofstede (1991). In this model, perception, symbolism and be-

haviour are the key elements that define any culture, and they help establish shared values and structured patterns of behaviour. Their major contribution consists in establishing a framework for studying cultural values that differ between countries, identifying specific website features related to these dimensions that can be quantitatively measured and compared. The variables used in website design relate to Hofstede's behaviour dimension: <sup>5</sup>

- Individualism-Collectivism: Related to self-perception as an individual or as part of a group or collective.
- Power distance: Related to the acceptance and expectations of unequal distribution of power.
- Uncertainty Avoidance: The importance of predictability, structure and order versus the willingness to take risks and accept ambiguity and limited structure.

- Masculinity-Femininity: The importance of achievements versus personal relationships.
- Low-High Context: The importance each culture assigns to the context as opposed to the message. <sup>6</sup>

These five dimensions are identified and linked to certain elements in websites that, used as quantifiable variables, allow comparisons of cultural differences between websites from different regions or countries. As expected, stark differences in these values emerge. For example, the United States and Australia rank high on the individualism-collectivism dimension, while most Latin American countries or Indonesia rank very low. The cultural adaptations that the authors recommend in high-collectivism cases are to enhance community relations, chats, add family themes; in the opposite case good privacy policies or personalization might be more

effective. In the case of uncertainty avoidance, Greece, Portugal or Japan rank high, while Singapore, Sweden or the United States rank very low. In this case, uncertainty avoidance can be controlled with customer service, guided navigation and testimonials.

The studies carried out by Singh use two possible methodologies: comparing original sites in every country and comparing degrees of adaptation in web localization to the specific values of the target countries. Using these cultural dimensions to quantify the degree to which websites are adapted in localization, the scholars proposed the notion of 'localization level', a notion clearly related to the monetary and time resources that are or can be devoted to any localization project.

## **Localization Levels and**

## **Cultural Adaptations**

Localization is clearly constrained by limited time, human and economic resources. Hence the localization level, or the extent to which the website is adapted to the receiving culture, normally depends on the importance or size of the local market or audience (Brooks 2000). The notion of localization level was defined in the context of software localization as:

The amount of translation and customization necessary to create different language editions. The levels, which are determined by balancing risk and return, range from translating nothing to shipping a completely translated product with customized features.

(Microsoft Corporation 2003: 15)

In a market environment such as localization, decisions about the localization level normally

depend on Return on Investment (ROI) issues: whether the potential benefits of the localization process outweigh the initial investment needed to produce the localized version(s). It is up to commissioners or initiators to request a specific localization level that can be set out in the localization commission or brief. The overall localization process therefore depends on, and is constrained by, the resources and guidelines laid out by the commissioners. In practice, web localization processes vary widely, from simply translating a small text box with contact info right up to a fully localized website.

The first mention of localization levels can be attributed to Microsoft (Brooks 2000: 49–50) and distinguished three distinct levels upon which Windows operating systems were localized:

1. Enabled products: those in which users can write and use their own language

and scripts, but the software and the accompanying help and guides appear in a different language.

2. Localized products: those in which the user interface and all help files are localized, but some language-specific tools such as spell-checkers and dictionaries are not available.
3. Adapted products: those in which all linguistic tools, functionalities and content are fully adapted to the target language/locale.

Initially, this classification was also applied to web localization processes, although only the second and third levels were applicable to web environments. The differences between software products and websites soon led scholars to propose different categorizations based on industry approaches (Yunker 2003), cultural studies applied to web design (Singh and Pereira

2005: 10–15) or Translation Studies (Jiménez-Crespo 2012a, 2012b). These three different proposals could be considered complementary, as they can offer different bases for empirical studies of the strategies surrounding localization practices. The categorization proposed by Singh and Pereira (2005), primarily based on the role of cultural adaptations, distinguishes five distinct levels:

1. Standardized websites: in which a multinational company simply offers a site in one language for all countries/markets.
2. Semi-localized websites: in which the only locale/specific content is a contact page in the target language with information about local branches, contacts, etc.
3. Localized websites: in which most content and pages are localized, but the original functionalities and back end are

not modified.

4. Extensively localized websites: in which there is a global localization and all content and site structure/functionalities are fully adapted to the target locale.
5. Culturally adapted websites. This is the most advanced level of localization, the one that the authors advocate, and in which there is a total immersion in the target locale. Sites are adapted to the levels of cultural descriptions proposed by Hofstede (1991): perception, symbolism and behaviour.

These different levels of adaptation entail different degrees of re-engineering of the deep structure of the website, the hidden structure that contains the programming and tagging. Normally, web localization operates on the structure that the user sees, the visual (Mata 2005), front-end (Cronin 2003) or surface

structure (Kersten *et al.* 2002), while higher localization levels also require adaptations and re-engineering in the underlying structure or deep structure. Lower localization levels only require the translation of the surface structure by means of replacing the textual strings in the website. Often, websites are not fully adapted to the receiving culture due to cost considerations, and hence, as Singh and Pereira (2005) point out, very few websites are fully localized to the highest level; the only example close to this level of adaptation was the IKEA website. Their categorization has been widely used for research studies into web localization. However, it cannot cover all possible cases, such as some localization processes undertaken from non-economic motives (volunteer translation, crowdsourcing, non-profit websites, self localization of personal websites, for instance). In Jiménez-Crespo (2012a, 2012b), I proposed a localization-level model derived from studying

the web presence of almost 2000 non-profit organizations in the US and their web strategies for disseminating information. These websites cannot be strictly understood in terms of resources available for localization or ROI issues, and, obviously, the localization strategies of non-profit organizations diverge considerably from corporate websites. The proposed categorization includes a 0 level for websites that included localized documents in .pdf or .doc format as well as machine translation, given that the organization does at least acknowledge the need for translation, even if this cannot be considered localization. This appears as a recurrent option for disseminating information (Gaspari 2007) in cases of economic or human-resource constraints within which these organizations operate. The categorization can be described as follows:

1. Level 0: Website offers translated .pdf

documents or MT engine links.

2. Level 1: Website offers a paragraph or page in a different language. Normally it is a brief description of the organization and basic contact information.
3. Level 2: Several localized web pages appear. All navigation menus are in English.
4. Level 3: Website offers several localized web pages with at least one navigation menu in the target language.
5. Level 4: Fully localized mirror website.

In any case, it should be mentioned that levels 0 and 1 might not be considered web localization *per se*, as they might not be cases of localized web content, but rather the posting of printed translations, or else simply writing the contact information from scratch. In these levels, translators might not work directly with any source texts or even adapted ones.

Finally, another interesting strategy that



determines the localization level in business scenarios is the centralized/decentralized model (Yunker 2003; O'Hagan and Ashworth 2003: 74). In centralized models the web localization process is controlled from a central location and stored in a common repository. The decentralized model implies offering a common 'shell' or visual structure for the sites, with the actual local websites controlled and produced in each country, often mixing localized and local content, but also creating a new full website from scratch.<sup>7</sup>

## **Localization and Web Usability**

The target-oriented perspective of web localization is closely related to the objectives of

web usability, which examines the reception of websites by means of empirical studies whose findings result in guidelines for web development. These guidelines are intended to improve user interaction, leading to higher user satisfaction and quality perception. Research focuses, from a cognitive perspective, on the basic patterns of interaction that guide relationships between users and websites and how websites are processed (Nielsen and Pernice 2010; Nielsen and Loranger 2006; Nielsen and Tahir 2002; Adkisson 2002; Brinck *et al.* 2002; Krug 2006). This type of research emerged from the challenges that new interactive on-screen hypertexts posed for developers and web users. Usability in general can be defined as 'a quality attribute that assesses how easy user interfaces are to use . . . [it] also refers to methods for improving ease-of-use during the design process' (Nielsen 2003).<sup>8</sup> It comprises five main dimensions: learnability, efficiency, memorabil-

ity, errors and user satisfaction.

The main basic premise behind usability is that on-screen texts are processed differently from printed ones. Research has shown that reading slows down by 25% to 50%, and users do not read web texts but rather scan the pages in search of the information that might draw their attention (Nielsen 2001: 101). If they find an item of interest, they focus on it and process it further. Since patterns of cognitive interaction with these on-screen multimodal texts are different, one of the goals of usability is therefore to research how best to adapt online texts to the new medium and screen presentation.

One important contribution of usability is to shift the focus from the static concept of readers as passive recipients of information<sup>9</sup> to 'users' who actively engage and interact with texts, charting their own reading path. This is one of the main differences between interactive digital texts and printed texts; the latter being

supposedly 'read' but not necessarily 'used'. In eye-tracking usability studies it is interesting to observe how users visually interact with a web page (Nielsen and Pernice 2010). The implications for web design are manifold, but the most critical is the fact that website success is measured by its so called 'stickiness':

[O]ne key benchmark of Web success is stickiness, the ability to attract new and repeat visitors and keep them on a site.

(LISA 2004: 35)

In web environments, users normally leave a web page or websites if some elements – texts, design, interaction – are too complex to process cognitively, moving on to search for similar information somewhere else (Nielsen and Loranger 2006).<sup>10</sup> The implications for web localization are clear: localized sites should be as clear, concise and efficient as possible.

The significance of web usability has been ac-

knowledged in web localization research (Pym and Windle 2011a; Jiménez-Crespo 2009a). For example, Pym and Windle (2011a) remind us that, as users scan pages, texts should be separated during localization according to their degree of risk within the site. Recently industry practices have been adopting this approach, for example differentiating texts within localization according to 'user sentiment' (O'Brien 2012) or creative segments that might require a different treatment. Pym focuses mostly on the significance of structural and design elements in usability research. However, despite the significance of these elements, translators are normally not in charge of any usability changes in the design or visual components. Usability publications, however, do offer guidelines on web writing styles that increase the usability of the site (Jiménez-Crespo 2011e), encouraging clear, concise and unambiguous text writing. The significance of good web writing style was

recognized in usability from the start:

Plain text is the foundation of most web information.

(Nielsen and Tahir 2002: 48)

Effective content writing in one of the most critical aspects of all web design.

(Nielsen and Tahir 2002: 14)

With these statements, usability guru Nielsen rated the importance of text production on a par with other elements of websites. Localizers are directly responsible for text production, so a sound knowledge of writing styles for on-screen reading is part of what has been called 'professional localization competence' (Jiménez-Crespo and Tercedor 2010). Web style guidelines developed by usability researchers therefore represent a key element for anyone involved in web localization (Jiménez-Crespo 2011e).

Another aspect that relates translation theo-

ries and usability research is the role of conventions. Functionalist approaches to translation (Reiss and Vermeer 1984; Nord 1997) highlight the replacement of source cultural conventions with target cultural conventions in instrumental translation as a key element of quality in translation. Similarly, the commonest mantra in usability publications is to follow established conventions at all levels, and some publications focus exclusively on this issue, for example Krug (2006). Users approach new websites, original or localized, with a conventional generic model that guides the interaction whenever they encounter anything new in the digital genre:

by the time a user arrives at your homepage for the first time, that user will already be carrying a large load of mental baggage, accumulated from prior visits to thousands of other homepages . . . by this time, users have accumulated a generic mental model of the way homepages are supposed to work, based on their experiences on these other sites.

(Nielsen and Tahir 2002: 37)

This generic mental model represents the matrix of expectations that guides the cognitive processing of the text, as happens with any other reading process. The underlying premise here is that users of websites have a lower tolerance of uncertainty, and presenting familiar or conventional features reduces the cognitive load needed to process web information (Nielsen and Loranger 2006; Spyridakis 2000). Some empirical studies have confirmed that following conventions has a clear effect on users' interactions with websites, and it has been proved that following structural, textual, lexical and pragmatic conventions improves comprehension, usability recall, satisfaction and navigation (Vaughan and Dillon 2006). However, different empirical studies have shown that professionally localized sites tend to not comply with the conventions found in spontaneously

translated or non-translated websites (i.e. Jiménez-Crespo 2009a).

## **Summary**

This chapter has outlined the global cycle of web localization within the larger GILT (Globalization, Internationalization, Localization and Translation) paradigm. A prototypical approach was adopted, as not all website localization processes follow the mainstream approach of large corporations (i.e. a small non-profit website). The overall web localization process was broken down into its constituents, and the main issues affecting the web localization process were discussed, such as localization levels, cultural adaptation, the communicative process and web usability.

## **Further Reading**

For an overview of the GILT cycle see Dunne (2006a) or Cadieux and Esselink (2002). Jiménez-Crespo (2010b) provides a critical overview of the impact of internationalization strategies. All the previously mentioned descriptive professional manuals on localization provide a breakdown of tasks during localization (Esselink 2001; Yunker 2003), as well as in Gouadec (2007: 38–45). For cultural adaptation in localization see Tercedor (2005), McDonough (2006a), Schäler (2002) and Singh and Pereira (2005). For localization levels and strategies see Brooks (2000), Singh and Pereira (2005: 10–15) and Jiménez-Crespo (2012a). See Karsch (2009) for a typical terminology process within localization or Sikes (2011) for the role of localiza-

tion managers. For the communicative context in which websites operate, see Janoschka (2003) and O'Hagan and Ashworth (2003). There is a massive amount of publications on web usability, some basic ones are Nielsen (2001), Nielsen and Loranger (2006), Nielsen and Pernice (2010) and Krug (2006). Jacob Nielsen's website [www.useit.com](http://www.useit.com) is an excellent resource for all types of web usability research.